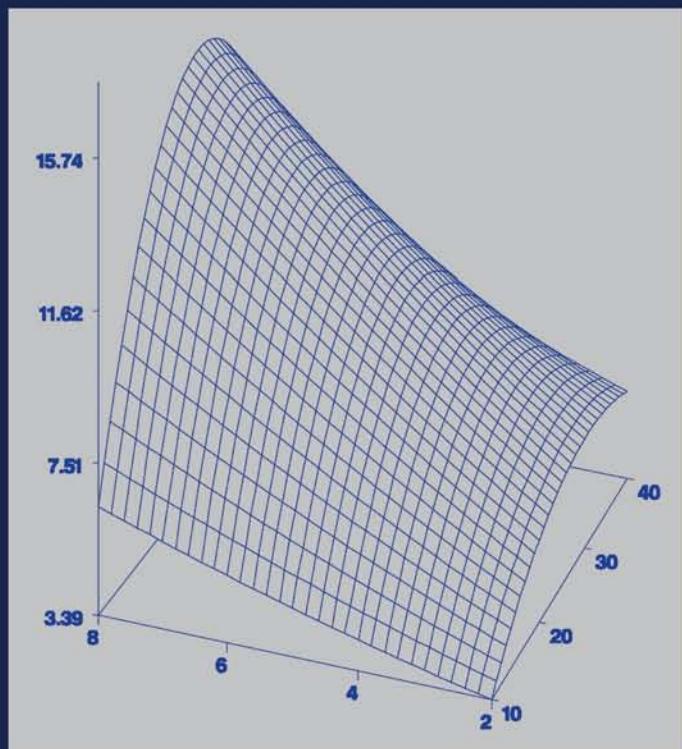


SECOND EDITION

ANALYSIS OF MESSY DATA

VOLUME 1

DESIGNED EXPERIMENTS



George A. Milliken
Dallas E. Johnson



CRC Press
Taylor & Francis Group

A CHAPMAN & HALL BOOK

**ANALYSIS
OF MESSY DATA
VOLUME 1
DESIGNED EXPERIMENTS**

SECOND EDITION

ANALYSIS OF MESSY DATA

VOLUME 1

DESIGNED EXPERIMENTS

SECOND EDITION

**George A. Milliken
Dallas E. Johnson**



CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A CHAPMAN & HALL BOOK

Chapman & Hall/CRC
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2009 by Taylor & Francis Group, LLC
Chapman & Hall/CRC is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works
Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-13: 978-1-58488-334-0 (Hardcover)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data

Milliken, George A., 1943-
Analysis of messy data / by George A. Milliken, Dallas E. Johnson. -- 2nd ed.
p. cm.
Includes bibliographical references and index.
Contents: v. 1. Designed experiments
ISBN-13: 978-1-58488-334-0 (v. 1)
ISBN-10: 1-58488-334-0 (v. 1)
1. Analysis of variance. 2. Experimental design. 3. Sampling (Statistics) I. Johnson, Dallas E., 1938- II. Title.

QA279.M48 2008
519.5'38--dc22

2008045111

Visit the Taylor & Francis Web site at
<http://www.taylorandfrancis.com>

and the CRC Press Web site at
<http://www.crcpress.com>

Contents

1. The Simplest Case: One-Way Treatment Structure in a Completely Randomized Design Structure with Homogeneous Errors	1
1.1 Model Definitions and Assumptions	1
1.2 Parameter Estimation	2
1.3 Inferences on Linear Combinations—Tests and Confidence Intervals	4
1.4 Example—Tasks and Pulse Rate	5
1.5 Simultaneous Tests on Several Linear Combinations	7
1.6 Example—Tasks and Pulse Rate (Continued)	9
1.7 Testing the Equality of All Means	11
1.8 Example—Tasks and Pulse Rate (Continued)	12
1.9 General Method for Comparing Two Models—The Principle of Conditional Error	13
1.10 Example—Tasks and Pulse Rate (Continued)	15
1.11 Computer Analyses	15
1.12 Concluding Remarks	16
1.13 Exercises	17
2. One-Way Treatment Structure in a Completely Randomized Design Structure with Heterogeneous Errors	21
2.1 Model Definitions and Assumptions	22
2.2 Parameter Estimation	22
2.3 Tests for Homogeneity of Variances	23
2.3.1 Hartley's <i>F</i> -Max Test	23
2.3.2 Bartlett's Test	24
2.3.3 Levene's Test	24
2.3.4 Brown and Forsythe's Test	24
2.3.5 O'Brien's Test	25
2.3.6 Some Recommendations	25
2.4 Example—Drugs and Errors	25
2.5 Inferences on Linear Combinations	29
2.6 Example—Drugs and Errors (Continued)	31

2.7	General Satterthwaite Approximation for Degrees of Freedom	33
2.8	Comparing All Means	34
2.9	Concluding Remarks	38
2.10	Exercises	39
3.	Simultaneous Inference Procedures and Multiple Comparisons	43
3.1	Error Rates	44
3.2	Recommendations	45
3.3	Least Significant Difference	46
3.4	Fisher's LSD Procedure	47
3.5	Bonferroni's Method	48
3.6	Scheffé's Procedure	48
3.7	Tukey–Kramer Method	49
3.8	Simulation Methods	50
3.9	Šidák Procedure	51
3.10	Example—Pairwise Comparisons	51
3.11	Dunnett's Procedure	53
3.12	Example—Comparing with a Control	54
3.13	Multivariate t	55
3.14	Example—Linearly Independent Comparisons	56
3.15	Sequential Rejective Methods	57
3.15.1	Bonferroni–Holm Method	58
3.15.2	Šidák–Holm Method	58
3.15.3	Benjamini and Hochberg Method to Control FDR	58
3.16	Example—Linearly Dependent Comparisons	59
3.17	Multiple Range Tests	61
3.17.1	Student–Newman–Keul's Method	61
3.17.2	Duncan's New Multiple Range Method	64
3.18	Waller–Duncan Procedure	65
3.19	Example—Multiple Range for Pairwise Comparisons	66
3.20	A Caution	68
3.21	Concluding Remarks	69
3.22	Exercises	69
4.	Basics for Designing Experiments	71
4.1	Introducing Basic Ideas	71
4.2	Structures of a Designed Experiment	77
4.2.1	Types of Design Structures	78
4.2.2	Types of Treatment Structures	80
4.3	Examples of Different Designed Experiments	83
4.3.1	Example 4.1: Diets	83
4.3.2	Example 4.2: House Paint	88
4.3.3	Example 4.3: Steel Plates	90
4.3.4	Example 4.4: Levels of N and K	92
4.3.5	Example 4.5: Blocks and Replications	94
4.3.6	Example 4.6: Row and Column Blocks	97
4.4	Concluding Remarks	98
4.5	Exercises	98

5. Multilevel Designs: Split-Plots, Strip-Plots, Repeated Measures and Combinations	101
5.1 Identifying Sizes of Experimental Units—Four Basic Design Structures	101
5.2 Hierarchical Design: A Multilevel Design Structure	114
5.3 Split-Plot Design Structures: Two-Level Design Structures	115
5.3.1 Example 5.1: Cooking Beans—The Simplest Split-Plot or Two-Level Design Structure	116
5.3.2 Example 5.2: Grinding Wheat—The Usual Split-Plot or Two-Level Design Structure	120
5.3.3 Example 5.3: Baking Bread—Split-Plot with Incomplete Block Design Structure	122
5.3.4 Example 5.4: Meat in Display Case—A Complex Split-Plot or Four-Level Design	125
5.4 Strip-Plot Design Structures—A Nonhierarchical Multilevel Design	131
5.4.1 Example 5.5: Making Cheese	131
5.5 Repeated Measures Designs	135
5.5.1 Example 5.6: Horse Feet—Basic Repeated Measures Design	136
5.5.2 Example 5.7: Comfort Study—Repeated Measures Design	139
5.5.3 Example 5.8: Crossover or Change-Over Designs	141
5.6 Designs Involving Nested Factors	142
5.6.1 Example 5.9: Animal Genetics	143
5.6.2 Example 5.10: Soybeans in Maturity Groups	143
5.6.3 Example 5.11: Engines on Aircraft	145
5.6.4 Example 5.12: Simple Comfort Experiment	147
5.6.5 Example 5.13: Multilocation Study with Repeated Measures	148
5.7 Concluding Remarks	150
5.8 Exercises	151
6. Matrix Form of the Model	155
6.1 Basic Notation	155
6.1.1 Simple Linear Regression Model	156
6.1.2 One-Way Treatment Structure Model	156
6.1.3 Two-Way Treatment Structure Model	158
6.1.4 Example 6.1: Means Model for Two-Way Treatment Structure	159
6.2 Least Squares Estimation	161
6.2.1 Least Squares Equations	162
6.2.2 Sum-to-Zero Restrictions	165
6.2.3 Set-to-Zero Restrictions	166
6.2.4 Example 6.2: A One-Way Treatment Structure	167
6.3 Estimability and Connected Designs	169
6.3.1 Estimable Functions	170
6.3.2 Connectedness	171
6.4 Testing Hypotheses about Linear Model Parameters	171
6.5 Population Marginal Means	173
6.6 Concluding Remarks	179
6.7 Exercises	179

7. Balanced Two-Way Treatment Structures	181
7.1 Model Definition and Assumptions	181
7.1.1 Means Model	181
7.1.2 Effects Model	182
7.2 Parameter Estimation	182
7.3 Interactions and Their Importance	183
7.4 Main Effects	183
7.5 Computer Analyses	184
7.6 Concluding Remarks	184
7.7 Exercises	184
8. Case Study: Complete Analyses of Balanced Two-Way Experiments	187
8.1 Contrasts of Main Effect Means	187
8.2 Contrasts of Interaction Effects	188
8.3 Paint-Paving Example	189
8.4 Analyzing Quantitative Treatment Factors	193
8.5 Multiple Comparisons	195
8.6 Concluding Remarks	195
8.7 Exercises	195
9. Using the Means Model to Analyze Balanced Two-Way Treatment Structures with Unequal Subclass Numbers	199
9.1 Model Definitions and Assumptions	199
9.2 Parameter Estimation	199
9.3 Testing Whether All Means Are Equal	201
9.4 Interaction and Main Effect Hypotheses	202
9.5 Population Marginal Means	204
9.6 Simultaneous Inferences and Multiple Comparisons	206
9.7 Concluding Remarks	207
9.8 Exercises	207
10. Using the Effects Model to Analyze Balanced Two-Way Treatment Structures with Unequal Subclass Numbers	209
10.1 Model Definition	209
10.2 Parameter Estimates and Type I Analysis	209
10.3 Using Estimable Functions in SAS	213
10.4 Types I-IV Hypotheses	217
10.5 Using Types I-IV Estimable Functions in SAS-GLM	222
10.6 Population Marginal Means and Least Squares Means	226
10.7 Computer Analyses	226
10.8 Concluding Remarks	228
10.9 Exercises	229
11. Analyzing Large Balanced Two-Way Experiments Having Unequal Subclass Numbers	231
11.1 Feasibility Problems	231
11.2 Method of Unweighted Means	232
11.3 Simultaneous Inference and Multiple Comparisons	233
11.4 An Example of the Method of Unweighted Means	235

11.5 Computer Analyses	236
11.6 Concluding Remarks	237
11.7 Exercises	238
12. Case Study: Balanced Two-Way Treatment Structure with Unequal Subclass Numbers	239
12.1 Fat-Surfactant Example	239
12.2 Concluding Remarks	243
12.3 Exercises	243
13. Using the Means Model to Analyze Two-Way Treatment Structures with Missing Treatment Combinations	245
13.1 Parameter Estimation	245
13.2 Hypothesis Testing and Confidence Intervals	247
13.2.1 Example 13.1	247
13.3 Computer Analyses	249
13.4 Concluding Remarks	253
13.5 Exercises	253
14. Using the Effects Model to Analyze Two-Way Treatment Structures with Missing Treatment Combinations	255
14.1 Type I and II Hypotheses	255
14.2 Type III Hypotheses	257
14.3 Type IV Hypotheses	259
14.4 Population Marginal Means and Least Squares Means	261
14.5 Computer Analyses	262
14.6 Concluding Remarks	262
14.7 Exercises	263
15. Case Study: Two-Way Treatment Structure with Missing Treatment Combinations	265
15.1 Case Study	265
15.2 Concluding Remarks	268
15.3 Exercise	268
16. Analyzing Three-Way and Higher-Order Treatment Structures	271
16.1 General Strategy	271
16.2 Balanced and Unbalanced Experiments	273
16.3 Type I and II Analyses	273
16.4 Concluding Remarks	274
16.5 Exercises	274
17. Case Study: Three-Way Treatment Structure with Many Missing Treatment Combinations	277
17.1 Nutrition Scores Example	277
17.2 An SAS-GLM Analysis	277
17.3 A Complete Analysis	282
17.4 Concluding Remarks	284
17.5 Exercises	285

18. Random Effects Models and Variance Components	287
18.1 Introduction	287
18.1.1 Example 18.1: Random Effects Nested Treatment Structure	289
18.2 General Random Effects Model in Matrix Notation	290
18.2.1 Example 18.2: One-Way Random Effects Model	290
18.3 Computing Expected Mean Squares	292
18.3.1 Algebraic Method	293
18.3.2 Computing Using Hartley's Method of Synthesis	295
18.4 Concluding Remarks	307
18.5 Exercises	308
19. Methods for Estimating Variance Components	309
19.1 Method of Moments	309
19.1.1 Applications. Example 19.1: Unbalanced One-Way Model	312
19.1.2 Example 19.2: Wheat Varieties in a One-Way Random Effects Model	313
19.1.3 Example 19.3: Data for Two-Way Design in Table 18.2	315
19.2 Maximum Likelihood Estimators	318
19.2.1 Example 19.4: Maximum Likelihood Solution for Balanced One-Way Model	319
19.3 Restricted or Residual Maximum Likelihood Estimation	322
19.3.1 Example 19.5: REML Solution for Balanced One-Way Model	323
19.4 MIVQUE Method	325
19.4.1 Description of the Method	325
19.4.2 Application. Example 19.6: MIVQUE for the Unbalanced One-Way Design	327
19.5 Estimating Variance Components Using JMP	329
19.6 Concluding Remarks	334
19.7 Exercises	334
20. Methods for Making Inferences about Variance Components	337
20.1 Testing Hypotheses	337
20.1.1 Using the Analysis of Variance Table	338
20.1.2 Example 20.1: Two-Way Random Effects TS in a CR DS	338
20.1.3 Example 20.2: Complex Three-Way Random Effects TS	339
20.1.4 Likelihood Ratio Test	344
20.1.5 Example 20.3: Wheat Varieties—One-Way Random Effects Model	345
20.1.6 Example 20.4: Unbalanced Two-Way	347
20.2 Constructing Confidence Intervals	347
20.2.1 Residual Variance σ_e^2	348
20.2.2 General Satterthwaite Approximation	348
20.2.3 Approximate Confidence Interval for a Function of the Variance Components	349
20.2.4 Wald-Type Confidence Intervals for Variance Components	353
20.2.5 Some Exact Confidence Intervals	353
20.2.6 Example 20.5: Balanced One-Way Random Effects Treatment Structure	356

20.2.7 Example 20.6	357
20.2.8 Example 20.6 Continued	360
20.3 Simulation Study	361
20.4 Concluding Remarks	362
20.5 Exercises	363
21. Case Study: Analysis of a Random Effects Model	365
21.1 Data Set	365
21.2 Estimation	367
21.3 Model Building	368
21.4 Reduced Model	370
21.5 Confidence Intervals	374
21.6 Computations Using JMP	379
21.7 Concluding Remarks	380
21.8 Exercises	382
22. Analysis of Mixed Models	385
22.1 Introduction to Mixed Models	385
22.2 Analysis of the Random Effects Part of the Mixed Model	387
22.2.1 Method of Moments	387
22.2.2 Method of Maximum Likelihood	390
22.2.3 Method of Residual Maximum Likelihood	392
22.2.4 MINQUE Method	394
22.3 Analysis of the Fixed Effects Part of the Model	394
22.3.1 Estimation	394
22.3.2 Construction of Confidence Intervals	395
22.3.3 Testing Hypotheses	396
22.4 Best Linear Unbiased Prediction	397
22.5 Mixed Model Equations	397
22.6 Concluding Remarks	399
22.7 Exercises	399
23. Case Studies of a Mixed Model	401
23.1 Two-Way Mixed Model	401
23.2 Unbalanced Two-Way Mixed Model	409
23.3 JMP Analysis of the Unbalanced Two-Way Data Set	415
23.4 Concluding Remarks	417
23.5 Exercises	418
24. Methods for Analyzing Split-Plot Type Designs	421
24.1 Introduction	421
24.1.1 Example 24.1: Bread Recipes and Baking Temperatures	422
24.1.2 Example 24.2: Wheat Varieties Grown in Different Fertility Regimes	426
24.2 Model Definition and Parameter Estimation	428
24.3 Standard Errors for Comparisons among Means	430
24.4 A General Method for Computing Standard Errors of Differences of Means	434
24.5 Comparison via General Contrasts	436

24.6	Additional Examples	440
24.6.1	Example 24.3: Moisture and Fertilizer	440
24.6.2	Example 24.4: Regression with Split-Plot Errors	444
24.6.3	Example 24.5: Mixed-Up Split-Plot Design	448
24.6.4	Example 24.6: Split-Split-Plot Design	451
24.7	Sample Size and Power Considerations	455
24.8	Computations Using JMP—Example 24.7	459
24.9	Concluding Remarks	466
24.10	Exercises	467
25.	Methods for Analyzing Strip-Plot Type Designs	471
25.1	Description of the Strip-Plot Design and Model	471
25.2	Techniques for Making Inferences	475
25.3	Example: Nitrogen by Irrigation	477
25.4	Example: Strip-Plot with Split-Plot 1	480
25.5	Example: Strip-Plot with Split-Plot 2	482
25.6	Strip-Plot with Split-Plot 3	483
25.7	Split-Plot with Strip-Plot 4	485
25.8	Strip-Strip-Plot Design with Analysis via JMP7	486
25.9	Concluding Remarks	491
25.10	Exercises	493
26.	Methods for Analyzing Repeated Measures Experiments	499
26.1	Model Specifications and Ideal Conditions	500
26.2	The Split-Plot in Time Analyses	502
26.2.1	Example 26.1: Effect of Drugs on Heart Rate	502
26.2.2	Example 26.2: A Complex Comfort Experiment	510
26.2.3	Example 26.3: Family Attitudes	516
26.3	Data Analyses Using the SAS-Mixed Procedure	519
26.3.1	Example 26.1	521
26.3.2	Example 26.2	524
26.3.3	Example 26.3	525
26.4	Concluding Remarks	529
26.5	Exercises	529
27.	Analysis of Repeated Measures Experiments When the Ideal Conditions Are Not Satisfied	535
27.1	Introduction	535
27.2	MANOVA Methods	537
27.3	<i>p</i> -Value Adjustment Methods	547
27.4	Mixed Model Methods	553
27.4.1	Maximum Likelihood Method	557
27.4.2	Restricted Maximum Likelihood Method	557
27.5	Summary	568
27.6	Exercises	570
28.	Case Studies: Complex Examples Having Repeated Measures	573
28.1	Complex Comfort Experiment	573
28.2	Family Attitudes Experiment	583

28.3 Multilocation Experiment	592
28.4 Exercises	597
29. Analysis of Crossover Designs	599
29.1 Definitions, Assumptions, and Models	599
29.2 Two Period/Two Treatment Designs	600
29.3 Crossover Designs with More than Two Periods	609
29.4 Crossover Designs with More than Two Treatments	616
29.5 Summary	625
29.6 Exercises	625
30. Analysis of Nested Designs	627
30.1 Definitions, Assumptions, and Models	627
30.1.1 Example 30.1: Companies and Insecticides	627
30.1.2 Example 30.2: Comfort Experiment Revisited	629
30.1.3 Example 30.3: Coffee Price Example Revisited	629
30.2 Parameter Estimation	629
30.2.1 Example 30.1: Continuation	630
30.2.2 Example 30.2: Continuation	631
30.2.3 Example 30.3: Continuation	632
30.3 Testing Hypotheses and Confidence Interval Construction	633
30.3.1 Example 30.1: Continuation	634
30.4 Analysis Using JMP	636
30.5 Concluding Remarks	640
30.6 Exercises	640
Appendix	643
References	651
Index	655

1

The Simplest Case: One-Way Treatment Structure in a Completely Randomized Design Structure with Homogeneous Errors

Suppose an experimenter wants to compare the effects of several different treatments, such as the effects of different drugs on people's heart rates or the yields of several different varieties of wheat. Often the first step in analyzing the data from such experiments is to use a statistical method, known as a one-way analysis of variance model, to describe the data. The model on which the one-way analysis of variance is based is one of the most useful models in the field of statistics. Many experimental situations are simply special cases of this model. Other models that appear to be much more complicated can often be considered as one-way models. This chapter is divided into several sections. In the first two sections, the one-way model is defined and the estimation of its parameters is discussed. In Sections 1.3 and 1.5, inference procedures for specified linear combinations of the treatment effects are provided. In Sections 1.7 and 1.9, we introduce two basic methods for developing test statistics. These two methods are used extensively throughout the remainder of the book. Finally, in Section 1.11, we discuss readily available computer analyses that use the above techniques. An example is used to demonstrate the concepts and computations described in each section.

1.1 Model Definitions and Assumptions

Assume that a sample of N experimental units is selected completely at random from a population of possible experimental units. An experimental unit is defined as the basic unit to which a treatment will be applied and independently observed. A more complete description of experimental units can be found in Chapters 4 and 5.

In order to compare the effects of t different treatments, the sample of N experimental units is randomly divided into t groups so that there are n_i experimental units in the i th

group, where $i = 1, 2, \dots, t$, and $N = \sum_{i=1}^t n_i$. Grouping the experimental units at random into t groups should remove any systematic biases. That is, randomness should ensure that the t groups of experimental units are similar in nature before the treatments are applied. Finally, one of the t treatments should be randomly assigned to each group of experimental units. Equivalently, the experimental units could be randomly assigned to the t treatment groups using some randomization device such as placing n_1 tags in a bowl with treatment 1, n_2 tags in a bowl with treatment 2, ..., n_t tags in a bowl with treatment t , mixing the tags and then randomly selecting tags from the bowl to determine the treatment assigned to each experimental unit. This process of using tags in a bowl can obviously be carried out using software that has random number generation possibilities.

Let y_{ij} denote a response from the j th experimental unit assigned to the i th treatment. The values $y_{11}, y_{12}, \dots, y_{1n_1}$ can be thought of as being a random sample of size n_1 from a population with mean μ_1 and variance σ_1^2 , the values $y_{21}, y_{22}, \dots, y_{2n_2}$ can be thought of as being a random sample of size n_2 from a population with mean μ_2 and variance σ_2^2 , and similarly for $i = 3, 4, \dots, t$. The parameters μ_i and σ_i^2 represent the population mean and population variance if one applied treatment i to the whole population of experimental units.

The simplest case is considered in this chapter in that the variances are assumed to be homogeneous or equal across treatments or $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_t^2$. That is, it is assumed that the application of the i th treatment to the experimental units may affect the mean of the responses but not the variance of the responses. The equal variance assumption is discussed in Chapter 2 as well as the analysis of variance with unequal variances.

The basic objectives of a good statistical analysis are to estimate the parameters of the model and to make inferences about them. The methods of inference usually include testing hypotheses and constructing confidence intervals.

There are several ways to write a model for data from situations like the one described above. The first model to be used is called the μ_i model or the means model. The means model is:

$$y_{ij} = \mu_i + \varepsilon_{ij} \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, n_i$$

where it is assumed that

$$\varepsilon_{ij} \sim i.i.d. N(0, \sigma^2) \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, n_i \quad (1.1)$$

The notation $\varepsilon_{ij} \sim i.i.d. N(0, \sigma^2)$ is used extensively throughout this book. It means that the ε_{ij} ($i = 1, 2, \dots, t$; $j = 1, 2, \dots, n_i$) are independently and identically distributed and that the sampling distribution of each ε_{ij} is the normal distribution with mean equal to zero and variance equal to σ^2 .

1.2 Parameter Estimation

The most important aspect of a statistical analysis is to get a good estimate of the error variance per experimental unit, namely σ^2 . The error variance measures the accuracy of an experiment—the smaller the σ^2 , the more accurate the experiment. One cannot make any

statistically valid inferences in any experiment or study without some knowledge of the experimental error variance.

In the above situation, the i th sample, $i = 1, 2, \dots, t$, provides an estimate of σ^2 when $n_i > 1$. The estimate of σ^2 obtained from the data from the i th treatment is

$$\hat{\sigma}_i^2 = \sum_{j=1}^{n_i} \frac{(y_{ij} - \bar{y}_{i\cdot})^2}{n_i - 1}$$

which is an unbiased estimate of σ^2 where

$$\bar{y}_{i\cdot} = \frac{\sum_{j=1}^{n_i} y_{ij}}{n_i}$$

The estimate of σ^2 from the i th treatment is $\hat{\sigma}_i^2$, which is based on $n_i - 1$ degrees of freedom, and the sampling distribution of $(n_i - 1)\hat{\sigma}_i^2/\sigma^2$ is a chi-square distribution with $n_i - 1$ degrees of freedom.

A weighted average of these t independent estimates of σ^2 provides the best estimate for σ^2 possible for this situation, where each estimate of the variance is weighted by its corresponding degrees of freedom. The best estimate of σ^2 is

$$\hat{\sigma}^2 = \sum_{i=1}^t (n_i - 1) \hat{\sigma}_i^2 / \sum_{i=1}^t (n_i - 1)$$

For computational purposes, each variance times its weight can be expressed as

$$(n_i - 1)\hat{\sigma}_i^2 = \sum_{i=1}^t (y_{ij} - \bar{y}_{i\cdot})^2 = \sum_{i=1}^t y_{ij}^2 - n_i \bar{y}_{i\cdot}^2 = \sum_{i=1}^t y_{ij}^2 - (\bar{y}_{i\cdot})^2 / n_i = SS_i$$

where $y_{i\cdot} = \sum_{j=1}^{n_i} y_{ij}$. Then the pooled estimate of the variance is

$$\hat{\sigma}^2 = \frac{SS_1 + SS_2 + \dots + SS_t}{(n_1 - 1) + (n_2 - 1) + \dots + (n_t - 1)} = \frac{\sum_{i=1}^t SS_i}{N - t}$$

The pooled estimate of the variance $\hat{\sigma}^2$ is based on $N - t$ degrees of freedom and the sampling distribution of $(N - t)\hat{\sigma}^2/\sigma^2$ is a chi-square distribution with $N - t$ degrees of freedom; that is, $(N - t)\hat{\sigma}^2/\sigma^2 \sim \chi_{N-t}^2$.

The best estimate of each μ_i is $\hat{\mu}_i = \bar{y}_{i\cdot}$, $i = 1, 2, \dots, t$.

Under the assumption given in Equation 1.1, the sampling distribution of $\hat{\mu}_i$ is normal with mean μ_i and variance σ^2/n_i . That is,

$$\hat{\mu}_i \sim N\left(\mu_i, \frac{\sigma^2}{n_i}\right) \quad i = 1, 2, \dots, t \quad (1.2)$$

Using the sampling distributions of $\hat{\mu}_i$ and $\hat{\sigma}_i^2$ then

$$t_i = \frac{\hat{\mu}_i - \mu_i}{\sqrt{\hat{\sigma}_i^2/n_i}} \sim t_{N-t} \quad i = 1, 2, \dots, t \quad (1.3)$$

That is, the sampling distribution of t_i is the t -distribution with $N - t$ degrees of freedom. In addition, $\hat{\mu}_1, \hat{\mu}_2, \dots, \hat{\mu}_t$ and $\hat{\sigma}_i^2$ are statistically independent.

1.3 Inferences on Linear Combinations—Tests and Confidence Intervals

This section provides tests of hypotheses and confidence intervals for linear functions of the parameters in the means model. The results in the previous section can be used to test hypotheses about the individual μ_i . Those results can also be used to test hypotheses about linear combinations of the μ_i or to construct confidence intervals for linear combinations of the μ_i .

For an experiment involving several treatments, the investigator selects the treatments to be in the study because there are interesting hypotheses that need to be studied. These interesting hypotheses form the objectives of the study. The hypotheses involving the treatment means most likely will involve specific linear combinations of the means. These linear combinations will enable the investigator to compare the effects of the different treatments or, equivalently, the means of the different treatments or populations. The hypotheses about the means the experimenter has selected can be of the following types of hypotheses:

$$H_{01}: \sum_{i=1}^t c_i \mu_i = a \text{ vs } H_{a1}: (\text{not } H_{01})$$

for some set of known constants c_1, c_2, \dots, c_t and a ,

$$H_{02}: \mu_1 = \mu_2 = \dots = \mu_t \text{ vs } H_{a2}: (\text{not } H_{02})$$

and

$$H_{03}: \mu_i = \mu_{i'}, \text{ for some } i \neq i' \text{ vs } H_{a3}: (\text{not } H_{03})$$

For a linear combination such as that given in H_{01} , one can show that

$$\frac{\sum_{i=1}^t c_i \hat{\mu}_i - \sum_{i=1}^t c_i \mu_i}{\sqrt{\hat{\sigma}^2 \sum_{i=1}^t c_i^2 / n_i}} \sim t_{(N-t)} \quad (1.4)$$

This result can be used to make inferences about linear combinations of the form $\sum_{i=1}^t c_i \mu_i$. Since the hypothesis in H_{03} can be written as $H_{03}: \mu_i - \mu_{i'} = 0$, it is a special case of H_{01} with

$c_i = 1$, $c_{i'} = -1$, and $c_k = 0$ if $k \neq i$ or i' . A test for H_{02} is given in Section 1.5. The estimated standard error of $\sum_{i=1}^t c_i \hat{\mu}_i$ is given by

$$\widehat{s.e.}(\sum c_i \hat{\mu}_i) = \sqrt{\hat{\sigma}^2 \sum \frac{c_i^2}{n_i}} \quad (1.5)$$

To test $H_{01}: \sum_{i=1}^t c_i \mu_i = a$ vs H_{a1} : (not H_{01}) compute the t -statistic

$$t_c = \frac{\sum c_i \hat{\mu}_i - a}{\widehat{s.e.}(\sum c_i \hat{\mu}_i)} \quad (1.6)$$

If $|t_c| > t_{\alpha/2, v}$ where $v = N - t$, then H_{01} is rejected at the $\alpha = 100\%$ significance level, where $t_{\alpha/2, v}$ is the upper $\alpha/2$ critical point of a t -distribution with v degrees of freedom. A $(1 - \alpha)$ 100% confidence interval for $\sum_{i=1}^t c_i \mu_i$ is provided by

$$\sum c_i \hat{\mu}_i \pm t_{\alpha/2, v} \widehat{s.e.}(\sum c_i \hat{\mu}_i) \quad (1.7)$$

1.4 Example—Tasks and Pulse Rate

The data in Table 1.1 came from an experiment that was conducted to determine how six different kinds of work tasks affect a worker's pulse rate. In this experiment, 78 male workers were assigned at random to six different groups so that there were 13 workers in each group. Each group of workers was trained to perform their assigned task. On a selected day after training, the pulse rates of the workers were measured after they had performed their assigned tasks for 1 h. Unfortunately, some individuals withdrew from the experiment during the training process so that some groups contained fewer than 13 individuals. The recorded data represent the number of heart pulsations in 20 s where there are $N = 68$ observations and the total is $y = 2197$.

For the tasks data, the best estimate of σ^2 is

$$\hat{\sigma}^2 = \sum_{i=1}^6 SS_i / (N - t) = 1,916.0761 / 62 = 30.9045$$

which is based on 62 degrees of freedom. The best estimates of the μ_i are $\hat{\mu}_1 = 31.923$, $\hat{\mu}_2 = 31.083$, $\hat{\mu}_3 = 35.800$, $\hat{\mu}_4 = 38.000$, $\hat{\mu}_5 = 29.500$, and $\hat{\mu}_6 = 28.818$.

For illustration purposes, suppose the researcher is interested in answering the following questions about linear combinations of the task means:

- a) Test $H_0: \mu_3 = 30$ vs $H_{a3}: \mu_3 \neq 30$.
- b) Find a 95% confidence interval for μ_1 .

TABLE 1.1
Pulsation Data and Summary Information for Six Tasks

	Task					
	1	2	3	4	5	6
27	29	34	34	28	28	
31	28	36	34	28	26	
26	37	34	43	26	29	
32	24	41	44	35	25	
39	35	30	40	31	35	
37	40	44	47	30	34	
38	40	44	34	34	37	
39	31	32	31	34	28	
30	30	32	45	26	21	
28	25	31	28	20	28	
27	29			41	26	
27	25			21		
34						
y_i	415	373	358	380	354	317
n_i	13	12	10	10	12	11
\bar{y}_i	31.9231	31.0833	35.8000	38.0000	29.5000	28.8182
SS_i	294.9231	352.9167	253.6000	392.0000	397.0000	225.6364

- c) Test $H_0: \mu_4 = \mu_5$ vs $H_a: \mu_4 \neq \mu_5$.
- d) Test $H_0: \mu_1 = (\mu_2 + \mu_3 + \mu_4)/3$ vs $H_a: \mu_1 \neq (\mu_2 + \mu_3 + \mu_4)/3$.
- e) Obtain a 90% confidence interval for $4\mu_1 - \mu_3 - \mu_4 - \mu_5 - \mu_6$.

These questions can be answered by applying the results of this section.

Part a result: A t -statistic for testing $H_0: \mu_3 = 30$ is obtained by substituting into Equation 1.6 to obtain

$$t_c = \frac{\hat{\mu}_3 - 30}{\widehat{s.e.}(\hat{\mu}_3)} = \frac{\hat{\mu}_3 - 30}{\sqrt{\hat{\sigma}^2/n_3}} = \frac{35.8 - 30.0}{\sqrt{30.9045/10}} = 3.30$$

The significance probability of this calculated value of t is $\hat{\alpha} = \Pr\{|t_c| > 3.30\} = 0.0016$ where $\Pr\{|t_c| > 3.30\}$ is the area to the right of 3.30 plus the area to the left of -3.30 in a t -distribution with 62 degrees of freedom. The above value of $\hat{\alpha}$ was obtained from computer output, but it can also be obtained from some special hand-held calculators. Readers of this book who lack access to a computer or a calculator should compare $t_c = 3.30$ to $t_{\alpha/2, 62}$ for their choice of α .

Part b result: A 95% confidence interval for μ_1 is given by

$$\begin{aligned}\hat{\mu}_1 \pm t_{0.025, 62} \widehat{s.e.}(\hat{\mu}_1) &= 31.923 \pm 2.00 \sqrt{30.9045/13} \\ &= 31.923 \pm 2.00 \times 1.542\end{aligned}$$

Thus the 95% confidence interval about μ_1 is $28.839 < \mu_1 < 35.007$ and we are 95% confident that this interval contains the true, but unknown value of μ_1 .

Part c result: To test $H_0: \mu_4 = \mu_5$, let $l_1 = \mu_4 - \mu_5$, then $\hat{l}_1 = \hat{\mu}_4 - \hat{\mu}_5 = 38.0 - 29.5 = 8.5$ and

$$\widehat{s.e.}(\hat{l}_1) = \sqrt{\hat{\sigma}^2 \sum_{i=1}^6 c_i^2 / n_i} = \sqrt{30.9045 \left(\frac{1}{10} + \frac{1}{12} \right)} = 2.380$$

since $c_1 = c_2 = c_3 = c_6 = 0$, $c_4 = 1$, and $c_5 = -1$.

The t -statistic for testing $H_0: \mu_4 = \mu_5$ is

$$t_c = \frac{8.5}{2.380} = 3.57$$

The significance probability for this test is $\hat{\alpha} = 0.0007$.

Part d result: A test of $H_0: \mu_1 = (\mu_2 + \mu_3 + \mu_4)/3$ is equivalent to testing $H_0: \mu_1 - \frac{1}{3}\mu_2 - \frac{1}{3}\mu_3 - \frac{1}{3}\mu_4 = 0$ or testing $H_0: 3\mu_1 - \mu_2 - \mu_3 - \mu_4 = 0$. By choosing the last version, the computations are somewhat easier and the value of the t_c test statistic is invariant with respect to a constant multiplier.

Let $l_2 = 3\mu_1 - \mu_2 - \mu_3 - \mu_4$, then

$$\hat{l}_2 = 3\hat{\mu}_1 - \hat{\mu}_2 - \hat{\mu}_3 - \hat{\mu}_4 = 3(31.923) - 31.083 - 35.8 - 38.0 = -9.114$$

The estimate of the standard error of \hat{l}_2 is

$$\widehat{s.e.}(\hat{l}_2) = \sqrt{30.9045 \left(\frac{9}{13} + \frac{1}{12} + \frac{1}{10} + \frac{1}{10} \right)} = 5.491$$

A t -statistic for testing $H_0: 3\mu_1 - \mu_2 - \mu_3 - \mu_4 = 0$ is

$$t_c = \frac{-9.114}{5.491} = -1.66$$

The significance probability corresponding to t_c is $\hat{\alpha} = 0.1020$.

Part e result: Let $l_3 = 4\mu_1 - \mu_3 - \mu_4 - \mu_5 - \mu_6$. Then $\hat{l}_3 = -4.426$ and $\widehat{s.e.}(\hat{l}_3) = 7.0429$. A 90% confidence interval for l_3 is

$$\hat{l}_3 \pm t_{0.05, 62} \widehat{s.e.}(\hat{l}_3) = -4.426 \pm 1.671 \times 7.043 = -4.426 \pm 11.769$$

Thus, a 90% confidence interval is $-16.195 < 4\mu_1 - \mu_3 - \mu_4 - \mu_5 - \mu_6 < 7.343$.

1.5 Simultaneous Tests on Several Linear Combinations

For many situations the researcher wants to test a simultaneous hypothesis about several linear combinations of the treatment's effects or means. For example, the general

hypothesis involving k linearly independent linear combinations of the treatment means can be expressed as

$$\begin{aligned} c_{11}\mu_1 + c_{12}\mu_2 + \cdots + c_{1t}\mu_t &= a_1 \\ H_0: \quad c_{21}\mu_1 + c_{22}\mu_2 + \cdots + c_{2t}\mu_t &= a_2 \quad \text{vs} \quad H_a: (\text{not } H_0) \\ &\vdots \\ c_{k1}\mu_1 + c_{k2}\mu_2 + \cdots + c_{kt}\mu_t &= a_k \end{aligned} \quad (1.8)$$

The results presented in this section are illustrated using vectors and matrices. However, knowledge of vectors and matrices is not really necessary for readers having access a computer with matrix manipulation software, since most computers allow even novice users to easily carry out matrix computations.

The hypothesis in Equation 1.8 can be written in matrix notation as

$$H_0: C\boldsymbol{\mu} = \mathbf{a} \text{ vs } H_a: C\boldsymbol{\mu} \neq \mathbf{a} \quad (1.9)$$

where

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1t} \\ c_{21} & c_{22} & \cdots & c_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ c_{k1} & c_{k2} & \cdots & c_{kt} \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_t \end{bmatrix}, \quad \text{and} \quad \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{bmatrix} \quad (1.10)$$

It is assumed that the k rows in C were chosen such that they are linearly independent, which means that none of the rows in C can be expressed as a linear combination of the remaining rows. If the k rows in C are not linearly independent, a subset of the rows that are linearly independent can always be selected so that they contain all the necessary information about the required hypothesis.

For example, suppose you have three treatments and you wish to test

$$H_0: \mu_1 - \mu_2 = 0, \mu_1 - \mu_3 = 0 \quad \text{and} \quad \mu_2 - \mu_3 = 0$$

the corresponding C matrix is

$$C = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}$$

but the third row of C is the difference between the second row and the first row, hence the three rows are not linearly independent. In this case, an equivalent hypothesis can be stated as $H_0: \mu_1 - \mu_2 = 0$ and $\mu_1 - \mu_3 = 0$, since if $\mu_1 - \mu_2 = 0$ and $\mu_1 - \mu_3 = 0$, then $\mu_2 - \mu_3$ must be equal to 0. The following discussion uses the assumption that the rows of C are linearly independent.

Denote the vector of sample means by $\hat{\mu}$, then the sampling distribution of $\hat{\mu}$ in matrix notation is

$$\hat{\mu} \sim N_t(\mu, \sigma^2 D) \quad \text{where } D = \begin{bmatrix} 1/n_1 & 0 & \cdots & 0 \\ 0 & 1/n_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/n_t \end{bmatrix}$$

This equation is read as follows: The elements of the $t \times 1$ vector $\hat{\mu}$ have a joint sampling distribution that is the t -variate normal distribution with means given by the vector μ and with variances and covariances given by the elements in the matrix $\sigma^2 D$. The i th diagonal element of $\sigma^2 D$ is the variance of $\hat{\mu}_i$ and the (i, j) th $i \neq j$ off-diagonal element gives the covariance between $\hat{\mu}_i$ and $\hat{\mu}_j$.

The sampling distribution of $C\hat{\mu}$ is

$$C\hat{\mu} \sim N_k(C\mu, \sigma^2 CDC')$$

The sum of squares due to deviations from H_0 or the sum of squares for testing $H_0: C\mu = a$ is given by

$$SS_{H_0} = (C\hat{\mu} - a)'(CDC')^{-1}(C\hat{\mu} - a) \quad (1.11)$$

and is based on k degrees of freedom, the number of linearly independent rows of C . Using the assumption of normality, the sampling distribution of SS_{H_0}/σ^2 is that of a noncentral chi-square with k degrees of freedom. If H_0 is true, then $SS_{H_0}/\sigma^2 \sim \chi_k^2$. The statistic for testing H_0 is

$$F_c = \frac{SS_{H_0}/k}{\hat{\sigma}^2}$$

The hypothesis $H_0: C\mu = a$ is rejected at the significance level of α if $F_c > F_{\alpha, k, N-t}$ where $F_{\alpha, k, N-t}$ is the upper α critical point of the F -distribution with k numerator degrees of freedom and $N-t$ denominator degrees of freedom. The result given here is a special case of Theorem 6.3.1 in Graybill (1976).

When H_0 is true, then SS_{H_0}/k is an unbiased estimate of σ^2 , which is then compared with $\hat{\sigma}^2$, which in turn is an unbiased estimate of σ^2 regardless of whether H_0 is true or not. Thus the F -statistic given above should be close to 1 if H_0 is true. If H_0 is false, the statistic SS_{H_0}/k is an unbiased estimate of

$$\sigma^2 + \frac{1}{k}(C\mu - a)'(CDC')^{-1}(C\mu - a)$$

Thus, if H_0 is false, the value of the F -statistic should be larger than 1. The hypothesis H_0 is rejected if the calculated F -statistic is significantly larger than 1.

1.6 Example—Tasks and Pulse Rate (Continued)

The following is a summary of the information from the example in Section 1.4 with the sample size and mean for each of the six tasks.

Task <i>i</i>	1	2	3	4	5	6
n_i	13	12	10	10	12	11
\bar{y}_i	31.9231	31.0833	35.8000	38.0000	29.5000	28.8182

The pooled estimate of the variance is $\hat{\sigma}^2 = 30.9045$ and it is based on 62 degrees of freedom. The D matrix associated with the sampling distribution of vector of estimated means is

$$D = \begin{bmatrix} \frac{1}{13} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{12} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{10} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{10} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{12} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{11} \end{bmatrix}$$

Suppose the researcher is interested in simultaneously testing the following hypothesis involving two linear combinations of the task means:

$$H_0: \mu_4 - \mu_5 = 4 \text{ and } 3\mu_1 - \mu_2 - \mu_3 - \mu_4 = 0 \text{ vs } H_a: (\text{not } H_0)$$

The C matrix consists of two rows, one for each of the linear combinations in H_0 , and the vector a has two elements as

$$C = \begin{bmatrix} 0 & 0 & 0 & 1 & -1 & 0 \\ 3 & -1 & -1 & -1 & 0 & 0 \end{bmatrix} \text{ and } a = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$$

Preliminary computations needed to provide the value of SS_{H_0} are:

$$C\hat{\mu} - a = \begin{bmatrix} 8.5 - 4 \\ -9.114 - 0 \end{bmatrix} = \begin{bmatrix} 4.500 \\ -9.114 \end{bmatrix}$$

$$\begin{aligned} CDC' &= \begin{bmatrix} \frac{1}{10} + \frac{1}{12} & -\frac{1}{10} \\ -\frac{1}{10} & \frac{9}{13} + \frac{1}{12} + \frac{1}{10} + \frac{1}{10} \end{bmatrix} \\ &= \begin{bmatrix} 0.1833 & -0.1000 \\ -0.1000 & 0.9756 \end{bmatrix} \end{aligned}$$

$$(CDC')^{-1} = \begin{bmatrix} 5.7776 & 0.5922 \\ 0.5922 & 1.0856 \end{bmatrix}$$

and

$$SS_{H_0} = (\mathbf{C}\hat{\boldsymbol{\mu}} - \mathbf{a})'(\mathbf{C}\mathbf{D}\mathbf{C}')^{-1}(\mathbf{C}\hat{\boldsymbol{\mu}} - \mathbf{a}) = 158.602$$

with 2 degrees of freedom. The test statistic is

$$F_c = \frac{158.602/2}{30.9045} = 2.566$$

The significance probability of this F -statistic is $\hat{\alpha} = \Pr[F > 2.566] = 0.0850$.

1.7 Testing the Equality of All Means

Often the first hypothesis of interest to most researchers is to test that the means are simultaneously equal. The hypothesis is $H_0: \mu_1 = \mu_2 = \dots = \mu_t$ vs H_a (not H_0). Two basic procedures are examined for testing the equal means hypothesis. For the particular situation discussed in this chapter, the two procedures give rise to the same statistical test. However, for most messy data situations (for treatment structures other than one-way), the two procedures can give rise to different tests. The first procedure is covered in this section, while the second is introduced in Section 1.9.

The equal means hypothesis, $H_0: \mu_1 = \mu_2 = \dots = \mu_t$ is equivalent to a hypothesis of the form, $H_0: \mu_1 - \mu_2 = 0, \mu_1 - \mu_3 = 0, \dots, \mu_1 - \mu_t = 0$, or any other hypothesis that involves $t - 1$ linearly independent linear combinations of the μ_i . The \mathbf{C} matrix and \mathbf{a} vector corresponding to the set of $t - 1$ pairwise differences are:

$$\mathbf{C} = \begin{bmatrix} 1 & -1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & -1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & -1 & \cdots & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & 0 & \cdots & -1 \end{bmatrix} \quad \text{and} \quad \mathbf{a} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

The \mathbf{C} matrix corresponding to following set of $t - 1$ linearly independent linear combinations of the μ_i ; $H_0: \mu_1 - \mu_2 = 0, \mu_1 + \mu_2 - 2\mu_3 = 0, \mu_1 + \mu_2 + \mu_3 - 3\mu_4 = 0, \dots, \mu_1 + \mu_2 + \dots + (t - 1)\mu_t = 0$ is:

$$\mathbf{C} = \begin{bmatrix} 1 & -1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & -2 & 0 & \cdots & 0 \\ 1 & 1 & 1 & -3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 1 & 1 & 1 & 1 & \cdots & t-1 \end{bmatrix} \quad \text{and} \quad \mathbf{a} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Many other matrices exist, so that $C\mu = \mathbf{0}$ if and only if $\mu_1 = \mu_2 = \dots = \mu_t$; however, all such matrices produce the same sum of squares for deviations from H_0 and the same degrees of freedom, $t - 1$, and hence the same F -statistic. For this special case Equation 1.11 always reduces to

$$SS_{H0:\mu_1=\mu_2=\dots=\mu_t} = \sum_{i=1}^t n_i (\bar{y}_{i\cdot} - \bar{y}_{..})^2 = \sum_{i=1}^t \left(\frac{y_{i\cdot}^2}{n_i} \right) - \frac{y_{..}^2}{N} \quad (1.12)$$

1.8 Example—Tasks and Pulse Rate (Continued)

For the task and pulse rate data in Section 1.4, the $SS_{H0:\mu_1=\mu_2=\dots=\mu_t}$ is computed using Equations 1.11 and 1.12.

Using the formula in Equation 1.12, provides

$$\begin{aligned} SS_{H0} &= \frac{415^2}{13} + \frac{373^2}{12} + \frac{358^2}{10} + \frac{380^2}{10} + \frac{354^2}{12} + \frac{317^2}{11} - \frac{2197^2}{68} \\ &= 694.4386 \end{aligned}$$

with $t - 1 = 5$ degrees of freedom. The value of the F_c statistic is

$$F_c = \frac{694.4386/5}{30.9045} = 4.49$$

and the significance probability is $\hat{\alpha} = 0.0015$.

Next, using Equation 1.11, the matrix C , vector a , and matrix D are

$$C = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}, \quad a = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

and

$$D = \begin{bmatrix} \frac{1}{13} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{12} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{10} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{10} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{12} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{11} \end{bmatrix}$$

Next compute the individual quantities in Equation 1.11 as

$$C\hat{\mu} - a = \begin{bmatrix} 0.844 \\ -3.877 \\ -6.077 \\ 2.423 \\ 3.105 \end{bmatrix} \quad \text{and} \quad CDC' = \begin{bmatrix} \frac{25}{156} & \frac{1}{13} & \frac{1}{13} & \frac{1}{13} & \frac{1}{13} \\ \frac{1}{13} & \frac{23}{130} & \frac{1}{13} & \frac{1}{13} & \frac{1}{13} \\ \frac{1}{13} & \frac{1}{13} & \frac{23}{130} & \frac{1}{13} & \frac{1}{13} \\ \frac{1}{13} & \frac{1}{13} & \frac{1}{13} & \frac{25}{156} & \frac{1}{13} \\ \frac{1}{13} & \frac{1}{13} & \frac{1}{13} & \frac{1}{13} & \frac{24}{143} \end{bmatrix}$$

The inverse of CDC' is

$$(CDC')^{-1} = \begin{bmatrix} 9.882 & -1.765 & -1.765 & -2.118 & -1.941 \\ -1.765 & 8.529 & -1.471 & -1.765 & -1.618 \\ -1.765 & -1.471 & 8.529 & -1.765 & -1.618 \\ -2.118 & -1.765 & -1.765 & 9.882 & -1.941 \\ -1.941 & -1.618 & -1.618 & -1.941 & 9.221 \end{bmatrix}$$

Finally, the value of the sum of squares is

$$SS_{H_0} = (C\hat{\mu} - a)' (CDC')^{-1} (C\hat{\mu} - a) = 694.4386$$

which is the same as the sum of squares computed using Equation 1.12.

Clearly, this formula is not easy to use if one must do the calculations by hand. However, in many messy data situations, formulas such as this one are necessary in order to obtain the statistic to test meaningful hypotheses. Fortunately, by utilizing computers, C matrices can be constructed for a specific hypothesis and then one can allow the computer to do the tedious calculations.

1.9 General Method for Comparing Two Models—The Principle of Conditional Error

A second procedure for computing a test statistic compares the fit of two models. In this section, the two models compared are $y_{ij} = \mu_i + \varepsilon_{ij}$, which is the general or unreduced model, and $y_{ij} = \mu + \varepsilon_{ij}$, which is the model one would have if $H_0: \mu_1 = \mu_2 = \dots = \mu_t = \mu$ (say) were true. The first model is called the full model or the unrestricted model, while the second model is called the reduced model or the restricted model.

The principle known as the principle of conditional error is used to compare two models where one model is obtained by placing restrictions upon the parameters of another model. The principle is very simple, requiring that one obtain the residual or error sums of squares for both the full model and the reduced model. Let ESS_F denote the error sum of squares after fitting the full model and ESS_R denote the error sum of squares after fitting the

reduced model. Then the sum of squares due to the restrictions given by the hypothesis or deviations from the null hypothesis is $SS_{H_0} = ESS_R - ESS_F$. The degrees of freedom for both ESS_R and ESS_F are given by the difference between the total number of observations in the data set and the number of (essential) parameters to be estimated (essential parameters will be discussed in Chapter 6). Denote the degrees of freedom corresponding to ESS_R and ESS_F by df_R and df_F , respectively. The number of degrees of freedom corresponding to SS_{H_0} is $df_{H_0} = df_R - df_F$. An F -statistic for testing H_0 is given by

$$F_c = \frac{SS_{H_0}/df_{H_0}}{ESS_F/df_F}$$

One rejects H_0 at the significance level if $F_c > F_{\alpha, df_{H_0}, df_F}$.

For the case discussed above, $y_{ij} = \mu_i + \varepsilon_{ij}$ is the full model and $y_{ij} = \mu + \varepsilon_{ij}$ is the reduced model. The error sum of squares for the full model is

$$ESS_F = \sum_{i=1}^t \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2 = (N-t)\hat{\sigma}^2$$

with $df_F = N - t$, and the error sum of squares for the reduced model is

$$ESS_R = \sum_{i=1}^t \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2$$

with $df_R = N - 1$. Thus the sum of squares due to deviations from H_0 is

$$SS_{H_0: \mu_1 = \mu_2 = \dots = \mu_t} = ESS_R - ESS_F = \sum_{i=1}^t n_i (\bar{y}_{i.} - \bar{y}_{..})^2$$

with $t - 1$ degrees of freedom. This is the same sum of squares as was obtained in Equation 1.12.

The sums of squares that are of interest in testing situations are often put in a table called an analysis of variance table. Such a table often has a form similar to that in Table 1.2. The entries under the column “Source of variation” are grouped into sets. In a given situation only one of the labels in each set is used, with the choice being determined entirely by the experimenter.

TABLE 1.2

Analysis of Variance Table for One-Way Model to Test Equality of the Means

Source of Variation	df	SS	MS	F-test
$H_0: \mu_1 = \mu_2 = \dots = \mu_t$ Treatments between samples	$t - 1$	SS_{H_0}	$\frac{SS_{H_0}}{t - 1}$	$\frac{SS_{H_0}/t - 1}{\hat{\sigma}^2}$
Error within samples	$N - t$	SS_F	$\hat{\sigma}^2 = \frac{ESS_F}{N - t}$	

Note: df = degrees of freedom, SS = sum of square, and MS = mean square. These standard abbreviations are used throughout the book.

The principle of conditional error is also referred to as the model comparison procedure and the process is quite flexible. For example, if you are interested in testing a hypothesis for the task and pulse rate data, like $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6$ vs H_a : (not H_0), then the model under the conditions of H_0 has the form

$$\begin{aligned}y_{ij} &= \mu_0 + \varepsilon_{ij} && \text{for } i = 1, 2, 3 \\y_{ij} &= \mu_i + \varepsilon_{ij} && \text{for } i = 4, 5, 6\end{aligned}$$

that is, the model has equal means for the first three tasks and different means for the last three treatments. Such a model can be fit using most software packages where a qualitative or class variable is defined to have the value of 0 for tasks 1, 2, and 3 and the value of 1 for tasks 4, 5, and 6.

1.10 Example—Tasks and Pulse Rate (Continued)

The principle of conditional error is applied to the task and pulse rate data of Section 1.4 to provide a test of the equal means hypothesis, $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6$ vs H_a : (not H_0). The error sum of squares for the full model is $ESS_F = 1916.076$ with $df_F = 62$. The error sum of squares for the reduced model is $ESS_R = 73,593 - (2197)^2/68 = 2610.545$ with $df_R = 67$. Hence $SS_{H_0} = 2610.545 - 1916.076 = 694.439$ with $df_{H_0} = 67 - 62 = 5$. The analysis of variance table summarizing these computations is displayed in Table 1.3.

1.11 Computer Analyses

This chapter concludes with some remarks about utilizing computers and statistical computing packages such as SAS®, BMDP®, SYSTAT®, JMP®, and SPSS®. All of the methods and formulas provided in the preceding sections can easily be used on most computers. If the computer utilizes a programming language such as MATLAB, SAS-IML, or APL, the required matrix calculations are simple to do by following the matrix formulas given in the preceding sections. SAS, JMP, BMDP, SYSTAT, and SPSS each contain procedures that enable users to generate their own linear combinations of treatment means about which to test hypotheses. In addition, these packages all provide an analysis of variance table, treatment means, and their standard errors. Table 1.4 contains SAS-GLM code with estimate and contrast statements needed to test hypotheses described for the task and pulse data. The estimate statement is used to evaluate one linear combination of the means and the

TABLE 1.3

Analysis of Variance Table for Test Equality of the Means for the Task and Pulse Rate Data

Source of Variation	df	SS	MS	F	α
Due to H_0	5	694.439	138.888	4.49	0.0015
Error	62	1,916.076	30.9045		

TABLE 1.4

Proc GLM Code to Fit the Task and Pulse Rate Data with Estimate and Contrast Statements Needed to Provide the Analysis Described in the Text

```
PROC GLM DATA=EX1; CLASS TASK;
MODEL PULSE20=TASK/NOINT SOLUTION E;
ESTIMATE 'Ho: M4=M5' TASK 0 0 0 1 -1 0;
ESTIMATE 'Ho: 3M1=M2+M3+M4' TASK 3 -1 -1 -1 0 0;
ESTIMATE 'Ho: 3M1=M2+M3+M4_mn' TASK 3 -1 -1 -1 0 0/DIVISOR=3;
ESTIMATE '4M1-M3-M4-M5-M6_mn' TASK 4 0 -1 -1 -1 -1/DIVISOR=4;
CONTRAST '4M1-M3-M4-M5-M6_mn' TASK 4 0 -1 -1 -1 -1;
CONTRAST 'M4=M5 & 3M1=M2+M3+M4' TASK 0 0 0 1 -1 0, TASK 3 -1 -1 -1 0 0;
CONTRAST 'EQUAL MEANS 1'
  TASK 1 -1 0 0 0 0, TASK 1 0 -1 0 0 0, TASK 1 0 0 -1 0 0,
  TASK 1 0 0 0 -1 0, TASK 1 0 0 0 0 -1;
```

TABLE 1.5

Proc IML Code to Carry Out the Computations for the Task and Pulse Data in Section 1.6

```
proc iml;
dd={13 12 10 10 12 11};
d=diag(dd);
c={0 0 0 1 -1 0, 3 -1 -1 -1 0 0};
muhat={31.9231 31.0833 35.8000 38.0000 29.5000 28.8182}`;
s2=30.90445;
a={4,0};
cmua=C*muhat - a;
cdc=c*inv(D)*c`;
cdci=inv(cdc);
ssho=cmua`*cdci*cmua;
f=ssho/(2*s2);al=1-probf(f,2,62);
print dd d cmua cdc cdci ssho f al;
```

provided results are the estimate of the contrast, its estimated standard error, and the resulting t -statistic with its corresponding significance level. The contrast statement is used to evaluate one or more linear combinations of the means and the provided results are the sums of squares, degrees of freedom, and the resulting F -statistic. For both the estimate and contrast statements in SAS-GLM, the only values of a in the hypotheses are zero, that is, one can only test the linear combinations of means that are equal to zero.

Table 1.5 contains SAS-IML code to provide the computations for the hypotheses being tested in Section 1.6. By constructing the code in a matrix language, one can obtain a test of any hypothesis of the form $C\mu = a$.

1.12 Concluding Remarks

In this chapter, the analysis of the one-way analysis of variance model was described. General procedures for making statistical inferences about the effects of different treatments were provided and illustrated for the case of homogeneous errors. Two basic procedures

for obtaining statistical analyses of experimental design models were introduced. These procedures are used extensively throughout the remainder of the book for more complex models used to describe designed experiments and for messier data situations. A test for comparing all treatment effect means simultaneously was also given. Such a test may be considered an initial step in a statistical analysis. The procedures that should be used to complete the analysis of a data set could depend on whether the hypothesis of equal treatment means is rejected.

1.13 Exercises

- 1.1 A company studied five techniques of assembling a part. Forty workers were randomly selected from the worker population and eight were randomly assigned to each technique. The worker assembled a part and the measurement was the amount of time in seconds required to complete the assembly. Some workers did not complete the task.

Data for Comparing Techniques of Assembling a Part for Exercise 1.1

Technique 1		Technique 2		Technique 3		Technique 4		Technique 5	
Worker	Time								
1	45.6	7	41.0	12	51.7	19	67.5	26	57.1
2	41.0	8	49.1	13	60.1	20	57.7	27	69.6
3	46.4	9	49.2	14	52.6	21	58.2	28	62.7
4	50.7	10	54.8	15	58.6	22	60.6		
5	47.9	11	45.0	16	59.8	23	57.3		
6	44.6			17	52.6	24	58.3		
				18	53.8	25	54.8		

- 1) Write down a model appropriate to describe the data. Describe each component of the model.
 - 2) Estimate the parameters of the model in part 1.
 - 3) Construct a 95% confidence interval about $\mu_1 - \mu_2$.
 - 4) Use a t -statistic to test $H_0: \mu_1 + \mu_2 - \mu_3 - \mu_4 = 0$ vs $H_a: (\text{not } H_0)$.
 - 5) Use a F -statistic to test $H_0: \mu_1 + \mu_2 - \mu_3 - \mu_5 = 0$ vs $H_a: (\text{not } H_0)$.
 - 6) Use a t -statistic to test $H_0: (\mu_1 + \mu_2 + \mu_3)/3 = (\mu_4 + \mu_5)/2$ vs $H_a: (\text{not } H_0)$.
 - 7) Use a F -statistic to test $H_0: \mu_1 = \mu_2 = \mu_3$ vs $H_a: (\text{not } H_0)$.
 - 8) Use a F -statistic to test $H_0: (\mu_1 + \mu_2 + \mu_3)/3 = (\mu_4 + \mu_5)/2, (\mu_1 + \mu_2 + \mu_6)/3 = (\mu_3 + \mu_4 + \mu_5)/3$, and $(\mu_1 + \mu_4 + \mu_5)/3 - (\mu_3 + \mu_6)/2$ vs $H_a: (\text{not } H_0)$.
- 1.2 Five rations were evaluated as to their ability to enable calves to grow. Thirty-one calves were used in the study. A mistake in the feeding of the rations produced unbalanced distributions of the calves to the rations. The data recorded was the number of pounds of weight gained over the duration of the study.
- 1) Write down a model appropriate to describe the data. Describe each component of the model.

Gain Data for Comparing Ratios of Exercise 1.2

Ration 1		Ration 2		Ration 3		Ration 4		Ration 5	
Calf	Gain	Calf	Gain	Calf	Gain	Calf	Calf	Calf	Gain
1	825	10	874	19	861	21	829	23	837
2	801	11	854	20	856	22	814	24	851
3	790	12	883					25	824
4	809	13	839					26	781
5	830	14	836					27	810
6	825	15	839					28	847
7	839	16	840					29	826
8	835	17	834					30	832
9	872	18	894					31	830

- 2) Estimate the parameters of the model in part 1.
- 3) Construct a 95% confidence interval about $\mu_1 + \mu_2 - 2\mu_5$.
- 4) Use a *t*-statistic to test $H_0: \mu_1 + \mu_2 - 2\mu_3 = 0$ vs H_a : (not H_0).
- 5) Use an *F*-statistic to test $H_0: 2\mu_2 - \mu_4 - \mu_5 = 0$ vs H_a : (not H_0).
- 6) Use a *t*-statistic to test $H_0: (\mu_1 + \mu_2 + \mu_3)/3 = (\mu_4 + \mu_5)/2$ vs H_a : (not H_0).
- 7) Use an *F*-statistic to test $H_0: \mu_1 = \mu_2$ and $\mu_3 = \mu_4$ vs H_a : (not H_0).
- 8) Use an *F*-statistic to test $H_0: \mu_1 + \mu_2 - 2\mu_3 = 0, 2\mu_2 - \mu_4 - \mu_5 = 0, (\mu_1 + \mu_2 + \mu_3)/3 = (\mu_4 + \mu_5)/2$ vs H_a : (not H_0).
- 1.3 A study was conducted to evaluate the effect of elevation on the lung volume of birds raised at specified elevations. Thirty-five environmental chambers which could simulate elevations by regulating the air pressure were used. The five effective elevations were each randomly assigned to seven chambers and 35 baby birds were randomly assigned to the chambers, one per chamber. When the birds reached adult age, their lung volumes were measured. The data table contains the effective elevations and the volumes of the birds. Three birds did not survive the study, thus producing missing data.

Lung Volumes for Birds Raised at Different Simulated Elevations

Elevation 1000 ft		Elevation 2000 ft		Elevation 3000 ft		Elevation 4000 ft		Elevation 5000 ft	
Bird	Volume								
1	156	8	160	15	156	22	168	29	177
2	151	9	160	16	173	23	167	30	170
3	161	12	154	18	165	24	171	31	169
4	152	13	152	19	172	25	173	32	176
5	164	14	153	20	169	26	167	33	183
6	153			21	168	27	167	34	178
7	163					28	173	35	174

- 1) Write down a model appropriate to describe the data. Describe each component of the model.
- 2) Estimate the parameters of the model in part 1.

- 3) Determine if there is a linear trend in the lung volume as elevation increases by testing $H_0: -2\mu_1 - \mu_2 - 0\mu_3 + \mu_4 + 2\mu_5 = 0$ vs H_a : (not H_0) (coefficients were obtained from a table of orthogonal polynomials for equally spaced values (Beyer, 1966, p. 367)).
- 4) Determine if there is a quadratic trend in the lung volume as elevation increases by testing $H_0: 2\mu_1 - \mu_2 - 2\mu_3 - \mu_4 + 2\mu_5 = 0$ vs H_a : (not H_0).
- 5) Determine if the assumption of a linear/quadratic response to elevation is appropriate by simultaneously testing the cubic and quadratic trends to be zero by testing $H_0: -1\mu_1 + 2\mu_2 + 0\mu_3 - 2\mu_4 + 1\mu_5 = 0, 1\mu_1 - 4\mu_2 + 6\mu_3 - 4\mu_4 + 1\mu_5 = 0$ vs H_a : (not H_0).
- 6) Use a t -statistic to test $H_0: (\mu_1 + \mu_2 + \mu_3)/3 = (\mu_4 + \mu_5)/2$ vs H_a : (not H_0).
- 7) Use a F -statistic to test $H_0: \mu_1 = \mu_2 = \mu_3$ and $\mu_4 = \mu_5$ vs H_a : (not H_0).

2

One-Way Treatment Structure in a Completely Randomized Design Structure with Heterogeneous Errors

In this chapter, the case is considered where the treatments assigned to the experimental units may affect the variance of the responses as well as the mean. Start with the one-way means model, $y_{ij} = \mu_i + \varepsilon_{ij}$, for $i = 1, 2, \dots, t$, $j = 1, 2, \dots, n_i$. In Chapter 1 it was assumed that the experimental errors all had the same variance; that is, the treatments were expected to possibly change the mean of the population being sampled, but not the variance. In this chapter, some methods are described for analyzing data when the treatments affect the variances as well as the mean. The types of questions that the experimenter should want to answer about the means in this setting are similar to those in Chapter 1. That is,

- 1) Are all means equal?
- 2) Can pairwise comparisons among the means be made?
- 3) Can a test of the hypothesis of the form $\sum_{i=1}^t c_i \mu_i = a$ be tested and can confidence intervals be constructed about $\sum_{i=1}^t c_i \mu_i$?

In addition, there are also questions about the variances that may be of interest, such as

- 1) Are all of the variances equal?
- 2) Are there groupings of the treatments where within a group the variances are equal and between groups the variances are not equal?

Before questions about the means of the model can be answered, an appropriate description of the variances of the treatments must be obtained.

Tests of homogeneity of variances are used to answer questions about the variances of the data from the respective treatments. If there are two treatments, the problem of comparing means when there are unequal variances is usually known as the Behrens-Fisher problem. Also, heterogeneous error variances pose a much more serious problem

when ignored than non-normality of the error variances. The procedures in Chapter 1 are robust with respect to non-normality, but not quite so robust with respect to heterogeneous error variances. In the analyses previously considered, it was assumed that the population variances were all equal, which is a reasonable assumption in many cases. One method for analyzing data when variances are unequal is simply to ignore the fact that they are unequal and calculate the same F -statistics or t -tests that are calculated in the case of equal variances. Surprisingly perhaps, simulation studies have shown that these usual tests are quite good, particularly if the sample sizes are all equal or almost equal. Also, if the larger sample sizes correspond to the treatments or populations with the larger variances, then the tests computed with the equal variance assumption are also quite good. The usual tests are so good, in fact, that many statisticians do not even recommend testing for equal variances. Others attempt to find a transformation that will stabilize the treatment variances, that is, transform the data such that the treatment variances are equal. When the variances are not equal, there are techniques to make comparisons about the means in the framework of the unequal variance model.

Procedures for testing the equality of treatment variances are described for the one-way model and procedures for analyzing the treatment means when the variances are unequal are described in the following sections. These procedures should be used when the usual techniques are suspect. The unequal variance model is described next.

2.1 Model Definitions and Assumptions

The unequal variance model is

$$y_{ij} = \mu_i + \varepsilon_{ij} \quad \text{for } i = 1, 2, \dots, t, \ j = 1, 2, \dots, n_i \text{ and } \varepsilon_{ij} \sim \text{independent } N(0, \sigma_i^2) \quad (2.1)$$

The notation ε_{ij} -independent $N(0, \sigma_i^2)$ means that the errors, ε_{ij} , are all independent, normally distributed and the variance of each normal distribution depends on i and may be different for each population or treatment.

2.2 Parameter Estimation

The best estimates of the parameters in the model are:

$$\hat{\mu}_i = \sum_{j=1}^{n_i} y_{ij}/n_i = \bar{y}_{i.}, \quad i = 1, 2, \dots, t$$

and

$$\hat{\sigma}_i^2 = \frac{\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2}{n_i - 1}, \quad i = 1, 2, \dots, t$$

The sampling distributions associated with the parameter estimates are

$$\hat{\mu}_i \sim \text{independent } N(\mu_i, \sigma_i^2/n_i), \quad i = 1, 2, \dots, t$$

and

$$\frac{(n_i - 1) \hat{\sigma}_i^2}{\sigma_i^2} \sim \text{independent } \chi_{n_i - 1}^2, \quad i = 1, 2, \dots, t$$

These sampling distributions are used as the basis for establishing tests for equality of variances and for providing the analysis of the means when the variances are unequal.

2.3 Tests for Homogeneity of Variances

In this section, five procedures are described for testing the equal variances hypothesis,

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_t^2 \text{ vs } H_a: (\text{not } H_0)$$

Before the analysis of the means is attempted, the equal variance hypothesis should be investigated. If there is not enough evidence to conclude the variances are not equal, then the equal variance model in Chapter 1 can be used to investigate the means. If there is sufficient evidence to believe the variances are unequal, then the procedures described in Section 2.5 should be used to provide an analysis of the means in the unequal variance framework. The recommendation is to use the unequal variance model when the equal variance hypothesis is rejected at $\alpha \leq 0.01$.

2.3.1 Hartley's F-Max Test

The first test described is known as Hartley's *F*-max test (1950). This test requires that all samples be of the same size, that is, $n_1 = n_2 = \dots = n_t$. The test is based on the statistic

$$F_{\max} = \frac{\max_i \{\hat{\sigma}_i^2\}}{\min_i \{\hat{\sigma}_i^2\}}$$

Percentage points of F_{\max} are provided in the Appendix in Table A.1 for $\alpha = 0.05$ and 0.01 . The null hypothesis, H_0 , is rejected if $F_{\max} > F_{\max, \alpha, v, k}$ where $v = n - 1$, the degrees of freedom associated with each of the k individual treatment variances. If the n_i are not all equal, a "liberal" test of H_0 vs H_a can be obtained by taking $v = \max_i \{n_i\} - 1$. This test is liberal in the sense that one is assuming all treatments have the same (maximum) sample size and so you are going to reject the null hypothesis more often than specified by the choice of α . When the sample sizes are not too unequal, this process provides a reasonable test. It also protects one from doing the usual analysis of variance when there is even a remote chance of it being inappropriate. An example illustrating the use of this test is found in Section 2.4.

2.3.2 Bartlett's Test

A second test for testing for homogeneity of variances is a test proposed by Bartlett (1937), which has the advantage of not requiring the n_i to be equal. Bartlett's test statistic is

$$U = \frac{1}{C} \left[v \log_e(\hat{\sigma}^2) - \sum_{i=1}^t v_i \log_e(\hat{\sigma}_i^2) \right] \quad (2.2)$$

where

$$v = n_i - 1, \quad v = \sum_{i=1}^t v_i, \quad \hat{\sigma}^2 = \sum_{i=1}^t v_i \hat{\sigma}_i^2 / v$$

and

$$C = 1 + \frac{1}{3(t-1)} \left[\sum_{i=1}^t \frac{1}{v_i} - \frac{1}{v} \right]$$

The hypothesis of equal variances is rejected if $U > \chi^2_{\alpha,t-1}$. One of the disadvantages of the preceding two tests for homogeneity of variance is that they are quite sensitive to departures from normality as well as to departures from the equal variances assumption. Most of the following tests are more robust to departures from normality.

2.3.3 Levene's Test

Levene (1960) proposed doing a one-way analysis of variance on the absolute values of the residuals from the one-way means or effects model. The absolute values of the residuals are given by $z_{ij} = |y_{ij} - \bar{y}_i|$, $i = 1, 2, \dots, t$; $j = 1, 2, \dots, n_i$. The F -test from the analysis of variance is providing a test of the equality of the treatment means of the absolute values of the residuals. If the means are different, then there is evidence that the residuals for one treatment are on the average larger than the residuals for another treatment. The means of the absolute values of the residuals can provide a guide as to which variances are not equal and a multiple comparison test (see Chapter 3) can be used to make pairwise comparisons among these means. One modification of Levene's test is to use the squared residuals in the analysis of variance.

2.3.4 Brown and Forsythe's Test

Brown and Forsythe (1974) used Levene's process and modified it by doing a one-way analysis of variance on the absolute values of the deviations of the observations from the median of each treatment. The absolute values of the deviations from the medians are given by $u_{ij} = |y_{ij} - y_{i\text{med}}|$, $i = 1, 2, \dots, t$; $j = 1, 2, \dots, n_i$. The F -test from the analysis of variance provides a test of the equality of the treatment means of the absolute values of the deviations. If the means are different, then there is evidence that the deviations for one treatment are on the average larger than the deviations for another treatment. The means of the absolute values of the deviations from the medians can provide a guide as to which variances are not equal as a multiple comparison tests can be used to make pairwise comparisons among these means. This use of the deviations from the medians provides more powerful tests than Levene's when the data are not symmetrically distributed.

2.3.5 O'Brien's Test

O'Brien (1979) computed scores as

$$r_{ij} = [(w + n_i - 2)n_i(y_{ij} - \bar{y}_i)^2 - w\hat{\sigma}_i^2(n_i - 1)]/[(n_i - 1)(n_i - 2)] \quad (2.3)$$

where w is a weight parameter. The procedure is to carry out an analysis of variance on the computed score values. When $w = 0.5$, the means of the scores are the sample variances, $\hat{\sigma}_i^2$, thus the comparison of the means of the scores is a comparison of the variances of the data.

There are several other procedures that can be used to test the equality of variances or the equality of scale parameters using parametric and nonparametric methods (Conover et al., 1981; Olejnik and Algina, 1987). McGaughey (2003) proposes a test that uses the concept of data depth and applies the procedure to univariate and multivariate populations. Data depth is beyond the scope of this book.

2.3.6 Some Recommendations

Conover et al. (1981) and Olejnik and Algina (1987) conducted simulation studies of homogeneity of variance tests that included the ones above as well as numerous others. The studies indicate that no test is robust and most powerful for all situations. Levene's test was one of the better tests studied by Conover et al. O'Brien's test seems to provide an appropriate size test without losing much power according to Olejnik and Algina. The Brown–Forsythe test seems to be better when distributions have heavy tails. Based on their results, we make the following recommendations:

- 1) If the distributions have heavy tails, use the Brown–Forsythe test.
- 2) If the distributions are somewhat skewed, use the O'Brien test.
- 3) If the data are nearly normally distributed, then any of the tests are appropriate, including Bartlett's and Hartley's tests.

Levene's and O'Brien's tests can easily be tailored for use in designed experiments that involve more than one factor, including an analysis of covariance (Milliken and Johnson, 2002). Levene's, O'Brien's and Brown–Forsythe's tests were shown to be nearly as good as Bartlett's and Hartley's tests for normally distributed data, and superior to them for non-normally distributed data. Conover et al. and Olejnik and Algina discuss some nonparametric tests, but they are more difficult to calculate and the above recommended tests perform almost as well. An example follows where each of the tests for equality of variances is demonstrated.

2.4 Example—Drugs and Errors

The data in Table 2.1 are from a paired-association learning task experiment performed on subjects under the influence of two possible drugs. Group 1 is a control group (no drug), group 2 was given drug 1, group 3 was given drug 2, and group 4 was given both drugs.

TABLE 2.1

Data from Paired-Association Learning Task Experiment

	No Drug	Drug 1	Drug 2	Drugs 1 and 2
	1	12	12	13
	8	10	4	14
	9	13	11	14
	9	13	7	17
	4	12	8	11
	0	10	10	14
	1	—	12	13
	—	—	5	14
<i>n</i>	7	6	8	8
Sum	32	70	69	110
Median	4	12	9	14
Mean	4.5714	11.6667	8.6250	13.750
Variance	16.2857	1.8667	9.6964	2.786

The sample sizes, sums, medians, means and variances of each group's data are included in Table 2.1.

The F -max statistic is $F_{\max} = 16.286/1.867 = 8.723$. The liberal 5% critical point is obtained from Table A.1 with $k = t = 4$ and $v = 7$. The critical point is 8.44 and since $8.723 > 8.44$, one rejects $H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_t^2$ versus H_a : (not H_0) with significance level 0.05, but cannot reject at the $\alpha = 0.01$ level.

The computations for Bartlett's test are:

$$C = 1 + \frac{1}{3 \times 3} \left(\frac{1}{6} + \frac{1}{5} + \frac{1}{7} + \frac{1}{7} - \frac{1}{25} \right)$$

and

$$\hat{\sigma}^2 = \frac{6(16.2857) + 5(1.8667) + 7(9.6964) + 7(2.7860)}{25} = 7.7769$$

Thus

$$\begin{aligned} U &= \frac{1}{C} \left(v \log_e \hat{\sigma}^2 - \sum_{i=1}^4 v_i \log_e \hat{\sigma}_i^2 \right) \\ &= \frac{1}{1.068} [25 \log_e (7.7769) - 6 \log_e (16.2857) - 5 \log_e (1.8667) \\ &\quad - 7 \log_e (9.6964) - 7 \log_e (2.7860)] \\ &= 7.8111 \end{aligned}$$

The asymptotic sampling distribution associated with U is that of a chi-square distribution based on three degrees of freedom. The significance level of the test is 0.0501

and one would again conclude that the variances are unequal at an approximate 5% significance level.

The computations for Levene's test begin with the computation of the residuals or the deviations of the observations from the treatment means. Next the absolute values of the residuals are computed as illustrated in Table 2.2. Finally, a one-way analysis of variance is carried out on these absolute values of the residuals. The value of the resulting F -statistic is 6.97, which is based on 3 and 25 degrees of freedom. The observed significance level of Levene's test is 0.0015. The squared deviations or squared residuals version of Levene's test can be obtained by squaring the items in Table 2.2 before doing the analysis of variance. In this case, the value of the F -statistic is 7.36 and the observed significance level is 0.0011 (also based on 3 and 25 degrees of freedom).

The Brown–Forsythe test statistic is obtained by computing the absolute value of the deviations of the observations from the treatment median (medians are in Table 2.1). Table 2.3 contains the absolute values of the deviations from the medians. Next, the one-way analysis of variance provides an F -statistic of 5.49 and the observed significance level is 0.0049 (also based on 3 and 25 degrees of freedom).

Table 2.4 contains the values of r_{ij} computed using Equation 2.3 with $w = 0.5$. The O'Brien test statistic is obtained by carrying out an analysis of variance. The value of the F -statistic

TABLE 2.2

Values of $z_{ij} = |y_{ij} - \bar{y}_{i\cdot}|$ for Computing Levene's Test Where
 y_{ij} Values are from Table 2.1

No Drug	Drug 1	Drug 2	Drugs 1 and 2
3.571	0.333	3.375	0.750
3.429	1.667	4.625	0.250
4.429	1.333	2.375	0.250
4.429	1.333	1.625	3.250
0.571	0.333	0.625	2.750
4.571	1.667	1.375	0.250
3.571	—	3.375	0.750
—	—	3.625	0.250

TABLE 2.3

Absolute Values of Deviations of the Observations
from the Treatment Medians

No Drug	Drug 1	Drug 2	Drugs 1 and 2
3	0	3	1
4	2	5	0
5	1	2	0
5	1	2	3
0	0	1	3
4	2	1	0
3	—	3	1
—	—	4	0

TABLE 2.4Scores Using $w = 0.5$ for O'Brien's Test

Obr1	Obr2	Obr3	Obr4
14.740	-0.083	13.295	0.464
13.457	3.517	25.676	-0.155
23.540	2.167	6.176	-0.155
23.540	2.167	2.461	12.845
-1.210	-0.083	-0.324	9.131
25.190	3.517	1.533	-0.155
14.740	—	13.295	0.464
—	—	15.461	-0.155

is 6.30 and the observed significance level is 0.0025. The value of the F -statistic using $w = 0.7$ (computations not shown) is 5.90 and the observed significance level is 0.0035. There are 3 and 25 degrees of freedom associated with each of O'Brien's tests.

Each of the test statistics indicates that there is sufficient evidence to conclude that the variances are not equal. The group means of the absolute values of the residuals are shown in Table 2.5. Pairwise comparisons among these treatment absolute residual means are shown in Table 2.6. The means of the absolute values of the residuals for no drug and drug 2 are not different, for drug 1 and drugs 1 and 2 are not different, but there are differences between these two sets. A simple model with two variances could be used to continue the analysis of the treatment means. Using a simple variance model will improve the power of some of the tests about the means. The two variance model and the corresponding comparisons of means will follow the discussion of the analysis using four variances.

TABLE 2.5

Means of the Absolute Values of the Residuals

Group	Estimate	Standard Error	df	t-Value	Pr > t
Both drugs	1.0625	0.4278	25	2.48	0.0201
Drug 1	1.1111	0.4940	25	2.25	0.0336
Drug 2	2.6250	0.4278	25	6.14	<0.0001
No drug	3.5102	0.4574	25	7.67	<0.0001

TABLE 2.6

Pairwise Comparisons between the Group Means of the Absolute Values of the Residuals

Group	_Group	Estimate	Standard Error	df	t-Value	Pr > t
Both drugs	Drug 1	-0.04861	0.6535	25	-0.07	0.9413
Both drugs	Drug 2	-1.5625	0.6050	25	-2.58	0.0161
Both drugs	No drug	-2.4477	0.6263	25	-3.91	0.0006
Drug 1	Drug 2	-1.5139	0.6535	25	-2.32	0.0290
Drug 1	No drug	-2.3991	0.6732	25	-3.56	0.0015
Drug 2	No drug	-0.8852	0.6263	25	-1.41	0.1699

2.5 Inferences on Linear Combinations

The problems of testing hypotheses about and constructing confidence intervals for an arbitrary linear combination of the treatment means, $\sum_{i=1}^t c_i \mu_i$, are discussed in this section when the variances σ_i^2 are too unequal to apply the tests and confidence intervals discussed in Chapter 1. It is recommended that you use the procedures in this section and the next if the equality of variance hypothesis is rejected at the 0.01 or 1% level. If there is not sufficient evidence to believe that the variances are unequal, then one can use the results in Chapter 1 to make inferences about the treatment means.

The best estimate of $\sum_{i=1}^t c_i \mu_i$ is $\sum_{i=1}^t c_i \hat{\mu}_i$ and the sampling distribution is

$$\sum_{i=1}^t c_i \hat{\mu}_i \sim N\left(\sum_{i=1}^t c_i \mu_i, \sum_{i=1}^t c_i^2 \sigma_i^2 / n_i\right)$$

and thus,

$$z = \frac{\sum_{i=1}^t c_i \hat{\mu}_i - \sum_{i=1}^t c_i \mu_i}{\sqrt{\sum_{i=1}^t c_i^2 \sigma_i^2 / n_i}} \sim N(0, 1)$$

An obvious statistic to use for making inferences about $\sum_{i=1}^t c_i \mu_i$, when the variances are not known and are unequal, is

$$z = \frac{\sum_{i=1}^t c_i \hat{\mu}_i - \sum_{i=1}^t c_i \mu_i}{\sqrt{\sum_{i=1}^t c_i^2 \hat{\sigma}_i^2 / n_i}}$$

If the n_i corresponding to nonzero c_i are all very large, one can reasonably assume that Z has an approximate $N(0, 1)$ distribution, and hence Z can be used to make inferences about $\sum_{i=1}^t c_i \mu_i$. In this case, an approximate $(1 - \alpha)100\%$ confidence interval for $\sum_{i=1}^t c_i \mu_i$ is provided by

$$\sum_{i=1}^t c_i \hat{\mu}_i \pm z_{\alpha/2} \sqrt{\sum_{i=1}^t c_i^2 \hat{\sigma}_i^2 / n_i}$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ critical point of the standard normal probability distribution.

To test $H_0: \sum_{i=1}^t c_i \mu_i = a$ vs $H_a: \sum_{i=1}^t c_i \mu_i \neq a$, where a is a specified constant, one could calculate

$$z = \frac{\sum_{i=1}^t c_i \hat{\mu}_i - a}{\sqrt{\sum_{i=1}^t c_i^2 \hat{\sigma}_i^2 / n_i}}$$

and if $|z| > z_{\alpha/2}$, then reject H_0 at a significance level of α .

In other instances, note that z can be written as

$$z = \frac{\left(\sum_{i=1}^t c_i \hat{\mu}_i - \sum_{i=1}^t c_i \mu_i \right) / \sqrt{\sum_{i=1}^t c_i^2 \sigma_i^2 / n_i}}{\sqrt{\sum_{i=1}^t c_i^2 \hat{\sigma}_i^2 / n_i} / \sqrt{\sum_{i=1}^t c_i^2 \sigma_i^2 / n_i}}$$

The numerator of z has a standard normal distribution and the numerator and denominator of z are independently distributed. The distribution of z could be approximated by a $t(v)$ distribution if v could be determined such that

$$V = v \times \frac{\sum_{i=1}^t c_i^2 \hat{\sigma}_i^2 / n_i}{\sum_{i=1}^t c_i^2 \sigma_i^2 / n_i}$$

is approximately distributed as $\chi^2(v)$. In order to get a good chi-square approximation to the distribution of V when the variances are unequal, select a chi-square distribution that has the same first two moments as V . That is, to find v for the case of unequal variances, find v so that the moments of V are equal to the first two moments of a $\chi^2(v)$ distribution (this is known as Satterthwaite's method). This results in determining that the approximate number of degrees of freedom is

$$v = \frac{\left(\sum_{i=1}^t c_i^2 \sigma_i^2 / n_i \right)^2}{\sum_{i=1}^t [c_i^4 \sigma_i^4 / n_i^2 (n_i - 1)]}$$

Unfortunately, since v depends on $\sigma_1^2, \sigma_2^2, \dots, \sigma_t^2$ it cannot be determined exactly. The usual procedure is to estimate v by

$$\hat{v} = \frac{\left(\sum_{i=1}^t c_i^2 \hat{\sigma}_i^2 / n_i \right)^2}{\sum_{i=1}^t [c_i^4 \hat{\sigma}_i^4 / n_i^2 (n_i - 1)]} \quad (2.4)$$

Summarizing, one rejects $H_0: \sum_{i=1}^t c_i \mu_i = a$ vs $H_a: \sum_{i=1}^t c_i \mu_i \neq a$, if

$$|t_c| = \frac{\left| \sum_{i=1}^t c_i \hat{\mu}_i - a \right|}{\sqrt{\sum_{i=1}^t c_i^2 \hat{\sigma}_i^2 / n_i}} > t_{\alpha/2, \hat{v}}$$

where \hat{v} is determined using Equation 2.4. An approximate $(1 - \alpha)100\%$ confidence interval for $\sum_{i=1}^t c_i \mu_i$ is given by

$$\sum_{i=1}^t c_i \hat{\mu}_i \pm t_{\alpha/2, \hat{v}} \sqrt{\sum_{i=1}^t c_i^2 \hat{\sigma}_i^2 / n_i}$$

Unfortunately, every time one wants to test a new hypothesis or construct another confidence interval, the degrees of freedom \hat{v} must be re-estimated. It can be shown that $n_* - 1 \leq \hat{v} \leq t(n^* - 1)$ where $n_* = \min\{n_1, n_2, \dots, n_t\}$ and $n^* = \max\{n_1, n_2, \dots, n_t\}$. Thus, if $|t_c| > t_{\alpha/2, n_* - 1}$, one can be assured that $|t_c| > t_{\alpha/2, \hat{v}}$, and if $|t_c| < t_{\alpha/2, t(n^* - 1)}$, one can be assured that $|t_c| < t_{\alpha/2, \hat{v}}$. In these cases, one can avoid calculating \hat{v} . When $t_{\alpha/2, t(n^* - 1)} < |t_c| < t_{\alpha/2, n^* - 1}$ the value of \hat{v} must be calculated in order to be sure whether one should reject or fail to reject the null hypothesis being tested. For confidence intervals, \hat{v} should always be calculated. Next, the preceding results are demonstrated with the drug errors example.

2.6 Example—Drugs and Errors (Continued)

Consider the data in Table 2.1, and suppose the experimenter is interested in answering the following questions:

- 1) On average, do drugs have any effect on learning at all?
- 2) Do subjects make more errors when given both drugs than when given only one?
- 3) Do the two drugs differ in their effects on the number of errors made?

To answer the first question, one might test the hypothesis that the mean of the three drug groups is equal to the control mean. That is, one would test

$$H_{01}: l_1 = \mu_1 - \frac{(\mu_2 + \mu_3 + \mu_4)}{3} = 0 \text{ vs } H_{a1}: l_1 \neq 0$$

The estimate of this linear combination is

$$\hat{l}_1 = \hat{\mu}_1 - \frac{\hat{\mu}_2 + \hat{\mu}_3 + \hat{\mu}_4}{3} = 4.571 - \frac{1}{3}(34.042) = -6.776$$

and the estimate of the corresponding standard error of \hat{l}_1 is

$$\begin{aligned} s.e.(\hat{l}_1) &= \sqrt{\sum_{i=1}^4 \left(\frac{c_i^2 \hat{\sigma}_i^2}{n_i} \right)} \\ &= \sqrt{\frac{\hat{\sigma}_1^2}{7} + \frac{1}{9} \left(\frac{\hat{\sigma}_2^2}{6} \right) + \frac{1}{9} \left(\frac{\hat{\sigma}_3^2}{8} \right) + \frac{1}{9} \left(\frac{\hat{\sigma}_4^2}{8} \right)} = \sqrt{2.535} = 1.592 \end{aligned}$$

The approximate degrees of freedom associated with this estimated standard error are obtained by using

$$\sum_{i=1}^4 \frac{c_i^4 \hat{\sigma}_i^4}{n_i^2(n_i - 1)} = 0.9052$$

so that

$$\hat{v} = \frac{(2.535)^2}{0.9052} = 7.10$$

The value of the test statistic is $t_c = -6.776/1.992 = -4.256$ with the observed significance level $\hat{\alpha} = 0.0038$.

A 95% confidence interval for l_1 is

$$\hat{l}_1 \pm t_{\alpha/2, \hat{v}} \times \widehat{s.e.}(\hat{l}_1) = -6.776 \pm (2.365)(1.592)$$

which simplifies to

$$-10.54 < \mu_1 - \frac{\mu_2 + \mu_3 + \mu_4}{3} < -3.01$$

Next test to see if the mean of the group given both drugs is equal to the mean of the average of the means of the two groups given a single drug. That is, test

$$H_{02}: l_2 = \mu_4 - \frac{\mu_2 + \mu_3}{2} = 0 \text{ vs } H_{a2}: l_2 \neq 0$$

The estimate of this linear combination is

$$\hat{l}_2 = \hat{\mu}_4 - \frac{\hat{\mu}_2 + \hat{\mu}_3}{2} = 3.6042$$

and its estimated standard error is

$$\begin{aligned} \widehat{s.e.}(\hat{l}_2) &= \sqrt{\sum_{i=1}^4 \left(\frac{c_i^2 \hat{\sigma}_i^2}{n_i} \right)} \\ &= \sqrt{\frac{1}{4} \left(\frac{\hat{\sigma}_2^2}{6} \right) + \frac{1}{4} \left(\frac{\hat{\sigma}_3^2}{8} \right) + \left(\frac{\hat{\sigma}_4^2}{8} \right)} = \sqrt{0.7290} = 0.8538 \end{aligned}$$

The value of the test statistic is $t_c = 3.6042/0.8538 = 4.221$, which is significant at $\alpha = 0.01$ since $|t_c| > t_{0.005, 5}$. In this case, the value of \hat{v} need not be computed using $n_* - 1$ as the approximating degrees of freedom. The computed value of \hat{v} is 16.8, which would be needed if one wanted to construct a confidence interval about l_2 .

Finally, to test the hypothesis to see if the two drug means differ, test $H_0: l_3 = \mu_2 - \mu_3 = 0$ vs $H_a: l_3 = \mu_2 - \mu_3 \neq 0$. The estimate of this linear combination is $\hat{l}_3 = \hat{\mu}_2 - \hat{\mu}_3 = 3.042$ and its estimated standard error is

$$\widehat{s.e.}(\hat{l}_3) = \sqrt{\sum_{i=1}^4 \left(\frac{c_i^2 \hat{\sigma}_i^2}{n_i} \right)} = \sqrt{\left(\frac{\hat{\sigma}_2^2}{6} \right) + \left(\frac{\hat{\sigma}_3^2}{8} \right)} = \sqrt{1.523} = 1.234$$

The approximate number of degrees of freedom is computed using

$$\sum_{i=1}^4 \frac{c_i^4 \hat{\sigma}_i^4}{n_i^2(n_i - 1)} = 0.229$$

so that

$$\hat{v} = \frac{(1.523)^2}{0.229} = 10.1$$

Thus, $t_c = 3.042/1.234 = 2.465$, which has an observed significance level of $\alpha = 0.0334$.

2.7 General Satterthwaite Approximation for Degrees of Freedom

The Satterthwaite approximation to the number of degrees of freedom associated with estimated standard error is obtained from

$$v = \frac{2 * (E\{\widehat{s.e.}(\hat{l})]^2\})^2}{\text{Var}\{\widehat{s.e.}(\hat{l})]^2\}}$$

where $\widehat{s.e.}(\hat{l})]^2$ is used to estimate $E[s.e.(\hat{l})]^2$ and the $\text{Var}[\widehat{s.e.}(\hat{l})]^2$ is estimated by $\sum_{i=1}^t c_i^4 \hat{\sigma}_i^4 / [n_i^2(n_i - 1)]$. For more complex models, $\text{Var}[\widehat{s.e.}(\hat{l})]^2$ can be approximated by using a first-order Taylor's series (Kendall and Stuart, 1952) as $\mathbf{q}' \mathbf{M} \mathbf{q}$ where \mathbf{M} is the estimated asymptotic covariance matrix of the estimates of the variances and the elements of the vector \mathbf{q} are the first derivatives of $E[s.e.(\hat{l})]^2$ with respect to the individual variances, that is,

$$q_i = \frac{\partial E[(s.e.(\hat{l})]^2}{\partial \sigma_i^2}, \quad i = 1, 2, \dots, t$$

The q_i are evaluated at the estimated values of each treatment's variances (Montgomery and Runger, 1993, 1994). When the data from each of the samples are normally distributed, then

$$\frac{(n_i - 1)\hat{\sigma}_i^2}{\sigma_i^2}$$

is distributed as a central chi-square random variable. Thus $E(\hat{\sigma}_i^2) = \sigma_i^2$ and $\text{Var}(\hat{\sigma}_i^2) = 2\sigma_i^4/(n_i - 1)$. Let the linear combination of interest be $l = \sum_{i=1}^t c_i \mu_i$, which has variance $\sigma_l^2 = \sum_{i=1}^t c_i^2 \sigma_i^2 / n_i$. The partial derivative of σ_l^2 with respect to σ_i^2 is

$$\frac{\partial \sigma_l^2}{\partial \sigma_i^2} = \frac{c_i^2}{n_i}$$

The approximate variance of σ_l^2 obtained using the Taylor's series first-order approximation is

$$\text{Var}(\sigma_l^2) = \sum_{i=1}^t \left[\left[\frac{c_i^2}{n_i} \right]^2 \left[\frac{2\sigma_i^4}{n_i - 1} \right] \right]^2$$

The next step is to replace the population variances with their corresponding sample estimates providing the approximating degrees of freedom

$$\hat{v} = \frac{2 * (E[\widehat{s.e.(\hat{l})}]^2))^2}{\text{Var}[\widehat{s.e.(\hat{l})}]^2} = \frac{\left(\sum_{i=1}^t c_i^2 \hat{\sigma}_i^2 \right)^2}{\sum_{i=1}^t c_i^4 \hat{\sigma}_i^4 / [n_i^2(n_i - 1)]}$$

the same as that provided by the Satterthwaite approximation above.

2.8 Comparing All Means

As previously stated, the usual F -test is very robust when the variances are unequal, provided that the sample sizes are nearly equal or provided that the larger sample sizes correspond to the samples from populations with the larger differences of variances. In this section, two additional tests of the hypothesis of equal means are provided. The first test of the equal means hypothesis, $H_0: \mu_1 = \mu_2 = \dots = \mu_t$ vs H_a : (not H_0), is given by Welch (1951), and is known as Welch's test. Define weights $W_i = n_i / \hat{\sigma}_i^2$, let $\bar{y}^* = \sum_{i=1}^t W_i \bar{y}_i / \sum_{i=1}^t W_i$ be a weighted average of the sample means, and let

$$\Lambda = \sum_{i=1}^t \frac{(1 - W_i/W.)^2}{n_i - 1}$$

where $W. = \sum_{i=1}^t W_i$. Then Welch's test statistic is

$$F_c = \frac{\sum_{i=1}^t W_i \frac{(\bar{y}_i - \bar{y}^*)^2}{(t-1)}}{1 + 2(t-1)\Lambda/(t^2 - 1)} \quad (2.5)$$

which has an approximate F -distribution with numerator and denominator degrees of freedom, $v_1 = t - 1$ and $v_2 = (t^2 - 1)/3\Lambda$, respectively. Thus, the null hypothesis $H_0: \mu_1 = \mu_2 = \dots = \mu_t$ is rejected if $F_c > F_{\alpha, v_1, v_2}$. The numerator of Equation 2.5 can also be computed as

TABLE 2.7
Quantities for Computing Welch's Test

<i>i</i>	Drug 1	Drug 2	Drug 3	Drug 4
n_i	7	6	8	8
\bar{y}_i	4.5714	11.6667	8.62500	13.7500
$\hat{\sigma}_i^2$	16.2857	1.8667	9.69643	2.7857
W_i	0.4298	3.2143	0.82505	2.8718

$[\sum_{i=1}^t (W_i \bar{y}_i^2) - W_i \bar{y}_i^{*2}] / (t - 1)$. The procedure is demonstrated using the data from Section 2.4 and the preliminary computations are provided in Table 2.7.

From the above information compute $W_i = 7.341$, $\bar{y}_i^* = 11.724$,

$$\Lambda = \frac{(1 - 0.430/7.341)^2}{6} + \frac{(1 - 3.214/7.341)^2}{5} + \frac{(1 - 0.825/7.341)^2}{7} + \frac{(1 - 2.872/7.341)^2}{7} = 0.376$$

and $\sum_{i=1}^t W_i \bar{y}_i^2 - \bar{W}_i \bar{y}_i^{*2} = 1050.8069 - 1009.0954 = 43.7114$.

The value of Welch's test statistic is

$$F_c = \frac{41.7114/3}{1 + 2 \times 2 \times 0.376/15} = \frac{13.9038}{1.1003} = 12.6355$$

with $v_1 = 3$ and $v_2 = 15/(3 \times 0.376) = 13.283$ degrees of freedom. The observed significance probability corresponding to F_c is $\hat{\alpha} = 0.00035$. For comparison purposes, the usual F -statistic is $F_c = 14.91$ with 3 and 25 degrees of freedom. Welch's test can be obtained using SAS®-GLM by specifying WELCH as an option on the MEANS statement. Table 2.8 contains the

TABLE 2.8

SAS-GLM Code to Provide the Brown–Forsythe's Test of Equality of Variances and to Provide the Welch Test of Equal Means with the Unequal Variance Model

```
proc glm data=task;
class group;
model errors=group;
means group/HOVTEST=BF WELCH;
format group druggrps.;
```

Welch's Test

Source	df	F-Value	Pr > F
Group	3	12.64	0.0003
Error	13.2830		

Brown and Forsythe's Test for Homogeneity of Errors Variance ANOVA of Absolute Deviations from Group Medians

Source	df	Sum of Squares	Mean Square	F-Value	Pr > F
Group	3	31.3762	10.4587	5.49	0.0049
Error	25	47.5893	1.9036		

GLM code used to provide BF test for equality of variances and Welch's test for equality of means. The important parts of the output are in the second part of Table 2.8. Other tests for equality of variances can be obtained by specifying O'Brien, Levene or Bartlett.

The second procedure for testing the equality of the treatment means is obtained from generalizing the process of testing a hypothesis about a set of linear combinations of the μ_i . Suppose a hypothesis is formed involving r independent linear combinations of the μ_i , such as $H_0: \sum_{i=1}^t c_{1i}\mu_i = 0, \sum_{i=1}^t c_{2i}\mu_i = 0, \dots, \sum_{i=1}^t c_{ri}\mu_i = 0$ vs H_a ; (not H_0). Let C be a $r \times t$ matrix where the k th row contains the coefficients of the k th linear combination. If one assumes the data from each of the populations or treatments are normally distributed, then the joint sampling distribution of the vector of treatment means is $\hat{\mu} \sim N[\mu, V]$ where V is a diagonal matrix whose i th diagonal element is σ_i^2/n_i . The joint sampling distribution of the set of linear combinations $C\mu$ is $C\hat{\mu} \sim N[C\mu, CVC']$. The sum of squares due to deviations from the null hypothesis is $SSH_0 = [C\hat{\mu}]' [C\hat{V}C']^{-1} [C\hat{\mu}]$, which is asymptotically distributed as a chi-square distribution with r degrees of freedom. An approximate small sample size statistic is $F_c = SSH_0/r$ with the approximating distribution being F with r and v degrees of freedom where v needs to be approximated (Fai and Cornelius, 1996; SAS Institute, Inc., 1999, p. 2118). The computation of the approximate degrees of freedom starts with carrying out a spectral decomposition on $C\hat{V}C' = QDQ'$ where D is an $r \times r$ diagonal matrix having the characteristic roots of $C\hat{V}C'$ as diagonal elements and where Q is a $r \times r$ orthogonal matrix of the corresponding characteristic vectors of $C\hat{V}C'$. Let z'_k be the k th row of QC , and let

$$v_k = \frac{2(d_k)^2}{b'_k M b_k}$$

where d_k is the k th diagonal element of D , b_k contains the partial derivatives of $z'_k V z_k$ with respect to each of the variance parameters in V evaluated at the estimates of the variances, and M is the asymptotic covariance of the vector of variances. Let

$$S = \sum_{k=1}^r \frac{v_k}{v_k - 2} I[v_k > 2]$$

where $I[v_k > 2]$ is an indicator function with the value of 1 when $v_k > 2$ and 0 otherwise. The approximate denominator degrees of freedom for the distribution of F_c are

$$v = \begin{cases} \frac{2S}{S-r} & \text{if } S > r \\ 0 & \text{if } S \leq r \end{cases}$$

The above process can be used to provide a test of the equal means hypothesis by selecting a set of $t - 1$ linearly independent contrasts of the μ_i .

The SAS-Mixed procedure implements a version of this approximation to the denominator degrees of freedom associated with an approximate F statistic with multiple degrees of freedom in the numerator. SAS-Mixed can be used to fit models with unequal variances per treatment group or unequal variances in some other prespecified pattern using the REPEATED statement and specifying the GROUP = option. The Mixed code in Table 2.9 was used to fit the unequal variance model to the data in Table 2.1. The REPEATED statement is used to specify that a different variance (each value of group) is to be estimated for

TABLE 2.9

SAS-Mixed Code to Fit the Unequal Variance Model to the Data in Table 2.1

```
proc mixed cl covtest data=task;
  class group;
  model errors=group/ddfm=kr;
  repeated/group=group;
  estimate "part (1)" group -1 -1 -1 3/divisor=3 cl alpha=0.05;
  estimate "part (2)" group 2 -1 -1 0/divisor=2 cl alpha=0.05;
  estimate "part (3)" group 0 1 -1 0/cl alpha=0.05;
  lsmeans group/diff cl;
```

each treatment. The three Estimate statements are used to provide the computations corresponding to the three questions in Section 2.6.

The results from the Mixed procedure are given in Table 2.10, where the Covariance Parameter Estimates are the estimates of the four treatment variances, AIC in the Fit Statistics is the Akaike Information Criteria (Akaike, 1974), the Null Model Likelihood Ratio Test provides a test of the equal variance hypothesis, the type III tests of fixed effects provides the test of the equal means hypothesis using the second statistic and the corresponding

TABLE 2.10

Results of Fitting the Unequal Variance Model to the Data in Table 2.1

Covariance Parameter Estimates

Covariance Parameter	Group	Estimate	Standard Error	Z-Value	Pr Z	α	Lower	Upper
Residual	Both drugs	2.7857	1.4890	1.87	0.0307	0.05	1.2178	11.5394
Residual	Drug 1	1.8667	1.1806	1.58	0.0569	0.05	0.7273	11.2286
Residual	Drug 2	9.6964	5.1830	1.87	0.0307	0.05	4.2388	40.1658
Residual	No drug	16.2857	9.4026	1.73	0.0416	0.05	6.7625	78.9710

Fit Statistics

AIC (smaller is better) 129.8

Null Model Likelihood Ratio Test

df	Chi-Square	Pr > Chi-Square
3	8.34	0.0394

Type III Tests of Fixed Effects

Effect	Num df	Den df	F-Value	Pr > F
group	3	11.8	12.53	0.0006

Estimates

Label	Estimate	Standard Error	df	t-Value	Pr > t	α	Lower	Upper
Part 1	-6.7758	1.5920	7.1	-4.26	0.0036	0.05	-10.5299	-3.0217
Part 2	3.6042	0.8538	16.8	4.22	0.0006	0.05	1.8011	5.4073
Part 3	3.0417	1.2342	10.1	2.46	0.0332	0.05	0.2962	5.7871

TABLE 2.11

Estimates of the Drug Group Means and Pair Wise Comparisons Using the Unequal Variance Model

Least Squares Means

Effect	Group	Estimate	Standard Error	df	t-Value	Pr > t	α	Lower	Upper
Group	Both drugs	13.7500	0.5901	7	23.30	<0.0001	0.05	12.3546	15.1454
Group	Drug 1	11.6667	0.5578	5	20.92	<0.0001	0.05	10.2329	13.1005
Group	Drug 2	8.6250	1.1009	7	7.83	0.0001	0.05	6.0217	11.2283
Group	No drug	4.5714	1.5253	6	3.00	0.0241	0.05	0.8392	8.3037

Differences of Least Squares Means

Effect	Group	_Group	Estimate	Standard		df	t-Value	Pr > t	α	Lower	Upper
				Error	df						
Group	Both drugs	Drug 1	2.0833	0.8120	11.9	2.57	0.0249	0.05	0.3117	3.8550	
Group	Both drugs	Drug 2	5.1250	1.2491	10.7	4.10	0.0018	0.05	2.3668	7.8832	
Group	Both drugs	No drug	9.1786	1.6355	7.78	5.61	0.0006	0.05	5.3886	12.9685	
Group	Drug 1	Drug 2	3.0417	1.2342	10.1	2.46	0.0332	0.05	0.2962	5.7871	
Group	Drug 1	No drug	7.0952	1.6241	7.55	4.37	0.0027	0.05	3.3109	10.8796	
Group	Drug 2	No drug	4.0536	1.8811	11.3	2.15	0.0536	0.05	-0.07507	8.1822	

approximate degrees of freedom for the denominator, and the Estimates contain the results corresponding to the three questions in Section 2.6, where t -statistics, approximate denominator degrees of freedom, and 95% confidence intervals are provided. Table 2.11 contains the estimated treatment means with their corresponding estimated standard errors. The denominator degrees of freedom are the degrees of freedom corresponding to their respective variances. The second part of Table 2.11 contains the pairwise comparisons of the treatment means including the approximate denominator degrees of freedom for each comparison. This model could be simplified by using one variance for drug 1 and both drugs and one variance for drug 2 and no drug. This can be accomplished by defining a variable, say T , to be 1 for drug 1 and both drugs and 0 for the other two treatments. Then place T in the class statement and use Repeated/Group = T; in the model specification. The estimates of the two variances are 2.4028 and 12.7376 and the AIC is 126.4, which is a smaller AIC value than that for the four variance model, indicating the two variance model is adequate to describe the data. Using a model with fewer variances in the model specification provides more degrees of freedom for the respective standard errors and thus provides more powerful tests of hypotheses concerning the fixed effects in the model.

2.9 Concluding Remarks

In summary, for comparing all means, the following are recommended:

- 1) If the homogeneity of variance test is not significant at the 1% level, do the usual analysis of variance test.

- 2) If the homogeneity of variance test is significant at the 1% level use either Welch's test or the mixed models test and the corresponding approximate denominator degrees of freedom.
- 3) If the homogeneity of variance is significant at the 1% level, use the AIC to determine if a simpler or fewer number of variances can be used to adequately describe the data in order to increase the power of tests concerning the means.

Many text books and articles have been written about using transformations on data in order to achieve equal treatment variances so that the usual analysis of variance can be used to compare the treatments. With the ability to fit an unequal variance model to provide estimated standard errors of means and comparisons of means, many situations will not require the use of transformations. One major benefit of not having to use a transformation to achieve equal variances is that the units of the means are in the units of measurement, thus simplifying interpretations.

This chapter contains discussion about the statistical analysis of a one-way analysis of variance model with heterogeneous errors. The discussion included several statistical tests for determining homogeneity of the error variances and recommendations on when to use each test. Procedures appropriate for making statistical inferences about the effects of different treatments upon discovering heterogeneous error variances as well as examples illustrating the use of these procedures were also reviewed.

2.10 Exercises

- 2.1 The following data are body temperatures of calves that were vaccinated and then challenged to determine if the vaccination protected the animal. Test the equality of variances of the treatment groups using two or more techniques. Based on the results of the test of equality of variances, test the equality of the treatment means using both Welch's and the mixed model F-statistics and make all pairwise comparisons.

Data for Exercise 2.1

Vaccine A	Vaccine B	Vaccine C	Vaccine D	Vaccine E	Vaccine F	Vaccine G
101.5	96.3	101.8	97.3	97.5	96.9	97.3
100.5	97.2	97.4	96.8	96.4	97.1	100.7
104.5	99.3	104.9	97.1	98.6	96.8	103.3
102.3	98.0	104.0	97.0	96.6	97.0	100.2
100.6	97.6	103.7	97.1		96.2	103.5
97.7	96.8	104.5	96.9		96.6	
	99.1	100.4	96.1			
	96.7	102.2	96.3			
	96.4	100.2	96.7			
			97.1			

- 2.2 Use the data in Table 1.1 and test the equality of variances using several of the methods described in Section 2.2. What is your conclusion?
- 2.3 The data in the following table are times required for a student to dissolve a piece of chocolate candy in their mouth. Each time represents one piece of candy

dissolved by one student. Provide a detailed analysis of the data set and provide tests of the following hypotheses:

- 1) The mean of the Blue Choc = the mean of the Red Choc.
- 2) The mean of the Buttons = the mean of the means of the Blue Choc and Red Choc.
- 3) The mean of the ChocChip = the mean of the WchocChip.
- 4) The mean of the Small Choc = $\frac{1}{2}$ the mean of the means of the Blue Choc and Red Choc.
- 5) The mean of the Blue Choc and Red Choc = the mean of the ChocChip and WchocChip.

Data for Exercise 2.3

Buttons	Blue Choc	Small Choc	ChocChip	WChocChip	Red Choc
69	57	28	52	35	47
76	41	27	50	37	70
59	70	28	60	38	48
55	66	30	55	40	51
68	48	29	57	34	42
34	62	28	49	35	
35		24		36	

- 2.4 The following data are the amount of force (kg) required to fracture a concrete beam constructed from one of three beam designs. Unequal sample sizes occurred because of problems with the pouring of the concrete into the forms for each of the designs.

- 1) Write out an appropriate model to describe the data and describe each component of the model.
- 2) Estimate the parameters of the model in part 1.
- 3) Use Levene's, O'Brien's, and Brown–Forsythe's methods to test the equality of the variances.
- 4) Use Welch's test to test $H_0: \mu_1 = \mu_2 = \mu_3$ vs H_a : (not H_0).
- 5) Use the mixed model F -test to test $H_0: \mu_1 = \mu_2 = \mu_3$ vs H_a : (not H_0).

Data for Exercise 2.4

Design	Beam 1	Beam 2	Beam 3	Beam 4	Beam 5	Beam 6	Beam 7	Beam 8	Beam 9	Beam 10
1	195	232	209	201	216	211	205			
2	231	215	230	221	218	227	218	219		
3	223	226	223	224	224	226	227	224	226	226

- 2.5 Four rations with different amounts of cellulose were evaluated as to the amount of feed required for a chicken to gain one pound during the trial. Twenty-four chickens were randomly assigned to the four rations (six chickens per ration) and the chickens were raised in individual cages.

- 1) Write out an appropriate model to describe the data and describe each component of the model.

- 2) Estimate the parameters of the model in part 1.
- 3) Use Levene's, O'Brien's, and Brown–Forsythe's methods to test the equality of the variances.
- 4) Use Welch's test to test $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ vs H_a : (not H_0).
- 5) Use the mixed model F test to test $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ vs H_a : (not H_0).
- 6) Construct 90% confidence intervals about c_1 , c_2 , and c_3 where $c_1 = \mu_1 - \mu_2 + \mu_3 - \mu_4$, $c_2 = \mu_1 + \mu_2 - \mu_3 - \mu_4$, and $c_3 = \mu_1 - \mu_2 - \mu_3 + \mu_4$.

Data for Exercise 2.5

	Chick 1	Chick 2	Chick 3	Chick 4	Chick 5	Chick 6
Ration 1	2.60	2.54	2.87	2.33	2.45	2.77
Ration 2	3.87	3.18	2.59	3.62	2.71	3.08
Ration 3	2.69	5.31	2.08	4.00	3.12	4.19
Ration 4	4.43	5.59	5.06	4.17	5.17	4.47

3

Simultaneous Inference Procedures and Multiple Comparisons

Often an experimenter wants to compare several functions of the μ_i in the same experiment, leading to a multiple testing situation. Experimenters *should* consider all functions of the μ_i that are of interest; that is, they should attempt to answer all questions of interest about relationships among the treatment means. The overriding reason to include more than two treatments in an experiment or study is to be able to estimate and/or test hypotheses about several relationships among the treatment means. Often the treatments are selected to provide a structure of comparisons of interest (see, for example, the drug experiment in Section 2.4). At other times, the experimenter may be interested in comparing each treatment to all other treatments, that is, making all pairwise comparisons. This would be the case, for example, when one is comparing the yields of several varieties of wheat or for any other set of treatments that have been selected for a comparative study.

One concern when making several comparisons in a single experiment is whether significant differences observed are due to real differences or simply to making a very large number of comparisons. Making a large number of comparisons increases the chance of finding differences that appear to be significant when they are not. For example, if an experimenter conducts 25 independent tests in an experiment and finds one significant difference at the 0.05 level, she should not put too much faith in the result because, on average, she should expect to find $(0.05)(25) = 1.25$ significant differences just by chance alone. Thus, if an experimenter is answering a large number of questions with one experiment (which we believe one should do), it is desirable to have a procedure that indicates whether the differences might be the result of chance alone. Fisher (1949) addressed this problem when he put forward the protected least significant difference (LSD) procedure. Since then, many authors have contributed to the area of multiple testing where procedures for numerous settings have been developed.

In this chapter, several well-known and commonly used procedures for making multiple inferences are discussed and compared. Some of the procedures are primarily used for testing hypotheses, while others can also be used to obtain simultaneous confidence intervals; that is, a set of confidence intervals for a set of functions of the μ_i can be derived for

which one can be 95% confident that all the confidence intervals simultaneously contain their respective functions of the μ_i .

3.1 Error Rates

One of the main ways to evaluate and compare multiple comparison procedures is to calculate error rates. If a given confidence interval does not contain the true value of the quantity being estimated, then an error occurs. Similarly, if a hypothesis test is used, an error is made whenever a true hypothesis is rejected or a false hypothesis is not rejected. Next four kinds of error rates are defined.

Definition 3.1: The comparisonwise error rate is equal to the ratio of the number of incorrect inferences made to the total number of inferences made in all experiments analyzed.

Definition 3.2: The experimentwise error rate (EER) is equal to the ratio of the number of experiments in which at least one error is made to the total number of experiments analyzed. It is the probability of making at least one error in an experiment when there are no differences between the treatments. The EER is also referred to as the experimentwise error rate under the complete null hypothesis (EERC).

Definition 3.3: The familywise error rate (FWER) (Westfall et al., 1999) is the probability of making at least one erroneous inference for a predefined set of k comparisons or confidence intervals. The set of k comparisons or confidence intervals is called the family of inferences.

Definition 3.4: The false discovery rate (FDR) (Benjamini and Hochberg, 1995) is the expected proportion of falsely rejected hypotheses among those that were rejected.

The EER controls the error rate when the null hypothesis is that all of the treatments are equally effective, that is, there are no differences among the treatment means. But many experiments involve a selected set of treatments where there are known differences among some treatments. Instead of an all means equal null hypothesis, there may be a collection of k null hypotheses, $H_{01}, H_{02}, \dots, H_{0k}$ about the set of t means. These k null hypotheses are called partial null hypotheses and the error rate is controlled by using a method that controls the FWER (Westfall et al., 1999). For example, the set of treatments in Exercise 2.3 are six candy types, buttons, blue choc, red choc, small choc, chocChip and WchocChip. It is known at the start of the study that the time required to dissolve the small choc is much less than the time required to dissolve any of the other candies. The null question could be: Is the time it takes to dissolve a small choc equal to one-half of the mean times to dissolve the red and blue chocs? In this case a method that controls the FWER is in order since the condition of using a method that controls the EER does not hold; that is, it is known that the mean times are not all equal from the start. The FDR is very useful in the context of microarray experiments in genetics.

In order to avoid finding too many comparisons significant by chance alone in a single experiment, one quite often attempts to fix the experimentwise error rate, when applicable, or the FWER when needed at some prescribed level, such as 0.05. Whenever an experimenter is trying to answer many questions with a single experiment, it is a good strategy to control the FWER.

3.2 Recommendations

There are five basic types of multiple comparison problems: 1) comparing a set of treatments to a control or standard; 2) making all pairwise comparisons among a set of t means; 3) constructing a set of simultaneous confidence intervals or simultaneous tests of hypotheses; 4) exploratory experiments where there are numerous tests being conducted; and 5) data snooping where the comparisons are possibly data-driven. In the first four situations, the number of comparisons or confidence intervals or family of inferences is known before the data are analyzed. In the last situation, there is no set number of comparisons of interest and the final number can be very large. The recommendations given in this chapter are based on information from Westfall et al. (1999), SAS Institute, Inc. (1999), and Westfall (2002).

- 1) If the experiment is an exploratory or discovery study and the results are going to be used to design a follow-up or confirmatory study, then possibly no adjustment for multiplicity is necessary, thus use t -tests or unadjusted confidence intervals based on LSD values.
- 2) Use Dunnett's procedure for comparing a set of treatments with a control. There are two-sided and one-side versions of Dunnett's procedure, so one can select a version to fit the situation being considered.
- 3) For pairwise comparisons, if there is an equal number of observations per treatment group, use Tukey's method. If the data are unbalanced, then use the method that simulates (Westfall et al., 1999) a percentage point, taking into account the pattern of unequal numbers of observations.
- 4) If the set of linear combinations is linearly independent, then the multivariate t can be used to construct confidence intervals or to test hypotheses. If the linear combinations are uncorrelated or orthogonal, the multivariate t works well. If the linear combinations are not uncorrelated, then a simulation method that incorporates the correlation structure should be used instead of the multivariate t . Most cases with unequal numbers of observations per treatment group provide correlated linear combinations and the simulation method should be used.
- 5) The Bonferroni method can be used to construct simultaneous confidence intervals or tests about a selected number of linear combinations of the means, but if the number of combinations of interest is large (say 20 or more), the Scheffé procedure can often produce shorter confidence intervals, so check it out. For a set of hypotheses, the methods of Šidák (1967), Holm (1979), or Šidák–Holm can be used effectively. When the linear combinations are uncorrelated these bounds are quite good, but when there are correlations among the linear combinations, the realized FWER can be much less than desired. SAS®-MULTTEST can be used carry out bootstrap and simulated percentage points for a given set of comparisons that takes into account the correlation among the comparisons within the set.
- 6) For data snooping or for data-driven comparisons or hypotheses, use Scheffé's procedure as one can make as many comparisons as one wants and still control the EER or FWER.
- 7) For studies such as genetic studies that involve thousands of comparisons, use a method that controls the FDR, such as the method suggested by Benjamini and Hochberg (1995).

- 8) For studies that involve evaluating the safety of a treatment as compared with a control or placebo for possible adverse effects, use a method that does not correct for multiple tests or comparisons. Adjustment for multiplicity may not be needed for safety studies, where it is much more serious to make a type II error than it is to make a type I error.
- 9) Once the type of comparison is determined and the desired level of error rate control is specified, select the method satisfying these conditions that provides the smallest p -values or smallest critical differences or shortest confidence interval widths.

Each of the recommended multiple comparison procedures as well as a few other popular procedures available for the one-way treatment structure of Chapter 1 are examined in the following discussion. Each of the procedures can also be used in much more complex situations, as will be illustrated throughout the remainder of this book. The parameter v used during the remainder of this book represents the degrees of freedom corresponding to the estimator of σ^2 . For the one-way case of Chapter 1, the error degrees of freedom are $v = N - t$.

3.3 Least Significant Difference

The LSD multiple comparison method has possibly been used more than any other method, perhaps because it is one of the easiest to apply. It is usually used to compare each treatment mean with every other treatment mean, but it can be used for other comparisons as well. The LSD at the 100% significance level for comparing μ_i to μ_j is

$$\text{LSD}_\alpha = t_{\alpha/2,v} \hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_j}} \quad (3.1)$$

One concludes that $\mu_i \neq \mu_j$ if $|\hat{\mu}_i - \hat{\mu}_j| > \text{LSD}_\alpha$. This procedure has a comparisonwise error rate equal to α . A corresponding $(1 - \alpha)100\%$ confidence interval for $\mu_i - \mu_j$ is

$$\hat{\mu}_i - \hat{\mu}_j \pm t_{\alpha/2,v} \hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_j}} \quad (3.2)$$

If all sample sizes are equal (to n , say), then a single LSD value can be used for all pairwise comparisons. In this case, the single LSD_α value is given by

$$\text{LSD}_\alpha = t_{\alpha/2,v} \hat{\sigma} \sqrt{\frac{2}{n}} \quad (3.3)$$

Suppose a study includes t treatment means and that all possible pairwise comparisons at the 5% significance level are going to be made. Comparisons of the comparisonwise and

TABLE 3.1

Simulated Error Rates for the LSD Procedure

Number of treatments	2	3	4	5	6	8	10	20
Comparisonwise error rate	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Experimentwise error rate	0.05	0.118	0.198	0.280	0.358	0.469	0.586	0.904

experimentwise error rates for experiments with different values of t are displayed in Table 3.1. The information in the table applies to cases where all treatment means are equal. Table 3.1 shows that, in an experiment involving six treatments, 35.8% of the time one would find at least one significant difference, even when all the treatment means were equal to one another. Obviously, using the LSD procedure could be very risky without some additional protection. When there is more than one test or parameter or linear combination of parameters of interest, meaning there is a multiplicity problem, some adjustment should be taken into account in order to eliminate discovering false results. Westfall et al. (1999) present an excellent discussion of all of the problems associated with the multiplicity problem and/or the multiple comparison problem. The following discussion attempts to describe those procedures that are useful or have been used in the analysis and interpretation of the results from designed experiments.

3.4 Fisher's LSD Procedure

Fisher's recommendation offers some protection for the LSD procedure discussed in the preceding section. In Fisher's procedure, LSD tests are made at the $\alpha 100\%$ significance level by utilizing Equation 3.1, but only if $H_0: \mu_1 = \mu_2 = \dots = \mu_t$ is first rejected at that level of α by the F -test discussed in Chapter 1.

This gives a rather large improvement over the straight LSD procedure since the experimentwise error rate is now approximately equal to α . However, it is possible to reject $H_0: \mu_1 = \mu_2 = \dots = \mu_t$ and not reject any of $H_0: \mu_i = \mu_j$ for $i \neq j$. It is also true that this procedure may not detect some differences between pairs of treatments when differences really exist. In other words, differences between a few pairs of treatments may exist, but equality of the remaining treatments may cause the F -test to be nonsignificant, and this procedure does not allow the experimenter to make individual comparisons without first obtaining a significant F -statistic. The other problem with this procedure is that many experiments contain treatments where it is known there are unequal means among some subsets of the treatments. In this case, one expects to reject the equal means hypothesis and the LSD would be used to make all pairwise comparisons. If a subset of the treatments has equal means, then more of the pairwise comparisons will be detected as being significantly different than expected. Thus the FWER is not maintained. Fisher's LSD can be recommended only when the complete null hypothesis is expected to be true.

These two LSD procedures are not recommended for constructing simultaneous confidence intervals on specified contrasts of the μ_i because the resulting confidence intervals obtained will generally be too narrow.

Each of the above LSD procedures can be generalized to include several contrasts of the treatment means. The generalization is: conclude that $\sum_{i=1}^t c_i \mu_i \neq 0$ if

$$\left| \sum_{i=1}^t c_i \hat{\mu}_i \right| > t_{\alpha/2, v} \hat{\sigma} \sqrt{\sum_{i=1}^t c_i^2 / n_i} \quad (3.4)$$

Examples are given in Sections 3.10, 3.12, 3.14, and 3.16.

3.5 Bonferroni's Method

Although this procedure may be the least used, it is often the best. It is particularly good when the experimenter wants to make a small number of comparisons. This procedure is recommended for planned comparisons whenever it is necessary to control the FWER. Suppose the experimenter wants to make p such comparisons. She would conclude that the q th comparison $\sum_{i=1}^t c_{iq} \mu_i \neq 0$, $q = 1, 2, \dots, p$, if

$$\left| \sum_{i=1}^t c_{iq} \hat{\mu}_i \right| > t_{\alpha/2p, v} \hat{\sigma} \sqrt{\sum_{i=1}^t c_{iq}^2 / n_i} \quad (3.5)$$

These p -tests will give a FWER less than or equal to α and a comparisonwise error rate equal to α/p . Usually the FWER is much less than α . Unfortunately, it is not possible to determine how much less. Values of $t_{\alpha/2p, v}$ for selected values of α , p , and v are given in the Appendix in Table A.2. For example, if $\alpha = 0.05$, $p = 5$, and $v = 24$, then from Table A.2 one gets $t_{\alpha/2p, v} = 2.80$. The examples in Sections 3.10, 3.12, 3.14, and 3.16 demonstrate the use of the Bonferroni method. The tables m is equivalent to our p .

Simultaneous confidence intervals obtained from the Bonferroni method, which can be recommended, have the form:

$$\sum_{i=1}^t c_{iq} \hat{\mu}_i \pm t_{\alpha/2p, v} \hat{\sigma} \sqrt{\sum_{i=1}^t \frac{c_{iq}^2}{n_i}}, \quad q = 1, 2, \dots, p \quad (3.6)$$

The Bonferroni method can be applied to any set of functions of the parameters of a model, including variances as well as means.

3.6 Scheffé's Procedure

This procedure is recommended whenever the experimenter wants to make a large number of “unplanned” comparisons. Unplanned comparisons are comparisons that the experimenter had not thought of making when planning the experiment. These arise frequently, since the results of an experiment frequently suggest certain comparisons to the experimenter. This procedure can also be used when there are a large number of planned comparisons, but the widths of the confidence intervals are generally wider than

for other procedures, although not always. Consider testing $H_0: \sum_{i=1}^t c_i \mu_i = 0$ for a given contrast vector c . It is true that

$$\Pr \left\{ \frac{\left(\sum_{i=1}^t c_i \hat{\mu}_i - \sum_{i=1}^t c_i \mu_i \right)^2}{\sum_{i=1}^t c_i^2 / n_i} \leq (t-1)F_{\alpha, t-1, v} \hat{\sigma}^2 \quad \text{for all contrast vectors } c \right\} = 1 - \alpha$$

Thus a procedure with an FWER equal to α for comparing all possible contrasts of the μ_i to zero is as follows: Reject $H_0: \sum_{i=1}^t c_i \mu_i = 0$ if

$$\left| \sum_{i=1}^t c_i \hat{\mu}_i \right| > \sqrt{(t-1)F_{\alpha, t-1, v}} \hat{\sigma} \sqrt{\sum_{i=1}^t c_i^2 / n_i} \quad (3.7)$$

This procedure allows one to compare an infinite number of contrasts to zero while maintaining an experimentwise error rate equal to α . However, most experimenters will usually not be interested in an infinite number of comparisons; that is, only a finite number of comparisons are of interest. Scheffé's procedure can still be used, but in this case, the FWER will generally be much smaller than α . Bonferroni's method or the multivariate t -method when appropriate will often be better (narrower confidence interval or more powerful test) than Scheffé's procedure for a finite number of comparisons. That is, a smaller value of $\sum_{i=1}^t c_i \hat{\mu}_i$ can often enable one to declare that $\sum_{i=1}^t c_i \mu_i$ is significantly different from zero using Bonferroni's method or the multivariate t -method than can be declared significant by Scheffé's method. However, if one is going to "muck around" in the data to see if anything significant turns up, then one should use Scheffé's method, since such comparisons are really unplanned comparisons rather than planned comparisons. It should be noted that Scheffé's method will not reveal any contrasts significantly different from zero unless the F -test discussed in Chapter 1 rejects $H_0: \mu_1 = \mu_2 = \dots = \mu_t$. Scheffé's procedure can also be used to obtain simultaneous confidence intervals for contrasts of the μ_i . The result required is that, for any set of contrasts c_1, c_2, \dots , one can be at least $(1 - \alpha)100\%$ confident that $\sum_{i=1}^t c_i \mu_i$ will be contained within the interval given by

$$\sum_{i=1}^t c_{iq} \hat{\mu}_i \pm \sqrt{(t-1)F_{\alpha, t-1, v}} \hat{\sigma} \sqrt{\sum_{i=1}^t c_{iq}^2 / n_i} \quad \text{for all } q = 1, 2, \dots \quad (3.8)$$

If one wants to consider all linear combinations of the μ_i rather than just all contrasts, then $\sqrt{(t-1)F_{\alpha, t-1, v}}$ must be replaced by $\sqrt{[t]F_{\alpha, t, v}}$ in Equations 3.7 and 3.8.

Examples can be found in Sections 3.10, 3.12, and 3.14.

3.7 Tukey–Kramer Method

The preceding procedures can be used regardless of the values of the n_i . Tukey's (Tukey, 1952, 1953; Kramer, 1956) honest significant difference (HSD) procedure was designed to

make all pairwise comparisons among a set of means. The procedure, however, requires equal n_i . Tukey (1953) and Kramer (1956) provided a modification for the case where one has unequal sample sizes. Hayter (1984) provided proof that the Tukey–Kramer method provides FWER protection, although an approximate procedure can be used if the n_i are not too unequal. The Tukey–Kramer method is to reject $H_0: \mu_i = \mu_{i'}$ for $i \neq i'$ if

$$|\hat{\mu}_i - \hat{\mu}_{i'}| > q_{\alpha, t, v} \sqrt{\frac{\hat{\sigma}^2}{2} \left(\frac{1}{n_i} + \frac{1}{n_{i'}} \right)} \quad (3.9)$$

where $q_{\alpha, t, v}$ is the upper percentile of the distribution of the Studentized range statistic. Values of $q_{\alpha, t, v}$ for selected values of α , t , and v are given in Appendix Table A.4.

If the sample sizes are all equal to n , then the decision is to reject $H_0: \mu_i = \mu_{i'}$ for $i \neq i'$ if

$$|\hat{\mu}_i - \hat{\mu}_{i'}| > q_{\alpha, t, v} \sqrt{\frac{\hat{\sigma}^2}{n}}$$

Tukey's general procedure for equal sample sizes is to reject $H_0: \sum_{i=1}^t c_i \mu_i = 0$ for a contrast if

$$\left| \sum_{i=1}^t c_i \hat{\mu}_i \right| > q_{\alpha, t, v} \frac{\hat{\sigma}}{\sqrt{n}} \left(\frac{1}{2} \sum_{i=1}^t |c_i| \right)$$

3.8 Simulation Methods

For unequal sample size problems, for problems where the comparisons are other than pairwise comparisons, and for problems where the comparisons are not linearly independent, the above methods provide FWER significance levels that are less than desired. In this case, the percentage points for the appropriate set of comparisons can be simulated.

Suppose you are interested in p linear combinations of the μ_i such as $\sum_{i=1}^t c_{iq} \mu_i$, $q = 1, 2, \dots, p$ and it is desired to provide a procedure that controls the FWER for either the set of hypotheses $H_0: \sum_{i=1}^t c_{iq} \mu_i = 0$, $q = 1, 2, \dots, p$ or a set of simultaneous confidence intervals for $\sum_{i=1}^t c_{iq} \mu_i$. The process is:

- 1) Generate a sample of data in the same structure of the data set at hand. If there are five treatments with sample sizes, 5, 9, 3, 6, and 7, generate data with those sample sizes.
- 2) Carry out the analysis of the generated data set as is to be done with the actual data set and compute the p t -statistics:

$$t_q = \frac{\sum_{i=1}^t c_{iq} \hat{\mu}_i}{\sqrt{\hat{\sigma}^2 \sum_{i=1}^t c_{iq}^2 / n_i}} \quad q = 1, 2, \dots, p$$

- 3) Compute the maximum of the absolute values of the t_q , $T_s = \max(|t_1|, |t_2|, \dots, |t_p|)$.
- 4) Repeat steps 1, 2 and 3 a very large number of times, keeping track of the computed values of T_s . Determine the upper $\alpha/100$ percentile of the distribution of the T_s and denote this percentile by T_α .
- 5) For the actual data set, compute t_q , $q = 1, 2, \dots, p$ and reject the q th hypothesis if $|t_q| > T_\alpha$, $q = 1, 2, \dots, p$ or construct simultaneous confidence intervals as

$$\sum_{i=1}^t c_{iq} \hat{\mu}_i \pm T_\alpha \sqrt{\hat{\sigma}^2 \sum_{i=1}^t c_{iq}^2 / n_i}, \quad q = 1, 2, \dots, p$$

The accuracy of the simulation can be specified by using the method of Edwards and Berry (1987). SAS-MULTTEST can be used to obtain simultaneous inferences using the bootstrap method (Westfall et al., 1999). Bootstrap methodology is beyond the scope of this book.

3.9 Šidák Procedure

Šidák (1967) provided a modification of the Bonferroni method by using a different percentage point for each of the comparisons. The process is to compute a t -statistic for each of the comparisons:

$$t_q = \frac{\sum_{i=1}^t c_{iq} \hat{\mu}_i}{\sqrt{\hat{\sigma}^2 \sum_{i=1}^t c_{iq}^2 / n_i}}, \quad q = 1, 2, \dots, p$$

Compute the significance level for each comparison and order the significance levels from smallest to largest as p_1, p_2, \dots, p_p . For a FWER of α , reject the individual comparison if $p_q \leq 1 - (1 - \alpha)^{1/p}$ or equivalently if $\alpha \geq 1 - (1 - p_q)^p$.

3.10 Example—Pairwise Comparisons

The task data in Section 1.6 is used to demonstrate the results of the above multiple comparisons procedures. Table 3.2 contains the SAS-Mixed code to fit the one-way means model and the LSMeans statements are used to extract several of the multiple comparison procedures. Table 3.3 contains the percentage points used to provide confidence differences or significant differences for the simulate, Tukey-Kramer, Bonferroni, Šidák, Scheffé, and t (unadjusted) methods. Excluding the unadjusted t , the other methods provide 0.05 type I FWER for all pairwise comparisons. The simulate and Tukey-Kramer methods use the smallest quantiles with the Šidák and Bonferroni methods in the middle, while the Scheffé method is largest. Table 3.4 contains the critical significant differences for each of

TABLE 3.2

SAS System Code Using Proc Mixed to Request the Computation of Several Multiple Comparisons Procedures for All Pairwise Comparisons

```
PROC mixed DATA=EX1; CLASS TASK;
MODEL PULSE20=TASK/NOINT SOLUTION;
LSMEANS TASK/ DIFF CL;
LSMEANS TASK/ DIFF ADJUST=TUKEY CL;
LSMEANS TASK/ DIFF ADJUST=BON CL;
LSMEANS TASK/ DIFF ADJUST=SCHEFFE CL;
LSMEANS TASK/ DIFF ADJUST=SIDAK CL;
LSMEANS TASK/ DIFF ADJUST=SIMULATE (REPORT SEED=4938371) CL;
```

TABLE 3.3

Percentage Points Used for All Pairwise Comparisons of the Six Task Means

Simulation Results

Method	95% Quantile	Estimated α	99% Confidence Limits
Simulate	2.932480	0.0500	0.0450 0.0550
Tukey–Kramer	2.940710	0.0486	0.0436 0.0535
Bonferroni	3.053188	0.0359	0.0316 0.0401
Šidák	3.044940	0.0370	0.0326 0.0413
Šcheffé	3.437389	0.0131	0.0105 0.0157
<i>t</i>	1.998972	0.3556	0.3446 0.3666

TABLE 3.4

Critical Differences Used to Compare the Differences between Pairs of Means for the Unadjusted *t* and Several Multiple Comparison Procedures

TASK	_TASK	Standard			Bonferroni	Tukey–Kramer	Šidák	Simulate
		Estimate	Error	<i>t</i>				
1	2	0.840	2.225	4.449	6.795	6.544	7.650	6.776 6.526
1	3	-3.877	2.338	4.674	7.139	6.876	8.038	7.120 6.857
1	4	-6.077	2.338	4.674	7.139	6.876	8.038	7.120 6.857
1	5	2.423	2.225	4.449	6.795	6.544	7.650	6.776 6.526
1	6	3.105	2.277	4.553	6.953	6.697	7.828	6.935 6.679
2	3	-4.717	2.380	4.758	7.267	7.000	8.182	7.248 6.980
2	4	-6.917	2.380	4.758	7.267	7.000	8.182	7.248 6.980
2	5	1.583	2.270	4.537	6.929	6.674	7.801	6.911 6.655
2	6	2.265	2.321	4.639	7.085	6.824	7.977	7.066 6.805
3	4	-2.200	2.486	4.970	7.591	7.311	8.546	7.570 7.291
3	5	6.300	2.380	4.758	7.267	7.000	8.182	7.248 6.980
3	6	6.982	2.429	4.855	7.416	7.143	8.349	7.396 7.123
4	5	8.500	2.380	4.758	7.267	7.000	8.182	7.248 6.980
4	6	9.182	2.429	4.855	7.416	7.143	8.349	7.396 7.123
5	6	0.682	2.321	4.639	7.085	6.824	7.977	7.066 6.805

TABLE 3.5

Adjusted Significance Levels to Test the Equality of All Pairwise Comparisons of TASK Minus $_TASK$ Obtained from Six Procedures Where t Corresponds to the Unadjusted t

TASK	$_TASK$	t	Bonferroni	Tukey-Kramer	Scheffé	Šidák	Simulate
1	2	0.7072	1.0000	0.9990	0.9996	1.0000	0.9990
1	3	0.1024	1.0000	0.5642	0.7378	0.8021	0.5646
1	4	0.0117	0.1751	0.1129	0.2552	0.1615	0.1111
1	5	0.2805	1.0000	0.8840	0.9446	0.9928	0.8804
1	6	0.1777	1.0000	0.7484	0.8661	0.9469	0.7501
2	3	0.0520	0.7795	0.3645	0.5642	0.5509	0.3657
2	4	0.0051	0.0761	0.0546	0.1506	0.0735	0.0545
2	5	0.4880	1.0000	0.9815	0.9923	1.0000	0.9813
2	6	0.3328	1.0000	0.9238	0.9651	0.9977	0.9234
3	4	0.3796	1.0000	0.9488	0.9772	0.9992	0.9474
3	5	0.0103	0.1543	0.1014	0.2364	0.1437	0.0985
3	6	0.0055	0.0831	0.0590	0.1596	0.0799	0.0584
4	5	0.0007	0.0104	0.0087	0.0366	0.0104	0.0090
4	6	0.0004	0.0053	0.0046	0.0219	0.0053	0.0052
5	6	0.7699	1.0000	0.9997	0.9999	1.0000	0.9998

the pairwise comparisons. The observed differences for task 1 to task 4, task 2 to task 4, task 3 to task 4, task 3 to task 5, task 3 to task 6, task 4 to task 5 and task 4 to task 6 all exceed the critical differences for the t or LSD, which controls the comparisonwise error rate, but not the experimentwise error rate. Only the comparisons of task 4 to task 5 and task 4 to task 6 exceed the critical differences for the other five methods, all of which provide experiment wise error rate protection. The magnitudes of the critical differences are smallest for the uncorrected t or LSD method. The simulate and Tukey-Kramer critical differences are similar in magnitude while the simulate values are a little smaller. The Šidák and Bonferroni differences are similar in magnitude, with the Šidák values slightly smaller. The Scheffé critical differences are largest, as is expected since they control the FWER for an infinite number of comparisons and only 15 pairwise comparisons are made. A set of simultaneous confidence intervals about all pairwise comparisons can be constructed by adding and subtracting the critical difference from the estimated difference. For example, the simultaneous 95% confidence interval about $\mu_1 - \mu_2$ using the simulate method is 0.840 ± 6.526 . Table 3.5 contains the adjusted p -values for each of the methods. The p -values provide the same decision as the 5% critical differences in Table 3.4.

3.11 Dunnett's Procedure

One really interesting case is that of comparing all treatments with a control. This type of inference is important in safety studies, where it is of interest to compare different doses of a treatment with the control or placebo. Dunnett's test is to declare a treatment mean μ_i to be significantly different from the mean of the control μ_0 if

$$|\hat{\mu}_i - \hat{\mu}_0| > d_{\alpha, t, v} \sqrt{\hat{\sigma}^2 \left(\frac{1}{n_i} + \frac{1}{n_0} \right)}$$

where $d_{\alpha,t,v}$ is the upper $\alpha/100$ percentile of the “many-to-one t -statistic” (Miller, 1967). Dunnett’s method controls the FWER. If the sample sizes are unequal, a simulate procedure can take into account the sample size structure and possibly provide a shorter bound.

3.12 Example—Comparing with a Control

The task data in Section 1.6 is used to demonstrate the process of comparing each treatment with a control. In this study, assume that task 2 is the control task and the other five tasks are the experimental tasks. Table 3.6 contains the SAS-Mixed code to use the unadjusted t , Bonferroni, Dunnett, Scheffé, Šidák, and Simulate methods to compare all of the other tasks with task 2. The option on the LSMean statement DIFF=CONTROL('2') requests that task 2 be considered as the control and is compared with each of the other tasks in the study. Table 3.7 contains the 95% quantiles for each of the methods. The Dunnett quantile is less than the others (except for the unadjusted t) with the simulate method very close. There are five comparisons being made, which dictates the magnitude of the Bonferroni and Šidák quantiles. The Scheffé quantile is the same as in Table 3.4, which is useful for an infinite number of comparisons. The only comparison where the observed difference exceeds the critical difference is for comparing task 4 to the control or task 2. A set of simultaneous confidence intervals about all differences between the treatment and control means can be constructed by adding and subtracting the critical difference in Table 3.8

TABLE 3.6

SAS System Code Using Proc Mixed to Request the Computation of Several Multiple Comparisons Procedures for Comparing Each Task to the Means of Task 2 (Control)

```
PROC mixed DATA=EX1; CLASS TASK;
MODEL PULSE20=TASK/NOINT;
LSMEANS TASK/ DIFF=CONTROL ('2') CL;
LSMEANS TASK/ DIFF=CONTROL ('2') ADJUST=BON CL;
LSMEANS TASK/ DIFF=CONTROL ('2') ADJUST=DUNNETT CL;
LSMEANS TASK/ DIFF=CONTROL ('2') ADJUST=SIDAK CL;
LSMEANS TASK/ DIFF=CONTROL ('2') ADJUST=SIMULATE (REPORT SEED=4938371) CL;
LSMEANS TASK/ DIFF=CONTROL ('2') ADJUST=scheffe CL;
```

TABLE 3.7

Percentage Points Used for Comparing Each Task Mean to the Mean of Task 2 (Control)

Simulation Results

Method	95% Quantile	Exact α
Simulate	2.590707	0.0494
Dunnett, two-sided	2.585505	0.0500
Bonferroni	2.657479	0.0418
Šidák	2.649790	0.0427
Scheffé	3.437389	0.0048
t	1.998972	0.1831

TABLE 3.8

Critical Differences for Comparing Each Treatment with the Control for the Unadjusted t and Five Multiple Comparisons

TASK	_TASK	Standard		t	Bonferroni	Dunnett	Scheffé	Šidák	Simulate
		Estimate	Error						
1	2	0.840	2.225	4.449	5.914	5.754	7.650	5.897	5.765
3	2	4.717	2.380	4.758	6.326	6.154	8.182	6.307	6.167
4	2	6.917	2.380	4.758	6.326	6.154	8.182	6.307	6.167
5	2	-1.583	2.270	4.537	6.031	5.868	7.801	6.014	5.880
6	2	-2.265	2.321	4.639	6.167	6.000	7.977	6.149	6.012

TABLE 3.9

Adjusted Significance Levels for Testing $H_0: \mu_i = \mu_2$ Obtained from Six Procedures Where Task 2 is the Control and t Corresponds to the Unadjusted t

TASK	_TASK	t	Bonferroni	Dunnett	Scheffé	Šidák	Simulate
1	2	0.7072	1.0000	0.9953	0.9996	0.9978	0.9940
3	2	0.0520	0.2598	0.1895	0.5642	0.2342	0.1899
4	2	0.0051	0.0254	0.0220	0.1506	0.0251	0.0223
5	2	0.4880	1.0000	0.9354	0.9923	0.9648	0.9299
6	2	0.3328	1.0000	0.7953	0.9651	0.8678	0.7938

from the estimated difference. For example, the simultaneous 95% confidence interval about $\mu_1 - \mu_2$ using the simulate method is 0.840 ± 5.765 . The adjusted p -values for each of the methods are included in Table 3.9.

3.13 Multivariate t

The multivariate t method is a good method to use when the experimenter wants to consider a linearly independent set of linear combinations of the μ_i . It is important that this procedure not be used for a set of linear combinations that are linearly dependent. This restriction prevents one from using this method for making all possible pairwise comparisons between a set of treatment means, since the set of all possible comparisons between means is not a linearly independent set. Discussion concerning linearly dependent linear combinations is given at the end of this section. If the experimenter wants to make p linearly independent comparisons, then she would conclude that the q th comparison $\sum_{i=1}^t c_{iq} \mu_i \neq 0$ if

$$\left| \sum_{i=1}^t c_{iq} \hat{\mu}_i \right| > t_{\alpha/2, p, v} \hat{\sigma} \sqrt{\sum_{i=1}^t \frac{c_{iq}^2}{n_i}} \quad (3.10)$$

where $t_{\alpha/2, p, v}$ is the upper $\alpha/2$ percentile of a p -variate multivariate t -distribution with v degrees of freedom and correlation matrix I_p . Values of $t_{\alpha/2, p, v}$ are given in Appendix Table A.3 for selected values of α , p , and v . Simultaneous confidence intervals based on the multivariate t method are appropriate and can be recommended where $m = p$ in Table A.3.

The multivariate t method has an FWER less than or equal to α . If the linear combinations $\sum_{i=1}^t c_{iq}\hat{\mu}_i$ are statistically independent or uncorrelated for $q = 1, 2, \dots, p$, then the FWER is exactly equal to α . For a specified set of linearly independent comparisons, the multivariate t method will often give the best results. If the set of comparisons of interest to the experimenter is linearly independent, then the multivariate t method will always be better than Bonferroni's method, that is, the resulting confidence intervals will be narrower.

The following result enables one to extend the application of the multivariate t procedure to a linearly dependent set of comparisons (however, the resulting tests or confidence intervals are not entirely satisfactory): Let l_1, l_2, \dots, l_p be a linearly independent set of linear combinations of the μ_i . If $|l_q| \leq c_q$ for $q = 1, 2, \dots, p$, then

$$\left| \sum_{i=1}^p \lambda_q l_q \right| \leq \sum_{i=1}^p |\lambda_q| c_q$$

To make use of this result, use the following procedure:

- 1) Let l_1, l_2, \dots, l_p be a linearly independent set of linear combinations of the μ_i . This set is denoted as the linear combinations of primary interest to the experimenter. For this set of comparisons, conclude that $l_q = \sum_{i=1}^t c_{iq}\mu_i$ is significantly different from zero if

$$|\hat{l}_q| > t_{\alpha/2, p, v} \hat{\sigma} \sqrt{\sum_{i=1}^t c_{iq}^2 / n_i} \quad (3.11)$$

- 2) Let l^* be any comparison that is of secondary importance, that is, a linear combination of the l_q , $q = 1, 2, \dots, p$. That is, $l^* = \sum_{q=1}^p \lambda_q l_q$ for some set of λ_q . One declares that l^* is significantly different from zero if

$$|\hat{l}^*| > t_{\alpha/2, p, v} \hat{\sigma} \sum_{q=1}^p \left(|\lambda_q| \sqrt{\sum_{i=1}^t c_{iq}^2 / n_i} \right) \quad (3.12)$$

An experimenter can make as many l^* -type comparisons as needed without increasing the FWER. This extension of the multivariate t method gives very powerful tests for those comparisons of primary importance but is less powerful for those comparisons of secondary importance. This is illustrated in an example in Section 3.16.

3.14 Example—Linearly Independent Comparisons

Table 3.10 contains the SAS-Mixed code to provide estimates, standard errors and confidence interval widths for five linearly independent contrasts of the task means. The five contrasts are:

$$\begin{aligned} \mu_1 - \frac{1}{2}(\mu_2 + \mu_3), \frac{1}{3}(\mu_1 + \mu_2 + \mu_3) - \frac{1}{3}(\mu_4 + \mu_5 + \mu_6) \\ \mu_6 - \frac{1}{2}(\mu_4 + \mu_5), \frac{1}{2}(\mu_1 + \mu_6) - \frac{1}{2}(\mu_4 + \mu_5) \\ \frac{1}{2}(\mu_1 + \mu_6) - \frac{1}{4}(\mu_2 + \mu_3 + \mu_4 + \mu_5) \end{aligned}$$

TABLE 3.10

Proc Mixed Code to Provide Estimates of Five Linearly Independent Contrasts

```
PROC mixed DATA=EX1; CLASS TASK;
MODEL PULSE20=TASK/NOINT SOLUTION;
estimate '1-0.5*(2+3)' task 2 -1 -1 0 0 0/divisor=2 cl;
estimate '1+2+3-4-5-6' task 1 1 1 -1 -1 -1/divisor=3 cl;
estimate '6-.5*(4+5)' task 0 0 0 -1 -1 2/divisor=2 cl;;
estimate '1+6-4-5' task 1 0 0 -1 -1 1/divisor=2 cl;;
estimate '1+6-.5*(2+3+4+5)' task 2 -1 -1 -1 -1 2/divisor=4 cl;
```

TABLE 3.11Widths of Confidence Intervals for the Unadjusted t and Three Adjusted Methods

Label	Estimate	Standard Error	t	Multivariate t	Bonferroni t	Scheffé
1 - 0.5 * (2 + 3)	-1.519	1.948	7.787	10.319	10.352	13.390
1 + 2 + 3 - 4 - 5 - 6	0.829	1.355	5.416	7.178	7.200	9.314
6 - 0.4 * (4 + 5)	-4.932	2.056	8.219	10.891	10.926	14.133
1+6-4-5	-3.379	1.647	6.585	8.727	8.755	11.324
1 + 6 - 0.5 * (2 + 3 + 4 + 5)	-3.225	1.416	5.661	7.502	7.526	9.734

Table 3.11 contains the results of computing the widths of the confidence intervals about each of the five contrasts using the multivariate t , Bonferroni t and Scheffé methods. The unadjusted t is also provided. The widths of the confidence intervals are computed by $2 \times Q \times (\text{stderr})$ where Q represents the quantile of the respective procedure. The quantiles areas follows: For the unadjusted t , $Q = t_{0.025,62} = 1.999$, for the multivariate t , $Q = t_{0.025,5,60} = 2.649$ (60 degrees of freedom were used from the table instead of 62 degrees of freedom), for the Bonferroni t , $Q = t_{0.02515,62} = 2.657$, and for Scheffé, $Q = \sqrt{(5F_{0.05,5,62})} = 3.437$. The widths of the confidence intervals for the multivariate t are shorter than either of those given by the Bonferroni or Scheffé confidence intervals. The Bonferroni t confidence intervals are just a little wider than those of the multivariate t . The Bonferroni t confidence intervals increase in width as the number of comparisons increases. If the number of comparisons is 47 or less, the Bonferroni confidence intervals are shorter than those of the Scheffé method. For more than 47 comparisons, the Scheffé confidence intervals are shorter than the Bonferroni.

3.15 Sequential Rejective Methods

The previous methods are called single-step methods as only one step is needed to determine the proper critical value for all tests. Each individual test or confidence level is evaluated without reference to the others. In the stepwise or sequential methods, the result of a given test depends on the results of previous tests (Westfall et al., 1999). These methods

can be applied to a set of tests of hypotheses. Assume there are p hypotheses of interest, H_{0q} : $\sum_{i=1}^t c_{iq} \mu_i = 0$ vs H_{0q} : $\sum_{i=1}^t c_{iq} \mu_i \neq 0$, $q = 1, 2, \dots, p$. Compute the p t -statistics

$$t_q = \frac{\sum_{i=1}^t c_{iq} \hat{\mu}_i}{\sqrt{\hat{\sigma}^2 \sum_{i=1}^t c_{iq}^2 / n_i}}, \quad q = 1, 2, \dots, p$$

and let p_1, p_2, \dots, p_p denote the observed significance levels. Next order the p_q from smallest to largest as $p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(p)}$.

3.15.1 Bonferroni–Holm Method

The Bonferroni–Holm method starts with the Bonferroni adjustment for the first test, but increases the significance level for tests that follow (Holm, 1979). If $p_{(1)} > \alpha/p$, then fail to reject all of the hypotheses. If $p_{(1)} < \alpha/p$, then reject the hypothesis corresponding to $p_{(1)}$ and compare $p_{(2)}$ to $\alpha/(p - 1)$. If $p_{(2)} > \alpha/(p - 1)$, then fail to reject all of the remaining hypotheses. If $p_{(2)} < \alpha/(p - 1)$, then reject the hypothesis corresponding to $p_{(2)}$ and compare $p_{(3)}$ to $\alpha/(p - 2)$. Continue the procedure until for some q , $p_{(q)} > \alpha/(p - q + 1)$, where one fails to reject hypothesis corresponding to $p_{(q)}$ and all of the remaining hypotheses. This method controls the FWER and is more powerful than the Bonferroni method since the ordered p -values are compared with larger values, except for the first one.

3.15.2 Šidák–Holm Method

The Holm modification used in the Bonferroni–Holm method can be used with the Šidák method where the adjusted p -values are

$$\begin{aligned}\tilde{p}_{(1)} &= 1 - (1 - p_{(1)})^k \\ \tilde{p}_{(2)} &= \min[\tilde{p}_{(1)}, 1 - (1 - p_{(2)})^{k-1}] \\ &\vdots \\ \tilde{p}_{(j)} &= \min[\tilde{p}_{(j-1)}, 1 - (1 - p_{(j)})^{k-j+1}] \\ &\vdots \\ \tilde{p}_{(k)} &= \min(\tilde{p}_{(k-1)}, p_{(k)})\end{aligned}$$

This method provides control of the FWER when the comparisons are independent and for most situations involving dependent comparisons (Holland and Copenhaver, 1987).

3.15.3 Benjamini and Hochberg Method to Control FDR

Assume there are p comparisons of interest, $\sum_{i=1}^t c_{iq} \mu_i$, $q = 1, 2, \dots, p$. Compute the p t -statistics

$$t_q = \frac{\sum_{i=1}^t c_{iq} \hat{\mu}_i}{\sqrt{\hat{\sigma}^2 \sum_{i=1}^t c_{iq}^2 / n_i}}, \quad q = 1, 2, \dots, p$$

and let p_1, p_2, \dots, p_p denote the observed significance levels. Next order the p_q from smallest to largest as $p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(p)}$. The adjusted p -values that control the FDR (but not the FWER) are

$$\begin{aligned}\tilde{p}_{(k)} &= p_{(k)} \\ \tilde{p}_{(k-1)} &= \min\{\tilde{p}_{(k)}, [k/(k-1)]p_{(k-1)}\} \\ &\vdots \\ \tilde{p}_{(k-j)} &= \min\{\tilde{p}_{(k-j+1)}, [k/(k-j)]p_{(k-j)}\} \\ &\vdots \\ \tilde{p}_{(1)} &= \min(\tilde{p}_{(2)}, kp_{(1)})\end{aligned}$$

3.16 Example—Linearly Dependent Comparisons

Eight linearly dependent comparisons were selected and the unadjusted t , Bonferroni t , Bonferroni–Holm, Šidák, Šidák–Holm, and Benjamini–Hochberg methods were used to compute the adjusted significance levels. The eight comparisons are:

$$\begin{aligned}&\mu_1 - \mu_2, \mu_1 - \frac{1}{2}(\mu_2 + \mu_3), \mu_1 - \mu_3 \\ &\frac{1}{3}(\mu_1 + \mu_2 + \mu_3) - \frac{1}{3}(\mu_4 + \mu_5 + \mu_6), \mu_4 - \mu_5 \\ &\mu_6 - \frac{1}{2}(\mu_4 + \mu_5), \frac{1}{2}(\mu_1 + \mu_6) - \frac{1}{2}(\mu_4 + \mu_5) \\ &\frac{1}{2}(\mu_1 + \mu_6) - \frac{1}{4}(\mu_2 + \mu_3 + \mu_4 + \mu_5)\end{aligned}$$

Table 3.12 contains the SAS-MULTTEST code to provide the adjusted p -values for these comparisons. The adjusted p -values are displayed in Table 3.13. The column labeled Raw

TABLE 3.12

Proc MULTTEST Code to Obtain Simultaneous Test Information for Eight Contrasts

```
proc multtest data=ex1 pval bonferroni holm sidak stepsid FDR;
class task;
test mean(pulse20);
contrast '1-2' 1 -1 0 0 0 0;
contrast '1-.5*(2+3)' 2 -1 -1 0 0 0;
contrast '1-3' 1 0 -1 0 0 0;
contrast '1+2+3-4-5-6' 1 1 1 -1 -1 -1;
contrast '4-5' 0 0 0 1 -1 0;
contrast '6-.5*(4+5)' 0 0 0 -1 -1 2;
contrast '1+6-4-5' 1 0 0 -1 -1 1;
contrast '1+6-.5*(2+3+4+5)' 2 -1 -1 -1 -1 2;
```

TABLE 3.13Raw and Adjusted p -Values for the Eight Contrasts

Contrast	Raw	Bonferroni	Stepdown Bonferroni	Šidák	Stepdown Šidák	False Discovery Rate
1 – 2	0.7072	1.0000	1.0000	0.9999	0.8230	0.7072
1 – 0.5 * (2 + 3)	0.4386	1.0000	1.0000	0.9901	0.8230	0.5847
1 – 3	0.1024	0.8190	0.4095	0.5785	0.3508	0.1638
1 + 2 + 3 – 4 – 5 – 6	0.5426	1.0000	1.0000	0.9981	0.8230	0.6202
4 – 5	0.0007	0.0056	0.0056	0.0055	0.0055	0.0056
6 – 0.4 * (4 + 5)	0.0195	0.1557	0.1362	0.1455	0.1285	0.0699
1 + 6 – 4 – 5	0.0444	0.3555	0.2222	0.3048	0.2033	0.0889
1 + 6 – 0.5 * (2 + 3 + 4 + 5)	0.0262	0.2096	0.1572	0.1914	0.1473	0.0699

contains the unadjusted t p -values. The Bonferroni column contains the Bonferroni adjusted p -values. The column Stepdown Bonferroni contains the Bonferroni–Holm adjusted p -values. The column Šidák contains the Šidák adjusted p -values. The column Stepdown Šidák contains the Šidák–Holm adjusted p -values. The column False Discovery Rate includes the Benjamini–Hochberg adjusted p -values. The Šidák adjusted p -values are generally smaller than those for the Bonferroni t . The stepdown Šidák adjusted p -values are generally smaller than those of the stepdown Bonferroni, and are smaller than the Šidák adjusted p -values. The Bonferroni, stepdown Bonferroni, Šidák, and stepdown Šidák methods provide values that are larger than those for the false discovery rate method. The false discovery rate would usually be used when there are numerous tests being made and it is not of interest to control the FWER.

To demonstrate the application of the extension of the multivariate t to dependent comparisons, a confidence interval will be computed about $\mu_1 - \frac{1}{2}(\mu_2 + \mu_3)$ given that the linearly independent or primary set contains $\mu_1 - \mu_2$ and $\mu_1 - \mu_3$. In this case, $\mu_1 - \frac{1}{2}(\mu_2 + \mu_3) = \frac{1}{2}(\mu_1 - \mu_2) + \frac{1}{2}(\mu_1 - \mu_3)$, which in terms of the results in Section 3.13, $l^* = \mu_1 - \frac{1}{2}(\mu_2 + \mu_3) = \frac{1}{2}(\mu_1 - \mu_2) + \frac{1}{2}(\mu_1 - \mu_3)$, and $\lambda_1 = \lambda_2 = \frac{1}{2}$. The critical difference used to evaluate l^* (from Equation 3.12) is

$$\begin{aligned}
t_{\alpha/2,p,v} \hat{\sigma} \sum_{q=1}^p \left(|\lambda_q| \sqrt{\sum_{i=1}^t c_{iq}^2 / n_i} \right) &= (2.649)(5.559) \left(\left| \frac{1}{2} \right| \sqrt{\frac{1}{13} + \frac{1}{12}} + \left| \frac{1}{2} \right| \sqrt{\frac{1}{13} + \frac{1}{10}} \right) \\
&= \frac{1}{2}(5.895) + \frac{1}{2}(6.194) \\
&= 6.044
\end{aligned}$$

Multiplying this result by 2 provides the width of a confidence interval as 12.088. The width of the confidence interval when $\mu_1 - \frac{1}{2}(\mu_2 + \mu_3)$ was one of the linearly independent set was 10.319, 10.352, and 13.390 for the multivariate t , Bonferroni t and Scheffé methods, respectively. The set of linearly independent comparisons is called the primary comparisons and those that are linear functions of those in the primary set are called secondary comparisons. Using Equation 3.12 to provide critical differences for secondary comparisons yields values that are considerably larger than would be obtained as a primary comparison, but in this case the width is still shorter than that provided by the Scheffé method. Thus, comparisons that are of secondary importance when using the multivariate

t method are much less sensitive than those using any of the other methods, except for the Scheffé method. Similar results hold for all other pairs of means.

3.17 Multiple Range Tests

Duncan's and Student–Newman–Keul's multiple range tests do not control the FWER and thus cannot be recommended. See Westfall et al. (1999, p. 154) for a discussion. These procedures are described next since they have been used extensively by researchers in many areas.

3.17.1 Student–Newman–Keul's Method

This method requires that the n_i be equal to one another; however, a good approximation can be obtained provided that the sample sizes are not too unequal. In this case, the variable n in the following formulas is replaced by the harmonic mean of the n_i . The harmonic mean is given by

$$\tilde{n} = t \left(\frac{1}{n_1} + \frac{1}{n_2} + \cdots + \frac{1}{n_t} \right)^{-1}$$

The procedure controls the EERC under the all means equal hypothesis, but not if some means are equal and some means are not equal.

To apply this method, rank the t means in ascending order as $\bar{y}_{(1)} \leq \bar{y}_{(2)} \leq \cdots \leq \bar{y}_{(t)}$. Next, the Studentized range across the t means is

$$\frac{\bar{y}_{(t)} - \bar{y}_{(1)}}{\hat{\sigma}/\sqrt{n}}$$

and this is compared with the critical point $q_{\alpha,t,v}$. If

$$\bar{y}_{(t)} - \bar{y}_{(1)} \leq \frac{\hat{\sigma}}{\sqrt{n}} q_{\alpha,t,v}$$

then one concludes that there are no significant differences among the t treatment means and no further tests are considered. If

$$\bar{y}_{(t)} - \bar{y}_{(1)} > \frac{\hat{\sigma}}{\sqrt{n}} q_{\alpha,t,v}$$

then one concludes that $\mu_{(t)} > \mu_{(1)}$ or there is a significant range over the t means. Next, carry out comparisons using the two subsets of means $\{\bar{y}_{(t)}, \bar{y}_{(t-1)}, \dots, \bar{y}_{(2)}\}$ and $\{\bar{y}_{(t-1)}, \bar{y}_{(t-2)}, \dots, \bar{y}_{(1)}\}$. One compares the range of each set with $[\hat{\sigma}/\sqrt{n}]q_{\alpha,t-1,v}$. If

$$\bar{y}_{(t)} - \bar{y}_{(2)} \leq \frac{\hat{\sigma}}{\sqrt{n}} q_{\alpha,t-1,v}$$

then one concludes there are no differences among the treatment means in the subset given by $\{\mu_{(t)}, \mu_{(t-1)}, \dots, \mu_{(2)}\}$. Likewise if

$$\bar{y}_{(t-1)} - \bar{y}_{(1)} \leq \frac{\hat{\sigma}}{\sqrt{n}} q_{\alpha, t-1, v}$$

then one concludes there are no differences among the treatment means in the subset given by $\{\mu_{(t-1)}, \mu_{(t-2)}, \dots, \mu_{(1)}\}$. If both of the above conclusions hold, then one stops the process of making comparisons among means and one concludes that the means given in each of the two subsets of means are not significantly different.

If

$$\bar{y}_{(t)} - \bar{y}_{(2)} > \frac{\hat{\sigma}}{\sqrt{n}} q_{\alpha, t-1, v}$$

then one concludes that $\mu_{(t)} > \mu_{(2)}$. Likewise if

$$\bar{y}_{(t-1)} - \bar{y}_{(1)} > \frac{\hat{\sigma}}{\sqrt{n}} q_{\alpha, t-1, v}$$

one concludes that $\mu_{(t-1)} > \mu_{(1)}$. If either of these two conclusions hold, then one would carry out additional comparisons within those subsets where significant differences are found. For example, if both of the preceding sets contained significant differences, then one would consider the subsets given by $\{\bar{y}_{(t)}, \bar{y}_{(t-1)}, \dots, \bar{y}_{(3)}\}$, $\{\bar{y}_{(t-1)}, \bar{y}_{(t-2)}, \dots, \bar{y}_{(2)}\}$ and $\{\bar{y}_{(t-2)}, \bar{y}_{(t-3)}, \dots, \bar{y}_{(1)}\}$. The range of each would be compared with $(\hat{\sigma}/\sqrt{n})q_{\alpha, t-2, v}$. One continues to examine smaller subsets of means as long as the previous subset had a significant range. Each time a range proves nonsignificant, the means involved are included in a single group. No subset of means grouped in a nonsignificant group can later be deemed significant; that is, no further tests should be carried out on means that have previously been grouped into a common subgroup. When all range tests prove nonsignificant, the procedure is complete. Any two means grouped in the same group are not significantly different; otherwise, they are significantly different.

The method is illustrated using the task data from the example in Section 1.6 using $\alpha = 0.05$. Since the sample sizes are not equal, the harmonic mean of the sample sizes is 11.23. First rank the means in ascending order, as

Task	6	5	2	1	3	4
Mean	28.818	29.500	31.083	31.923	35.800	38.000
Rank	1	2	3	4	5	6

The first step is to compare across the range of six means or compare $38.000 - 28.818 = 9.182$ with

$$q_{0.05, 6, 62} \frac{\hat{\sigma}}{\sqrt{n}} = (4.16) \sqrt{\frac{30.9045}{11.23}} = 6.900$$

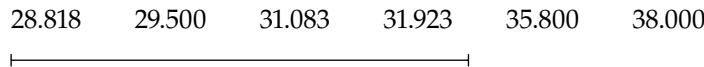
Since $9.819 > 6.90$, one examines the two subsets of $t - 1 = 5$ means. In this case, compare both $35.800 - 28.818 = 6.982$ and $38.000 - 29.500 = 8.500$ with

$$q_{0.05, 5, 62} \frac{\hat{\sigma}}{\sqrt{n}} = (3.98) \sqrt{\frac{30.9045}{11.23}} = 6.593$$

Since $7.619 > 6.60$ and $8.5 > 6.60$, next look at subsets of $t - 2 = 4$ means, of which there are three. Here one compares $31.923 - 28.818 = 3.105$, $35.800 - 29.500 = 6.300$, and $38.0 - 31.083 = 6.917$ with

$$q_{0.05, 4, 62} \frac{\hat{\sigma}}{\sqrt{n}} = (3.74) \sqrt{\frac{30.9045}{11.23}} = 6.195$$

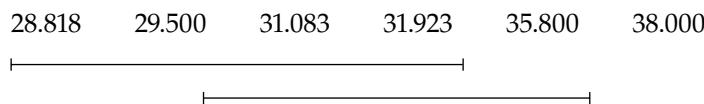
Since $3.105 < 6.195$, the first four means are grouped in a single group, and since $6.300 > 6.195$ and $6.917 > 6.195$, both of the remaining groups of four means must be further subdivided into groups of three means. Before proceeding with this next step, consider the following schematic diagram, which illustrates the present position where the line indicates the first four means form a group and are considered to not be different:



The first subset of four means, $\{29.500, 31.083, 31.923, 35.800\}$, contains two groups of three means that have already been grouped together, namely, $29.500-31.923$ and $28.818-31.083$. Hence, the ranges $31.923 - 29.500 = 2.423$ and $31.083 - 28.818 = 2.265$ are not compared with a critical point. The ranges that still must be compared are $35.800 - 31.083 = 4.717$ and $38.0 - 31.923 = 6.077$; these must be compared with

$$q_{0.05, 3, 62} \frac{\hat{\sigma}}{\sqrt{n}} = (3.40) \sqrt{\frac{30.9045}{11.23}} = 5.635$$

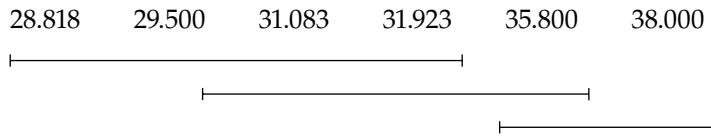
Since $4.717 < 5.635$, these three means are now grouped in a single group, whereas since $6.077 > 5.635$, the second group must be further subdivided into groups of two means each. The following diagram illustrates the present position:



There is only one subset of size two that has not already been combined into a common group, that being $\{35.8, 38.0\}$. The range $38.0 - 35.8 = 2.2$ is compared with

$$q_{0.05, 2, 62} \frac{\hat{\sigma}}{\sqrt{n}} = (2.83) \sqrt{\frac{30.9045}{11.23}} = 4.691$$

Since $2.2 < 4.69$, these final two means are grouped together. The final diagram is:



Another way to illustrate this information is to label means with the same letter if they occur in the same group and with different letters if they occur in different groups. Thus, the comparisons of the task data means can be represented as

Task	Mean
1	31.923 bc
2	31.083 bc
3	35.800 ab
4	38.000 a
5	29.500 c
6	28.818 c

3.17.2 Duncan's New Multiple Range Method

At present, this procedure is generally referred to as Duncan's method. It is one of the more popular methods, partly because it is often easier to find significant differences using this method than by any other, except perhaps Fisher's LSD method. But while Duncan's method controls the comparisonwise error rate, it does not control the FWER. This procedure also requires equal n_i , and as in the preceding section, the variable n in the following formulas can be replaced by \tilde{n} if the sample sizes are not too unequal for an approximate procedure.

Application of this procedure is similar to the application of the Student–Newman–Keul method except that the Studentized range critical point for comparing a group of p means, $q_{\alpha,p,v}$ is replaced by $q_{\alpha_p,p,v}$ where $\alpha_p = 1 - (1 - \alpha)^{p-1}$. Values of $q_{\alpha_p,p,v}$ are given in the Appendix Table A.5. For the data in the preceding section, this procedure is applied as follows:

- 1) Compare $38.000 - 28.818 = 9.182$ with

$$q_{\alpha_6,6,62} \frac{\hat{\sigma}}{\sqrt{\tilde{n}}} = (3.198) \sqrt{\frac{30.9045}{11.23}} = 5.303$$

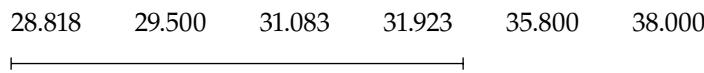
where $\alpha_6 = 1 - (1 - 0.05)^{6-1} = 0.226$. The range of six means is significant.

- 2) Compare $35.800 - 28.818 = 6.982$ and $38.000 - 29.500 = 8.500(3.143)(1.659) = 5.213$ with

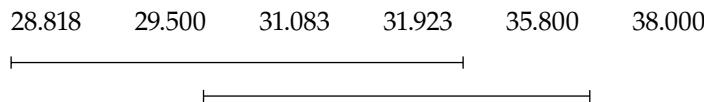
$$q_{\alpha_5,5,62} \frac{\hat{\sigma}}{\sqrt{\tilde{n}}} = (3.143) \sqrt{\frac{30.9045}{11.23}} = 5.214$$

where $\alpha_5 = 1 - (1 - 0.05)^{5-1} = 0.185$. Both ranges are significant.

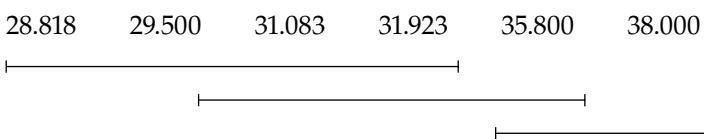
- 3) Compare $31.923 - 28.818 = 3.105$, $35.800 - 29.500 = 6.300$, and $38.000 - 31.083 = 6.917$ with $(3.073)(1.659) = 5.098$. The latter two ranges are significant, while the first range is not. The grouping at this point is



- 4) Compare $35.800 - 31.083 = 4.717$ and $38.000 - 31.923 = 6.077$ to $(2.976)(1.659) = 4.937$. The second range is significant, while the first range is not. The groupings are now:



- 5) Compare $38.000 - 35.800 = 2.200$ to $(2.829)(1.659) = 4.693$. The range is not significant. The final diagram is:



In this case the Student–Newman–Keul’s method and Duncan’s method yield the same diagram; however, often they do not since they use different quantiles of the Studentized range. The Duncan’s method uses smaller values to compare the ranges of means, which is the reason it does not control the FWER.

3.18 Waller–Duncan Procedure

This procedure is not applicable to most messy data situations because it also requires equal n_i , but is implemented using the harmonic mean of the sample sizes when they are unequal. The procedure is included for several reasons, but primarily it is included because the procedure is not well known and because it seems to have some desirable properties, which are discussed later. The Waller–Duncan procedure uses the sample data to help determine whether a conservative rule (like Tukey–Kramer) or a nonconservative rule (like Fisher’s LSD) is needed to make pairwise comparisons among the treatment means. The procedure makes use of the computed value of the F -test for testing $H_0: \mu_1 = \mu_2 = \dots = \mu_r$. If the F -value is small, then the sample data tend to indicate that the means are homogeneous. In this case, the Waller–Duncan procedure requires a large absolute difference in sample means in order to declare significance, so as to prevent declaring too many differences as being significant. If the F -value is large, the sample data would tend to indicate that the means are heterogeneous. In this case, the procedure requires a smaller absolute difference in the sample means in order to declare significance, so as to prevent declaring too few differences significant.

The Waller–Duncan procedure requires a constant K called the error rate ratio to be chosen, which designates the seriousness of a type I error relative to a type II error. The relationship between K and the probability of a type I error is approximately given below:

	Typical Value		
α	0.10	0.05	0.01
K	50	100	500

Thus, the critical points for this procedure depend on K , v , t , and the computed value of the F -statistic for testing $H_0: \mu_1 = \mu_2 = \dots = \mu_v$. Tables have not been included here; they are given in Ott (1988). To use the procedure, one calculates a Waller–Duncan LSD and compares all pairs of means to this single LSD value just as one would do with Fisher's LSD procedure or the Tukey–Kramer HSD procedure for the equal-sample-size problem. When one requests the Means/Waller option in SAS-GLM, the Waller–Duncan LSD value is computed and the means are grouped into diagrams similar to those given in the last two sections.

3.19 Example—Multiple Range for Pairwise Comparisons

The Duncan's, Student–Newman–Keul's, and Waller–Duncan methods can be accomplished using the Means statement of SAS-GLM. Table 3.14 contains the GLM code to fit the one-way model to the task data set and the Means statements are included to provide the three methods to compare the six task means. Table 3.15 contains the results for Duncan's multiple comparison method. The critical range values are computed and are presented to enable one to look at a range of means and determine if they are different. For example, when the group involves four means, the range is compared with 5.096. The means are listed from largest to smallest and the letters A, B, and C show the significant groupings. Table 3.16 contains similar results for the Student–Newman–Keul's method. As indicated above, the critical range values for the Student–Newman–Keul's method are larger than those for the Duncan's method. This increase in the value of the critical range values enables the Student–Newman–Keul's procedure to control the experimentwise error rate under the complete null hypotheses (all means equal) where the experimentwise error rate is not controlled with Duncan's method. Table 3.17 contains the results for the Waller–Duncan method. The minimum significant difference is 4.448, a value smaller than any of

TABLE 3.14
Proc GLM Code to Produce the Duncan,
Student–Newman–Keul's, and Waller–
Duncan Multiple Comparison Procedures

```
PROC GLM DATA=EX1; CLASS TASK;
MODEL PULSE20=TASK;
MEANS TASK/DUNCAN;
MEANS TASK/SNK;
MEANS TASK/WALLER;
```

TABLE 3.15

Results of Using Duncan's Multiple Comparison Procedure

α	0.05
Error degrees of freedom	62
Error mean square	30.90445
Harmonic mean of cell sizes	11.22547
<i>Note:</i> Cell sizes are not equal	
Number of means	2 3 4 5 6
Critical range	4.691 4.935 5.096 5.213 5.303

Means with the Same Letter are Not Significantly Different

Duncan Grouping	Mean	N	TASK
A	38.000	10	4
B	35.800	10	3
B	31.923	13	1
B	31.083	12	2
	29.500	12	5
C	28.818	11	6

Note: This test controls the type I comparisonwise error rate, not the experimentwise error rate.**TABEL 3.16**

Results of SNK Multiple Comparison Procedure

α	0.05
Error degrees of freedom	62
Error mean square	30.90445
Harmonic mean of cell sizes	11.22547
<i>Note:</i> Cell sizes are not equal	
Number of means	2 3 4 5 6
Critical range	4.691 5.635 6.195 6.593 6.900

Means with the Same Letter are Not Significantly Different

SNK Grouping	Mean	N	TASK
A	38.000	10	4
B	35.800	10	3
B	31.923	13	1
B	31.083	12	2
	29.500	12	5
C	28.818	11	6

Note: This test controls the type I experimentwise error rate under the complete null hypothesis but not under partial null hypotheses.

the unadjusted t values in Table 3.4. The minimum significant difference uses the harmonic mean of the sample sizes to provide one value for all of the comparisons. The minimum significant difference is small because the value of the F -statistic is large, indicating that there are likely to be differences among the means.

TABLE 3.17

Results of the Waller–Duncan Multiple Comparison Procedure

K ratio	100
Error degrees of freedom	62
Error mean square	30.90445
F-value	4.49
Critical value of t	2.03599
Minimum significant difference	4.7775
Harmonic mean of cell sizes	11.22547
<i>Note:</i> Cell sizes are not equal	

Means with the Same Letter are Not Significantly Different

Waller Grouping	Mean	N	TASK
A	38.000	10	4
B	35.800	10	3
B	31.923	13	1
B	31.083	12	2
C	29.500	12	5
C	28.818	11	6

Note: This test minimizes the Bayes risk under additive loss and certain other assumptions.

3.20 A Caution

Before concluding this chapter, it is noted that the underlining or grouping procedure can give rise to inconsistencies when the sample sizes are unequal. Consider an example where the estimate of the standard deviation is 2, which is based on 50 degrees of freedom. The sample sizes and sample means are:

Treatment	1	2	3	4
$\hat{\mu}_i$	39.3	40.1	42.0	43.0
n_i	2	25	25	2

The F -statistic needed to test the hypothesis of equal means has a value of 4.90, with a significance level of 0.005. The 5% LSD value for comparing μ_1 with μ_2 , μ_1 with μ_3 , μ_2 with μ_4 , and μ_3 with μ_4 is $(2.008)(2)\sqrt{(\frac{1}{2} + \frac{1}{25})} = 2.951$. The 5% LSD value for comparing μ_1 with μ_4 is $(2.008)(2)\sqrt{(\frac{1}{2} + \frac{1}{2})} = 4.016$ and the 5% LSD value for comparing μ_2 with μ_3 is $(2.008)(2)\sqrt{(\frac{1}{25} + \frac{1}{25})} = 1.136$. Thus, for these data the difference between the largest and smallest means, $\hat{\mu}_1 - \hat{\mu}_4 = 3.7$, is not significant, while the smaller difference between the two middle means, $\hat{\mu}_2 - \hat{\mu}_3 = 1.9$, is significant. One can explain this apparent inconsistency by noting that there is enough information (large enough sample sizes) to claim a statistically significant difference between μ_2 and μ_3 , but not enough information (sample sizes are too small) to claim a statistically significant difference between any other pairs of means.

3.21 Concluding Remarks

In this chapter, procedures for making many inferences from a single experimental data set were discussed. Such procedures are necessary to ensure that differences between treatments that are observed are due to real differences in the parameter functions being compared and not due to chance alone. Some procedures are more appropriate for planned comparisons, while other are more appropriate for data snooping. For studies with many variables, the F -test for equality of all means can be used to eliminate carrying out multiple comparisons on some variables. If the F -test is not significant at the desired α level, then one need not necessarily carry out further comparisons. But if the F -test is significant at the desired α level, then a multiple comparison method that adjusts for multiplicity should be used to compare the means (this is not Fisher's protected LSD). Finally, the multiple comparison and testing procedures discussed can be applied to the unequal variance models (discussed in Chapter 2) by using the LSMeans statement in SAS-Mixed where the REPEATED statement is used to specify the unequal variance structure. Recommendations about which procedure to use in a given circumstance were also given.

3.22 Exercises

- 3.1 Use the data in Exercise 1.1 and carry out the following multiple comparisons.
- 1) Carry out all pairwise comparisons among the five means using methods of Bonferroni, Scheffé, Tukey–Kramer, Šidák, simulate, and Fisher.
 - 2) Consider technique 4 as the control and use methods of Dunnett, Bonferroni, Šidák, Scheffé, multivariate t , and simulate to compare all other techniques to technique 4.
 - 3) Use the methods of Bonferroni, Bonferroni–Holm, Šidák, Šidák–Holm, simulate and Scheffé to provide adjusted p -values with familywise error rate protection for the following linear combinations of the means:

$$\begin{aligned} & \mu_1 + \mu_2 - \mu_3 - \mu_4, \mu_1 + \mu_2 - \mu_3 - \mu_5, (\mu_1 + \mu_2 + \mu_3)/3 - (\mu_3 + \mu_4)/2 \\ & (\mu_1 + \mu_2 + \mu_3)/3 - (\mu_3 + \mu_4)/2, (\mu_1 + \mu_4 + \mu_5)/3 - (\mu_2 + \mu_3 + \mu_5)/3 \\ & (\mu_1 + \mu_4 + \mu_5)/3 - (\mu_3 + \mu_4)/2 \end{aligned}$$

- 3.2 Use the data in Exercise 1.2 and carry out the following multiple comparisons.
- 1) Use ration 1 as the control and compare the other four rations to the control using the methods of Dunnett, Bonferroni, Šidák, Scheffé, multivariate t , and simulate.
 - 2) Use the false discovery rate to make all pairwise comparisons.
- 3.3 Use the data in Exercise 1.3 and carry out the following multiple comparisons.
- 1) Provide simultaneous tests that control the FWER to 0.05 for the following hypotheses:

$$\begin{aligned} H_0: -2\mu_1 - \mu_2 - 0\mu_3 + \mu_4 + 2\mu_5 = 0 \text{ vs } H_a: (\text{not } H_0) \\ H_0: 2\mu_1 - \mu_2 - 2\mu_3 - \mu_4 + 2\mu_5 = 0 \text{ vs } H_a: (\text{not } H_0) \end{aligned}$$

$$H_0: -1\mu_1 + 2\mu_2 + 0\mu_3 - 2\mu_4 + 1\mu_5 = 0 \text{ vs } H_a: (\text{not } H_0)$$

$$H_0: 1\mu_1 - 4\mu_2 + 6\mu_3 - 4\mu_4 + 1\mu_5 = 0 \text{ vs } H_a: (\text{not } H_0).$$

- 2) Use the methods of Bonferroni, Scheffé, Tukey–Kramer, Šidák, simulate, and Fisher to carry out all pairwise comparisons.
- 3) Consider $\mu_1 - \mu_2$, $\mu_1 - \mu_3$, $\mu_1 - \mu_4$, and $\mu_1 - \mu_5$ as primary comparisons and use the multivariate t to construct a confidence interval about the secondary comparison $\mu_1 - (1/4)(\mu_2 + \mu_3 + \mu_4 + \mu_5)$.
- 3.4 Use the data in Exercise 1.1 and construct simultaneous 95% confidence intervals about σ^2 , $\mu_1 - \mu_2$, $\mu_1 - \mu_3$, $\mu_1 - \mu_4$, $\mu_1 - \mu_5$, $\mu_2 - \mu_3$, $\mu_2 - \mu_4$, $\mu_2 - \mu_5$, $\mu_3 - \mu_4$, $\mu_3 - \mu_5$, $\mu_4 - \mu_5$.
- 3.5 For the data in Exercise 2.3 use the unequal variance model and the methods of Bonferroni, Bonferroni–Holm, Šidák, Šidák–Holm, simulate, and Scheffé to provide adjusted *p*-values and simultaneous confidence intervals about the following five comparisons:
- 1) The mean of the Blue Choc = the mean of the Red Choc.
 - 2) The mean of the Buttons = the mean of the means of the Blue Choc and Red Choc.
 - 3) The mean of the ChocChip = the mean of the WchocChip.
 - 4) The mean of the Small Choc = 1/2 the mean of the means of the Blue Choc and Red Choc.
 - 5) The mean of the Blue Choc and Red Choc = the mean of the ChocChip and WchocChip.

4

Basics for Designing Experiments

Properly designed and analyzed experiments provide the maximum amount of information about the conditions investigated for the resources used. This chapter presents concepts and methods for experimenters to use in designing and analyzing experiments. The basic concepts discussed in this chapter are *treatment structure* and *design structure* as well as the ideas of *replication*, *blocking*, and *experimental unit*. Examples of combining design and treatment structures are presented to demonstrate the concepts for complete block and incomplete block designs. These designs involve one size of experimental unit. Chapter 5 describes the concept of the size of the experimental units and presents various designs involving more than one size of an experimental unit. The design structures presented in this chapter include the completely randomized (CRD), randomized complete block (RCBD), incomplete block (IBD) and Latin square (LSD). The treatment structures considered include the one-way, two-way, two-way with controls, fractional factorial, and n -way structures. In this chapter, the models and analysis of variance tables with necessary sources of variation and degrees of freedom are presented. The discussion provides methods to determine the sources of variation used to compute the error sum of squares and algorithms to use to compute the resulting degrees of freedom. In general, the error sums of squares are obtained from comparisons of observations or linear combinations of *observations that are treated alike*. The computation of other sums of squares is discussed in later chapters. The basic approach in this chapter is to demonstrate the concepts with examples. The split-plot, repeated measures, strip-plot and crossover designs use the concept of different sizes of experimental units and are described in Chapter 5. Designs involving nesting are also discussed in Chapter 5.

4.1 Introducing Basic Ideas

Design of experiments is concerned with planning experiments in order to obtain the maximum amount of information from the available resources. The design of an experiment

should start with a statement of objectives to be attained. The primary response variable or variables should be specified and a connection between the experimental objectives and the response variables should be fully described. The entities to be used as experimental units should be explained and related to the objectives of the experiment. Often, when designing an experiment, the experimenter has control over certain factors called treatments, populations, or treatment combinations. The experimenter generally controls the choice of the experimental units to be used and whether those experimental units can be put into groups, called blocks. A typical experiment involves t treatments (or treatment combinations) that are to be compared or whose effects are to be studied.

Before an experiment can be carried out, several questions must be answered:

- 1) *What are the objectives of the experiment and what is (are) the response variable(s)?* This is a very important aspect of the design as it is imperative to understand the process that is to generate the data and how the data relate to the objectives.
- 2) *How many treatments are to be studied?* The number of treatments may already be specified, but sometimes a discussion as to how the choice of the treatments relates to the objectives of the study is in order. For example, a nutrition student wanted to design a study to determine the maximum amount of soy flour that can be used in place of wheat flour in a cookie product so that there is no soy flavor in the resulting cookies. The student selected cookie formulations that involved replacing 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100% of the wheat flour with soy flour. The first question asked was, "If you use 100% soy flour can you taste the soy flavor?" For if you cannot taste the soy flavor with 100% soy flour, then there is no need to run the experiment, and if you can taste the soy flavor there is no need to include 100% soy flour in the set of treatments. After discussing the set of treatments, it was decided that it was unknown if soy flavor could be tasted with products made from 20, 30, 40, and 50% soy flour. That is, it was determined that one could not taste soy flavor with 10% soy flour and one could taste the soy flavor with more than 50% soy flour. After relating the selection of the treatments to the objectives of the study, only five (0, 20, 30, 40, and 50% soy flour) of the initial 11 treatments were needed. The 0% was included as the wheat flour control. This process greatly reduced the number of samples needed for the study.
- 3) *How many times does each treatment need to be observed?* This question relates to the sample sizes needed to achieve the specific objectives.
- 4) *What are the experimental units?* A lot of experiments involve only one size of experimental unit. But the important idea involved is that of an independent replication. Some experiments involve more than one size of experimental unit, a concept described in Chapter 5. Often researchers think that only one size of experimental unit is involved in an experiment and fail to recognize situations where more than one size of an experimental unit is involved. This question needs to be carefully addressed, as discussed in Chapter 5.
- 5) *How does the experimenter apply the treatments to the available experimental units and then observe the responses?* This question relates to the use of randomization to assign treatments to experimental units as well as the use of randomization in other parts of the process providing the data. It is important that randomization be used to assign treatments to experimental units. It is also important to run samples through the laboratory in a random order or to use a random order to evaluate subjects in a medical study.

- 6) *Can the resulting design be analyzed or can the desired comparisons be made?* This is possibly the most important question for the researcher, but a major goal of this book is to enable the reader to use more complicated designs and still be able to carry out the analyses to estimate important parameters and test desired hypotheses.

The answers to these questions are not necessarily straightforward and the questions cannot be answered in a general way. Hopefully, the ideas and concepts discussed here will help the experimenter put together enough information to provide answers for their study.

To continue, consider an experiment involving t treatments in which each treatment is applied to r different experimental units. A mathematical model that can be used to describe y_{ij} , the response observed from the j th experimental unit assigned to the i th treatment, is

$$y_{ij} = \mu_i + \varepsilon_{ij} \quad \text{for } i = 1, 2, \dots, t, \text{ and } j = 1, 2, \dots, r \quad (4.1)$$

where μ_i is the true, but unknown, mean of the responses to the i th treatment and ε_{ij} is a random variable representing the noise resulting from natural variation and other possible sources of random and nonrandom error. Researchers should do their best to control nonrandom sources of error, which include model error, measurement error, observational error, and a misspecification of treatments errors. In order to conduct this experiment, the researcher must select rt experimental units and then randomly assign r of the experimental units to each treatment. The *randomization part of this process is very important in preventing bias* from entering into the treatment assignments. Just by the fact that experimental units are randomly assigned to treatments, a randomization or permutation analysis can be used to develop the theory for an appropriate analysis (Kempthorne, 1952). The very least that can be said about the use of randomization is that it prevents the introduction of systematic bias into the experiment. If the experimenter does not use randomization, then she cannot tell whether an observed difference is due to differences in response of the experimental units to the treatments or due to the systematic method used to assign the experimental units to the treatments.

The statistical objective of an experiment is to compare the observed response of treatments on experimental units. For example, if the researcher wants to compare the effect of a hypertension compound on human blood pressures, then using white mice as experimental units in the study will not enable one to make inferences to humans. Inferences can only be made to the population of experimental units from which those used in the study are a representative sample. It is very important to characterize the population of experimental units to which one wishes to make inferences. The sample of experimental units used in the study must be randomly selected from the population to make inferences to that population. Often it is not possible to carry out a random selection of experimental units from a population of experimental units to be included in study. But at a minimum, the sample of experimental units must be representative of the population or a conceptual population. For example, when sampling a production line, one hopes the items selected for the study will be representative of the yet to be produced population of items. If the production process is not changed, then it is reasonable to assume that the selected items will be representative of future items. It is very important to describe the population of experimental units represented by the experimental units used in the experiment. If the experimental units are not representative of the population of experimental units to which inferences are to be made, inferences cannot be made, but instead, one can make conjectures as to the effect of the treatments to an unsampled population. For example, information about hypertension compounds acquired from a study involving white mice can be used to make a conjecture (not make an inference) about the effects of the hypertension compounds on humans.

When selecting experimental units for a study, one obtains better comparisons between treatments when the set of experimental units used in the study is homogeneous or very nearly alike. In many experiments, it is impossible to select rt identical experimental units. The nonidentical experimental units contribute to the noise in the data through the ε_{ij} . When experimental units are not homogeneous or not alike, there are three methods than can be used to help account for variation among the experimental units. One method is to measure characteristics that describe differences among the experimental units, such as weight and age, and use analysis of covariance as described in Milliken and Johnson (2001). A second method is to group experimental units into sets of nearly alike experimental units. Experimental units that are nearly alike are called homogeneous. When this is the case, the treatments can be compared on the similar experimental units within a group where the group to group variation can be accounted for in the analysis. Groups of similar experimental units are called blocks. A third method of accounting for variability among experimental units is to use both blocking and covariate analysis. Let there be r blocks each with t experimental units where each treatment occurs once in each block. A model that represents the observed response of the i th treatment in the j th block is

$$y_{ij} = \mu_i + b_j + \varepsilon_{ij}^* \quad \text{for } i = 1, 2, \dots, t, \text{ and } j = 1, 2, \dots, r \quad (4.2)$$

For model (4.2), the ε_{ij} s in model (4.1) have been replaced by $\varepsilon_{ij} = b_j + \varepsilon_{ij}^*$; that is, the variation between groups or blocks of experimental units has been identified and isolated from ε_{ij}^* , which represents the variability of experimental units within a block. By isolating the block effect from the experimental units, the within-block variation can be used to compare treatment effects, which involves computing the estimated standard errors of contrasts of the treatments.

Two treatments (or any contrast of treatments) can be compared, free of block effects, by taking within-block differences of the responses of the two treatments as

$$\begin{aligned} y_{ij} - y_{i'j} &= (\mu_i + b_j + \varepsilon_{ij}^*) - (\mu_{i'} + b_j + \varepsilon_{i'j}^*) \\ &= \mu_i - \mu_{i'} + \varepsilon_{ij}^* - \varepsilon_{i'j}^* \end{aligned}$$

which does not depend on the block effect b_j . The result of this difference is that the variance of the difference of two treatment responses within a block depends on the within-block variation among the experimental units and not the between-block variation.

An objective of experimental design is to select and group the experimental material so that the noise or experimental error amongst the experimental units within groups in the experiment is reduced as much as possible. Thus, the experimental units on which the treatments are to be compared should be as much alike as possible so that a smaller difference between two treatments can be detected as a significant difference.

If there are t treatments and t experimental units, an experiment can be conducted and the mean of each treatment can be estimated from the observations. But an estimate of the error variance cannot be obtained. An estimate of the error variance (associated with ε_{ij} or ε_{ij}^*) can be obtained only when some or all of the treatments are replicated. A *replication* of a treatment is an independent observation of the treatment. Thus two replications of a treatment must involve two experimental units. An experimental unit is the entity to which the treatment has been applied. But the replications of the treatments must involve processes so that the treatment is applied and observed independently on each experimental unit. Therefore, suppose a researcher wants to compare the effect of two diets on the

growth rate of rabbits and he has 10 rabbits to use in the study. The process would be to randomly assign five rabbits to each of the treatments. But what if he puts the five rabbits assigned to one of the diets in one cage and the five rabbits assigned to the second diet in another cage where the rabbits within a cage are fed the diet from a common bowl. The individual rabbit was randomly assigned to the diet, but since all five rabbits are put into one cage and fed from a common bowl, the rabbits are not observed independently of each other. In this case, the cage of four rabbits becomes the experimental unit instead of the individual rabbit. The cages are treated independently of each other, but the rabbits within a cage are not treated independently. Thus the rabbits within the cage do not provide independent replications of the diet. It is very important to understand the complete process involved in carrying out a study in order to be able to see when independent replications occur and when they do not occur. It is imperative that this definition be observed to determine when replications are utilized during an experiment. Too often researchers use duplicate or split samples to generate two observations and call them replicates, when, in reality, they are actually sub-samples or repeated measures. Duplicates certainly do not provide the same information as independent replicates.

Suppose a researcher wanted to study the differences in the heights of male and female students at a specified university. Assume there are 22,000 students in the university. A process could be to randomly select 100 female students and 100 male students from the population of students at the university. The next step would be to find all of the students in the random sample and measure their heights, thus producing 100 measurements of heights of females and 100 measurements of heights of males. This seems like a lot of work to find all 200 students. Suppose, instead, that the researcher selects one female and one male and measures the height of each 100 times. This second process produces 100 measurements of heights of females (one female in this case) and 100 measurements of heights of males (also one male in this case). There are 200 observations in each of the data sets, but the variability among the 100 measurements within a set from the first process provides a measure of the variance among female students and among male students at the university. The variability of the 100 measurements within a set in the second case provides a measure of variance among the 100 measurements made on the same person. This second case provides information about the measurement process variability (repeated measurements of the same person), but not information about the variability among heights of females or of males at the specified university. The independent measurements of the height of one person do not provide a measure of the true variation in the heights of the population of people. The measurements of the heights of 100 females provide 100 replications where the 100 measurements of a single person provide repeated measurements on that person. These 100 measurements on the same person are called repeated measurements or sub-samples, but not replications.

A baker ran an experiment to compare the abilities of three preservatives to inhibit mold growth in a certain type of cake product. The baker mixed and baked one cake with each preservative. The number of mold spores per cubic centimeter of cake is measured after nine days of storage. The baker wanted 10 replications for the analysis so he split each cake into 10 slices and obtained the spore count on each slice. However, those 10 measurements did not result from 10 independent applications of the preservative. The variation measured by his sub-samples is an index of the within-cake variation and not an index of the experimental-unit-to-experimental-unit or cake-to-cake within a preservative variation. To obtain 10 replications of each preservative, the baker needs to bake 10 cakes with each preservative. These cakes need to be mixed and baked independently of each other. It might be easier to mix up one large batch of cake dough, mix in the preservative and then

pour the mixture into 10 cake pans. This process provides 10 cakes, but the cakes are not mixed independently of one another, so they are not independent replications. The baker needs to mix 10 batches of dough with preservative and then bake one cake from each batch in different ovens (or at different times) to obtain 10 replications of a preservative.

Another example of nonreplication involves what some researchers call a *strip trial*. In agronomy, the strip trial consists of planting all the seed of a given variety of plant in one row (or group of rows) and each row (or group of rows) is planted with a different variety. The rows (or groups of rows) are then partitioned into, say, eight parts, and the parts are called “replications.” The diagram in Figure 4.1 represents six strip plots of four rows each where the strip plots are split into eight parts denoted as replications. The advantage of using a strip trial instead of eight independent replications is that the researcher need not continually change the seed from one planter box position to another planter box position in the strip trial and the planting plan is very simple. Now, if she wants to run a well-designed experiment, she would have to change the planter boxes eight times, as dictated by a specific randomization scheme. Doing such a randomization would provide eight blocks in a randomized complete block design structure. If the experimenter analyzes the strip trial data as she would a randomized complete block design, her analysis will be incorrect. In fact, the strip trial experiment cannot be used to make inferences about variety differences since there is only one independent observation of each variety. The four rows or strip plot is the experimental unit to which the variety is applied. In the strip trial, the researcher could have just as easily partitioned the strip plots into 20 parts or 100 parts; after all, with more “replications,” one can detect smaller differences between two means as being significant. But these measurements made on each of these strips are not true replications; instead they are subsamples or repeated measurements of the strip plots. Thus, obtaining more and more parts does not aid in detecting differences between the means. One test for determining whether a part or an observation is a true replication is

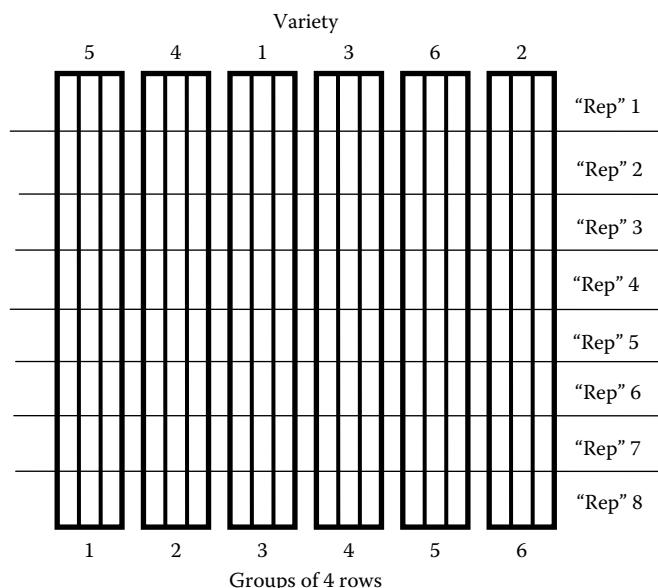


FIGURE 4.1 Schematic of a strip plot design with four row plots and six varieties arranged in eight pseudo-replications.

the following: If the researcher could have just as easily obtained more “replications” by splitting, then she is not obtaining true replications, but is obtaining subsamples or repeated measures. It is very *important to distinguish between a subsample and a replication* since the error variance estimated from between subsamples is in general considerably smaller than the error variance estimated from replications or between experimental units. The values of F -statistics constructed using the error variance computed from subsamples will be much larger than they should be, leading the experimenter to determine more differences as being statistically significant than she should.

4.2 Structures of a Designed Experiment

A designed experiment consists of two basic structures and it is vital to be able to identify and distinguish between each structure. Before an appropriate model and analysis can be constructed for a specific design, all of the factors used in a design of an experiment must be classified as belonging to either the treatment structure or the design structure. The following definitions and discussion are used to help delineate the differences between the two structures.

Definition 4.1: The *treatment structure* of a designed experiment consists of the set of treatments, factors, treatment combinations, or populations that the experimenter has selected to study and/or compare.

The treatment structure is constructed from those factors or treatments to be compared as measured by their effect on given response variables. The factors in the treatment structure must be selected to address the stated objectives of the experiment. The treatment structure could be a set of treatments, called a one-way treatment structure, or a set of treatment combinations, such as a two-way factorial arrangement or a higher-order factorial arrangement, plus any controls or standard treatments.

Definition 4.2: The *design structure* of a designed experiment consists of the grouping of the experimental units into homogeneous groups or blocks.

The design structure of a designed experiment involves the factors used to form groups of experimental units so that the conditions under which the treatments are observed are as uniform as possible. If all the experimental units are very homogeneous, then there need only be one group or block of observations and the experimental units can be assigned to the treatments completely at random. Such a design structure is called a *completely randomized design structure*.

If more than one group of experimental units is required so that the units within a group are much more homogeneous than experimental units between groups, then the design structure involves some type of a blocked design. There are several factors that can be used to form blocks of experimental units, but, as is discussed below, the factor or factors used to construct blocks must not interact with the factors in the treatment structure. Once the treatment structure and design structure have been selected, the designed experiment is specified by describing exactly the method of randomly assigning (randomizing) the treatments of the treatment structure to the experimental units in the design structure.

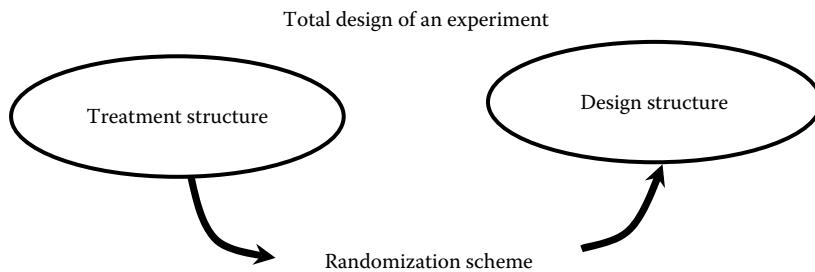


FIGURE 4.2 Graphical demonstration of the use of randomization to combine the treatment structure with the design structure to form the total design of an experiment.

Thus, the total designed experiment involves 1) the choice of the treatment structure, 2) the choice or the design structure, and 3) the method of randomly assigning the treatments or treatment combinations in the treatment structure to the experimental units in the design structure. Figure 4.2 represents the two parts of a designed experiment.

The total designed experiment dictates the appropriate model to be used to obtain an appropriate analysis. In constructing the model to describe the design, two basic assumptions are made about design and treatment structures. First, it is assumed that the components of the design structure are random effects, that is, the blocks used in the design are a random sample from the population of possible blocks of experimental units. This assumption implies there is a population of blocks of experimental units to which the researchers wish to make inferences. Second, it is assumed that there is no interaction between the components or factors of the design structure and the components or factors of the treatment structure. In other words it is assumed that the relationships existing between the treatments will be consistent from block to block (except for random variation) or, stated another way, that the blocking factors will not influence the relationship between the treatments. Many text books describe blocking factors as nuisance factors and do not address the possibility of the nuisance factors interacting with the factors in the treatment structure (Cobb, 1997). If such interactions can occur, those nuisance factors must either be included in the treatment structure or the values of the nuisance factors could be considered as covariates with the possibility of unequal slopes (Milliken and Johnson, 2001). The selection of factors for possible use of constructing blocks must be carefully evaluated to prevent probable interaction with factors in the treatment structure.

The design structure is selected by using all available knowledge of the experimental units and is chosen independently of the treatment structure (do not let the treatment structure influence the selection of a poor design structure). Likewise, the experimenter should select the treatment structure without any knowledge of the design structure (do not let the design structure hamper the selection of the necessary set of treatments). After the appropriate design structure is specified and the desired treatment structure selected, some compromises may be needed in either one or both of the structures to make them compatible with one another and to enable the experimenter to conduct an effective experiment.

4.2.1 Types of Design Structures

The design structure is determined by the type of blocking or grouping of the experimental units into homogeneous groups and is specified by the factors used to form the blocks.

The following are descriptions of some common design structures. There are two basic design structures which are described as complete block and incomplete block design structures. Several specific design structures are examined in more detail in Section 4.3.

- 1) *Completely randomized design structure.* In a completely randomized design structure, all experimental units are assumed to be homogeneous and the experimental units are assigned to the treatments completely at random. Generally, the treatments are assigned to an equal number of experimental units, although this is not required. This design structure may also be used when the experimental units are not homogeneous and the experimenter cannot find factors that allow them to be grouped into more homogeneous groups. Analysis of covariance could be used where the values of the blocking factors are used as covariates instead of using them to form blocks (Milliken and Johnson, 2001).
- 2) *Randomized complete block design.* If there are t treatments, then the randomized complete block design structure consists of having blocks of experimental units with t or more experimental units in each block. With the block size equal to or greater than the number of treatments, there is the possibility of having a complete set of treatments occurring within each block, thus the name, randomized *complete* block. If the block size is exactly equal to t for each block, then each treatment is randomly assigned to exactly one experimental unit within each block. If there are more than t experimental units within each block, then each treatment can be assigned to one experimental unit and some treatments (maybe all) can be assigned to more than one experimental unit. If $t = 5$ and the blocks are of size 8, then each treatment will be assigned to experimental units within each block and three of the treatments will be assigned to one additional experimental unit for a total of two replications of those treatments. One strategy would be to set out a pattern of treatments to blocks where the numbers of observations per treatment are as balanced as possible. Figure 4.3 consists of one way to assign five treatments to three blocks of size 8. Within each block, randomly assign the respective treatments to the

2	5	1	4	3	5	4	1		Block 1
3	5	3	2	1	2	5	4		Block 2
4	2	4	2	5	1	1	3		Block 3

FIGURE 4.3 Randomized complete block design structure with five treatments in blocks of size 8.

experimental units within that block. The arrangements shown are one set of the possible randomizations. If each block consists of $c \times t$ experimental units where c is an integer, then each treatment can be assigned to c experimental units within each block. This is also a randomized complete block design structure. A randomized complete block design structure is any blocking scheme in which the number of experimental units within a block is greater than or equal to the number of treatments, and thus a complete set of treatments can be assigned to experimental units in each block.

- 3) *Latin square design.* The Latin square design structure consists of blocking in two directions. For an experiment involving t treatments, t^2 experimental units are arranged into a $t \times t$ square where the rows are called *row blocks* and the columns are called *column blocks*. Thus the $t \times t$ arrangement of experimental units is blocked in two directions (row blocks and column blocks). To construct a Latin square design structure, the treatments are randomly assigned to experimental units in the square such that each treatment occurs once and only once in each row block and once and only once in each column block. See Cochran and Cox (1957) for various arrangements of treatments into row and column blocks. Blocking in two or more directions is common in many disciplines, in particular, blocking by rows and columns is useful when the experimental units occur in a rectangle (one of the dimensions could be time). Graeco-Latin squares can be used to form blocks in three directions (Cochran and Cox, 1957).
- 4) *Incomplete block designs.* Incomplete block designs occur when the number of treatments exceeds the number of experimental units in one or more blocks. When this occurs, then a complete set of treatments cannot occur within each block, hence the name “incomplete block.” There are several special incomplete block design structures such as balanced incomplete blocks and partially balanced incomplete blocks. A balanced incomplete block design structure is one where the assignment of treatments to blocks is such that every pair of treatments appears in the same block an equal number of times. A partially balanced incomplete block design structure occurs when sets of treatment occur together within blocks an equal number of times and other treatments occur a different number of times together within some blocks. The split-plot design structure, discussed in Chapter 5, is an example of a partially incomplete block design structure. Some incomplete block design structures are described in Example 4.5.
- 5) *Various combinations and generalizations.* There are various ways to group the experimental units. Sometimes a grouping does not satisfy the above definitions but still provides a valid design structure. An example is where the block sizes vary from block to block where some blocks are incomplete while others are complete. In any case, these other blocking schemes can provide an experimenter with very viable design structures with which to conduct effective experiments.

4.2.2 Types of Treatment Structures

The treatment structure consists of the various treatments or treatment combinations or factors and factor combinations that the experimenter wishes to study. The components of the treatment structure should be selected so as to relate to the objectives of the experiment which should be specified in the protocol or description of the study. Next, some common types of treatment structures are described, each of which are examined in more detail in Section 4.3.

- 1) *One-way treatment structure.* The one-way treatment structure consists of a set of t treatments or populations where there is no assumed structure among the treatments. There can be a relationship among the treatments such as using four temperatures, 120, 130, 150, and 160°C. If the treatments are constructed by combining two or more factors, the factors are not used in the representation of the treatments in the model. The one-way treatment structure can be used to represent any set of treatments and is often used to represent factorial treatment structures when some of the possible treatment combinations are missing. This approach is used in Chapter 13.
- 2) *Two-way treatment structure.* A two-way treatment structure consists of the set of treatments constructed by combining the levels or possibilities of two different factors. The resulting set of treatments, called treatment combinations, is generated by combining each possibility for one of the factors with each possibility for the other factor. If the first factor has s possibilities and the second factor has r possibilities, the combination produces sr treatment combinations. Figure 4.4 presents an example of a two-way treatment structure where factor A has three possibilities, factor B has four possibilities, and the crossing generates 12 treatment combinations.
- 3) *Factorial arrangement treatment structure.* A factorial arrangement treatment structure consists of the set of treatment combinations constructed by combining the levels of two or more factors. The two-way treatment structure is a two-way factorial arrangement. Three-way or up to an n -way treatment structures are also factorial arrangements. An n -way treatment structure is generated by combining the possibilities for n factors, where the factors have s_1, s_2, \dots, s_n possibilities, respectively, which generates $s_1 \times s_2 \times \dots \times s_n$ treatment combinations. Examples of factorial arrangement treatment structures are scattered throughout the text.
- 4) *Fractional factorial arrangement treatment structure.* A fractional factorial arrangement treatment structure consists of only a part, or fraction, of the possible treatment combinations in a factorial arrangement treatment structure. There are many systematic techniques for selecting an appropriate fraction, most of which depend on the assumptions the experimenter makes about interactions among the various

		Factor B			
		B_1	B_2	B_3	B_4
		A_1B_1	A_1B_2	A_1B_3	A_1B_4
Factor A	A_2	A_2B_1	A_2B_2	A_2B_3	A_2B_4
	A_3	A_3B_1	A_3B_2	A_3B_3	A_3B_4

FIGURE 4.4 Two-way treatment structure where factor A has three levels and factor B has four levels, generating 12 treatment combinations.

types of factors in the treatment structure. An experiment may involve eight different factors, each at two levels for a total of 2^8 treatment combinations. The researcher may want to look for important main effects and two-factor interactions. In that case, a one-fourth fraction or 64 treatment combinations could be used in the study. Such a design is often denoted as a 2^{8-4} fractional factorial arrangement. See Milliken and Johnson (1989) for more details. A Latin square arrangement treatment structure involves a three-way factorial arrangement with n row treatments, n column treatments, and n cell treatments. The Latin square arrangement consists of n^2 of the n^3 possible treatment combinations or is a $1/n^3 = (n^2/n^3)$ th fraction of the n^3 possible treatment combinations. One possible use of the Latin square arrangement is when it can be assumed there are no two-way or three-way interactions among the three factors.

- 5) *Optimal design treatment structures.* For many experimental situations all of the factors in the treatment structure are quantitative and the objective of the study is to collect data such that a particular linear or nonlinear model can be fit. The resulting treatment combinations selected using one of the design criteria (St. John and Draper, 1975) form an optimal design. This set of treatment combinations is called an optimal design treatment structure.
- 6) *Factorial arrangement with one or more controls.* The desired treatment structure used to satisfy the goals of the experiment can include combining more than one treatment structure. For example, a treatment structure for an experiment could consist of combining a one-way treatment structure of c controls with a two-way factorial arrangement treatment structure. Figure 4.5 contains one such treatment structure where the factorial arrangement consists of two levels of factor A and three levels of factor B combined with three controls.

All of the above treatment structures can always be considered as a one-way treatment structure for analysis purposes. In particular, when the treatment structure is a complex combination of two or more treatment structures, as in Figure 4.5, it is usually best to consider the set of treatments as a one-way treatment structure when analyzing the resulting data.

		Factor B						
		B_1	B_2	B_3				
Factor A		A_1	A_1B_1	A_1B_2	A_1B_3			
		A_2	A_2B_1	A_2B_2	A_2B_3			
						Control 1	Control 2	Control 3

FIGURE 4.5 Combination of one-way and two-way treatment structures to form the treatment structure with nine treatment combinations for an experiment.

Split-plot and repeated measures design structures are constructed from incomplete block design structures and factorial arrangement treatment structures involving two or more factors. In effect, the combination of the design structure and the treatment structure for the split-plot and repeated measures designs generates different sizes of experimental units, a topic that must be addressed in order to obtain an appropriate analysis. Such designs are discussed in Chapter 5.

4.3 Examples of Different Designed Experiments

There is a vast amount of published information about various types of designs used for many types of experiments, for example, see Cochran and Cox (1957), Davies (1954), Federer (1955), Hicks (1993), John (1971), Kirk (1968), Cornell (1990), Anderson and McLean (1974), Box et al. (1978), Cobb (1997), Kempthorne (1952), Laundsby and Weese (1993), Lentner and Bishop (1986), Meed (1988), Montgomery (1991), and Winer (1971). This section contains several examples that demonstrate the design structures and treatment structures described in Section 4.2. Hopefully, this discussion will help readers to apply these concepts to their own experiments. In most examples, the resulting designed experiment is named by specifying the type of design structure and the type of treatment structure. For example, a designed experiment could consist of a two-way treatment structure in a randomized complete block design structure. This method of describing a designed experiment differs from that generally used in the literature, but the authors think using the design and treatment structures is the best way to identify a designed experiment. In addition, this description also helps one to construct a suitable model and develop the appropriate analysis. For each experimental situation, the design structure and the treatment structure are specified and the corresponding model and the resulting analysis of variance table with the sources of variation and corresponding degrees of freedom are given. The formulas for computing sums of squares are not given in this section, but some examples with computations are included in other chapters. The emphasis of this chapter is determining the way the error sum of squares is computed and establishing the corresponding degrees of freedom.

4.3.1 Example 4.1: Diets

A nutritionist wants to study the effect of five diets on losing weight. The treatment structure of this experiment is a one-way classification involving a single factor, called diet, with five levels or five treatments. Many different design structures can be used to evaluate the diets. If there are 20 homogeneous people, then a completely randomized design structure can be used where each diet is randomly assigned to four people. One model for a one-way treatment structure in a completely randomized design structure is

$$y_{ij} = \mu_i + \varepsilon_{ij} \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, n_i \quad (4.3)$$

where μ_i denotes the mean of the i th treatment (diet) and ε_{ij} denotes the random error. The analysis of variance table for model (4.3), assuming the ideal conditions that $\varepsilon_{ij} \sim i.i.d. N(0, \sigma^2)$, is given in Table 4.1. Table 4.1 contains the different sources of variation and their respective degrees of freedom for the model in Equation 4.3. The 15 degrees of freedom for experimental error are obtained from the variation among persons treated alike. There are four persons

TABLE 4.1

Analysis of Variance Table for a One-Way Treatment Structure in a Completely Randomized Design Structure

Source of Variation	df
Diet	4
Error	15

given diet 1, providing four persons treated alike, and the variation among these four persons' weight loss values provides three degrees of freedom for error. There are four persons treated alike for each diet. Thus, there are three degrees of freedom available for error from the data for each of the five diets. If the variances are equal, then these five sets of three degrees of freedom can be pooled into an error term involving 15 degrees of freedom. The methods in Chapter 2 can be used to evaluate the plausibility of the equal variance assumption.

Assume the researcher could enroll and interview just five persons during a given time period. Let time period be a blocking factor and randomly assign the five diets to one person within each time period. The design is a one-way treatment structure in a randomized complete block design structure with four blocks of size five. A model that can be used to describe data from this design is

$$y_{ij} = \mu_i + b_j + \varepsilon_{ij} \quad i = 1, 2, 3, 4, 5, \quad j = 1, 2, 3, 4 \quad (4.4)$$

where μ_i denotes the mean of the i th treatment (diet), b_j denotes the effect of the j th block and ε_{ij} denotes the random error associated with the person assigned to the i th treatment in the j th block. The analysis of variance table for model (4.4) is displayed in Table 4.2. The error degrees of freedom are computed by using four orthogonal contrasts of the treatments within each block, such as,

$$\begin{aligned} q_{1j} &= y_{1j} - y_{2j} & j &= 1, 2, 3, 4 \\ q_{2j} &= y_{1j} + y_{2j} - 2y_{3j} & j &= 1, 2, 3, 4 \\ q_{3j} &= y_{1j} + y_{2j} + y_{3j} - 3y_{4j} & j &= 1, 2, 3, 4 \\ q_{4j} &= y_{1j} + y_{2j} + y_{3j} + y_{4j} - 4y_{5j} & j &= 1, 2, 3, 4 \end{aligned}$$

The q_{ij} values for the same value of i all have the same mean, indicating they are all treated alike. Thus the variance of each set of four values of q_{ij} provides three degrees of freedom

TABLE 4.2

Analysis of Variance Table for a One-Way Treatment Structure in a Randomized Complete Block Design Structure with One Replication per Treatment in Each Block

Source of Variation	df
Block (date)	3
Diet	4
Error	12

for error. Thus there are four sets of three degrees of freedom that can be pooled (if the variances are equal) into the error sum of squares with 12 degrees of freedom. However, the variances of the q_{ij} are not equal and must be rescaled before pooling. The variances are $\text{Var}(q_{1j}) = 2\sigma^2$, $\text{Var}(q_{2j}) = 6\sigma^2$, $\text{Var}(q_{3j}) = 12\sigma^2$, and $\text{Var}(q_{4j}) = 20\sigma^2$ for $j = 1, 2, 3, 4$, so the variances of the q_{ij} need to be divided by the coefficient of σ^2 before pooling. In effect, these pooled sums of squares provide the block by diet interaction sum of squares. Therefore, a second way of obtaining the variance of things treated alike is to use the block-by-treatment interaction to obtain the experimental error estimate for a randomized complete block design structure.

Next, suppose there are not 20 homogeneous persons available, but there are 10 homogeneous males and 10 homogeneous females. One strategy would be to use sex of person as a blocking factor where there are two blocks of size 10. A randomized complete block design structure could be used where each diet would be randomly assigned to two males and two females, so there are two replications of each diet within each block. The model for a one-way treatment structure in a randomized complete block design structure is

$$y_{ijk} = \mu_i + b_j + \varepsilon_{ijk} \quad i = 1, 2, 3, 4, 5, \quad j = 1, 2, \quad k = 1, 2 \quad (4.5)$$

where μ_i denotes the mean of the i th treatment (diet) effect, b_j denotes the effect of the j th block and ε_{ijk} denotes the random error associated with the k th person assigned the i th treatment in the j th block. The analysis of variance table for model (4.5) is displayed in Table 4.3. There is one degree of freedom associated with the design structure or sex of person that has been removed from the error term of Table 4.1. The error term for a one-way treatment structure in a randomized complete block design structure where there is one observation per treatment in each block is computed from the block by treatment interaction. This design involves blocking and multiple observations per treatment in each block. The number of degrees of freedom for the block by treatment interaction is equal to $(2 - 1) \times (5 - 1)$ or four degrees of freedom. The variability of the two observations of each diet within each block provides one degree of freedom. Thus there are five degrees of freedom for error from the comparisons of persons treated alike within each block, or 10 degrees of freedom pooled across the two blocks. Pooling the block by treatment interaction with the within block variability provides 14 degrees of freedom for error.

In most cases sex of person is not a good choice for a blocking factor since the treatments (diets in this case) might interact with sex of person. If the factor (or factors) selected to construct blocks can possibly interact with the treatments, then that factor needs to be included in the treatment structure either as a stratification factor or in an analysis of covariance with possibly unequal slopes to carry out an appropriate analysis (Milliken and Johnson, 2001). In this example, sex of the person must be combined with the five diets

TABLE 4.3

Analysis of Variance Table for a One-Way Treatment Structure
in a Randomized Complete Block Design Structure with Two
Replications per Treatment in Each Block

Source of Variation	df
Block (sex of person)	1
Diet	4
Error	14

to form a two-way factorial arrangement or a two-way treatment structure. This two-way treatment structure consists of the 10 treatment combinations generated by combining the two levels of sex of person with the five levels of diet. By switching sex of the person from the design structure to the treatment structure, the resulting design structure is a completely randomized design with two replications of each treatment combination. The randomization scheme is to randomly assign each diet to two males and to two females. This is the same randomization scheme as for model (4.5), but with different treatment and design structures. One model for a two-way treatment structure in a completely randomized design structure is the means model:

$$y_{ijk} = \mu_{ij} + \varepsilon_{ijk} \quad i = 1, 2, \dots, 5, \quad j = 1, 2, \quad k = 1, 2 \quad (4.6)$$

where μ_{ij} denotes the mean of the ij th treatment combination (sex of person by diet) and ε_{ijk} denotes the random error. Sometimes the mean μ_{ij} is expressed as an effects model:

$$\mu_{ij} = \mu + \tau_i + \beta_j + \gamma_{ij}$$

where μ is the overall mean, τ_i is the effect of the i th diet, β_j is the effect of the j th sex of person and γ_{ij} is the interaction effect. The analysis of variance tables for model (4.6) for both expressions of μ_{ij} are given in Table 4.4.

Next, suppose that the diets have a structure consisting of a control diet and four diets made up of the four combinations of two protein levels and two carbohydrate levels, as shown in Figure 4.6. The diet treatment structure is a two-way factorial arrangement with a control that, when crossed with sex of person, generates a three-way treatment structure (protein \times carbohydrate \times sex) with two controls where there is one control for males and one control for females. The design structure is completely randomized where each treatment combination is to be assigned to two persons. A model that can be used to describe this data is

$$y_{ijk} = \mu_{ij} + \varepsilon_{ijk} \quad i = 0, 1, 2, 3, 4, \quad j = 1, 2, \quad k = 1, 2 \quad (4.7)$$

where μ_{01} and μ_{02} denote the means of the controls and the $\mu_{ij}, i = 1, 2, 3, 4$ and $j = 1, 2$ denote the means of the diet by sex of person treatment combinations. The analysis of variance

TABLE 4.4

Analysis of Variance Table for a Two-Way Treatment Structure in a Completely Randomized Design Structure for Both Means and Effects Models

Source of Variation	df
μ_{ij} model	
Sex \times diet	9
Error	10
$\mu + \tau_i + \beta_j + \gamma_{ij}$ model	
Sex	1
Diet	4
Sex \times diet	4
Error	10

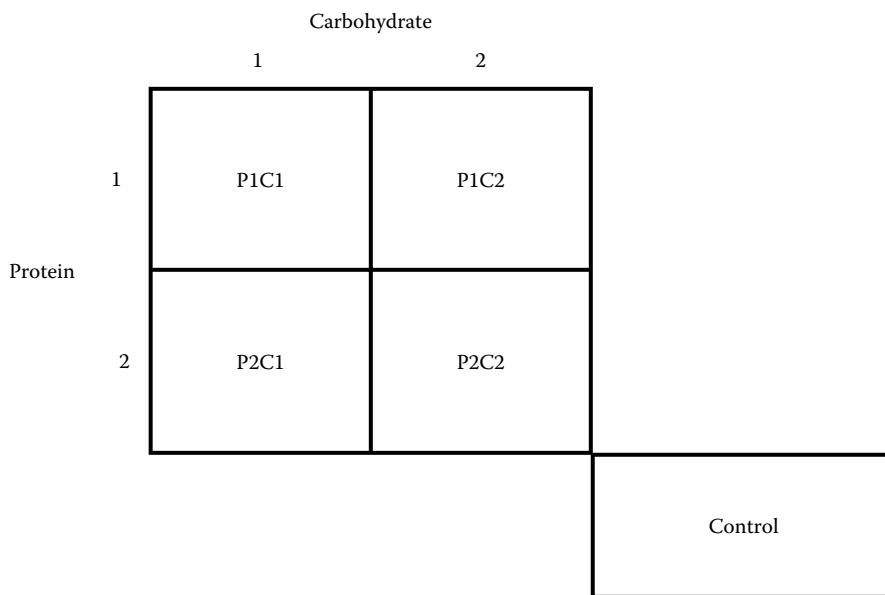


FIGURE 4.6 Structure of diets with a two-way factorial arrangement plus one control for the treatment structure.

table for model (4.7) is in Table 4.5, where “Control vs 2²” denotes a comparison between the control diet and the average of the four protein by carbohydrate treatment combinations. The complete analysis would most likely be accomplished using the two-way treatment structure of diet by sex of person with appropriate contrasts of the treatment means to provide the comparisons involving protein, carbohydrate, sex of person, and the controls. The construction of such contrasts is discussed in Chapters 6 and beyond.

TABLE 4.5

Analysis of Variance Table for a Treatment Structure Consisting of Three-Way Factorial Arrangement Combined with Two Controls in a Completely Randomized Design Structure

Source of Variation	df
Sex	1
Diet	4
Control vs 2 ²	1
Protein	1
Carbohydrate	1
Protein × carbohydrate	1
Sex × Diet	4
Sex × control vs 2 ²	1
Sex × protein	1
Sex × carbohydrate	1
Sex × protein × carbohydrate	1
Error	10

4.3.2 Example 4.2: House Paint

A paint company wants to compare the abilities of four white house paints to withstand environmental conditions. Four square houses, each with one side facing exactly north, were available for the experiment, thus houses can be used as a blocking factor. Each side of a house is possibly exposed to different types of weather, thus the sides (indicated here by directions north, south, east, and west) of the houses can also be used as a blocking factor. Since the number of treatments (the four paints) was the same as the number of levels of both blocking factors, a Latin square design structure can be used to study the paints. Here the paints can be assigned to sides of houses where each paint can occur once and only once on each house and once and only once in each direction. There are three basic Latin square arrangements (see Cochran and Cox, 1957). The randomization process is to randomly select one of the three possible arrangements, then randomly assign the rows to the houses, randomly assign the directions to the columns and randomly assign the types of paint to the letters in the square. One such arrangement of assigning paints to houses and directions is shown in Table 4.6.

The design of the experiment is a one-way treatment structure in a Latin square design structure. A model that can be used to describe data from this experiment is

$$y_{ijk} = \mu_i + h_j + d_k + \varepsilon_{ijk} \quad i = 1, 2, 3, 4, \quad j = 1, 2, 3, 4, \quad k = 1, 2, 3, 4 \quad (4.8)$$

where μ_i denotes the mean wearability score for the i th paint, h_j denotes the effect of the j th house, d_k denotes the effect of the k th direction, and ε_{ijk} denotes the experimental unit error. The analysis of variance table for model (4.8) is given in Table 4.7. The error for the Latin square design structure consists of contrasts that measure the paint by house by direction interaction.

TABLE 4.6

Assignment of a Set of Treatments from a One-Way Treatment Structure to a Latin Square Design Structure

House	Directions			
	North	South	East	West
1	A	B	C	D
2	D	A	B	C
3	C	D	A	B
4	B	C	D	A

TABLE 4.7

Analysis of Variance Table for a One-Way Treatment Structure in a Latin Square Treatment Structure

Source of Variation	df
House	3
Direction	3
Paint	3
Error	6

TABLE 4.8

Two-Way Treatment Structure for House Paint Example

		Additive II	
Additive I		None	Some
None		Base	Base + II
Some		Base + I	Base + I + II

Next, suppose the paints have a structure as given by 1) base paint, 2) base paint plus additive I, 3) base paint plus additive II, and 4) base paint plus both additive I and additive II. This structure of the paints provides a two-way treatment structure where one factor is additive I at two levels (zero and some) and the second factor is additive II at two levels (zero and some). The resulting four treatment combinations are shown in Table 4.8.

One model for a two-way treatment structure in a Latin square design structure is

$$y_{ijkm} = \mu + \gamma_i + \beta_j + (\gamma\beta)_{ij} + h_k + d_m + \varepsilon_{ijkm} \quad i = 1, 2, \quad j = 1, 2, \quad k = 1, 2, 3, 4, \quad m = 1, 2, 3, 4 \quad (4.9)$$

where γ_i denotes the effect of additive I, β_j denotes the effect of additive II, and $(\gamma\beta)_{ij}$ denotes the interaction between the two additives. The analysis of variance table for model (4.9) is given in Table 4.9. The only difference between analyzing models (4.8) and (4.9) is that in model (4.9) the paints have a structure that is used to partition the paint effect into effects due to additive I, additive II, and the interaction of additives I and II. The part of the analysis corresponding to the design structure remains unaffected even though the analysis of the treatment structure has changed.

Finally, suppose eight houses were available so that the experiment could be conducted by using two different Latin square design structures. Table 4.10 shows one possible assignment of paints to the house-direction combinations. If the paints have the two-way treatment structure in Table 4.8, then a model is given by

$$y_{ijkmn} = \mu + \gamma_i + \beta_j + (\gamma\beta)_{ij} + s_k + h_{n(k)} + d_m + \varepsilon_{ijkmn} \quad i = 1, 2, \quad j = 1, 2, \quad k = 1, 2, \quad m = 1, 2, 3, 4, \quad n = 1, 2, 3, 4 \quad (4.10)$$

TABLE 4.9

Analysis of Variance Table for a Two-Way Treatment Structure in a Latin Square Treatment Structure

Source of Variation	df
House	3
Direction	3
Paint	3
Additive I	1
Additive II	1
Additive I \times Additive II	1
Error	6

TABLE 4.10

Arrangement Showing a One-Way Treatment Structure in a Replicated Latin Square Design Structure

Direction	House							
	Square 1				Square 2			
	1	2	3	4	5	6	7	8
N	C	A	B	D	D	C	A	B
S	D	B	C	A	C	B	D	A
E	A	C	D	B	A	D	B	C
W	B	D	A	C	B	A	C	D

TABLE 4.11

Analysis of Variance Table for a Two-Way Treatment Structure in a Repeated Latin Square Treatment Structure

Source of Variation	df
Houses	7
Squares	1
Houses (square)	6
Direction	3
Paint	3
Additive I	1
Additive II	1
Additive I \times additive II	1
Error	18

where s_k denotes the effect of square k and $h_{n(k)}$ denotes the effect of house n in square k . The analysis of variance table for model (4.10) is given in Table 4.11.

4.3.3 Example 4.3: Steel Plates

A Latin square design structure is very useful when there is a need to block in two directions, but every Latin square arrangement used by experimenters is not a Latin square design structure. This example is used to demonstrate the consequences of using a Latin square arrangement treatment structure. Two types of paint additives are to be combined and steel plates are to be painted. The objective of the experiment is to study the ability of the paint combinations to protect steel from heat. There are five levels of each paint additive and five temperatures at which to check the protecting ability. This experiment is suited for a Latin square array where the levels of additive I are assigned to the rows, the levels of additive II are assigned to the columns, and the levels of temperature are assigned to the cells within the square. This arrangement generates 25 treatment combinations. The experimental units are 25 sheets of steel 0.2 cm thick and 1 m² in area. The randomization process is to randomly assign one of the 25 treatment combinations to each of the 25 sheets of steel. In this case, the treatment structure is a fraction of a 5³ factorial arrangement (as it consists of 25 of the 125 possible treatment combinations of additive I \times additive

$\text{II} \times \text{temperature}$), called a Latin square arrangement or Latin square treatment structure. The design structure is a completely randomized design, as the treatment combinations are assigned completely at random to the sheets of steel. Since this is a fractional factorial, each main effect is partially aliased (see Cochran and Cox, 1957, p. 245) with the two-factor interaction of the other two factors and the three-factor interaction. In order to properly analyze this experimental design, some assumptions must be made about the parameters in the model. The usual assumptions are that there are no two-way interactions and no three-way interaction. However, one should be very careful not to make such assumptions without having some prior information (which can come from other experiments, existing literature, etc.) showing that the interactions are, in fact negligible. One such Latin square arrangement is given in Table 4.12, and model for a Latin square treatment structure in a completely randomized design structure is

$$y_{ijk} = \mu + \text{AI}_i + \text{AII}_j + T_k + \varepsilon_{ijk}, \quad (i, j, k) \in \text{Index} \quad (4.11)$$

where AI_i denotes the effect of the i th level of additive I, AII_j denotes the effect of the j th level of additive II, T_k denotes the effect of the k th level of temperature, and Index denotes an index set consisting of the 25 treatment combinations observed in the experiment.

If one ignores the application of the levels of temperature, the resulting data table is that of a two-way treatment structure in a completely randomized design structure, as shown in Table 4.13. The analysis of variance table for this two-way treatment structure is in Table 4.14. The design consists of one observation per treatment combination, so there are

TABLE 4.12

Latin Square Arrangement Treatment Structure, Where T_i Denotes the i th Level of Temperature

Level of Additive I	Level of Additive II				
	1	2	3	4	5
1	T_1	T_2	T_3	T_4	T_5
2	T_5	T_1	T_2	T_3	T_4
3	T_4	T_5	T_1	T_2	T_3
4	T_3	T_4	T_5	T_1	T_2
5	T_2	T_3	T_4	T_5	T_1

TABLE 4.13

The Two-Way Treatment Structure for the Levels of Additive I by the Levels of Additive II, Ignoring the Levels of Temperature in the Latin Square Arrangement

Level of Additive I	Level of Additive II				
	1	2	3	4	5
1	(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)
2	(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)
3	(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)
4	(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)
5	(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)

TABLE 4.14

Analysis of Variance Table for Two-Way Treatment Structure Part of the Latin Square Arrangement

Source	<i>df</i>
Additive I	4
Additive II	4
Additive I \times additive II	16
Error	0

no degrees of freedom available for estimating the error as there no sheets of steel treated alike. The interaction between the levels of AI and the levels of AII is associated with 16 degrees of freedom. Now, when temperature is included in the structure, the sum of squares due to testing the equality of temperature means is part of the AI by AII interaction sum of squares, as denoted in Table 4.15. This means that part of the AI by AII interaction is identical to the temperature effect, or four degrees of freedom associated with temperature are *aliased* with four degrees of freedom associated with the AI by AII interaction. Likewise, the four degrees of freedom associated with the AI effect are aliased with four degrees of freedom of the AII by temperature interaction and the four degrees of freedom associated with the AII effect are aliased with four degrees of freedom of the AI by temperature interaction. The analysis of variance table for the Latin square treatment structure using model (4.11) is given in Table 4.16. The term “residual” is used rather than “error” since the corresponding sum of squares involves error plus any interaction effects that may not be zero. If the assumption of zero interactions is not correct, then the residual mean square will be too large and the resulting *F*-tests will be too small. Consequently, if there is interaction in the experiment, it cannot be discovered and any other detectable treatment effects may be masked.

4.3.4 Example 4.4: Levels of *N* and *K*

A model and the resulting analysis consist of three basic components, i) the treatment structure, ii) the design structure, and iii) the error structure(s). This example demonstrates how the three basic components can be used to construct the model. A plant breeder wants

TABLE 4.15

Analysis of Variance Table for Two-Way Treatment Structure Part of the Latin Square Arrangement with the Variation for Temperatures Partitioned from the Additive I by Additive II Interaction

Source	<i>df</i>
Additive I	4
Additive II	4
Additive I \times additive II	16
Temperature	4
Residual	12
Error	0

TABLE 4.16

Analysis of Variance Table for the Latin Square Arrangement Treatment Structure

Source	<i>df</i>
Additive I	4
Additive II	4
Temperature	4
Residual	12

to study the effect of combining three levels of nitrogen (N) and four levels of potassium (K) on the yield of his new variety of corn. His treatment structure is a two-way factorial arrangement with 12 (3 levels of $N \times 4$ levels of K) treatment combinations. He has three parcels of land on which to carry out the experiment and he uses these parcels of land as blocks. Each of the blocks is partitioned into 12 parts called plots. Each treatment combination is randomly assigned to one plot within each block. Thus, the design structure is a randomized complete block design since each treatment combination occurs once in each block. The total design of the experiment is called a two-way treatment structure in a randomized complete block design structure. (Blocks in a randomized complete block design are called replications by some authors; however, we prefer to call them blocks instead of replications in order to distinguish them from replications in the completely randomized design structure. The discussion in Example 4.5 describes the important differences between blocks and replications.) The model for this example is

$$y_{ijk} = \mu_{ij} + b_k + \varepsilon_{ijk} \quad i = 1, 2, 3, \quad j = 1, 2, 3, 4, \quad k = 1, 2, 3 \quad (4.12)$$

where μ_{ij} is the mean of the i th level of N with the j th level of K , b_k is the effect of the k th block, and ε_{ijk} denotes the random error associated with the plots within each block. In general, the general model is constructed by summing the models for each of the three structures as

$$Y = \text{treatment structure} + \text{design structure} + \text{error structure(s)} \quad (4.13)$$

Likewise, the corresponding analysis of variance table has three parts. The general analysis of variance table for model (4.13) is given in Table 4.17. The analysis of variance table for model (4.12) is given in Table 4.18. In general, allowance must be made for the possibility of more than one error term. For example, split-plot and repeated measures models have more than one error term (see Chapter 5).

TABLE 4.17

Analysis of Variance Table for the General Model

Source of Variation	<i>df</i>
Design structure	df_{DS}
Treatment structure	df_{TS}
Error structure(s)	$df_{\text{ERROR}(S)}$

TABLE 4.18

Analysis of Variance Table for a Two-Way Treatment Structure
in a Randomized Complete Block Design Structure

Source of Variation	df
<i>Design structure</i>	
Blocks	2
<i>Treatment structure</i>	
N	2
K	3
$N \times K$	6
<i>Error structure</i>	
Block \times treatment	22

4.3.5 Example 4.5: Blocks and Replications

In many of the textbooks on the design of experiments, there is either no distinction made between blocks and replications or, at the very least, there is confusion about the distinction. This set of examples is included to demonstrate the difference between the two concepts as well as to illustrate that the combination of a treatment structure with a design structure can result in more than one total design of the experiment. Suppose the researcher wants to study the effect of four treatments in a one-way treatment structure using a design structure with six blocks of size two (she has only two homogeneous experimental units per block). In this case, the required design structure is an incomplete block design. If there are enough blocks so that every pair of treatments can occur together in a block the same number of times, then it is possible to use a balanced incomplete block design structure (Cochran and Cox, 1957). For example, the four treatments could be assigned to blocks, as shown in Table 4.19. In this case, there are six blocks in the design structure, and each treatment is replicated three times. This example is used to point out that the concepts of blocks and replications are different and to emphasize that blocks and replications should always be kept separate. Blocks and replications are equivalent only for the case of a randomized complete block design structure where each treatment is observed once and only once within each block. The randomization process for the incomplete block design structure consists of assigning blocks to block numbers and then randomly assign the two treatments assigned to that block to the two experimental units within the block. For the example in Table 4.19, the design structure is associated with the six blocks (see Table 4.19), not the three replications that just happen to occur because of the assignment process. The model for the arrangement in Table 4.19 is

$$y_{ij} = \mu_i + b_j + \varepsilon_{ij}, \quad \text{for } (i, j) \in \text{Index} \quad (4.14)$$

TABLE 4.19

First Assignment of Four Treatments to Six Blocks of Two Experimental Units, Providing a Balanced Incomplete Block Design Structure

Block 1	Block 2	Block 3	Block 4	Block 5	Block 6
A	A	A	B	B	C
B	C	D	C	D	D

where Index = $\{(A, 1), (B, 1), (A, 2), (C, 2), (A, 3), (D, 3), (B, 4), (C, 4), (B, 5), (D, 5), (C, 6), (D, 6)\}$ and the pair (i, j) can take on only those values of treatment \times block combinations that are observed as indicated by the Index set. The analysis of variance table for model (4.14) is given in Table 4.20. The degrees of freedom associated with this connected block-treatment arrangement are computed from the degrees of freedom associated with the block by treatment interaction as if all combinations were observed minus the number of empty cells. Table 4.21 is a display of the observed block-treatment combinations (denoted by the "X"). There are six blocks and four treatments, so if all combinations were present, the block by treatment interaction would be based on $(6 - 1)(4 - 1) = 15$ degrees of freedom. There are 12 missing cells, so the number of degrees of freedom associated with the error term is $15 - 12 = 3$.

Table 4.22 contains a second assignment pattern for assigning the four treatments to the six blocks of size two. Treatment A occurs in all blocks and is replicated six times. Treatments B, C, and D occur in two blocks providing two replications of each. Model (4.14) can be used to describe the data where the index set is

$$\text{Index} = \{(A, 1), (B, 1), (A, 2), (C, 2), (A, 3), (D, 3), (A, 4), (B, 4), (A, 5), (C, 5), (A, 6), (D, 6)\}$$

TABLE 4.20

Analysis of Variance Table for the Balanced Incomplete Block Design Structure in Table 4.19

Source	df
Block	5
Treatments	3
Error	3

TABLE 4.21

Two-Way Structure of Blocks by Treatments Where X Denotes Observed Combination Showing That 12 Cells Are Filled and 12 Cells Empty

Blocks	A	B	C	D
1	X	X		
2	X		X	
3	X			X
4		X	X	
5		X		X
6			X	X

TABLE 4.22

Second Assignment of Four Treatments to Six Blocks of Two Experimental Units, Providing an Unbalanced Incomplete Block Design Structure

Block 1	Block 2	Block 3	Block 4	Block 5	Block 6
A	A	A	A	A	A
B	C	D	B	C	D

The analysis of variance table in Table 4.20 is also appropriate for the arrangement in Table 4.22. The arrangement in Table 4.19 is optimal (using D -optimality criteria; St. John and Draper, 1975) if one wishes to compare all four treatments to one another, that is, test $\mu_A = \mu_B = \mu_C = \mu_D$. The arrangement in Table 4.22 is optimal if one wishes to compare treatment A to each of the other three treatments; that is, test $\mu_A = \mu_B, \mu_A = \mu_C, \mu_A = \mu_D$.

A third way to assign the four treatments to six blocks is presented in Table 4.23. There are two groups of treatments where A and B occur together in three blocks and treatments C and D occur together in three blocks. Treatment A does not occur in a block with either treatment C or D . The structure between the treatments and blocks is not connected because of the separation of the treatments into these two groups. Define a new variable, "group," to indicate the two groups of treatments. The comparison between the two groups is a comparison between the mean of blocks 1, 3, and 5 and the mean blocks 2, 4, and 6, or is a between-block comparison. The comparisons of treatments A with B and of treatments C with D are within-block comparisons. The number of degrees of freedom associated with the block by treatment interaction for treatments A and B is two. Likewise, the number of degrees of freedom associated with the block by treatment interaction for treatments C and D is also two. The error sum of squares is obtained by pooling these two block by treatment interaction sums of squares (as well as their degrees of freedom). The analysis of variance table for the arrangement in Table 4.23 is displayed in Table 4.24. The sum of squares due to blocks is partitioned into the sum of squares for groups and the sum of squares for blocks nested within groups. The sum of squares due to groups is used as the error to test the hypothesis ($\mu_A + \mu_B = \mu_C + \mu_D$), which has one of the degrees of freedom due to treatments. The two groups of treatments are confounded with the blocks; that is, if there is a difference between the two groups' means, you do not know if it is due to the difference between the sets of treatments or to the differences among the two groups of blocks. The concept of confounding is similar to the concept of aliasing, but aliasing involves two (or more) terms being indistinguishable where both terms are from the treatment structure and confounding involves two terms being indistinguishable where one term is from the treatment structure and one term is from the design structure.

TABLE 4.23

Third Assignment of Four Treatments to Six Blocks of Two Experimental Units, Providing an Unconnected Incomplete Block Design Structure

Block 1	Block 2	Block 3	Block 4	Block 5	Block 6
A	C	A	C	A	C
B	D	B	D	B	D

TABLE 4.24

Analysis of Variance Table for Incomplete Block Design in Table 4.25

Source	df
Groups ($\mu_A + \mu_B = \mu_C + \mu_D$)	1
Blocks (groups)	4
$\mu_A = \mu_B$	1
$\mu_C = \mu_D$	1
Error	4

The differences between the three designs in Tables 4.19, 4.22, and 4.23 are in the assignment of the treatments to the blocks. Remember the design structures and the treatment structures are identical for all three designs, so the design and treatment structures do not describe the total design of the experiment. One must also specify the method of randomly assigning the treatments from the treatment structure to the experimental units in the design structure.

4.3.6 Example 4.6: Row and Column Blocks

Blocking can occur in many ways and the Latin square design structure is one design where there are both row blocks and column blocks. There are various alterations of Latin square design structures where there are fewer rows (or columns) or more rows (or columns) than there are treatments (Cochran and Cox, 1957). This example consists of the experimental units being blocked by rows and columns where the intersection of each row and column contains several experimental units. The treatment structure is a two-way with factor A having two levels and factor B having two levels, thus generating four treatment combinations. There are 24 experimental units arranged in three columns and two rows where each row-column combination contains four experimental units. Randomly assign the four treatment combinations to the four experimental units within each of the row-column groups. Table 4.25 is a display of the assignment of treatment combinations (not randomized) to the experimental units. This design structure essentially consists of six blocks of size four, but it may be of interest to evaluate the variability between the row blocks and among column blocks, so a model can be constructed to include those factors as:

$$y_{ijkm} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + r_k + c_m + (rc)_{km} + \varepsilon_{ijkm} \quad \text{for } i = 1, 2, \\ j = 1, 2, k = 1, 2, m = 1, 2, 3 \quad (4.15)$$

where $\mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$ denotes the main effects and interaction of factors A and B , r_k denotes the row block effect, c_m denotes the column block effect, $(rc)_{km}$ denotes the interaction among the row and column blocks and ε_{ijkm} denotes the experimental unit error. The analysis of variance table corresponding to model (4.15) is in Table 4.26. The analysis could just involve six blocks and the sums of squares due to row blocks, column blocks and their interaction could be pooled to provide five degrees of freedom for the design structure. The error sum of squares is obtained from the treatment structure by design structure interaction; that is, by pooling $A \times$ row block, $A \times$ column block, $A \times$ row block \times column block, $B \times$ row block, $B \times$ column block, $B \times$ row block \times column block, $A \times B \times$ row block, $A \times B \times$ column block, and $A \times B \times$ row block \times column block sums of squares.

TABLE 4.25

Design with Row and Column Blocks in the Design Structure and a Two-Way Arrangement in the Treatment Structure (Nonrandomized Form)

	Column Block 1		Column Block 2		Column Block 3	
Row block 1	A_1B_1	A_1B_2	A_1B_1	A_1B_2	A_1B_1	A_1B_2
	A_2B_1	A_2B_2	A_2B_1	A_2B_2	A_2B_1	A_2B_2
Row block 2	A_1B_1	A_1B_2	A_1B_1	A_1B_2	A_1B_1	A_1B_2
	A_2B_1	A_2B_2	A_2B_1	A_2B_2	A_2B_1	A_2B_2

TABLE 4.26

Analysis of Variance Table for Row–Column Design Structure
with Two-Way Treatment Structure

Source	<i>df</i>
<i>Design structure</i>	5
Row blocks	1
Column blocks	2
Row blocks \times column blocks	2
<i>Treatment structure</i>	3
<i>A</i>	1
<i>B</i>	1
<i>A \times B</i>	1
Error = design structure \times treatment structure	15

There are many other ways to construct designed experiments by combining various design structures and treatment structures. Hopefully, the above examples will enable the experimenter to construct the desired designed experiment, construct an appropriate model, and develop the corresponding analysis.

4.4 Concluding Remarks

This chapter presented concepts and methods experimenters can use in designing and analyzing experiments. The basic concepts for good designed experiments were also introduced. All designed experiments consist of two basic features: the treatment structure and the design structure. These concepts are generally not used in many other statistical analysis books. Understanding the difference between these two features of a designed experiment will help data analysts select appropriate analyses for their experiments. The choice of blocking factors is discussed where it is imperative that they do not interact with the factors in the treatment structure. Finally, it is mandatory that one is able to identify when one has true replications and when one merely has subsamples.

4.5 Exercises

- 4.1 Find two published research papers in the literature which use a two-way or higher order treatment structure within a designed experiment which has only one size of experimental unit. For each paper, describe in detail the treatment structure, the design structure and the experimental unit used in the experiment. Comment as to the appropriateness of the design and its analysis [at least as far as the information provided by the author(s) is concerned].
- 4.2 The Tire Co. researcher wanted to determine if her new design of tire wears better than the existing designs. She selected three existing designs to be

evaluated with her new design. She had four tires from each design, all of the same size. She had available four cars and four drivers that could be used during the driving test. The test was to measure the tread wear during the 25,000 mile driving test. Describe how you would design an appropriate experiment for her. Describe in detail the treatment structure and the design structure. Write down the appropriate model and key out the corresponding analysis of variance table; include sources and degrees of freedom.

- 4.3 A food scientist wants to develop a healthy yet enjoyable muffin by changing some of the ingredients. He has three main factors to vary, *oil* at three levels, *sugar* at two levels, and *egg white powder* at four levels. In order to have a reference point, he compared the experimental muffins with the *standard* muffin recipe. He has a pan in which he can bake 25 muffins at a time. Design an experiment for him, describe the design and treatment structures, write down an appropriate model, and key out the corresponding analysis of variance table.
- 4.4 A plant breeder wants to evaluate how well corn plants of selected varieties grow in a high temperature-low humidity environment. A growth chamber is available for the study that can be used to control the temperature and humidity. She has four cultivars (or treatments) that should be tolerant to the hot-dry environment. The growth chamber can hold up to seven pots, each pot consisting of plants from a single cultivar. The growth chamber can be used up to four times. The growth of the plants (increase in dry matter per plant) is the measurement of interest. Design an experiment for her, describe the design and treatment structures, write down an appropriate model, and key out the corresponding analysis of variance table.
- 4.5 Discuss the changes in the designed experiment that must take place for the house paint example if the directions cause differences in the relationships of the paint means.
- 4.6 A researcher wants to set up a study to evaluate four methods of teaching statistics. One possible design would be to teach one class of students with each of the teaching methods and use the students in the class as replications of the teaching method. Discuss the implications of using this design.

5

Multilevel Designs: Split-Plots, Strip-Plots, Repeated Measures and Combinations

Consulting statisticians do not always get the chance to design the experiments for which they must help construct appropriate analyses. Instead, the statistician must first identify the type of designed experiment the researcher has employed. The first and most important step in the identification process is to determine if more than one size of experimental unit has been used, and if so, to identify each size of experimental unit. As will become evident in this section, each size of experimental unit will have an associated design structure and treatment structure. After the different sizes of the experimental units have been identified, the model for carrying out an appropriate analysis can be constructed by combining the models used to describe the design structure and treatment structure corresponding to each size of experimental unit.

5.1 Identifying Sizes of Experimental Units—Four Basic Design Structures

The design structures that involve more than one size of experimental unit are called multilevel designs structures and include split-plot type design structures, strip-plot design structures, repeated measures design structures, hierarchical or nested types of design structures and design structures involving various combinations of the above. Repeated measures and split-plot type design structures are similar, although the assumptions used to develop the analyses can be different. Split-plot and strip-plot design structures evolved from the agricultural sciences, but are used in many other disciplines including engineering and manufacturing, and their analyses are discussed in Chapters 24 and 25. Repeated measures designs are used extensively in the social and biological sciences, but are applicable in most areas when it is of interest to evaluate treatment effects over time, and their analyses are presented in Chapters 26–28. Nested treatment structures are different from the treatment structures used with repeated measures and split-plot designs and the nested treatment structures are described in Chapter 30.

There are four basic design structures and most complex design structures are combinations of these. The four basic design structures are the completely randomized design structure, the randomized complete block design structure, the split-plot design structure and the strip-plot design structure. Each of these basic design structures has its own analysis with a unique process for computing the required error sums of squares. The basic design structures are described in this section along with the process needed to compute the error sum of squares. An example is used in the discussion where the treatment structure is a two-way with one factor at two levels and the other factor is at three levels, and the design structure involves 18 experimental units.

The experiment consists of evaluating the volume of cupcakes after baking where there are three recipes and two cooking temperatures; thus the treatment structure consists of the six combinations of three recipes with the two cooking temperatures. It is desired to have three replications of each of the treatment combinations; thus 18 cupcakes need to be baked. The diagram in Figure 5.1 is used to demonstrate the process of using the completely randomized design structure. The process is to completely at random assign the six treatment combinations in the treatment structure to the 18 experimental units (cupcakes) in the design structure. The arrows indicate that each treatment combination is assigned to three cupcakes. Often the process is to make one cupcake at a time, thus the order in which the cupcakes are mixed and baked corresponds the experimental units. The completely randomized design structure is completed by mixing a batch of a given recipe, filling a cupcake form, and then baking that cupcake in an oven set to the specific temperature. Each cupcake must be made from its own batch, that is, only one cupcake per batch of cake recipe, and each cupcake must be baked by itself in the oven set to the specified temperature. This process requires the use of 18 batches of cake mix and 18 uses of one or more ovens to bake the cupcakes. The error associated with the completely randomized design structure is computed from the variability among cupcakes treated alike. There are three cupcakes from each recipe by temperature combination, thus the variability in the volumes of the three cupcakes provides two degrees of freedom to measure cupcake error. One should test the equality of the six treatment combination variances (see Chapter 2) and, if possible, the estimates of error from the six treatment combinations are pooled together to provide 12 degrees of freedom for error that measures the variability among cupcakes treated alike. A model that can be used to describe the volumes of the cupcakes is

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}, \quad i = 1, 2, \quad j = 1, 2, 3, \quad \text{and } k = 1, 2, 3$$

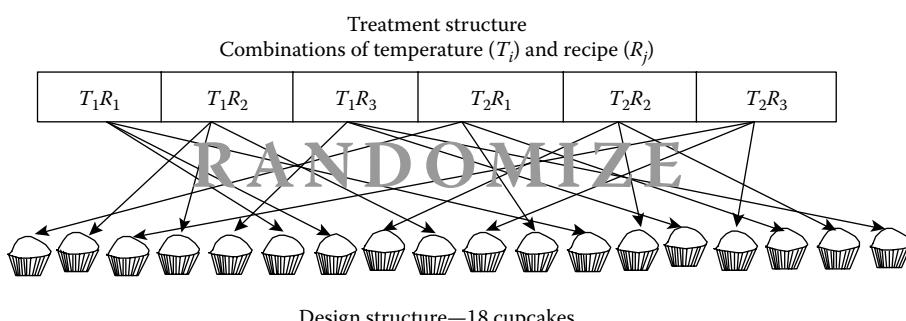


FIGURE 5.1 Randomization for a two-way treatment structure in a completely randomized design structure.

TABLE 5.1

Analysis of Variance Table for the Two-Way Treatment Structure in a Completely Randomized Design Structure

Source	df	EMS
Temperature	1	$\sigma_e^2 + \phi^2(\tau)$
Recipe	2	$\sigma_e^2 + \phi^2(\beta)$
Temperature \times recipe	2	$\sigma_e^2 + \phi^2(\tau\beta)$
Error	12	σ_e^2

where y_{ijk} denotes the volume of the k th cupcake made with the j th recipe and baked at the i th temperature, μ denotes the overall mean, τ_i denotes the effect of the i th temperature, β_j denotes the effect of the j th recipe, $(\tau\beta)_{ij}$ is the temperature by recipe interaction, and ε_{ijk} denotes the variability associated with the batches, the cupcakes within a batch, and the variability from oven bake to oven bake. Table 5.1 contains the analysis of variance table for the model for the two-way treatment structure in a completely randomized design structure where there are 12 degrees of freedom computed from the variability of experimental units or cupcakes treated alike. This analysis has five degrees of freedom associated with the treatment structure and there are 12 degrees of freedom associated with the design structure, all of which are assigned to the error term. The column of Table 5.1 labeled EMS gives the forms of the expected mean squares for the respective rows of the ANOVA table. The functions $\phi^2(\tau)$, $\phi^2(\beta)$, and $\phi^2(\tau\beta)$ represent quadratic functions in the temperature main effect means, the recipe main effect means, and the interaction effects, respectively. These functions are non-negative and equal to zero when the corresponding effects do not exist. Similar interpretations can be used throughout the remainder of this chapter.

The second basic design structure is the randomized complete block design and the diagram in Figure 5.2 is a display of the process of assigning treatments from the treatment structure to the experimental units in the design structure. Suppose that the researcher can make and bake six cupcakes per day, so the experiment must be spread

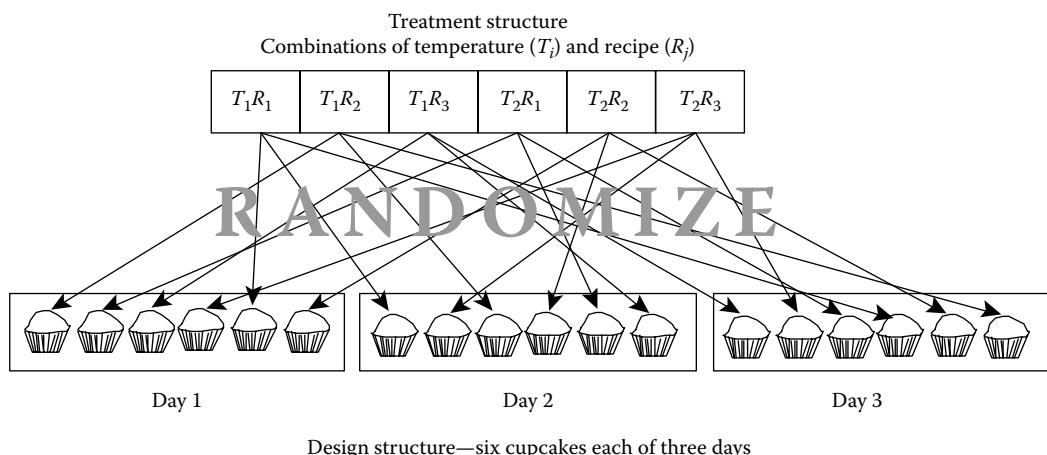


FIGURE 5.2 Randomization scheme for two-way treatment structure in randomized complete block design structure.

over three days in order to achieve three replications of each of the treatment combinations. The experimental units of the design structure are divided into three groups or blocks of size six. Next, the six treatments from the treatment structure are randomly assigned to the six experimental units within each of the blocks, as indicated by the arrows in Figure 5.2. As discussed in Chapter 4, the error sum of squares for the randomized complete block design structure is obtained by computing the treatment structure by design structure or treatment by block interaction. This design involves six treatments and three blocks, so the treatment by block interaction provides $(3 - 1)(6 - 1) = 10$ degrees of freedom associated with experimental error. The 10 degrees of freedom consist of pooling the degrees of freedom associated with the block by temperature, the block by recipe, and the block by temperature by recipe interactions. A model that can be used to describe the volumes of cupcakes for this two-way treatment structure in a randomized complete block design structure is

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + d_k + \varepsilon_{ijk}, \quad i = 1, 2, j = 1, 2, 3, \text{ and } k = 1, 2, 3$$

where d_k denotes the effect of the k th day, the blocking factor, and ε_{ijk} denotes the variability associated with the batches, the cupcakes within a batch, and the variability from oven bake to oven bake within a day. Table 5.2 contains the analysis of variance table for the model for the two-way treatment structure in a randomized complete block design structure where there are 10 degrees of freedom computed from the variability of experimental units treated alike as measured by the block by treatment combination interaction. This analysis has the same five degrees of freedom associated with the treatment structure as the completely randomized design structure, but now, the 12 degrees of freedom for the design structure are distributed between days (blocks) and the error. The term σ_{day}^2 in Table 5.2 represents the variance of d_k , $k = 1, 2, 3$. Similar interpretations can be made throughout the remainder of this chapter.

The third basic design structure is the split-plot design structure. Here the 18 cupcakes are divided into six blocks of size three, as shown in Figure 5.3. Since there are six treatment combinations in the treatment structure, all six treatments cannot occur within a block, thus this is an incomplete block design structure. Three cupcakes, one from each of the recipes, will be baked at a specified temperature within the same oven (only one cupcake was baked in an oven for the first two basic design structures). The diagram in Figure 5.3 shows that the treatment structure has been separated into two parts, one designated as the cupcake part and one for the oven part. The blocks of size three or the three cupcakes assigned to each oven form the experimental units for the levels of temperature. Thus the oven is the experimental unit for temperature. The first part of the randomization

TABLE 5.2

Analysis of Variance Table for the Two-Way Treatment Structure
in a Randomized Complete Block Design Structure

Source	df	EMS
Day	2	$\sigma_{\varepsilon}^2 + 6\sigma_{\text{day}}^2$
Temperature	1	$\sigma_{\varepsilon}^2 + \phi^2(\tau)$
Recipe	2	$\sigma_{\varepsilon}^2 + \phi^2(\beta)$
Temperature \times recipe	2	$\sigma_{\varepsilon}^2 + \phi^2(\tau\beta)$
Error	10	σ_{ε}^2

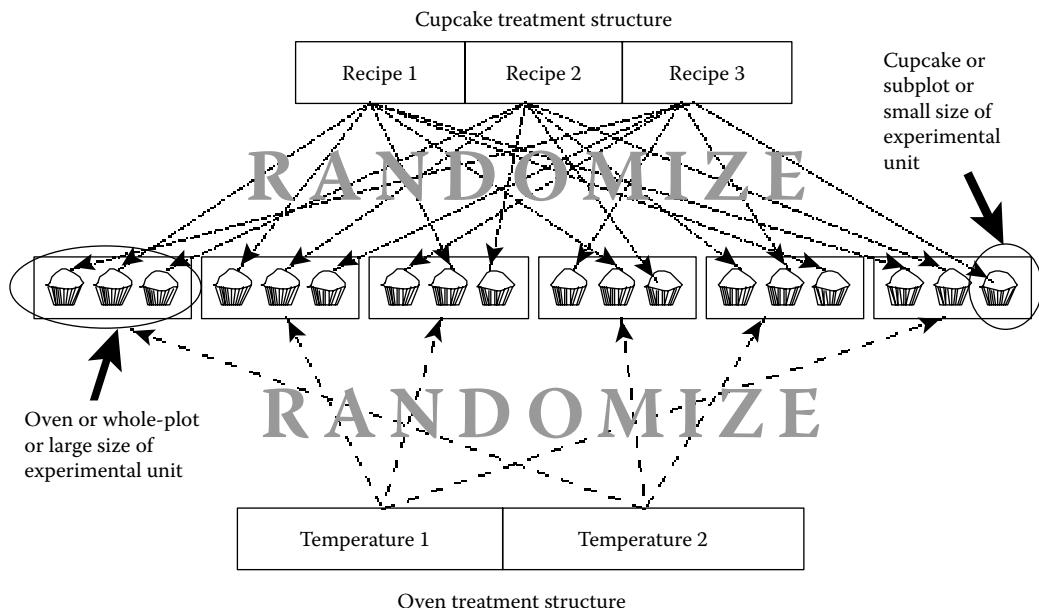


FIGURE 5.3 Randomization scheme for the split-plot with completely randomized whole-plot design structure.

procedure is to randomly assign each temperature to three of the ovens or blocks, as demonstrated by the arrows going from the levels of temperature to the ovens in Figure 5.3. The cupcakes within the ovens are the experimental units for the levels of recipe and the randomization procedure is to randomly assign the levels of recipe to one cupcake within each of the ovens. There are two sizes of experimental units and there are two design and treatment structures. The treatment and design structures for the oven experimental units consists a one-way treatment structure (two levels of temperature) in a completely randomized design structure with six ovens. For the individual cupcakes, the treatment and design structures consist of a one-way treatment structure (three levels of recipe) in a randomized complete block design structure where the ovens represent blocks of similar experimental units.

This design structure is a nested or hierarchical structure as the cupcakes are nested within the ovens. Thus, the split-plot design is also a hierarchical design structure. Since there are two sizes of experimental units, this is called a multilevel design. The oven is the larger size of experimental unit and is often called the whole-plot. The cupcake is the smaller size of experimental unit and is often called the subplot or split-plot.

The first step in the analysis of this split-plot design is to ignore the individual cupcakes or recipes and just consider the two temperatures and the six ovens. The display in Figure 5.4 indicates that the design corresponding to the oven size of experimental unit is a one-way treatment structure in a completely randomized design structure. A model to describe a response measured on each of these ovens is

$$y_{ik}^* = \mu + \tau_i + e_{ik}, \quad i = 1, 2, \text{ and } k = 1, 2, 3$$

where y_{ik}^* denotes the response measured on the k th oven assigned to the i th temperature and e_{ik} represents the error term associated with the ovens. Table 5.3 contains the analysis

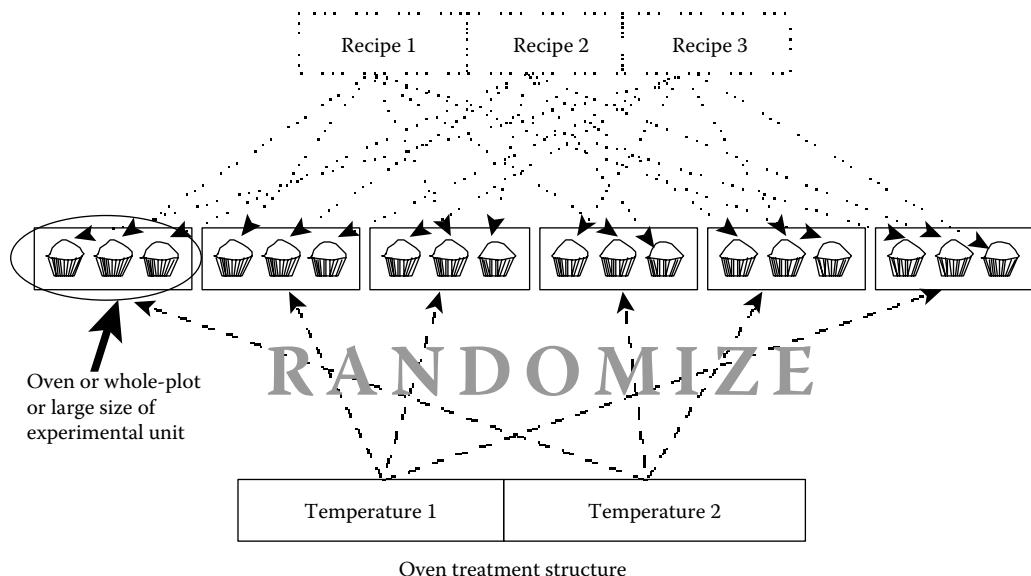


FIGURE 5.4 Oven design and treatment structures for split-plot design.

of variance table for the oven model. For each level of temperature there are three ovens that are treated alike, thus there are two degrees of freedom available from each temperature measuring how ovens vary when treated alike. If the variances of the ovens between the two temperatures are equal, then the two variances can be pooled to provide the error term for ovens with four degrees of freedom.

The next step in the analysis is to ignore the levels of temperature providing a design that is a one-way treatment structure (three recipes) in a randomized complete block design structure (six blocks). The randomization process is displayed in Figure 5.5. A model that can be used to describe the volume of a cupcake is

$$y_{ijk} = \mu + \beta_j + o_{ik} + \varepsilon_{ijk}^*, \quad i = 1, 2, j = 1, 2, 3, \text{ and } k = 1, 2, 3$$

where o_{ik} denotes the effect of the block or oven effect and ε_{ijk}^* denotes the error associated with a cupcake. The analysis of variance table is in Table 5.4 where the residual sum of squares consists of the recipe by oven interaction. If all ovens were treated alike, the residual sum of squares would provide an estimate of the cupcake to cupcake variability. However, some ovens are subjected to one temperature and the other ovens are subjected to a second temperature; hence, the recipe by oven interaction includes the temperature by

TABLE 5.3

Analysis of Variance Table for the One-Way Treatment Structure in a Completely Randomized Design Structure for the Ovens as Experimental Units Where the Levels of Recipe Are Ignored

Source	df	EMS
Temperature	1	$\sigma_e^2 + \phi(\tau)$
Error (ovens)	4	σ_e^2

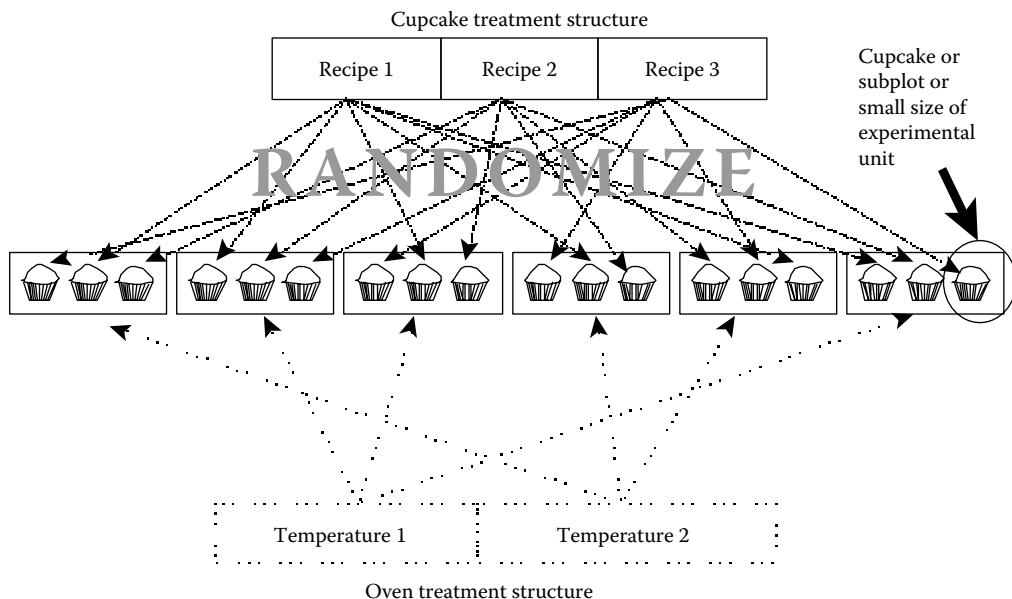


FIGURE 5.5 Randomized complete block design structure for the cupcake experimental unit.

recipe interaction. The expected mean squares in Table 5.4 are not fully determined by this model, and that is denoted by the "?" in the table. One additional reduction in the design is required in order to obtain the cupcake to cupcake variability. Consider only those ovens subjected to temperature 1, as displayed by the dark lines in Figure 5.6. The reduced design is that of a one-way treatment structure in a randomized complete block design structure with three blocks all treated alike (temperature 1), and the analysis of variance table is given in Table 5.5. The recipe by oven interaction provides the measure of how cupcakes within an oven treated alike will vary. The process continues by considering the other three ovens at the second temperature where the recipe by oven interaction provides an additional four degrees of freedom for measuring how cupcakes vary when treated alike within an oven. When the variances of the cupcakes from the two temperatures are equal (see Chapter 2 for tests), one can pool the two sources together to provide the error sum of squares for measuring how cupcakes vary when treated alike with eight degrees of freedom. The cupcake error sum of squares is computed by computing the oven by recipe interaction within a temperature pooled across temperatures. The final complete split-plot

TABLE 5.4

Analysis of Variance Table for the One-Way Treatment Structure in a Randomized Complete Block Design Structure Where the Levels of Temperature Are Ignored

Source	df	EMS
Ovens	5	$\sigma_e^2 + 3\sigma_{\text{oven}}^2 + ?$
Recipe	2	$\sigma_e^2 + \phi^2(\beta) + ?$
Residual	10	$\sigma_e^2 + ?$

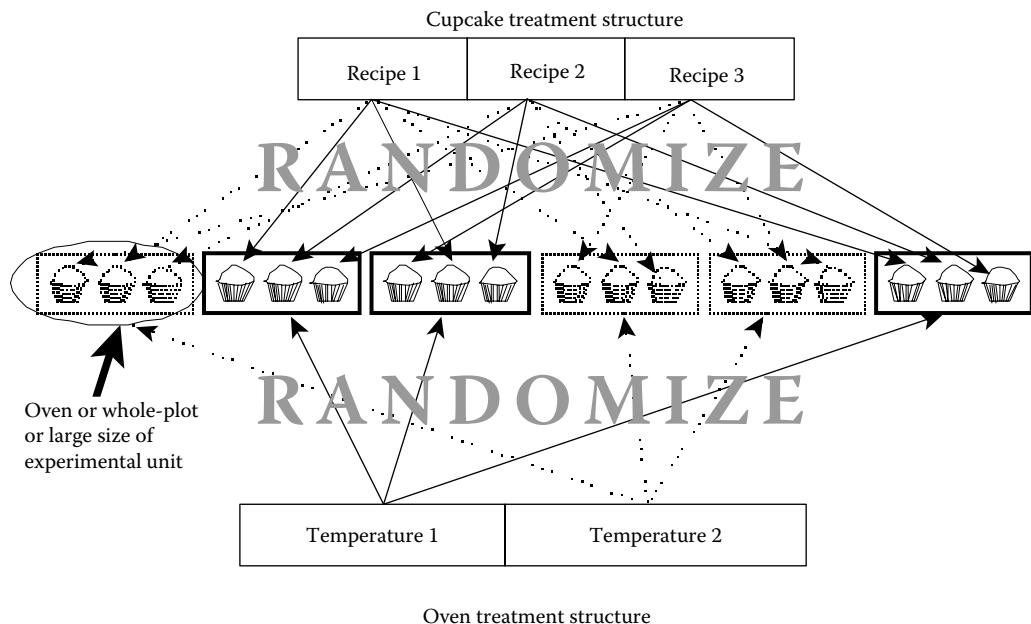


FIGURE 5.6 Part of the design structure for temperature 1 oven runs only.

TABLE 5.5

Analysis of Variance Table for Temperature 1 Data for the One-Way Treatment Structure in a Randomized Complete Block Design Structure

Source	df	EMS
Ovens	2	$\sigma_e^2 + 3\sigma_{\text{oven}}^2$
Recipe	2	$\sigma_e^2 + \phi^2(\beta)$
Residual	4	σ_e^2

analysis of variance table with the recipe by temperature interaction separated from the cupcake residual is given in Table 5.6. A model to describe the data from a two-way treatment structure in a split-plot design structure with a completely randomized design whole-plot design structure is

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + o_{ik} + \varepsilon_{ijk}, \quad i = 1, 2, \quad j = 1, 2, 3, \quad \text{and } k = 1, 2, 3$$

where o_{ik} denotes oven variation within a temperature and ε_{ijk} denotes the variability of cupcakes within an oven. There are still five degrees of freedom associated with the treatment structure in Table 5.6, but the 12 degrees of freedom associated with the design structure are distributed between the two error terms where there are four degrees of freedom associated with the oven error component and eight degrees of freedom associated with the cupcake error component.

As a variation on the split-plot design structure, suppose that the researcher can only do two oven runs within a day; thus, the ovens are separated into three sets of two ovens where the study will take three days to complete. The diagram in Figure 5.7 depicts the

TABLE 5.6

Analysis of Variance Table for the Two-Way Treatment Structure in a Split-Plot Design with a Completely Randomized Whole-Plot Design Structure

Source	df	EMS
Temperature	1	$\sigma_e^2 + 3\sigma_{\text{oven}}^2 + \phi^2(\tau)$
Error (oven)	4	$\sigma_e^2 + 3\sigma_{\text{oven}}^2$
Recipe	2	$\sigma_e^2 + \phi^2(\beta)$
Temperature \times recipe	2	$\sigma_e^2 + \phi^2(\tau\beta)$
Error (cupcake)	8	σ_e^2

random assignment of the levels of temperature to the ovens within each day and then randomly assigning the levels of recipes within each of the ovens. The oven design structure is a randomized complete block design, so the oven error term is computed by the day by temperature interaction. Imposing a blocking structure on the ovens or whole-plots does not change the design structure for the cupcakes or subplots. There still are six blocks of size three, so the cupcake part of the analysis does not change. A model to describe the volume of cupcakes for a two-way treatment structure in a split-plot design structure with a randomized complete block whole-plot design structure is

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + d_k + o_{ik} + \varepsilon_{ijk}, \quad i = 1, 2, \ j = 1, 2, 3, \text{ and } k = 1, 2, 3$$

where d_k denotes the day effect, o_{ik} denotes the oven variation within a day and ε_{ijk} denotes the variability of cupcakes within an oven within a day. The analysis of variance table corresponding to the above model is given in Table 5.7 which includes rows for days, error (oven) and error (cupcake).

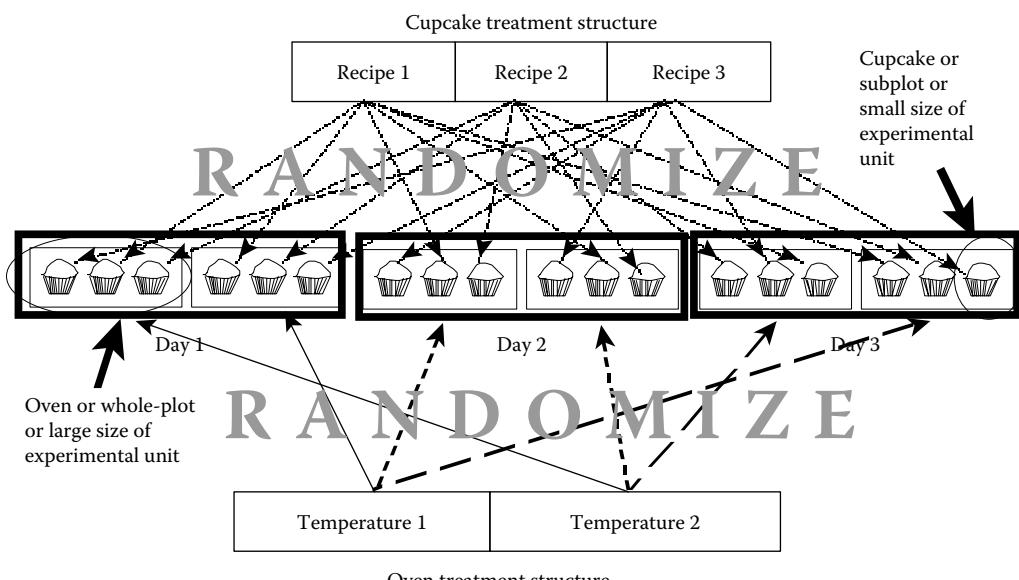


FIGURE 5.7 Diagram of split-plot design structure with a randomized complete block whole-plot (oven) design structure.

TABLE 5.7

Analysis of Variance Table for the Two-Way Treatment Structure in a Split-Plot Design with a Randomized Complete Block Whole-Plot Design Structure

Source	df	EMS
Day	2	$\sigma_e^2 + 3\sigma_{\text{oven}}^2 + 6\sigma_{\text{day}}^2$
Temperature	1	$\sigma_e^2 + 3\sigma_{\text{oven}}^2 + \phi^2(\tau)$
Error (oven)	2	$\sigma_e^2 + 3\sigma_{\text{oven}}^2$
Recipe	2	$\sigma_e^2 + \phi^2(\beta)$
Temperature \times recipe	2	$\sigma_e^2 + \phi^2(\tau\beta)$
Error (cupcake)	8	σ_e^2

As mentioned above, the split-plot design structures are incomplete block designs. Figure 5.8 contains the display of the assignment of treatment combinations to the ovens or blocks within each of the days. Only three treatment combinations can occur within each of the blocks and, in fact, within a block only those treatment combinations with the same level of temperature can occur. The resulting design is a partially balanced incomplete block where some treatment combinations always occur together within a block and some treatment combinations never occur together within a block.

The fourth basic design structure is a strip-plot design. The strip-plot design structure is constructed by first arranging the experimental units into rectangles with rows and columns, as displayed in Figure 5.9. In this experiment, a batch of cake dough is mixed using one of the recipes and two cupcakes are extracted from the batch. The batch is the entity being made from a recipe, so the batch is the experimental unit for the levels of recipe. One of the cupcakes is to be baked at temperature 1 and the other cupcake is to be baked at temperature 2. Hence, each oven will include three cupcakes, one from each recipe. The oven is the entity to which a level of temperature is assigned and is the experimental unit for the levels of temperature.

The columns of the rectangles in Figure 5.9 correspond to the batches of cake dough and the three recipes are randomly assigned to the columns; that is, both cupcakes in a column

Oven 1			Oven 2			Day 1
Temp 2 Recipe 1	Temp 2 Recipe 2	Temp 2 Recipe 3	Temp 1 Recipe 1	Temp 1 Recipe 2	Temp 1 Recipe 3	
<hr/>						
Oven 3			Oven 4			Day 2
Temp 1 Recipe 1	Temp 1 Recipe 2	Temp 1 Recipe 3	Temp 2 Recipe 1	Temp 2 Recipe 2	Temp 2 Recipe 3	
<hr/>						
Oven 5			Oven 6			Day 3
Temp 2 Recipe 1	Temp 2 Recipe 2	Temp 2 Recipe 3	Temp 1 Recipe 1	Temp 1 Recipe 2	Temp 1 Recipe 3	

FIGURE 5.8 Split-plot design structure expressed as an incomplete block design structure.

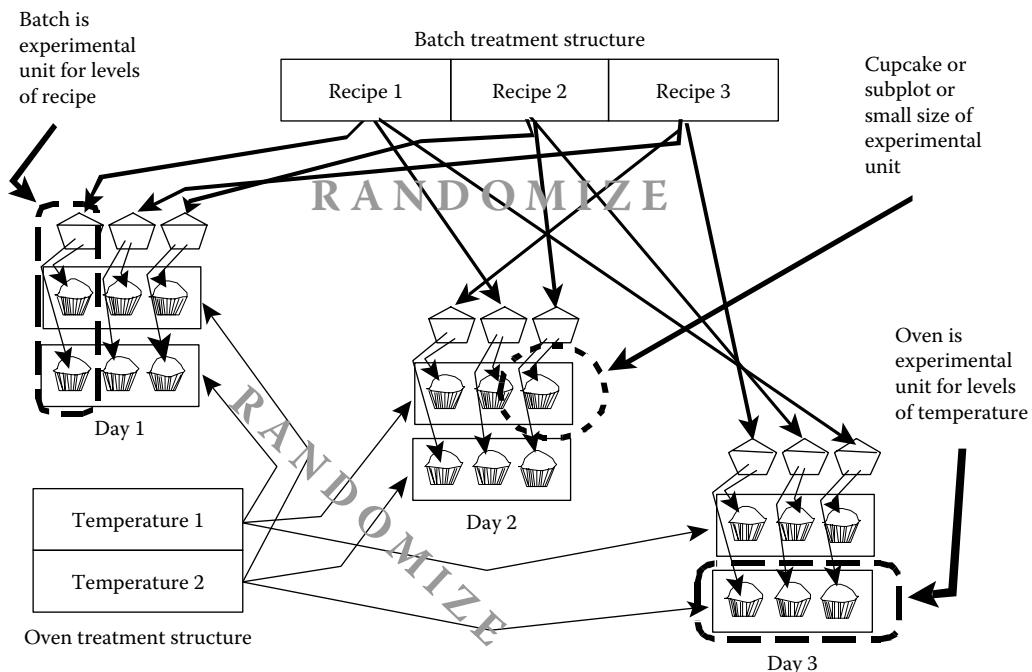


FIGURE 5.9 Diagram of strip-plot design structure for cupcake experiment.

are from the same batch and the batch is the experimental unit for the levels of recipe. The rows consist of three cupcakes and correspond to the ovens, which are the experimental units for the levels of temperature. The analysis of this design can be constructed by considering the design and treatment structures for each size of experimental unit. First, ignore the recipes and only consider the rows of the rectangles, then the oven design is a one-way treatment structure in a randomized complete block design structure. The error term corresponding to ovens is computed from the rectangle (or day) by temperature interaction, which provides two degrees of freedom for error (oven), as displayed in Table 5.8. Next, ignore the temperatures and only consider the columns of the rectangles, then the batch design is a one-way treatment structure in a randomized complete block design structure. The error term corresponding to the batches is computed from the rectangle (or day) by recipe interaction which provides four degrees of freedom for error (batch) as displayed in Table 5.9. Finally, the interactions between the recipes and temperatures are contrasts that are free of the row effects and free of the column effects, leaving the cupcake

TABLE 5.8

Analysis of Variance Table for the Temperature Part of the Treatment Structure in the Strip-Plot Design Structure

Source	df	EMS
Day	2	$\sigma_{\text{oven}}^{2*} + 2\sigma_{\text{day}}^{2*}$
Temperature	1	$\sigma_{\text{oven}}^{2*} + \phi^2(\tau)$
Error (oven)	2	$\sigma_{\text{oven}}^{2*}$

TABLE 5.9

Analysis of Variance Table for the Recipe Part of the Treatment Structure in a Strip-Plot Design Structure

Source	df	EMS
Day	2	$\sigma_{\text{batch}}^{2*} + 3\sigma_{\text{day}}^{2*}$
Recipe	2	$\sigma_{\text{batch}}^{2*} + \phi^2(\beta)$
Error (batch)	4	$\sigma_{\text{batch}}^{2*}$

to be the experimental unit for interaction comparisons. The interaction between recipes, temperatures, and rectangles provides the cupcake error term, which in this case has four degrees of freedom. This design involves three sizes of experimental units and a model that can be used to describe data from this structure is

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + d_k + o_{ik} + b_{jk} + \varepsilon_{ijk}, \quad i = 1, 2, \ j = 1, 2, 3, \text{ and } k = 1, 2, 3$$

where d_k denotes the rectangle or day effect, o_{ik} denotes the oven effect within a rectangle, b_{jk} denotes the batch effect within a rectangle, and ε_{ijk} denotes the cupcake effect within a batch, oven and rectangle. The analysis of variance table corresponding to the strip-plot model is displayed in Table 5.10 where there are three error terms. There are still five degrees of freedom associated with the treatment structure, but the 12 degrees of freedom associates with the design structure are distributed as two for rectangles, two for error (oven), four for error (batch) and four for error (cupcake). The strip-plot design structure is a multilevel design, but it is not a hierarchical design structure since the rows are not nested within the columns and the columns are not nested within the rows. The rows and columns are nested within a rectangle, but that is where the nesting stops.

As will be discussed in Chapter 20, the expected mean squares corresponding to an effect will indicate which error term is used in evaluating hypotheses associated that effect. Some expected mean squares involve variances with “*” superscripts to indicate that there are other factors influencing the variability than are indicated by the subscript.

The process of writing the error terms in the analysis of variance table by including the size of experimental unit associated with the error in parentheses is a convention used throughout the rest of this book whenever there is more than one size of experimental unit

TABLE 5.10

Analysis of Variance Table for the Two-Way Treatment Structure in a Strip-Plot Design Structure

Source	df	EMS
Day	2	$\sigma_{\varepsilon}^2 + 3\sigma_{\text{oven}}^2 + 2\sigma_{\text{batch}}^2 + 6\sigma_{\text{day}}^2$
Temperature	1	$\sigma_{\varepsilon}^2 + 3\sigma_{\text{oven}}^2 + \phi^2(\tau)$
Error (oven)	2	$\sigma_{\varepsilon}^2 + 3\sigma_{\text{oven}}^2$
Recipe	2	$\sigma_{\varepsilon}^2 + 2\sigma_{\text{batch}}^2 + \phi^2(\beta)$
Error (batch)	4	$\sigma_{\varepsilon}^2 + 2\sigma_{\text{batch}}^2$
Temperature \times recipe	2	$\sigma_{\varepsilon}^2 + \phi^2(\tau\beta)$
Error (cupcake)	4	σ_{ε}^2

involved in a study. This convention enables researchers to easily identify the sources of variation in their studies.

More complex design structures are often more conserving in terms of the material and time required to carry out the studies. For the cupcake baking experiment, the completely randomized and randomized complete block design structures require 18 batches of cupcake dough to be made and an oven to be used 18 times for 18 bakes. The split-plot design structures involve 18 batches of cake dough (one for each cupcake), but only six baking times. Hence the split-plot design uses only one-third of the time for baking the cupcakes as that required for the completely randomized design. The strip-plot requires nine batches of cake dough and six baking times. The strip-plot design structure uses only one-half of the batches of cake dough and one-third of the time for baking. Using the more complex design structures is often time- and resource-conserving as well as being a more convenient way to perform experiments, provided one can resolve the necessary error terms as discussed above.

The degrees of freedom associated with the error corresponding to the smallest experimental unit size are reduced as more structure is imposed on the design. However, in the previous example, the error associated with a completely randomized design includes variation due to ovens, batches and cupcakes within a batch. The split-plot design has two error terms where part of the error is designated as due to variability among ovens and the other part is the variability among cupcakes that also includes batch to batch variability. The strip-plot design has three error terms where the variance in the study is split into variability among ovens, variability among batches, and variability among cupcakes within a batch. Using the more complex design structures provides fewer degrees of freedom for the cupcake error term, but that error term is refined down to the cupcake to cupcake variability in the strip-plot design structure, where it involves oven and batch variability in the completely randomized and randomized complete block and batch variability for the split-plot. Therefore, the fact that there are fewer degrees of freedom does not necessarily mean that there is less power for the comparisons among the treatment factors as the magnitude of the important variance components can also decrease, providing an increase in power. The more complex designs have error terms with fewer sources of variability than do the error terms of the simpler design structures. A summary of the sources of variability associated with each of the error terms for the cupcake designs is given in Table 5.11.

TABLE 5.11

Sources of Variability Attributed to Each of the Error Terms for the Various Designs Associated with the Cupcake Examples

Design Structure	Error Term	Source of Variance
Completely randomized	Error	Day, batch, oven, cupcake
Randomized complete block	Error	Batch, oven, cupcake
Split-plot CR whole-plot design structure	Error (oven) Error (cupcake)	Day, oven Batch, cupcake
Split-plot RCB whole-plot design structure	Day Error (oven) Error (cupcake)	Day Oven Batch, cupcake
Strip-plot design structure	Rectangle or day Error (oven) Error (batch) Error (cupcake)	Day Oven Batch Cupcake

The following sections provide examples of several multilevel designs, each of which can be expressed as one or a combination of several of the four basic design structures described above. Being able to identify the components of a design that correspond to one of the four basic design structures provides a method for determining the source of error for each size of the experimental units in the study. It is important to be able to describe all of the error terms in a model as those descriptions are needed when using software to extract an appropriate analysis. When there are repeated measures or unequal variances at the residual or smallest size of experimental unit level of a model, the expression for the residual sum of squares is needed in order for the software to model the appropriate variances and covariances. The modeling of repeated measures experiments will require knowing how to compute the residual sum of squares. Most authors compute the residual sum of squares by subtraction, but the subtraction method is not sufficient when one needs to model the residual variances and covariances.

Multilevel designs have two important characteristics. First, the treatment structure consists of at least a two-way set of treatment combinations. The second characteristic that distinguishes the multilevel designs from those in Chapter 4 is that more than one size of experimental unit is used in an experiment. There is one size of experimental unit for each level in the design structure. Each size of experimental unit has its own design and treatment structures and the model can be constructed by combining the models from each size of experimental unit. Since there is more than one size of experimental unit, there is more than one error term used in the analysis; that is, there is one error term for each size of experimental unit in the experiment, which is also reflected in the model.

This chapter presents several examples to demonstrate the principles needed to use the four basic design structures discussed in Section 5.2 to properly identify the designed experiment employed in a study. Once the experimenter is able to use these principles to identify the designed experiments discussed in this chapter, she will be able to use them to identify the characteristics of other designs.

Multilevel designs can be structured in many different ways. The next series of examples is used to demonstrate the process of identifying the different sizes of experimental units and then this information is used to construct an appropriate model on which to base the analysis. Each example includes an analysis of variance table that lists the sources of variation, the degrees of freedom, and expected mean squares (see Chapter 18 for a discussion on computing expected mean squares). The design structures are related to the four basic design structures so that the form of the error terms can be determined. It is important to list the appropriate sources of variation and the corresponding degrees of freedom for an analysis before subjecting the data to computer analysis, as it provides an excellent check on whether appropriate model specification code was used to describe the data.

5.2 Hierarchical Design: A Multilevel Design Structure

Hierarchical designs are often used in the social sciences where groups of individuals form the larger size of experimental unit and the individuals within the group are the smaller size of experimental unit. For example, a study to evaluate methods of teaching mathematics to fifth graders involved selecting six classes of fifth graders from a school system and randomly assigning each of two methods to three of the classes. The classes

TABLE 5.12

Analysis of Variance Table for the Teaching Method Study Using a Hierarchical or Split-Plot Design Structure

Source	df	EMS
Method	1	$\sigma_e^2 + k_2 \sigma_{\text{class}}^2 + \phi^2(\tau)$
Classes (method) = error (class)	4	$\sigma_e^2 + k_1 \sigma_{\text{class}}^2$
Sex	1	$\sigma_e^2 + \phi^2(\beta)$
Sex \times method	1	$\sigma_e^2 + \phi^2(\tau\beta)$
Error (student)	n... -8	σ_e^2

are the experimental units for teaching methods. It is of interest to determine if the teaching methods have different effects on male and female students. The student is the experimental unit for sex of student. The individuals are nested within sex of student within a class and the classes are nested within teaching method. This study involves a nested design structure and a two-way treatment structure with two teaching methods by two sexes of students. A model that can be used to describe a student's score on a math test after being taught by one of the teaching methods is

$$y_{ijkm} = \mu + \tau_i + c_{ij} + \beta_k + (\tau\beta)_{jk} + \varepsilon_{ijkm}, \quad i = 1, 2, j = 1, 2, 3, k = 1, 2, m = 1, 2, \dots, n_{ijk}$$

where y_{ijkm} is the score from the m th student of the k th sex in the j th class taught by the i th method, μ denotes the mean score, τ_i denotes the teaching method effect, c_{ij} denotes the effect of the j th class taught by the i th method, β_k denotes the k th sex effect, $(\tau\beta)_{jk}$ denotes the teaching method by sex interaction, and ε_{ijkm} denotes the student effect within a sex of a class room taught by a teaching method. The analysis of variance table for the above model is given in Table 5.12 where the classes and the students are assumed to be random effects (see Chapter 18). The coefficients k_1 and k_2 depend on the numbers of students of each sex within the class rooms (see Chapter 18 for the evaluation of expected mean squares). This model has two error terms, one for classes and one for students, and this is similar to the structure in the split-plot design structure. Hence, the hierarchical design is identical to the split-plot basic design structure.

5.3 Split-Plot Design Structures: Two-Level Design Structures

Split-plot designs are used mainly in agricultural, industrial, and biological research, but they can also be used effectively in most other areas of research. A split-plot design structure is a multilevel design with two or more levels. Split-plot designs involve two- or higher-way treatment structures with an incomplete block design structure and at least two different sizes of experimental units. The feature that distinguishes split-plot designs from repeated measures designs is that the levels of each of the factors in the treatment structure can be randomly applied to the various sizes of experimental units. In contrast, repeated measures designs involve a step where the levels of at least one of the factors in the treatment structure (usually time) cannot be assigned at random to the respective experimental units. The following examples demonstrate the uses of split-plot designs and to provide guides for identifying the appropriate design.

5.3.1 Example 5.1: Cooking Beans—The Simplest Split-Plot or Two-Level Design Structure

An experimenter wants to study how five varieties of beans respond to three cooking methods. The dependent variables of interest are the tenderness and flavor of the beans after being cooked. The experimenter has a field consisting of 15 homogeneous rows. He randomly assigns each one of the five varieties of the one-way treatment structure to three rows, thus generating a one-way treatment structure in a completely randomized design structure. The varieties are assigned to the rows, as shown in Figure 5.10; hence, the rows are the experimental units associated with the varieties. At harvest time, the beans from each row are put into a box. For some measurement made on a row of beans (or a box), the model for the row experimental unit is

$$y_{ij} = \mu_i + r_{ij}, \quad i = 1, 2, 3, 4, 5, \quad j = 1, 2, 3$$

where μ_i represents the mean of the i th variety and r_{ij} denotes the error associated with the i th variety being assigned to the j th row. An analysis or variance table for the row model used to compare the mean response of the varieties is given in Table 5.13. There are three rows assigned to each variety, thus, the variability among rows treated alike within a variety provides two degrees of freedom for error (row). If the variances of the rows within the five varieties are equal, then pool the variances together to provide the 10 degrees of freedom associated with error (row). The row design is a completely randomized design structure.

Next, the experimenter wants to examine the cooking methods. There are several possible ways to carry out this part of the experiment, of which two are discussed. First, the experimenter could assign a different cooking method to each of the three rows planted

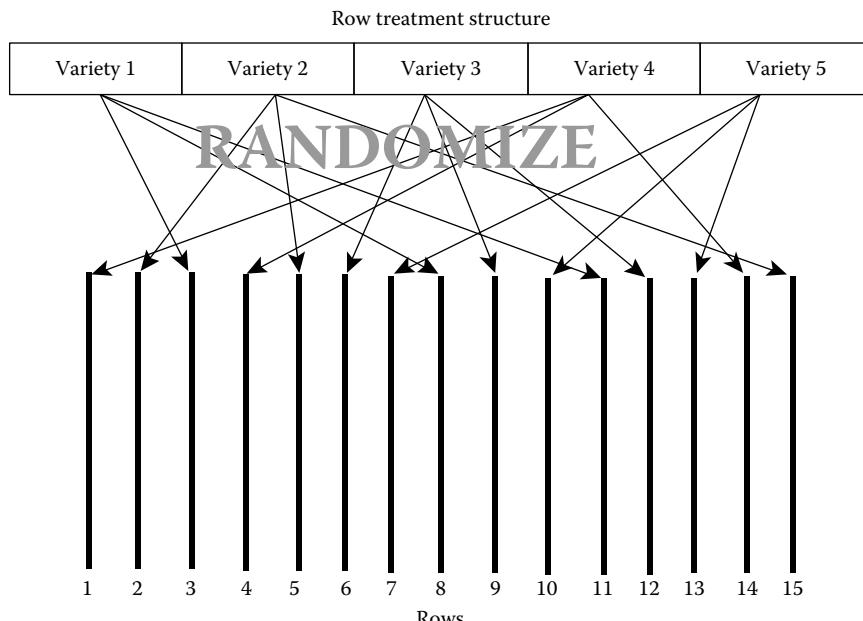


FIGURE 5.10 Randomization scheme for assigning the varieties to the rows for the cooking beans experiment.

TABLE 5.13

Analysis of Variance Table for the Row Analysis to Compare Varieties for the Cooking Beans Example

Source of Variation	df	EMS
Variety	4	$\sigma_{\text{row}}^{2*} + \phi^2(\mu_i)$
Error (row)	10	σ_{row}^{2*}

with a given variety, as shown in Figure 5.11. The arrangement in Figure 5.11 produces a two-way treatment structure in a completely randomized design structure where the rows are the experimental units. However, there is only one replication or row for each variety by cooking method combination. Hence there are no rows treated alike, which means there is no measure of the experimental error or row variance. The resulting analysis of variance table is shown in Table 5.14.

A design with zero degrees of freedom for error is not very desirable (although some analyses can be done using the two-way non-replicated experiment techniques discussed in Milliken and Johnson, 1989; Milliken and Graybill, 1970; and Johnson and Graybill, 1972). Another way of assigning the cooking methods avoids the zero degrees of freedom problem, but does make the experimental design and analysis more complex.

The alternative method is to split each box of beans (one box is obtained from each row) into three batches and then randomly assign each of the cooking methods to one of the

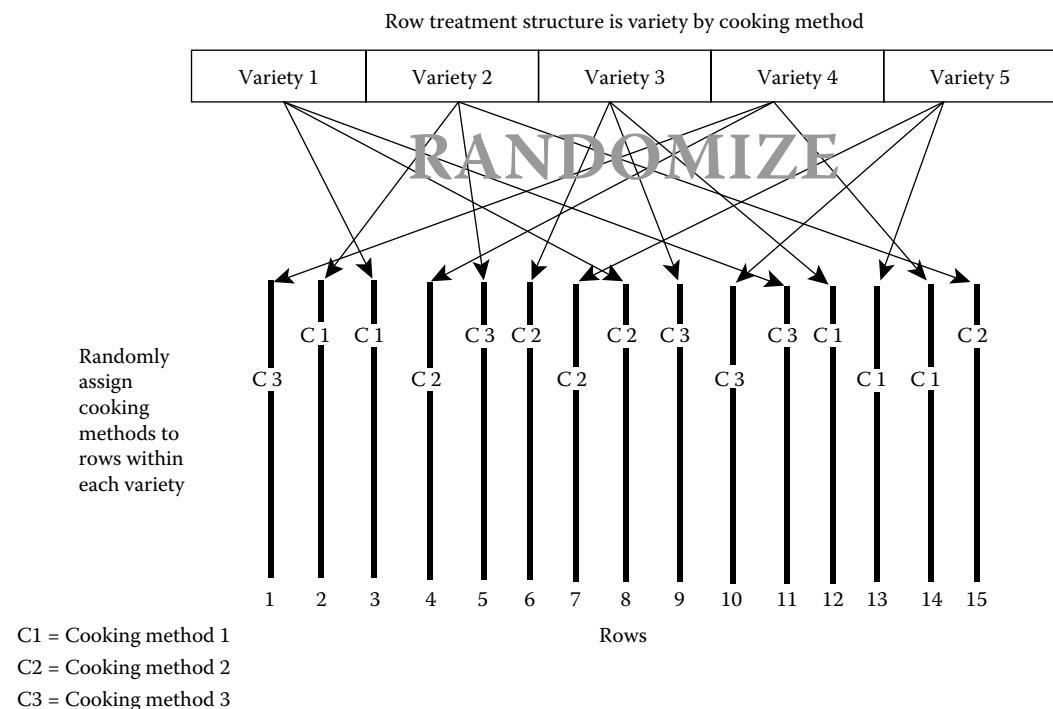


FIGURE 5.11 Randomization scheme of assigning cooking methods to rows within each variety for the cooking beans example.

TABLE 5.14

Analysis of Variance Table for the Row Analysis to Compare Varieties by Cooking Methods for the Cooking Beans Example

Source of Variation	df	EMS
Variety	4	$\sigma_{\text{row}}^{2*} + \phi^2$ (variety)
Cooking	2	$\sigma_{\text{row}}^{2*} + \phi^2$ (cook)
Variety \times cooking method	8	$\sigma_{\text{row}}^{2*} + \phi^2$ (variety \times cook)
Error (row)	0	0

three batches within a row. Since a cooking method is assigned to a batch, the experimental unit for the cooking method is a batch. Thus, there are two sizes of experimental units for this experiment; the row (large size) is the experimental unit for varieties and the batch (smaller size) is the experimental unit for cooking treatments. Such an assignment is displayed in Figure 5.12.

The treatment and design structures for the batch experimental units are a one-way treatment structure in a randomized complete block design structure where the rows (or boxes) are the blocks. The analysis of variance table for the batch part of the design is given in Table 5.15.

The sum of squares for rows consists of the sum of squares for variety plus the sum of squares for error (row) from Table 5.13. If all rows were treated alike, the cooking method by row interaction would provide the batch error term, but some rows are planted to

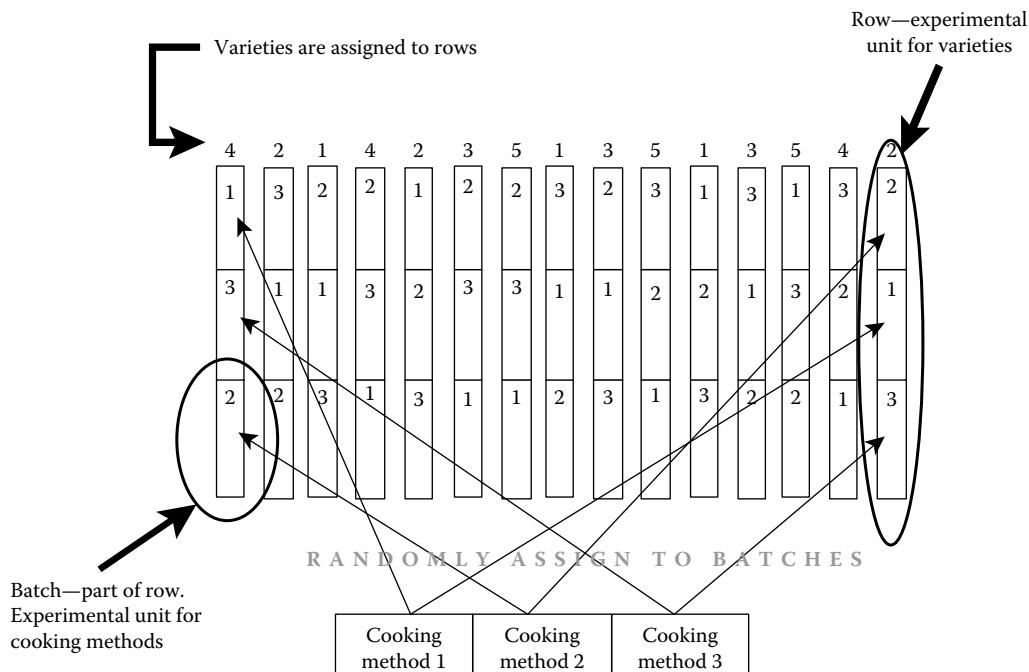


FIGURE 5.12 Diagram showing randomization process of assigning cooking methods to batches (lines are shown for rows 1 and 15 only).

TABLE 5.15

Analysis of Variance Table for the Batch Analysis, Ignoring Varieties, for the Cooking Beans Example

Source of Variation	df	EMS
Rows	14	$\sigma_e^2 + 3\sigma_{\text{row}}^2 + ?$
Cooking methods	2	$\sigma_e^2 + \phi^2 (\text{cooking method})$
Residual = rows \times cooking method	28	$\sigma_e^2 + ?$

variety 1, others to variety 2, and so on. Thus the cooking method by row interaction includes the cooking method by variety interaction. So, consider only those rows planted to variety 1. The row by cooking method interaction within variety 1 provides $(3 - 1) \times (3 - 1) = 4$ degrees of freedom for measuring how batches vary when treated alike within a row. The batch error sum of squares is obtained by pooling the row by cooking method interaction sum of squares within a variety across the five varieties yielding 20 degrees of freedom for error (batch). A model to describe data from the cooking beans example is

$$y_{ijk} = \mu_{ik} + r_{ij} + \varepsilon_{ijk}, \quad i = 1, 2, 3, 4, 5, \quad j = 1, 2, 3, \quad k = 1, 2, 3$$

where μ_{ik} denotes the mean of the i th variety cooked by the k th method, r_{ij} is the random effect of the j th row assigned to the i th variety that is assumed to be distributed as $N(0, \sigma_{\text{row}}^2)$, and ε_{ijk} denotes the random effect of the batch from the j th row of the i th variety cooked with the k th method that is assumed to be distributed as $N(0, \sigma_{\text{batch}}^2)$. It is also assumed that ε_{ijk} and r_{ij} are independent random variables.

The mean μ_{ik} can be expressed in an effects model as

$$\mu_{ik} = \mu + v_i + \omega_k + (v\omega)_{ik}$$

where μ denotes the overall mean, v_i denotes the effect of the i th variety, ω_k denotes the effect of the k th cooking method, and $(v\omega)_{ik}$ denotes the variety by cooking method interaction. The above model can be expressed with terms representing the two sizes of experimental units as

$$\begin{aligned} y_{ijk} &= \mu + v_i + r_{ij} && \} \text{row part of model} \\ &+ \omega_k + (v\omega)_{ik} + \varepsilon_{ijk} && \} \text{batch part of model} \end{aligned}$$

where the row part of the model is also the blocking structure for the batch part of the model. A split-plot design can be analyzed in two steps by carrying out the row analysis and then carrying out the batch part of the analysis. When the data set is balanced, identical results will be obtained where one fits a single model to the data or where one carries out the two step analysis process. The analysis of variance table for the cooking beans example is given in Table 5.16, which partitions the analysis for each of the experimental unit sizes, the row size and the batch size. The row analysis is also the blocking structure for the batch part of the analysis, as indicated by the arrows.

The whole-plot or row design structure for the cooking beans experiment is a completely randomized design structure; hence, this is called the simplest split-plot or two-level

TABLE 5.16

Analysis of Variance Table for the Cooking Bean Experiment Showing the Analysis for Each Size of Experimental Unit

Source	df	EMS
<i>Row Analysis</i>		
Variety	4	$\sigma_{\text{batch}}^2 + 3\sigma_{\text{row}}^2 + \phi^2(v)$
Error (row)	10	$\sigma_{\text{batch}}^2 + 3\sigma_{\text{row}}^2$
<i>Batch Analysis</i>		
Row	14	
Cooking method	2	$\sigma_{\text{batch}}^2 + \phi^2(\omega)$
Variety \times cooking method	8	$\sigma_{\text{batch}}^2 + \phi^2(v\omega)$
Error (batch)	20	σ_{batch}^2

design structure. The usual whole-plot design structure generally involves a randomized complete block design structure as demonstrated by the next example.

5.3.2 Example 5.2: Grinding Wheat—The Usual Split-Plot or Two-Level Design Structure

A grain milling experiment consists of evaluating the properties of various varieties of wheat after the wheat kernels are milled or ground into wheat flour. The experiment consists of setting the gap between the grinding rollers (called roll gap) to a value and then grinding a batch of each of the varieties (in a random order). Next the gap between the grinding rollers is changed to another value and new batches are ground. Suppose the researcher wishes to evaluate three roll gaps and five varieties. Thus, one replication of the two-way treatment structure of three roll gaps by five varieties requires 15 runs of the flour mill. One replication or 15 runs can be accomplished during one work day, so four days are required to obtain four replications. The environmental conditions such as humidity can have an effect on the milling process and these conditions can be different from day to day. Thus, to help control these conditions, day is used as a blocking factor where one replication of the 15 treatment combinations is obtained during one day. The randomization process is to randomly assign the order of the three roll gaps to a roll gap run within each day, then, within each of the roll gaps, to randomly assign the order of the varieties to batches to be milled. Figure 5.13 contains the diagram displaying the randomization process, where a day corresponds to a block, a group of five runs within a day is the whole-plot experimental unit, and a single run is the subplot experimental unit. Not all of the arrows showing the assignment of varieties to runs are shown as the complete set of lines would clutter the display. The whole-plot design structure is a randomized complete block with four blocks of size three. The whole-plot model can be expressed as

$$y_{ij} = \mu + R_i + d_j + e_{ij}, \quad i = 1, 2, 3, \quad j = 1, 2, 3, 4, \quad \text{where } d_j \sim \text{i.i.d. } N(0, \sigma_{\text{day}}^2) \text{ and } e_{ij} \sim \text{i.i.d. } N(0, \sigma_{\text{run}}^2)$$

where R_i denotes the roll gap effect, d_j denotes the random day effect and e_{ij} denotes the random roll gap run effect or whole-plot error. The whole-plot analysis of variance table is given in Table 5.17, where the run or whole plot error is computed from the day by roll gap interaction; that is, the error for a randomized complete block design structure is the treatment structure by design structure interaction. The next step in the analysis is to determine

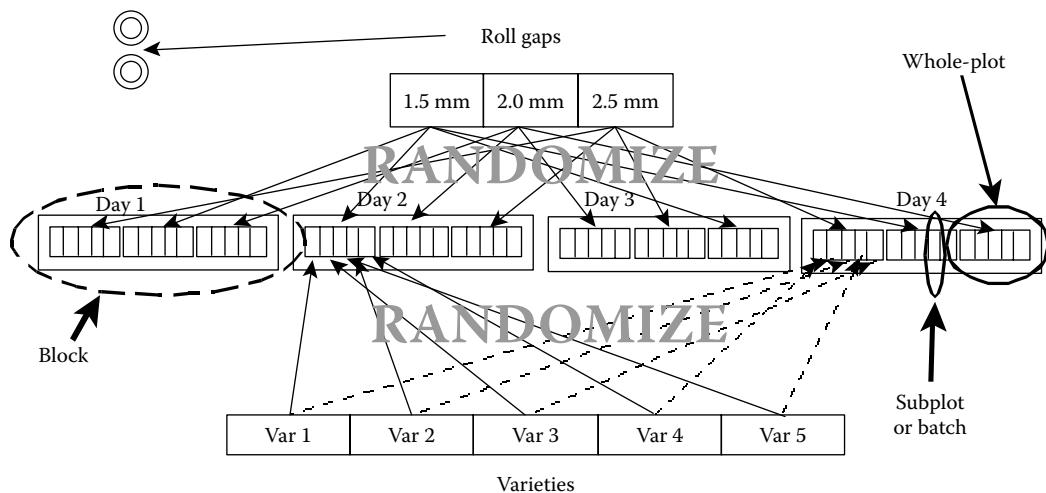


FIGURE 5.13 Display of randomization process for flour milling experiment.

the source of the batch error. This is accomplished by considering only those runs with one roll gap, say 1.5 mm. The analysis of variance table to compare the varieties at a roll gap of 1.5 mm is displayed in Table 5.18. The error term is computed by the variety by day (or run within a day since only one run is used for a given day) interaction, which provides 12 degrees of freedom to measure how batches treated alike vary within a run. This process is carried out for the other two roll gap settings, each providing 12 degrees of freedom for batch error. If these three variances are equal, then they can be pooled into the batch error term with 36 degrees of freedom. The batch error term can be expressed as variety \times day (roll gap) or is the variety by day interaction pooled across the levels of roll gap.

A model that includes the run and batch parts of the model can be expressed as

$$y_{ijk} = \mu + R_i + d_j + e_{ij} + V_k + (RV)_{ik} + \varepsilon_{ijk}, \quad i = 1, 2, 3, \quad j = 1, 2, 3, 4, \quad k = 1, 2, \dots, 5,$$

where

$$d_j \sim i.i.d. N(0, \sigma_{\text{day}}^2), \quad e_{ij} \sim i.i.d. N(0, \sigma_{\text{run}}^2), \quad \text{and} \quad \varepsilon_{ijk} \sim i.i.d. N(0, \sigma_{\text{batch}}^2)$$

In the above model R_i denotes the i th roll gap effect, V_k denotes the k th variety effect, $(RV)_{ik}$ denotes the roll gap by variety interaction and ε_{ijk} denotes the random batch effect. $\mu + R_i + d_j + e_{ij}$ is the whole plot or run part of the model and it is also the blocking structure for the subplot or batch part of the model. The batch part of the model is $V_k + (RV)_{ik} + \varepsilon_{ijk}$.

TABLE 5.17
Whole-Plot Analysis for the Flour Milling Experiment

Source	df	EMS
Day or block	3	$\sigma_{\text{run}}^{2*} + 3\sigma_{\text{day}}^{2*}$
Roll gap	2	$\sigma_{\text{run}}^{2*} + \phi^2(R_i)$
Error (run) = day \times roll gap	6	σ_{run}^{2*}

TABLE 5.18

Analysis of Variance for Comparing Varieties at Roll Gap of 1.5 mm for the Flour Milling Experiment

Source	df	EMS
Day or block	3	$\sigma_{\text{batch}}^2 + 5\sigma_{\text{run}}^{2*}$
Variety	4	$\sigma_{\text{batch}}^2 + \phi^2(V)$
Error (batch) = day \times variety	12	σ_{batch}^2

TABLE 5.19

Analysis of Variance Table for the Flour Grinding Experiment Showing the Analysis for Each Size of Experimental Unit

Source	df	EMS
<i>Run Analysis</i>		
Day or block	3	$\sigma_{\text{batch}}^2 + 5\sigma_{\text{run}}^2 + 15\sigma_{\text{run}}^2$
Roll gap	2	$\sigma_{\text{batch}}^2 + 5\sigma_{\text{run}}^2 + \phi^2(R)$
Error (run) = day \times roll gap	6	$\sigma_{\text{batch}}^2 + 5\sigma_{\text{run}}^2$
<i>Batch Analysis</i>		
Run	11	
Variety	4	$\sigma_{\text{batch}}^2 + \phi^2(V)$
Variety \times roll gap	8	$\sigma_{\text{batch}}^2 + \phi^2(VR)$
Error (batch) = variety \times day (roll gap)	36	σ_{batch}^2

The final analysis of variance table that includes both the run and batch analyses is displayed in Table 5.19. The brackets and arrows indicate that the run or whole plot part of the model is the blocking structure for the batch part of the model. As indicated previously, a split-plot design with a randomized complete block whole-plot design structure is the usual split-plot design.

5.3.3 Example 5.3: Baking Bread—Split-Plot with Incomplete Block Design Structure

A bakery scientist designed a study to evaluate the effect of temperature on the volume of loaves of bread made from two different recipes and baked at three different temperatures. Figure 5.14 displays the randomization process for this experiment. Only two ovens were available on a given day and they could be used just once during that day. Each oven can hold one loaf of bread from each recipe, but only two of the three temperatures could be observed on a given day. The researcher wanted to have four replications of each temperature, meaning that the study needed to be conducted on six different days. Table 5.20 contains the assignment of temperatures to the six days. Using Figure 5.14, the randomization of temperatures to ovens is displayed where day is the block and the oven is the experimental unit for the levels of temperature. The oven design is a one-way treatment structure (levels of temperature) in an incomplete block design structure. A model that could be used to describe data from each oven is

$$y_{ij} = \mu + T_i + d_j + o_{ij} \quad (i, j) \in \{(1, 1), (2, 1), (1, 2), (3, 2), (2, 3), (3, 3), (1, 4), (3, 4), (2, 5), (3, 5), (1, 6), (2, 6)\}$$

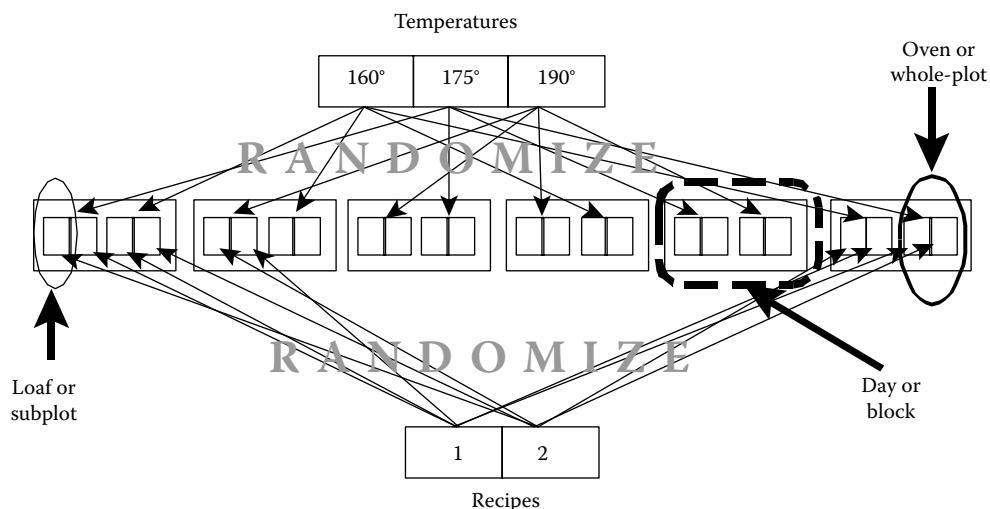


FIGURE 5.14 Assignments of temperatures to ovens within a day and recipes to loaves within an oven using incomplete block whole-plot design structure.

where

$$d_j \sim i.i.d. N(0, \sigma_{\text{day}}^2) \quad \text{and} \quad o_{ij} \sim i.i.d. N(0, \sigma_{\text{oven}}^2)$$

The subscripts (i, j) belong to the index set detailed above and the analysis of variance table is given in Table 5.21. The expected mean squares are computed for the type III sums of squares (see Chapter 10). The oven error term is computed from the design structure by treatment structure interaction or the day by temperature interaction. If all cells in Table 5.20 were observed, there would be $(3 - 1)(6 - 1) = 10$ degrees of freedom for the oven error, but six of the cells are empty leaving $10 - 6 = 4$ degrees of freedom for oven error. The loaf design is a one-way treatment structure (levels of recipes) in a randomized complete block design structure where the ovens are the blocks. Not all blocks are treated alike, so consider those ovens with a common temperature. The part of the treatment structure for the temperature 160°C is displayed in Table 5.22. The design in Table 5.22 is

TABLE 5.20

Assignment of Temperatures to Days for Incomplete Block Whole-Plot Design Structure

Days	Temperatures		
1	160°C	175°C	X
2	160°C	X	190°C
3	X	175°C	190°C
4	160°C	X	190°C
5	X	175°C	190°C
6	160°C	175°C	X

Note: X denotes temperature was not observed during that day.

TABLE 5.21

Analysis of Variance Table for the Oven Analysis, Ignoring Recipes, for the Incomplete Block Design Structure with EMS Computed Using Type III Sums of Squares

Source of Variation	df	EMS
Days	5	$\sigma_{\text{oven}}^{2*} + 1.8\sigma_{\text{Day}}^{2*}$
Temperatures	2	$\sigma_{\text{oven}}^{2*} + \phi^2(T)$
Error (ovens) = days \times temperatue	4	$\sigma_{\text{oven}}^{2*}$

TABLE 5.22

Design for Recipes Baked with Temperature 160°C

Days	Recipes	
1	1	2
2	1	2
4	1	2
5	1	2

TABLE 5.23

Analysis of Variance Table for the Loaf Analysis, Ignoring Temperatures for the Incomplete Block Design Structure

Source of Variation	df	EMS
Days	3	$\sigma_{\text{loaf}}^2 + 2\sigma_{\text{oven}}^2$
Recipes	1	$\sigma_{\text{loaf}}^2 + \phi^2(R)$
Error (loaves) = days \times recipes	3	σ_{loaf}^2

a one-way treatment structure (two recipes) in a randomized complete block design structure with four blocks or ovens or days. There are three degrees of freedom for the loaf error term from the 160°C temperature data, which corresponds to the recipe by day interaction. The analysis of variance table for the loaf analysis at 160°C is in Table 5.23. The loaf error can be computed for each of the temperatures and, if the variances are equal, the three variances are pooled to provide nine degrees of freedom for error (loaf). A model that can be used to represent all of the data is

$$y_{ijk} = \mu + T_i + d_j + o_{ij} + R_k + (TR)_{ik} + \varepsilon_{ijk} \\ (i, j) \in \{(1, 1), (2, 1), (1, 2), (3, 2), (2, 3), (3, 3), (1, 4), (3, 4), (2, 5), (3, 5), (1, 6), (2, 6)\}, k = 1, 2$$

where

$$d_j \sim N(0, \sigma_{\text{day}}^2), \quad o_{ij} \sim i.i.d. N(0, \sigma_{\text{oven}}^2) \quad \text{and} \quad \varepsilon_{ijk} \sim i.i.d. N(0, \sigma_{\text{loaf}}^2)$$

The final analysis of variance table for this model is in Table 5.24, where the expected mean squares are computed from the type III sums of squares. The type III sums of squares are used because of the incomplete block whole plot design structure. The oven part of the

TABLE 5.24

Complete Analysis of Variance Table for the Bread Baking Study with the Incomplete Block Whole-Plot Design Structure with EMS Computed for Type III Sums of Squares (See Chapter 10)

Source of Variation	df	EMS
Days	5	$\sigma_{\text{loaf}}^2 + 2\sigma_{\text{oven}}^2 + 3.6\sigma_{\text{day}}^2$
Temperatures	2	$\sigma_{\text{loaf}}^2 + 2\sigma_{\text{oven}}^2 + \phi^2(T)$
Error (ovens)	4	$\sigma_{\text{loaf}}^2 + 2\sigma_{\text{oven}}^2$
Recipes	1	$\sigma_{\text{loaf}}^2 + \phi^2(R)$
Temperatures \times recipes	2	$\sigma_{\text{loaf}}^2 + \phi^2(TR)$
Error (loaves) = days \times recipes (temperatures)	9	σ_{loaf}^2

model is $\mu + T_i + d_j + o_{ij}$, which is also the blocking structure for the loaf part of the model. The loaf part of the model is $R_k + (TR)_{ik} + \varepsilon_{ijk}$. This example demonstrates that any type of design structure can be used at each of the levels or sizes of experimental units. One of the exercises involves an incomplete block subplot or smallest size of experimental unit design structure.

5.3.4 Example 5.4: Meat in Display Case—A Complex Split-Plot or Four-Level Design

A meat scientist wants to study the effects of temperature (T) with three levels, types of packaging (P) with two levels, types of lighting (L) with four levels, and intensity of light (I) with four levels on the color of meat stored in a meat cooler for seven days. Six coolers are available for the experiment and the three temperatures ($1, 3$, and 5°C) are each assigned at random to two coolers, as shown in Figure 5.15.

Each cooler is partitioned into 16 compartments on a 4×4 grid (Figure 5.16). The light intensities are regulated by their distance above the cooler surface, thus, all partitions in a column are assigned the same light intensity. The four types of light are randomly assigned to the four partitions within each column. Finally, the two types of packaging are assigned to the steaks and both types of packaging are put into each partition. Figure 5.16 shows how one such cooler is arranged.

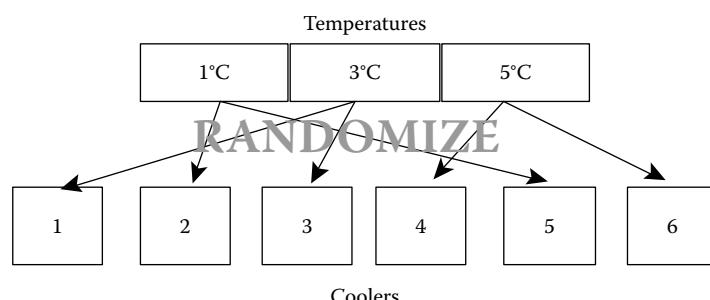


FIGURE 5.15 Assignments of temperatures to coolers for meat in display case study.

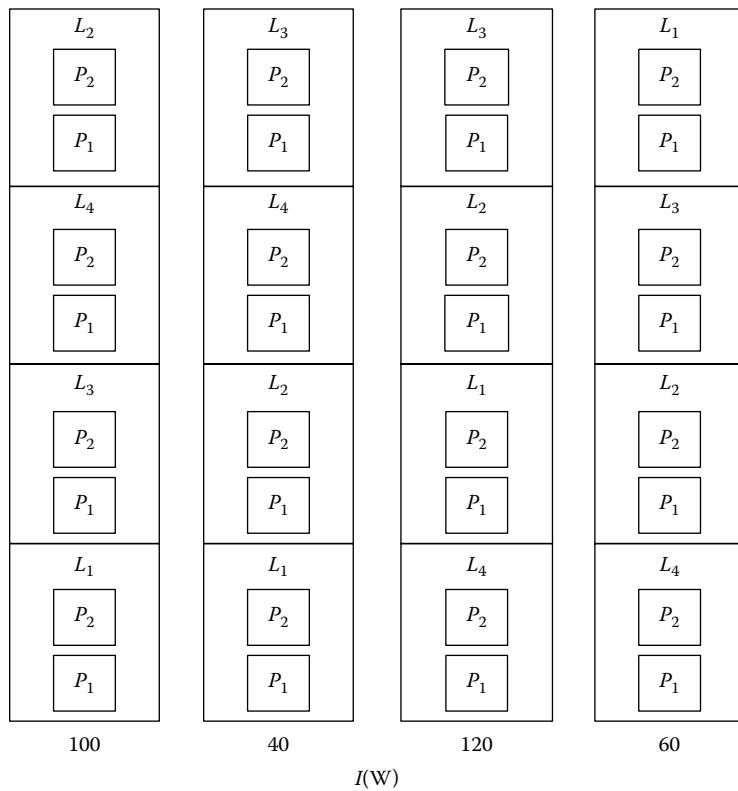


FIGURE 5.16 Assignments of intensities to columns, types of lighting to partitions, and types of packaging to half-partitions in each cooler for the meat in display case study.

One must first correctly identify the different sizes of experimental units or levels of the experiment before an appropriate analysis can be constructed. The experimental units for the levels of temperature are the coolers. The cooler design is a one-way treatment structure (levels of T or temperature) in a completely randomized design structure. If one measurement is made on each cooler, the response could be modeled by

$$y_{ij} = \mu + T_i + c_{ij}, \quad i = 1, 2, 3, \quad j = 1, 2, \text{ and } c_{ij} \sim i.i.d. N(0, \sigma_{\text{cooler}}^2)$$

The analysis of variance table for the cooler model is given in Table 5.25 where the cooler error term is computed from the variation of the two coolers within a temperature,

TABLE 5.25
Analysis of Variance Table for the Cooler Experimental Unit Part of the Model

Source	df
Temperature	2
Error (cooler)	3

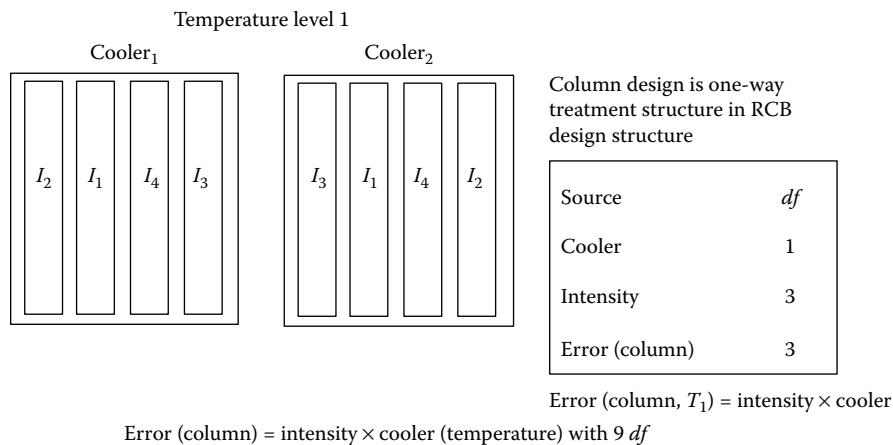


FIGURE 5.17 Comparisons of levels of intensity at temperature 1 to start column analysis for the meat in display case study.

pooled across the three temperatures. Thus there are three degrees of freedom for Error (cooler).

The experimental units for the levels of intensity are the columns of four partitions in a cooler. The column design consists of one-way treatment structure (levels of I or intensity) in a randomized complete block design structure with six blocks or coolers. If all coolers were treated alike, the column error term would be computed from the intensity by cooler interaction. But there are three different temperatures, so restrict the analysis to the two coolers assigned to 1°C. The two coolers assigned to 1°C are displayed in Figure 5.17. At this point, the design is a one-way treatment structure (four levels of I) in a randomized complete block design structure (two coolers). If a measurement is made on each of the columns of these two coolers, a model that can be used to describe the response is

$$y_{1jk} = \mu + I_k + c_{1j} + d_{1jk}, \quad j = 1, 2, k = 1, 2, 3, 4, \quad c_{1j} \sim i.i.d. N(0, \sigma_{\text{cooler}}^2) \text{ and } d_{1jm} \sim i.i.d. N(0, \sigma_{\text{column}}^2)$$

The analysis of variance table for the column model is shown in Figure 5.17, where the column error term is computed as the intensity by cooler interaction, providing three degrees of freedom. This process is repeated for the other two temperatures, each providing column error terms with three degrees of freedom. If these three variances are equal, they can be pooled into the error (column) with nine degrees of freedom. The error (column) can be represented by intensity by cooler (temperature), read as intensity by cooler interaction pooled across the levels of temperature.

The experimental units for the levels of lighting type are the partitions of a column. The partition design is a one-way treatment structure (levels of L) in a randomized complete block design structure with 24 blocks, consisting of the four columns from the six coolers. If all of the columns were treated alike, the partition error term would be computed as the type of lighting by column interaction. But all columns are not treated alike as there are three temperatures by four levels of intensity. Restrict the structure to involve only those columns from temperature 1°C and intensity I_1 , as shown in Figure 5.18. The design associated with Figure 5.18 is a one-way treatment structure (four levels of type of lighting) in a randomized complete block design structure (two columns, but each column is from a

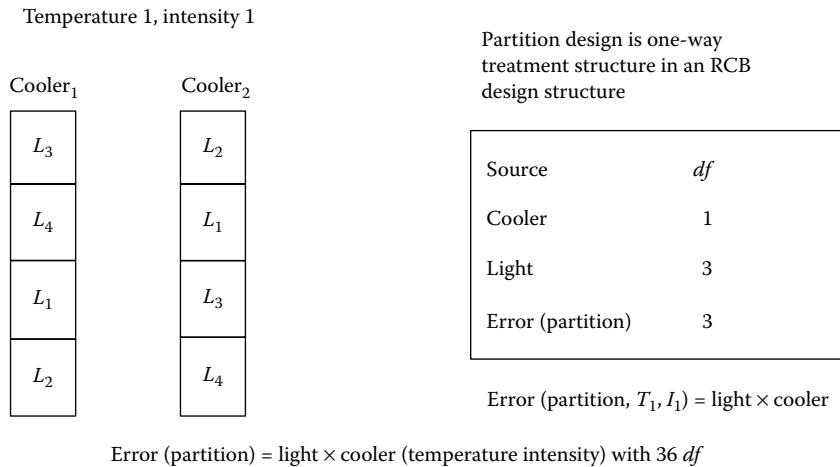


FIGURE 5.18 Comparison of the levels of lighting at temperature 1 and intensity 1 to start the partition analysis for the meat in display case study.

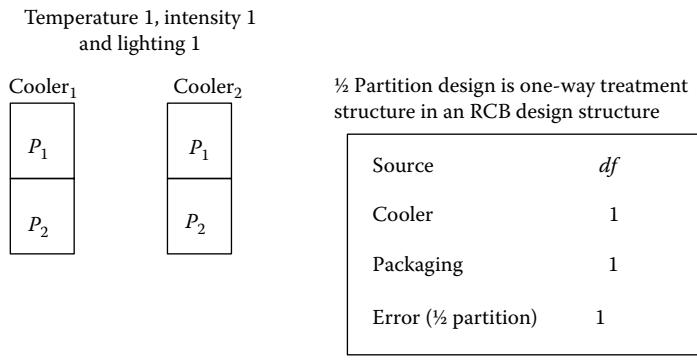
different cooler). If one measurement is made on each partition in Figure 5.18, a model that could be used to describe the responses is

$$y_{1j1m} = \mu + L_m + d_{1j1}^* + p_{1j1m}, \quad j = 1, 2, \quad m = 1, 2, 3, 4, \quad d_{1j1}^* \sim i.i.d. N(0, \sigma_{\text{column}}^{2*}) \text{ and} \\ p_{1j1m} \sim i.i.d. N(0, \sigma_{\text{partition}}^2)$$

where d_{1j1}^* denotes the combination of cooler and column of that cooler effect. The analysis of variance table for this partition model is shown in Figure 5.18, where the partition error term is computed as the lighting by cooler interaction providing three degrees of freedom. This process needs to be carried out for the 12 combinations of temperature by intensity, each providing three degrees of freedom for error (partition). If these 12 variances are equal, they can be pooled into one term providing the error (partition) with 36 degrees of freedom. The error (partition) term can be represented as light \times intensity \times cooler (temperature).

Finally, the experimental units for the levels of packaging are the half-partitions (or steaks). The half-partition design is a one-way treatment structure (levels of packaging or P) in a randomized complete block design structure with 96 blocks, consisting of the four partitions within each of the four columns from the six coolers. If all partitions were treated alike, the half-partition error term would be computed by the packaging by partition interaction. But all partitions are not treated alike as they are assigned to three levels of temperature, four levels of intensity and four levels of lighting. Select those partitions assigned to temperature 1°C, intensity I_1 , and lighting L_1 , as shown in Figure 5.19. If one measurement is made on each steak or half-partition in Figure 5.19, a model that can be used to describe the response is

$$y_{1j1n} = \mu + P_n + p_{1j1n}^* + \varepsilon_{1j1n}, \quad j = 1, 2, \quad n = 1, 2, \quad p_{1j1n}^* \sim i.i.d. N(0, \sigma_{\text{partition}}^{2*}) \text{ and} \\ \varepsilon_{1j1n} \sim i.i.d. N(0, \sigma_{\frac{1}{2}\text{partition}}^2)$$



$$\text{Error (½ partition, } T_1, I_1, L_1) = \text{packing} \times \text{cooler}$$

$$\text{Error (½ partition)} = \text{packing} \times \text{cooler (temperature, intensity, lighting)} \text{ with } 48 \text{ } df$$

FIGURE 5.19 Comparisons of the levels of packaging at temperature 1, intensity 1, and lighting 1 to start the half-partition analysis for the meat in display case study.

The analysis of variance table for this half-partition model is shown in Figure 5.19, where the half-partition error term is computed as the packaging by cooler interaction, providing one degree of freedom. This process needs to be carried out for the 48 combinations of temperature, intensity and lighting. These 48 sums of squares are pooled together to provide 48 degrees of freedom for error (half-partition). The error (half-partition) term can be represented as packaging \times light \times intensity \times cooler (temperature).

The above discussion provides models for each of the four levels in the study, the cooler model, the column model, the partition model, and the half-partition model. These models can be combined into a single model where the interactions between the factors are added together. Since the basic design structure is a split-plot at each level, the treatment structures at each level and interaction with the treatment structure from the levels above it are included in that level part of the model. A model that can be used to describe the data is

$$\begin{aligned}
 y_{ijkmn} = & \mu + T_i + c_{ij} && \{\text{cooler part of the model} \\
 & + I_k + (TI)_{ik} + d_{ijk} && \{\text{column part of the model} \\
 & + L_m + (TL)_{im} + (IL)_{km} + (TIL)_{ikm} + p_{ijkm} && \{\text{partition part of the model} \\
 & + P_n + (TP)_{in} + (IP)_{kn} + (TIP)_{ikn} + (LP)_{mn} \\
 & + (TLP)_{imn} + (ILP)_{kmn} + (TILP)_{ikmn} + \varepsilon_{ijkmn} \} && \{\text{half-partition part of the model}
 \end{aligned}$$

$$i=1, 2, 3, \quad j=1, 2, \quad k=1, 2, 3, 4, \quad m=1, 2, 3, 4, \quad n=1, 2,$$

$$c_{ij} \sim i.i.d. N(0, \sigma_{\text{cooler}}^2), \quad d_{ijk} \sim i.i.d. N(0, \sigma_{\text{column}}^2), \quad p_{ijkm} \sim i.i.d. N(0, \sigma_{\text{partition}}^2)$$

$$\text{and } \varepsilon_{ijkmn} \sim i.i.d. N(0, \sigma_{\frac{1}{2}\text{partition}}^2)$$

The analysis of variance table for the above model is given in Table 5.26 where the larger experimental unit analysis is also the blocking structure for the next smaller size of

TABLE 5.26

Complete Analysis of Variance Table for the Complex Split-Plot Design for Meat in Display Case

Source	df	EMS
<i>Cooler Analysis</i>		
Temperature (T)	2	$\sigma_{\frac{1}{2}\text{partition}}^2 + 2\sigma_{\text{partition}}^2 + 8\sigma_{\text{column}}^2 + 32\sigma_{\text{cooler}}^2 + \phi^2(T)$
Error (cooler)	3	$\sigma_{\frac{1}{2}\text{partition}}^2 + 2\sigma_{\text{partition}}^2 + 8\sigma_{\text{column}}^2 + 32\sigma_{\text{cooler}}^2$
<i>Column Analysis</i>		
Block	5	
Intensity (I)	3	$\sigma_{\frac{1}{2}\text{partition}}^2 + 2\sigma_{\text{partition}}^2 + 8\sigma_{\text{column}}^2 + \phi^2(I)$
$T \times I$	6	$\sigma_{\frac{1}{2}\text{partition}}^2 + 2\sigma_{\text{partition}}^2 + 8\sigma_{\text{column}}^2 + \phi^2(TI)$
Error (column)	9	$\sigma_{\frac{1}{2}\text{partition}}^2 + 2\sigma_{\text{partition}}^2 + 8\sigma_{\text{column}}^2$
<i>Partition Analysis</i>		
Block	23	
Lighting (L)	3	$\sigma_{\frac{1}{2}\text{partition}}^2 + 2\sigma_{\text{partition}}^2 + \phi^2(L)$
$T \times L$	6	$\sigma_{\frac{1}{2}\text{partition}}^2 + 2\sigma_{\text{partition}}^2 + \phi^2(TL)$
$I \times L$	9	$\sigma_{\frac{1}{2}\text{partition}}^2 + 2\sigma_{\text{partition}}^2 + \phi^2(IL)$
$T \times I \times L$	18	$\sigma_{\frac{1}{2}\text{partition}}^2 + 2\sigma_{\text{partition}}^2 + \phi^2(TIL)$
Error (partition)	36	$\sigma_{\frac{1}{2}\text{partition}}^2 + 2\sigma_{\text{partition}}^2$
<i>Half-partition Analysis</i>		
Blocks	95	
Packaging (P)	1	$\sigma_{\frac{1}{2}\text{partition}}^2 + \phi^2(P)$
$T \times P$	2	$\sigma_{\frac{1}{2}\text{partition}}^2 + \phi^2(TP)$
$I \times P$	3	$\sigma_{\frac{1}{2}\text{partition}}^2 + \phi^2(IP)$
$T \times I \times P$	6	$\sigma_{\frac{1}{2}\text{partition}}^2 + \phi^2(TIP)$
$L \times P$	3	$\sigma_{\frac{1}{2}\text{partition}}^2 + \phi^2(LP)$
$T \times L \times P$	6	$\sigma_{\frac{1}{2}\text{partition}}^2 + \phi^2(TLP)$
$I \times L \times P$	9	$\sigma_{\frac{1}{2}\text{partition}}^2 + \phi^2(ILP)$
$T \times I \times L \times P$	18	$\sigma_{\frac{1}{2}\text{partition}}^2 + \phi^2(TILP)$
Error ($\frac{1}{2}$ -partition)	48	$\sigma_{\frac{1}{2}\text{partition}}^2$

Note: The brackets and arrows indicate which effects form blocks for the next smaller size of experimental unit.

experimental unit. This is a model for a split-split-split-plot experiment consisting of four levels or sizes of experimental units and involves four error terms, one for each level or size of experimental unit. This is also a hierarchical design structure where the half-partitions are nested within the partitions, which are nested within the columns, which are nested within the coolers.

5.4 Strip-Plot Design Structures—A Nonhierarchical Multilevel Design

The process of constructing a strip-plot design structure is to arrange the experimental units into rectangles. The levels of one set of factors are randomly assigned to the rows of each rectangle and the levels of the other set of factors are randomly assigned to the columns of each rectangle. Thus, the rows are experimental units associated with the first set of factors and the columns are the experimental units associated with the second set of factors. But as a consequence, the cell or intersection of a row and a column is the experimental unit associated with the interaction comparisons between the two sets of factors. An example is used to demonstrate some of the uses of the strip-plot design structure.

5.4.1 Example 5.5: Making Cheese

A dairy scientist designed a cheese manufacturing study that involved two levels of fat, three types of cheese, two storage temperatures, and two levels of storage humidity. The experiment is a two-step process where the first step involves making batches of cheese with each type of cheese using the two levels of fat. Four one-pound packages of cheese are made from each of the batches. The second step is to store the cheese in a set of environmental conditions to allow it to cure for four weeks. The storage part of the study involves putting one package of cheese from each batch into a chamber assigned one of the storage temperatures and one of the levels of humidity. This process is equivalent to arranging 24 packages of cheese into a rectangle with four rows and six columns. The column corresponds to a batch of cheese and the row corresponds to an environmental chamber, as shown in Figure 5.20. The dairy scientist has four chambers available for use

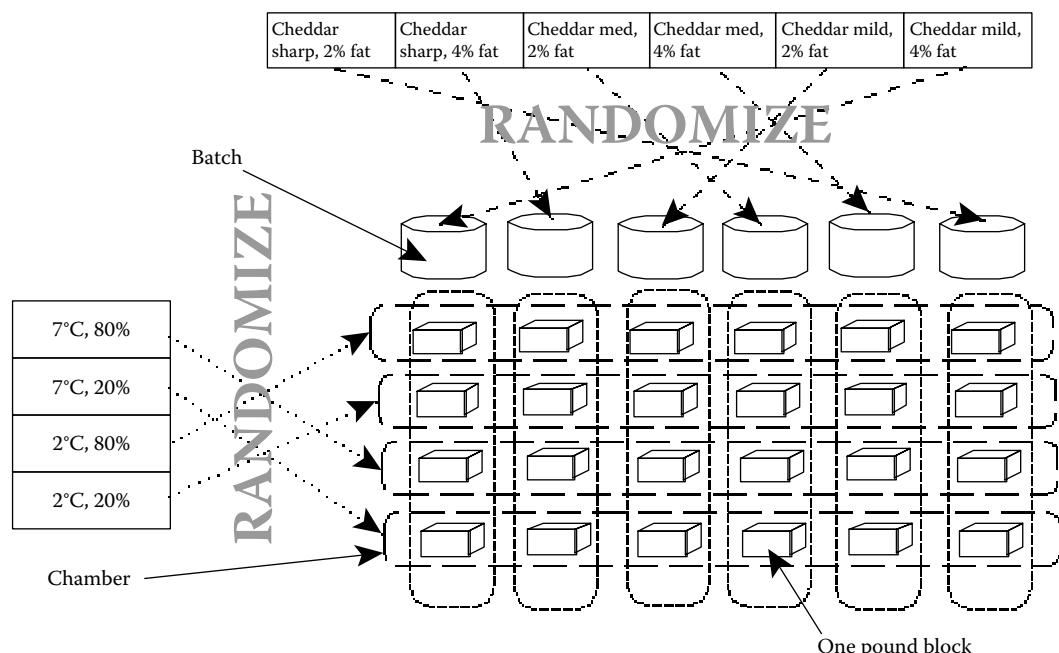


FIGURE 5.20 Randomization scheme for the cheese-making experiment using a strip-plot design structure.

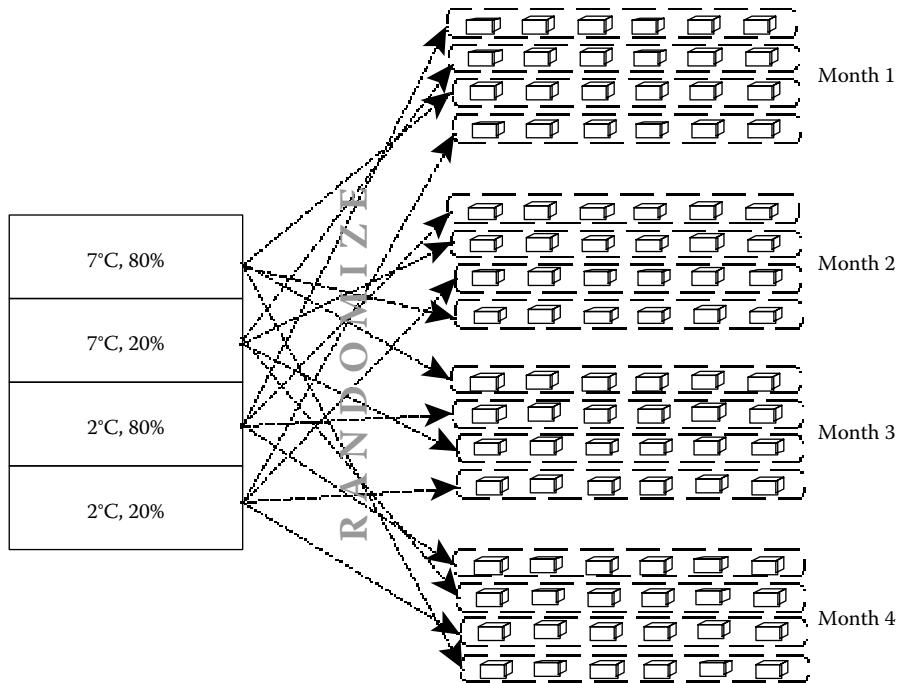


FIGURE 5.21 Randomization scheme for the chamber design for the cheese-making experiment.

in the study, but he wishes to have four replications. Thus, the four replications are obtained by carrying out the experiment in four different months. A model and analysis can be constructed by evaluating the treatment and design structures for each size of experimental unit. The chamber design is displayed in Figure 5.21, where only the temperature and humidity levels are shown for each of the four months. The chamber design is a two-way treatment structure (levels of temperature by levels of humidity) in a randomized complete block design structure where each month is a block. A model that can be used to describe a measurement made on each chamber is

$$y_{ijk} = \mu + T_j + H_k + (TH)_{jk} + m_i + c_{ijk}, \quad i = 1, 2, 3, 4, \quad j = 1, 2, \quad k = 1, 2$$

$$m_i \sim i.i.d. N(0, \sigma_{\text{month}}^2), \text{ and } c_{ijk} \sim i.i.d. N(0, \sigma_{\text{chamber}}^2)$$

Table 5.27 contains the analysis of variance table for the chamber model where the chamber error term is the treatment structure by design structure interaction. There are four treatment combinations in the treatment structure and four blocks in the design structure; thus the chamber error term is based on nine degrees of freedom. The error (chamber) term is the sum of the temperature by month interaction, humidity by month interaction, and the temperature by humidity by month interaction terms. If one uses the three-way interaction term in the SAS®-Mixed procedure, the three interaction terms will be pooled together into one error term. The variance component for chamber is denoted with an asterisk since it also includes the variation due to blocks of cheese. The variance component for month is denoted with an asterisk since it also includes the variation due to batches and blocks of cheese.

TABLE 5.27

Analysis of Variance for the Chamber Experimental Unit Design for the Making Cheese Experiment

Source	df	EMS
Month	3	$\sigma_{\text{chamber}}^2 + 4\sigma_{\text{month}}^2$
Temperature	1	$\sigma_{\text{chamber}}^2 + \phi^2(T)$
Humidity	1	$\sigma_{\text{chamber}}^2 + \phi^2(H)$
Temperature \times humidity	1	$\sigma_{\text{chamber}}^2 + \phi^2(TH)$
Error (chamber)	9	$\sigma_{\text{chamber}}^2$

The batch design is displayed in Figure 5.22, where only the cheese type and percentage fat are considered with the four months. The batch design is a two-way treatment structure (levels of types of cheese by levels of fat) in a randomized complete block design structure where each month is a block. A model that can be used to describe a measurement made on each batch is

$$y_{imn} = \mu + F_m + C_n + (FC)_{mn} + m_i + b_{imn}, \quad i = 1, 2, 3, 4, \quad m = 1, 2, \quad n = 1, 2, 3,$$

$$m_i \sim \text{i.i.d. } N(0, \sigma_{\text{month}}^2) \text{ and } b_{imn} \sim \text{i.i.d. } N(0, \sigma_{\text{batch}}^2)$$

Table 5.28 contains the analysis of variance table for the chamber model where the chamber error term is the treatment structure by design structure interaction. There are six treatment combinations in the treatment structure and four blocks in the design structure, thus the batch error term is based on 15 degrees of freedom. The error (batch) term is the sum of the fat by month interaction, cheese type by month interaction, and the fat by cheese type by month interaction terms. If one uses the three-way interaction term in the SAS-Mixed procedure, the three interaction terms will be pooled together into one error term. The variance component for batch is denoted with an asterisk since it also includes

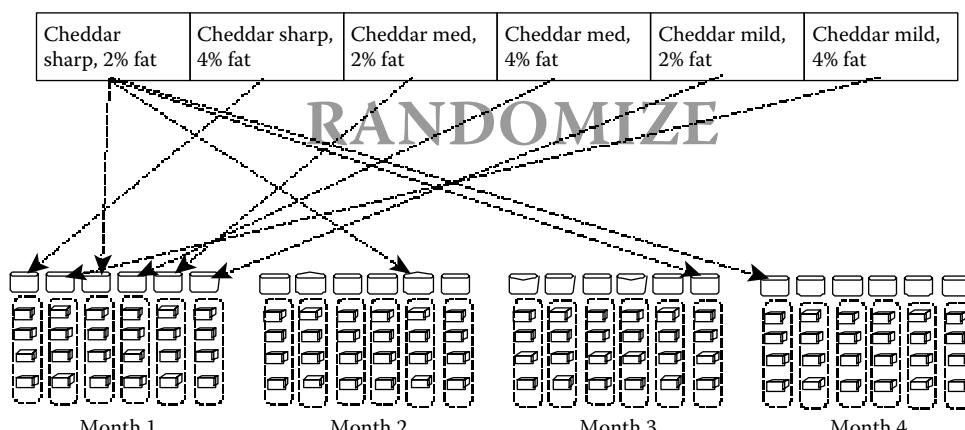


FIGURE 5.22 Randomization scheme for the batch design for the cheese-making experiment. Some of the lines are not included to simplify the drawing.

TABLE 5.28

Analysis of Variance for the Batch Experimental Unit Design for the Making Cheese Experiment

Source	df	EMS
Month	3	$\sigma_{\text{batch}}^{2*} + 6\sigma_{\text{month}}^2$
Fat	1	$\sigma_{\text{batch}}^{2*} + \phi^2(F)$
Cheese type	2	$\sigma_{\text{batch}}^{2*} + \phi^2(C)$
Fat \times cheese type	2	$\sigma_{\text{batch}}^{2*} + \phi^2(FC)$
Error (batch)	15	$\sigma_{\text{batch}}^{2*}$

the variation due to blocks of cheese. The variance component for month is denoted with an asterisk since it also includes the variation due to chambers and blocks of cheese.

The 1 pound cheese block is the experimental unit for interactions between the factors assigned to the chambers and factors assigned to the batches. The cheese block part of the model is

$$(TF)_{jm} + (TC)_{jn} + (TFC)_{jmn} + (HF)_{km} + (HC)_{kn} + (HFC)_{kmn} + (THFC)_{jkmn} + \varepsilon_{ijkmn}$$

$$\varepsilon_{ijkmn} \sim i.i.d. N(0, \sigma_{\text{block}}^2)$$

The model for this strip-plot design is obtained by combining the models for the three experimental units into one model as

$$y_{ijk} = \mu + m_i \quad \left. \begin{array}{l} \text{\{ blocking part of the model} \\ + T_j + H_k + (TH)_{jk} + c_{ijk} \text{\{ chamber part of the model} \\ + F_m + C_n + (FC)_{mn} + b_{imn} \text{\{ batch part of the model} \\ + (TF)_{jm} + (TC)_{jn} + (TFC)_{jmn} + (HF)_{km} \\ + (HC)_{kn} + (HFC)_{kmn} + (THFC)_{jkmn} + \varepsilon_{ijkmn} \end{array} \right\} \text{cheese block part of the model}$$

$$i = 1, 2, 3, 4, \quad j = 1, 2, \quad k = 1, 2, \quad m = 1, 2, \quad n = 1, 2, 3$$

$$m_i \sim i.i.d. N(0, \sigma_{\text{month}}^2), \quad c_{ijk} \sim i.i.d. N(0, \sigma_{\text{chamber}}^2), \quad b_{imn} \sim i.i.d. N(0, \sigma_{\text{batch}}^2),$$

$$\text{and } \varepsilon_{ijkmn} \sim i.i.d. N(0, \sigma_{\text{block}}^2)$$

The analysis of variance table for the cheese experiment is given in Table 5.29, where the rows are segregated by size of experimental unit. The cheese block error term is computed as the chamber treatment structure by batch treatment structure by design structure interaction providing $(4 - 1)(6 - 1)(4 - 1) = 45$ degrees of freedom. This design involves three sizes of experimental units, the chamber size, the batch size, and the block of cheese size. The blocks of cheese are nested within the batch and the blocks of cheese are nested within the chamber, but the batches are not nested within the chambers, nor are chambers nested within batches. Hence this multilevel design structure is not a hierarchical design structure.

This example demonstrates that the design structure associated with a given size of experimental unit can involve any needed treatment structure. The treatment structures

TABLE 5.29

Complete Analysis of Variance Table for the Cheese Making Experiment

Unit	Source	df	EMS
Blocking Structure			
	Months	3	$\sigma_{\text{block}}^2 + 4\sigma_{\text{batch}}^2 + 6\sigma_{\text{chamber}}^2 + 24\sigma_{\text{month}}^{2*}$
Chamber			
	Temperature	1	$\sigma_{\text{block}}^2 + 6\sigma_{\text{chamber}}^2 + \phi^2(T)$
	Humidity	1	$\sigma_{\text{block}}^2 + 6\sigma_{\text{chamber}}^2 + \phi^2(H)$
	Temperature \times humidity	1	$\sigma_{\text{block}}^2 + 6\sigma_{\text{chamber}}^2 + \phi^2(TH)$
	Error (chamber)	9	$\sigma_{\text{block}}^2 + 6\sigma_{\text{chamber}}^2$
Batch			
	Fat	1	$\sigma_{\text{block}}^2 + 4\sigma_{\text{batch}}^2 + \phi^2(F)$
	Cheese type	2	$\sigma_{\text{block}}^2 + 4\sigma_{\text{batch}}^2 + \phi^2(C)$
	Fat \times cheese type	2	$\sigma_{\text{block}}^2 + 4\sigma_{\text{batch}}^2 + \phi^2(FC)$
	Error (batch)	15	$\sigma_{\text{block}}^2 + 4\sigma_{\text{batch}}^2$
Block of cheese			
	Temperature \times fat	1	$\sigma_{\text{block}}^2 + \phi^2(TF)$
	Temperature \times cheese	2	$\sigma_{\text{block}}^2 + \phi^2(TC)$
	Temperature \times fat \times cheese	2	$\sigma_{\text{block}}^2 + \phi^2(TFC)$
	Humidity \times fat	1	$\sigma_{\text{block}}^2 + \phi^2(HF)$
	Humidity \times cheese	2	$\sigma_{\text{block}}^2 + \phi^2(HC)$
	Humidity \times fat \times cheese	2	$\sigma_{\text{block}}^2 + \phi^2(HFC)$
	Temperature \times humidity \times fat	1	$\sigma_{\text{block}}^2 + \phi^2(THF)$
	Temperature \times humidity \times cheese	2	$\sigma_{\text{block}}^2 + \phi^2(THC)$
	Temperature \times humidity \times fat \times cheese	2	$\sigma_{\text{block}}^2 + \phi^2(THFC)$
	Error (cheese block)	45	σ_{block}^2

for the chamber and batch experimental units are both two-way factorial arrangements. One could have a two-way factorial arrangement with one control as the treatment structure for one of the experimental units. This process of identifying the experimental units and then specifying the design and treatment structures of each provides a general method of identifying an appropriate design and corresponding model for complex experiments.

5.5 Repeated Measures Designs

Repeated measures designs are used effectively in many areas of study. These designs involve a treatment structure with at least two factors, an incomplete block design structure and at least two sizes of experimental units. The repeated measures design has the same type of design structure as a split-plot design, but a repeated measures design differs from the split-plot type of design in that the levels of one or more factors cannot be randomly assigned to the corresponding experimental units. Most often, time is the factor where its levels cannot be randomly assigned to the time intervals of a subject. Repeated measures designs involving time are designs used for longitudinal studies. Thus, repeated measures

designs involve a step or steps where it is not possible to randomly assign the levels of some of the factors to their experimental units, whereas in split-plot type designs it is possible to use randomization at each step. The following examples demonstrate some of the structures of repeated measures designs and provide a guide for proper identification of designed experiments.

5.5.1 Example 5.6: Horse Feet—Basic Repeated Measures Design

This particular experiment leads to the idea of a repeated measures design using two different sizes of experimental units. However, if the experimenter is not careful, he or she may inadvertently miss the fact that there are two different sizes of experimental units, resulting in an incorrect analysis.

A veterinarian has two techniques that can be used to fuse the joint of a horse's foot after it is broken and she wishes to determine if one technique is better than the other. The experiment consists of taking some horses, breaking a joint on each horse, repairing it with one of the two techniques, and determining the strength of the fused joint four months later. She also wants to determine if the same technique works equally well for front feet and back feet. Because horses are scarce and expensive, she plans to break the joint on a front foot, wait until it heals, and then break the joint on a rear foot or vice versa on each horse. This healing process also introduces a time factor into the design. Thus, the treatment structure is a 2^3 factorial arrangement generated from two fusion techniques (F), two positions (P), and two healing times (T). The design structure is an incomplete block design where each horse is a block and there are two observations per block (horse). Since the blocks are incomplete, some of the treatment structure information will be confounded with block or horse effects (Cochran and Cox, 1957). There are various ways of assigning the treatment combinations to the two feet of a horse. Suppose there are four horses. One process is to assign a fusion technique to each horse where both a front and a rear foot are treated with the same level of fusion, as shown in Table 5.30.

Let y_{ijkm} denote the observation obtained from i th fusion technique, j th position, k th time, and m th horse. The model used to describe the data is

$$y_{ijkm} = \mu_{ijk} + h_m + \varepsilon_{ijkm}, \quad i = 1, 2, j = 1, 2, k = 1, 2, m = 1, 2, \dots, 4$$

where μ_{ijk} denotes the mean response for the i th fusion technique, j th position, k th time, and h_m denotes the effect of the m th horse and ε_{ijkm} denotes the response error of a measurement made on a foot of a horse. Two types of comparison can be made among the treatment combination means, the μ_{ijk} . One type of comparison is the intra-horse (or within-horse)

TABLE 5.30

First Assignment of Treatment Combinations for Horse Feet Experiment

Horse	1	2	3	4
$F_1 P_1 T_1$	$F_1 P_1 T_2$	$F_2 P_1 T_1$	$F_2 P_1 T_2$	
$F_1 P_2 T_2$	$F_1 P_2 T_1$	$F_2 P_2 T_2$	$F_2 P_2 T_1$	

Note: The two fusing techniques are F_1 and F_2 , the two times are T_1 and T_2 , and the two positions are P_1 and P_2 .

comparison and the second type of comparison is the inter-horse (or between-horse) comparison. The factorial effects can be defined as

$$\begin{aligned}\text{Mean} &= \bar{\mu}_{...} \\ F &= \bar{\mu}_{1..} - \bar{\mu}_{2..}, P = \bar{\mu}_{.1} - \bar{\mu}_{.2}, T = \bar{\mu}_{..1} - \bar{\mu}_{..2} \\ F \times P &= \bar{\mu}_{11} - \bar{\mu}_{12} - \bar{\mu}_{21} + \bar{\mu}_{22} \\ F \times T &= \bar{\mu}_{1..1} - \bar{\mu}_{1..2} - \bar{\mu}_{2..1} + \bar{\mu}_{2..2} \\ P \times T &= \bar{\mu}_{.11} - \bar{\mu}_{.12} - \bar{\mu}_{.21} + \bar{\mu}_{.22} \\ F \times P \times T &= \mu_{111} - \mu_{112} - \mu_{121} + \mu_{122} - \mu_{211} + \mu_{212} + \mu_{221} - \mu_{222}\end{aligned}$$

The best estimator of μ_{ijk} is $\bar{y}_{ijk..}$. The best estimator of the main effect P can be expressed as

$$\begin{aligned}\hat{P} &= \bar{y}_{.1..} - \bar{y}_{.2..} \\ &= \frac{1}{4}[(y_{1111} - y_{1221}) + (y_{1122} - y_{1212}) + (y_{2113} - y_{2223}) + (y_{2124} - y_{2214})]\end{aligned}$$

which is an intra-horse or within-horse comparison. The intra-horse comparison is easily justified by substituting the right-hand side of the above model for y_{ijkm} in P , which gives

$$\begin{aligned}\hat{P} &= \frac{1}{4}[(\mu_{111} + h_1 + \varepsilon_{1111} - \mu_{122} - h_1 - \varepsilon_{1221}) + (\mu_{112} + h_2 + \varepsilon_{1122} - \mu_{121} - h_2 - \varepsilon_{1212}) \\ &\quad + (\mu_{211} + h_3 + \varepsilon_{2113} - \mu_{222} - h_3 - \varepsilon_{2223}) + (\mu_{212} + h_4 + \varepsilon_{2124} - \mu_{221} - h_4 - \varepsilon_{2214})] \\ &= \bar{\mu}_{.1..} - \bar{\mu}_{.2..} + \frac{1}{4}[(\varepsilon_{1111} - \varepsilon_{1221}) + (\varepsilon_{1122} - \varepsilon_{1212}) + (\varepsilon_{2113} - \varepsilon_{2223}) + (\varepsilon_{2124} - \varepsilon_{2214})]\end{aligned}$$

Note that the horse effects h_m subtract out of the expression. The variance of \hat{P} depends on the variance of the ε_{ijkm} and not the variance of the h_m . The variance of \hat{P} is $\text{Var}(P) = \sigma_\varepsilon^2/2$.

The best estimator of F can be expressed as

$$\begin{aligned}\hat{F} &= \bar{y}_{1...} - \bar{y}_{2...} \\ &= \frac{1}{4}[(y_{1111} + y_{1221}) + (y_{1122} + y_{1212}) - (y_{2113} + y_{2223}) - (y_{2124} + y_{2214})]\end{aligned}$$

This estimator is a between-horse comparison; that is, it is a comparison of the mean of horses 1 and 2 with the mean of horses 3 and 4, and thus it depends on the horse effects. The dependency on the horse effects is observed by expressing \hat{F} in terms of the right-hand side of the model as

$$\begin{aligned}\hat{F} &= \frac{1}{4}[(\mu_{111} + h_1 + \varepsilon_{1111} + \mu_{122} + h_1 + \varepsilon_{1221}) + (\mu_{112} + h_2 + \varepsilon_{1122} + \mu_{121} + h_2 + \varepsilon_{1212}) \\ &\quad - (\mu_{211} + h_3 + \varepsilon_{2113} + \mu_{222} + h_3 + \varepsilon_{2223}) - (\mu_{212} + h_4 + \varepsilon_{2124} + \mu_{221} + h_4 + \varepsilon_{2214})] \\ &= [\bar{\mu}_{.1..} - \bar{\mu}_{.2..} + \frac{1}{4}[(\varepsilon_{1111} + \varepsilon_{1221}) + (\varepsilon_{1122} + \varepsilon_{1212}) - (\varepsilon_{2113} + \varepsilon_{2223}) - (\varepsilon_{2124} + \varepsilon_{2214})]] \\ &\quad + \frac{1}{2}[h_1 + h_2 - h_3 - h_4]\end{aligned}$$

which involves the h_m . The variance of \hat{F} depends on the variance of the h_m and the variances of the ε_{ijkm} or $\text{Var}(\hat{F}) = \frac{1}{2} [\sigma_e^2 + 2\sigma_{\text{horse}}^2]$. Similarly, it can be shown that $P, T, F \times T$ and $F \times P$ are within-horse effects while $F, P \times T$ and $F \times P \times T$ are between-horse effects. Since the between-horse effects involve the h_m , they are said to be confounded with horse effects; that is, $F, P \times T$ and $F \times P \times T$ are confounded with horse effects.

This design consists of a three-way treatment structure in an incomplete block design structure where each horse is a block. Since some of the comparisons are comparisons between horses and others are comparisons between feet "within" a horse, this experiment involves two sizes of experimental units. The feet are the smaller-size experimental units, while the horses are the larger-size experimental units. The h_m term in the model represents the horse error (that is, variation due to differences among horses treated alike), while the ε_{ijkm} term represents the foot error (that is, variation due to differences among feet treated alike on the same horse). When designing experiments involving different sizes of experimental units, it is desirable to choose those effects that are most important so that they involve comparisons between the smaller experimental units and to let those effects that are least important involve comparisons between larger experimental units. However, this type of arrangement is not always possible. For example, if the horse foot experiment involved two types of horses (say, racing and working), it would be impossible for types of horse to be other than a between-horse comparison. In the horse foot experiment, which has only three factors, the experimenter is most interested in comparing the two fusion techniques. The design given in Table 5.30 is such that the fusion effect is confounded with horses resulting in less precision for comparing the fusion technique means than is desired. The design given in Table 5.31 yields a comparison between the fusion technique means that is not confounded with horses and thus achieves the goal of having the most important effect being compared using the variability associated with the smaller experimental unit, the horse feet.

The model for the first arrangement can also be used to represent data from this second assignment of treatment combinations to the horses or blocks. Using the same techniques as for the first arrangement, it can be shown that the F, P, T , and $F \times P \times T$ effects are within-horse or intra-horse comparisons and that the $F \times T, F \times P, P \times T$ effects are inter-horse comparisons and these effects are confounded with the horse effects. Neither of the above two designs yield enough observations to provide any degrees of freedom for estimating the two error terms. In order, to obtain some degrees of freedom for the two types of error variances, one could repeat the design in Table 5.31 using eight horses where two horses are randomly assigned to each set of treatment combinations. The analysis would still consist of two parts, a between-horse analysis and a within-horse or feet-within-horse analysis. There would be eight within-horse comparisons (one from each horse) that can be partitioned into estimates of the F, T, P , and $F \times P \times T$ effects and an estimate of the within-horse error variance, denoted by error (feet). There are seven between-horse

TABLE 5.31

Second Assignment of Treatment Combinations for Horse Feet Experiment

Horse	1	2	3	4
$F_1P_1T_1$	$F_2P_1T_1$	$F_1P_1T_2$	$F_2P_1T_2$	
$F_2P_2T_2$	$F_1P_2T_2$	$F_2P_2T_1$	$F_1P_2T_1$	

TABLE 5.32

Between-Horse and Within-Horse Analysis of Variance Table

Source	df
Between-Horse	7
$F \times P$	1
$F \times T$	1
$P \times T$	1
Error (Horse)	4
Within-Horse	8
F	1
T	1
P	1
$F \times P \times T$	1
Error (foot)	4

} Horse is experimental unit
 } Foot of horse is experimental unit

comparisons which can be partitioned into estimates of the $F \times T$, $F \times P$, and $P \times T$ effects and an estimate of the between-horse error variance, denoted by error (horse). The resulting analysis of variance table is displayed in Table 5.32.

This designed experiment falls into the class of repeated measures designs since there are two measurements (repeated) on each horse; that is, a front foot and a rear foot are measured on each horse, and the levels of position cannot be randomly assigned to its experimental units.

5.5.2 Example 5.7: Comfort Study—Repeated Measures Design

An experimenter wants to study the effect of six environmental conditions on the comfort of people. He has six environmental chambers, and each can be set with a different environmental condition. The experiment consists of putting one person in a chamber and then measuring the person's comfort after 1, 2, and 3 hours of exposure. There are 36 subjects in the study where six subjects are randomly assigned to each environment. The researcher can obtain data from six environmental conditions during one day, thus days are used as a blocking factor where each environmental condition is assigned to one of the six chambers each day. Time of exposure is an important factor in the study, thus the treatment structure is a two-way with the levels of environment crossed with the three exposure times. The subjects were randomly assigned a number from 1 to 36, where the first six persons were involved in the experiment during the first day, etc. An assignment of environments and persons to chambers is displayed in Figure 5.23. Each rectangle in Figure 5.23 represents a person and T_1 , T_2 , and T_3 represent the measures of comfort after 1, 2, and 3 hours of exposure, respectively. The experimental unit for environments is a person or chamber and the experimental unit for time is a 1 h interval "within" that person. In effect, the person is "split" into three parts, but this design structure is called a repeated measures design rather than a split-plot design since the levels of time of exposure cannot be randomly assigned to the three 1 h exposure times within a person. The structure of this design is identical to the usual split-plot with a two-way treatment structure in a randomized complete block whole-plot design structure. For this repeated measures design structure the larger-sized experimental unit or person or whole plot design consists of a one-way treatment structure (six levels of environment) in a randomized complete block design

	Environment					
	1	2	3	4	5	6
Day 1	(1) $T_1 T_2 T_3$	(2) $T_1 T_2 T_3$	(3) $T_1 T_2 T_3$	(4) $T_1 T_2 T_3$	(5) $T_1 T_2 T_3$	(6) $T_1 T_2 T_3$
Day 2	(7) $T_1 T_2 T_3$	(8) $T_1 T_2 T_3$	(9) $T_1 T_2 T_3$	(10) $T_1 T_2 T_3$	(11) $T_1 T_2 T_3$	(12) $T_1 T_2 T_3$
Day 3	(13) $T_1 T_2 T_3$	(14) $T_1 T_2 T_3$	(15) $T_1 T_2 T_3$	(16) $T_1 T_2 T_3$	(17) $T_1 T_2 T_3$	(18) $T_1 T_2 T_3$
Day 4	(19) $T_1 T_2 T_3$	(20) $T_1 T_2 T_3$	(21) $T_1 T_2 T_3$	(22) $T_1 T_2 T_3$	(23) $T_1 T_2 T_3$	(24) $T_1 T_2 T_3$
Day 5	(25) $T_1 T_2 T_3$	(26) $T_1 T_2 T_3$	(27) $T_1 T_2 T_3$	(28) $T_1 T_2 T_3$	(29) $T_1 T_2 T_3$	(30) $T_1 T_2 T_3$
Day 6	(31) $T_1 T_2 T_3$	(32) $T_1 T_2 T_3$	(33) $T_1 T_2 T_3$	(34) $T_1 T_2 T_3$	(35) $T_1 T_2 T_3$	(36) $T_1 T_2 T_3$

FIGURE 5.23 Data arrangement for the comfort study.

structure (six days). The smaller-sized experimental unit or 1 h time interval or split-plot design is a one-way treatment structure (three times of exposure) in a randomized complete block design structure where each person or chamber is a block. A model that can be used to describe data from this experiment is

$$y_{ijk} = \mu + E_i + d_j + c_{ij} + T_k + (ET)_{ik} + \varepsilon_{ijk}, \quad i = 1, 2, \dots, 6, \quad j = 1, 2, \dots, 6, \quad k = 1, 2, 3$$

where d_j denotes the j th day effect, c_{ij} denotes the effect of the chamber or person used for the i th environment during the j th day and ε_{ijk} is the time interval effect. If there is one observation made on each person, the resulting data set would be from a one-way treatment structure in a randomized complete block design structure. The chamber or person error term is computed from the day by environment interaction. If one compares the times at the first environmental condition, the resulting design is a one-way treatment structure in a randomized complete block design structure where the person or chamber or day is the blocking factor. The time interval error term is computed from the time by day interaction. An error term for time intervals can be computed for each of the environments and can be pooled into the time interval error, if the variances are equal. The analysis of variance table for this repeated measures design is given in Table 5.33. This basic repeated measures structure is identical to that of the usual split-plot design. The assumptions made for the split-plot model are that the d_j , c_{ij} , and ε_{ijk} are independent random variables with variances σ^2_{day} , $\sigma^2_{\text{chamber}}$, and σ^2_{ε} . The assumptions for the repeated measures design are that d_j , c_{ij} , and ε_{ijk} are independent random variables, but since the ε_{ijk} are measured on the same person, they can be correlated and that correlation structure may need to be modeled. The modeling of the covariance structure is discussed in Chapters 26 and 27.

An alternative to assigning one person to each environmental chamber is to assign a group of six people to a single environmental chamber and complete the study in one day. This assignment process seems like a good idea, but the resulting data set does not provide any measure as to how chambers vary when treated alike. The nonexistence of the chamber

TABLE 5.33

Analysis of Variance Table for the Comfort Study

Source	df	EMS
Days	5	$\sigma_e^2 + 3\sigma_{\text{chamber}}^2 + 18\sigma_{\text{day}}^2$
Person Analysis		
Environment	5	$\sigma_e^2 + 3\sigma_{\text{chamber}}^2 + \phi^2(E)$
Error (person)	25	$\sigma_e^2 + 3\sigma_{\text{chamber}}^2$
Hour-Interval Analysis		
Time	2	$\sigma_e^2 + \phi^2(T)$
Environment \times time	10	$\sigma_e^2 + \phi^2(ET)$
Error (hour-interval)	60	σ_e^2

error term is because the group of six people forms the experimental unit for environments, and thus there is only one independent observation for each environment. Without an error term for chambers, there is no way to assess the effects of environment on comfort.

5.5.3 Example 5.8: Crossover or Change-Over Designs

A useful method for comparing treatments to be administered to subjects (or objects) is to allow the subject to serve as its own control by applying two or more of the treatments to the same subject. If there are two treatments, the process is to assign treatment *A* to be applied in the first time period and measure the response, allow the effect of treatment *A* to diminish or wash out, and then apply treatment *B* to the subject in the second time period and observe the response to treatment *B*. The randomization process is to construct sequences of the treatments (*A* followed by *B* and *B* followed by *A*, here) and then randomly assign subjects or animals to the sequences. This approach can also be used on plants, plots of land, or other objects where after subjecting an experimental unit to a treatment, the experimental unit can at least partially recover from the effects of the treatment given in the first time period.

In this method of comparing two treatments, there are two sequences of treatment assignments for an animal—*A* followed by *B* and *B* followed by *A*. The two sequences are often denoted by the *AB* sequence and the *BA* sequence. The treatment structure is a one-way set of treatments with two levels (*A* and *B*), but since the treatments are applied in sequence to each experimental unit, the generated sequence becomes another type of treatment. Thus the design of this experiment involves a two-way treatment structure with treatments crossed with sequences.

This crossover design is a repeated measures design with two sizes of experimental units. The treatment sequence is assigned to a subject so that the subject is the larger sized experimental unit. The time periods or times during which the treatments are observed are the smaller-sized experimental unit. The design for the large-sized experimental units is a one-way treatment structure with two levels of sequence (the two possible sequences) in a completely randomized design structure. Although any design structure can be used, a completely randomized design is used most often. The design for the small-sized experimental units (time interval) is a one-way treatment structure {two time periods or two treatments} in a randomized complete block design structure where the subjects are

TABLE 5.34

Data Arrangement for a Two-Period Two-Treatment Crossover Design

Animal				
	1	2	...	n_i
Sequence 1				
A (time 1)	y_{11A}	y_{12A}	...	y_{1n_1A}
B (time 2)	y_{11B}	y_{12B}	...	y_{1n_1B}
Sequence 2				n_2
B (time 1)	y_{21B}	y_{22B}	...	y_{2n_2B}
A (time 2)	y_{21A}	y_{22A}	...	y_{2n_2A}

the blocks. The data can be arranged as shown in Table 5.34 and a model that can be used to describe the data is

$$y_{ijk} = \mu_{ik} + s_{ij} + \varepsilon_{ijk}, \quad \text{where } (i, j, k) \in I_D, \quad s_{ij} \sim i.i.d. N(0, \sigma_{\text{subject}}^2), \quad \text{and} \quad \varepsilon_{ijk} \sim i.i.d. N(0, \sigma_{\text{time}}^2)$$

where I_D is an index set with the collection of (i, j, k) triples actually used in the experiment, μ_{ik} denotes the effect of the k th treatment in the i th sequence, s_{ij} denotes the random effect of the j th subject assigned to the i th sequence, and ε_{ijk} denotes the random error of a measurement in the time period of the i th sequence to which the k th treatment was applied. The analyses of some crossover designs are presented in Chapter 29.

5.6 Designs Involving Nested Factors

In a given design, it is possible to have nested effects in the design structure, in the treatment structure, or in both structures. Nesting occurs most often in the design structure of an experiment where a smaller-sized experimental unit is nested within a larger sized one. One size of experimental unit is nested within a larger size if the smaller experimental units are different for each large experimental unit. When the design structure consists of several different sizes of experimental units and there is an ordering where the smallest size is nested within the next smallest size, on up to the next to largest is nested within the largest size experimental units, the design is also called a hierarchical design structure. The split-plot and repeated measures designs discussed in the previous sections of this chapter are good examples of design structures where the smaller-sized experimental units are nested within the larger-sized experimental units. Nesting occurs in the treatment structure when the levels of one factor occur with only one level of a second factor. In that case, the levels of the first factor are said to be nested within the levels of the second factor. Table 5.35 illustrates nesting in the design structure for the second part of Example 4.2 where houses are nested within squares (an "X" indicates that a house is included in the indicated square). Each square is the large experimental unit or block of houses, and the house is the smaller size of experimental unit. Since the houses for the first square are different from the houses for the second square, houses are nested within squares. Such a nested effect is often expressed in a model by $s_k + h_{m(k)}$ where s_k denotes the effect of the k th

TABLE 5.35

Design Showing Houses Nested within Squares

Square	Houses							
	1	2	3	4	5	6	7	8
1	X	X	X	X				
1					X	X	X	X

square and $h_{m(k)}$ denotes the effect of the m th house nested in the k th square. The sums of squares for squares and houses within squares are denoted by SSSQUARE and SSHOUSES(SQUARES), respectively. If houses and not squares are included in the model, then there is only a single sum of squares due to houses (SSHUSES) which can be partitioned as $SSHUSES = SSSQUARES + SSHUSES(SQUARES)$. This process can be carried on for one more step where the sides of the houses are nested within houses that are nested within squares. Thus these three sizes of experimental units form a hierarchical design structure.

5.6.1 Example 5.9: Animal Genetics

An animal scientist wants to study the effect of genetics on the growth rate of lambs. She has four males (sires) and 12 females (dams). The breeding structure is shown in Table 5.36 (an "X" denotes a mating). For this example, each sire is mated to three dams where the three dams are different for each sire. Thus, the levels of dam are called a nested factor where dams are nested within sires. If the animal scientist is interested the effect of this set of sires with this set of dams, then the levels of dams nested within the levels of sires forms a nested treatment structure.

When nesting occurs in the treatment structure, the treatment structure must consist of at least two factors. In this case, each level of the nested factor occurs just once with a level or levels of the other factor. The next two examples demonstrate nesting in a treatment structure.

5.6.2 Example 5.10: Soybeans in Maturity Groups

The agricultural crop of soybeans provides a great example of nesting in the treatment structure. Varieties of soybeans are classified into maturity groups. In the Midwest region

TABLE 5.36

Breeding Structure Showing Dams Nested within Sires

Sires	Dams											
	1	2	3	4	5	6	7	8	9	10	11	12
1	X	X	X									
2				X	X	X						
3							X	X	X			
4										X	X	X

TABLE 5.37

Treatment Structure of the Soybean Study with Varieties Nested within Maturity Groups

Maturity Group	Varieties							
	1	2	3	4	5	6	7	8
4	X	X						
5			X	X	X	X		
6							X	X

of the United States, varieties of soybeans in maturity groups 4, 5, and 6 are grown, but varieties from one maturity group may be better for a particular region than a variety of another maturity group. A study was designed to evaluate eight soybean varieties where two were from maturity group 4, four were from maturity 5 and two were from maturity group 6. Table 5.37 is a display of the treatment structure, indicating the levels of varieties are nested within the levels of maturity. A model that can be used to describe data from this nested treatment structure in a randomized complete block design structure with four blocks is

$$y_{ijk} = \mu + M_i + V(M)_{j(i)} + b_k + \varepsilon_{ijk}, \quad i = 1, 2, 3, \quad j = 1, \dots, n_i, \quad k = 1, 2, 3, 4$$

$$b_k \sim i.i.d. N(0, \sigma_{\text{row}}^2), \quad \varepsilon_{ijk} \sim i.i.d. N(0, \sigma_{\varepsilon}^2), \quad n_1 = n_3 = 2, \text{ and } n_2 = 4.$$

where M_i denotes the effect of the i th maturity group, $V(M)_{j(i)}$ denotes the effect of the j th variety nested within the i th maturity group, b_k denotes the effect of the k th block and ε_{ijk} denotes the experimental unit error. The analysis of variance table for the above model is given in Table 5.38 where the error term is computed as the variety by block interaction. The variety by block interaction can be partitioned as the maturity group by block interaction plus the variety nested within maturity group by block interaction.

One problem with the above design is that the plots need to be harvested at different times as maturity group 4 beans are ready for harvest before maturity group 5 beans which are ready to be harvested before maturity group 6 beans. An alternative to the randomized complete block design is to use a split-plot design where the whole plots are formed by the maturity groups and the subplots are formed by the varieties within a maturity group. The advantage of the split-plot design is that all varieties within a whole-plot could be harvested at the same time. The randomization process is to randomly assign the maturity group levels to sets of plots within each whole plot and then randomly assign

TABLE 5.38

Analysis of Variance Table for Soybean Varieties Nested within Maturity Groups in a Randomized Complete Block Design Structure

Source	df	
Blocks	3	$\sigma_{\varepsilon}^2 + 8\sigma_{\text{block}}^2$
Maturity groups	2	$\sigma_{\varepsilon}^2 + \phi^2(M)$
Varieties (maturity groups)	5	$\sigma_{\varepsilon}^2 + \phi^2[V(M)]$
Error	21	σ_{ε}^2

the respective varieties to the plots within a maturity group. A model for this split-plot design structure is

$$y_{ijk} = \mu + M_i + b_k + w_{ik} + V(M)_{j(i)} + \varepsilon_{ijk}, \quad i = 1, 2, 3, \quad j = 1, \dots, n_i, \quad k = 1, 2, 3, 4$$

$$b_k \sim i.i.d. N(0, \sigma_{\text{row}}^2), \quad w_{ik} \sim i.i.d. N(0, \sigma_{\text{wp}}^2), \quad \text{and} \quad \varepsilon_{ijk} \sim i.i.d. N(0, \sigma_{\varepsilon}^2)$$

where w_{ik} denotes the whole-plot effect and ε_{ijk} denotes the subplot effect. The whole plot design is a one-way treatment structure (three levels of maturity) in a randomized complete block design structure where the whole-plot error is computed from the block by maturity interaction. The subplot design consists of three separate one-way treatment structures in randomized complete block design structures, one for each maturity group. The whole-plots are of different size since there are different numbers of varieties within each maturity group. The analysis of variance table for this two-way nested treatment structure in a split-plot design structure is in Table 5.39.

5.6.3 Example 5.11: Engines on Aircraft

An aircraft company wants to evaluate the performance of seven engine types with three aircraft types. Because of certain mechanical characteristics, only certain engines can be used with each type of aircraft. The possible engine–aircraft configurations (marked by an “X”) are displayed in Table 5.40.

As shown in Table 5.40, each engine type can occur with one and only one aircraft type, thus the levels of engines are nested within the levels of aircraft. The aircraft company made three aircraft for each of the seven treatment combinations (engine types by aircraft types). The data collection process is to randomly order the 21 engine type–aircraft type configurations, have a test pilot fly the planes in the random order, and measure a performance characteristic. Let y_{ijk} denote a performance measure of the k th airplane made from the i th aircraft type with the j th engine type. Models that can be used to describe the performance measures expressed as a means model and as an effects model are

$$y_{ijk} = \mu_{ij} + p_{ijk}, \quad \text{for } (i, j) \in \Theta \text{ and } p_{ijk} \sim i.i.d. N(0, \sigma_{\text{plane}}^2)$$

or

$$y_{ijk} = \mu + A_i + E_{j(i)} + p_{ijk}, \quad \text{for } (i, j) \in \Theta,$$

TABLE 5.39

Analysis of Variance Table for Soybean Varieties Nested within Maturity Groups in a Split-Plot Design Structure

Source	df	
Blocks	3	$\sigma_{\varepsilon}^2 + 2.5\sigma_{\text{wp}}^2 + 7.2\sigma_{\text{block}}^2$
Maturity groups	2	$\sigma_{\varepsilon}^2 + 2.5\sigma_{\text{wp}}^2 + \phi^2(M)$
Error (whole-plot)	6	$\sigma_{\varepsilon}^2 + 2.5\sigma_{\text{wp}}^2$
Varieties (maturity groups)	5	$\sigma_{\varepsilon}^2 + \phi^2[V(M)]$
Error (subplot)	15	σ_{ε}^2

TABLE 5.40

Observable Engine–Aircraft Configurations of the Nested Treatment Structure

Aircraft Type	Engine Type						
	A	B	C	D	E	F	G
1	X	X	X				
2				X	X		
3						X	X

where

$$\Theta = \{(1, A), (1, B), (1, C), (2, D), (2, E), (3, F), (3, G)\}$$

An analysis of variance table for the effects model is displayed in Table 5.41. The plane error term is computed from the variation of the performance scores of the three planes made with the same engine type and aircraft type configuration pooled across the seven configurations. In terms of the means model, the aircraft sum of squares tests the equality of the aircraft means or tests the null hypothesis

$$\bar{\mu}_{1\cdot} = \bar{\mu}_{2\cdot} = \bar{\mu}_{3\cdot}$$

or

$$\frac{\mu_{1A} + \mu_{1B} + \mu_{1C}}{3} = \frac{\mu_{2D} + \mu_{2E}}{2} = \frac{\mu_{3F} + \mu_{3G}}{2}$$

The engine nested within aircraft sum of squares tests the null hypothesis

$$\mu_{1A} = \mu_{1B} = \mu_{1C}, \quad \mu_{2D} = \mu_{2E}, \quad \text{and} \quad \mu_{3F} = \mu_{3G}$$

One has to be quite careful when nesting occurs in the treatment structure because there is a tendency to carry out the analysis as if there is a nested or hierarchical design structure. The individual airplanes are the only experimental units for this study, that is, there is just one size of experimental unit and only one error term. When the nesting occurs in the design structure, there is more than one size of experimental unit and thus more than one error term in the model. The next example illustrates nesting in the design structure.

TABLE 5.41

Analysis of Variance Table for Two-Way Nested Treatment Structure in a Completely Randomized Design Structure

Source	df	EMS
Aircraft type	2	$\sigma_{\text{plane}}^2 + \phi^2(A)$
Engine type (aircraft type)	4	$\sigma_{\text{plane}}^2 + \phi^2[E(A)]$
Error (plane)	14	σ_{plane}^2

5.6.4 Example 5.12: Simple Comfort Experiment

A comfort experiment was conducted to study the effects of temperature (three levels, 18, 21, and 24°C) and sex of person [two levels, male (M) and female (F)] in a two-way treatment structure on a person's comfort. There are several methods to measure a person's comfort, and so for the discussion here, assume that there is one comfort measurement made on each person. The three temperatures were each randomly assigned to three of the nine available environmental chambers. A chamber is the experimental unit for the levels of temperature and the chamber design is a one-way treatment structure in a completely randomized design structure.

Eighteen males and 18 females were randomly assigned to chambers so that two males and two females were assigned to each of the nine chambers. The experimental unit for sex of person is a person, and the person design is a one-way treatment structure in a randomized complete block design structure where the chambers are the blocks. There are two replications on each level of sex within each block. The assignment of chambers to temperatures and of persons to chambers is displayed in Figure 5.24.

After the people were subjected to the environmental condition for 3 h, their comfort was measured. A means model and an effects model that can be used to describe these data are

$$y_{ijkm} = \mu_{ik} + c_{j(i)} + p_{m(jk)} \quad i = 1, 2, 3, \quad j = 1, 2, 3, \quad k = 1, 2, \quad m = 1, 2$$

for

$$c_{j(i)} \sim i.i.d. N(0, \sigma_{\text{chamber}}^2) \quad \text{and} \quad p_{m(jk)} \sim i.i.d. N(0, \sigma_{\text{person}}^2)$$

or

$$y_{ijkm} = \mu + T_i + c_{j(i)} + S_k + (TS)_{ik} + p_{m(jk)}$$

Temperature	Chamber 1	Chamber 2	Chamber 3
18°C	F 1 2 M 1 2	F 3 4 M 3 4	F 5 6 M 5 6
21°C	F 7 8 M 7 8	F 9 10 M 9 10	F 11 12 M 11 12
24°C	F 13 14 M 13 14	F 15 16 M 15 16	F 17 18 M 17 18

Numbers within boxes denote person numbers of a given sex

FIGURE 5.24 Assignments of persons and temperatures to chambers for simple comfort experiment.

TABLE 5.42

Analysis of Variance Table for the Simple Comfort Experiment with a Nested Design Structure

Source	df	EMS
Temperature	2	$\sigma_{\text{person}}^2 + 4\sigma_{\text{chamber}}^2 + \phi^2(T)$
Error (chamber)	6	$\sigma_{\text{person}}^2 + 4\sigma_{\text{chamber}}^2$
Sex	1	$\sigma_{\text{person}}^2 + \phi^2(S)$
Sex \times temperature	2	$\sigma_{\text{person}}^2 + \phi^2(ST)$
Error (person)	24	σ_{person}^2

where μ_{ik} denotes the mean of the i th temperature and k th sex, $c_{j(i)}$ is the random effect of the j th chamber assigned to the i th temperature, $p_{m(jk)}$ denotes the random effect of the m th person of the k th sex assigned to the j th chamber assigned to the i th temperature, and T_i , S_k , and $(TS)_{ik}$ denote the effects of temperature, sex, and temperature by sex interaction, respectively.

There are two levels of nesting involved in this experiment. First, environmental chambers are nested within temperatures. Second, persons of the same sex are nested within chambers. The analysis of variance table for this experiment is shown in Table 5.42. The effects of sex and the sex \times temperature interaction are between-person comparisons and, thus the person error term is used for making comparison among those means. The chamber error term is used to make comparisons between temperature means.

5.6.5 Example 5.13: Multilocation Study with Repeated Measures

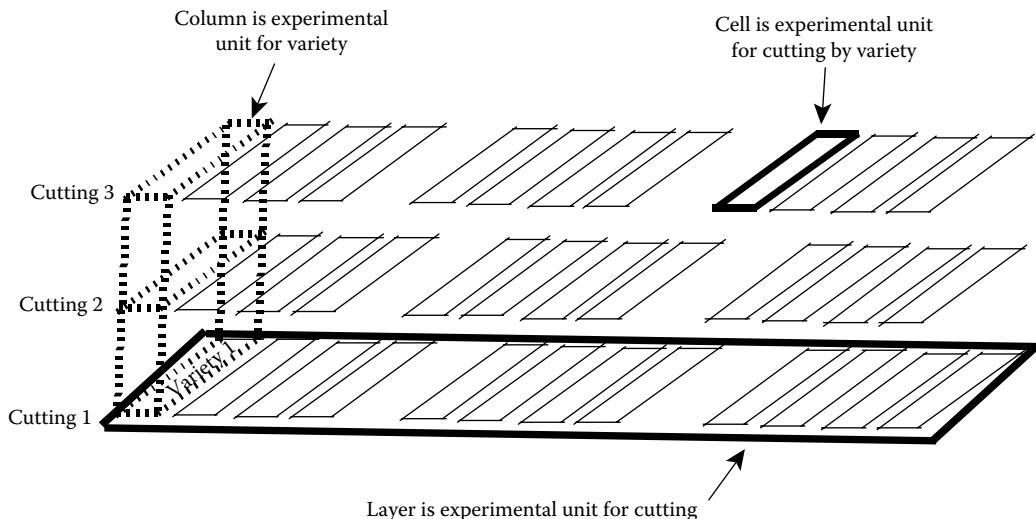
A typical agricultural study involves evaluating treatments at several locations and often the response is measured over time. This study is to evaluate four varieties of alfalfa using three locations. Alfalfa is a crop that is harvested three or more times during the growing season, so the data set contains repeated measures. At each location, the design is a one-way treatment structure in a randomized complete block design structure with three replications. The assignment of varieties to plots at each location is displayed in Figure 5.25. Figure 5.26 is a graphic of the experimental units at each of the locations. Since all of the plots at a location are cut or harvested at the same time, all of the plots at a location are subjected to the same environmental conditions that happen to occur between cuttings. Thus, it is the whole experiment at a given location that is being repeatedly measured and not the individual plots. A column in Figure 5.26 represents the entity to which a variety is applied and harvested three times and is the experimental unit for varieties. A layer of plots represents the entity which is subjected to the environmental conditions and is the experimental unit for the levels of cutting. The levels of cutting form a strip-plot structure across all of the plots at a given location. The cell denotes the unit of a variety being harvested at one cutting time.

The analysis for the columns can be constructed by ignoring the layers, which provides a one-way treatment structure in randomized complete block design structure at each location. Thus the column error is computed by the variety by block interaction pooled across locations. The analysis for the layers can be constructed by ignoring the columns, which provides a one-way treatment structure (levels of cutting) in a randomized complete block design structure where locations are the blocks. The layer error term is computed as the cutting by location interaction. Finally, the cell is the unit for the cutting by

				Location 1		
		Rep 1(1)	Rep 2(1)	Rep 3(1)		
Variety	2	1	4	3		
Rep 1(1)	Variety 2	Variety 1	Variety 4	Variety 3		
Rep 2(1)	Variety 1	Variety 3	Variety 2	Variety 4		
Rep 3(1)	Variety 2	Variety 4	Variety 1	Variety 3		

				Location 2		
		Rep 1(2)	Rep 2(2)	Rep 3(2)		
Variety	3	1	2	4		
Rep 1(2)	Variety 3	Variety 1	Variety 4	Variety 2		
Rep 2(2)	Variety 1	Variety 2	Variety 3	Variety 4		
Rep 3(2)	Variety 3	Variety 2	Variety 4	Variety 1		

				Location 3		
		Rep 1(3)	Rep 2(3)	Rep 3(3)		
Variety	4	1	2	3		
Rep 1(3)	Variety 4	Variety 1	Variety 3	Variety 2		
Rep 2(3)	Variety 3	Variety 1	Variety 4	Variety 2		
Rep 3(3)	Variety 4	Variety 1	Variety 3	Variety 2		

FIGURE 5.25 Assignment of alfalfa varieties to plots for multi-location experiment.**FIGURE 5.26** Display of experimental units at one location for multi-location study.

variety interaction and the cell error term is computed as the cutting by variety by block interaction pooled across locations plus the cutting by block interaction nested within locations. A model that can be used to describe a response from each cell is

$$y_{ijkl} = \mu + l_i + b_{j(i)} + V_k + (lV)_{ik} + (bV)_{jk(i)} + C_m + (lC)_{im} + (VC)_{km} + (lVC)_{ikm} + \epsilon_{ijkl}$$

TABLE 5.43

Analysis of Variance Table for the Multilocation Study with Repeated Measures

Source	df	EMS
Locations	2	$\sigma_{\text{cell}}^2 + 3\sigma_{B k(\text{Loc})}^2 + 12\sigma_{V^*B k(\text{Loc})}^2 + 3\sigma_{V^*V^*\text{Loc}}^2 + 12\sigma_{C^*\text{Loc}}^2 + 9\sigma_{V^*\text{Loc}}^2 + 36\sigma_{\text{Loc}}^2$
Blocks (location)	6	$\sigma_{\text{cell}}^2 + 3\sigma_{B k(\text{Loc})}^2 + 12\sigma_{V^*B k(\text{Loc})}^2$
Varieties	3	$\sigma_{\text{cell}}^2 + 3\sigma_{B k(\text{Loc})}^2 + 3\sigma_{V^*C^*\text{Loc}}^2 + 9\sigma_{V^*\text{Loc}}^2 + \phi^2(V)$
Locations \times varieties	6	$\sigma_{\text{cell}}^2 + 3\sigma_{B k(\text{Loc})}^2 + 3\sigma_{V^*C^*\text{Loc}}^2 + 9\sigma_{V^*\text{Loc}}^2$
Varieties \times blocks (varieties) = error (column)	18	$\sigma_{\text{cell}}^2 + 3\sigma_{B k(\text{Loc})}^2$
Cuttings	2	$\sigma_{\text{cell}}^2 + 3\sigma_{V^*C^*\text{Loc}}^2 + 12\sigma_{C^*\text{Loc}}^2 + \phi^2(C)$
Location \times cuttings = error (layer)	4	$\sigma_{\text{cell}}^2 + 3\sigma_{V^*C^*\text{Loc}}^2 + 12\sigma_{C^*\text{Loc}}^2$
Varieties \times cuttings	6	$\sigma_{\text{cell}}^2 + 3\sigma_{V^*C^*\text{Loc}}^2 + \phi^2(VC)$
Location \times varieties \times cuttings	12	$\sigma_{\text{cell}}^2 + 3\sigma_{V^*C^*\text{Loc}}^2$
Error (cell)	48	σ_{cell}^2

where

$$l_i \sim N(0, \sigma_{\text{Loc}}^2), \quad b_{j(i)} \sim N(0, \sigma_{B|k(\text{Loc})}^2), \quad (IV)_{ik} \sim N(0, \sigma_{V^*\text{Loc}}^2), \quad (bV)_{jk(i)} \sim N(0, \sigma_{V^*B|k(\text{Loc})}^2), \\ (lC)_{im} \sim i.i.d. N(0, \sigma_{C^*\text{Loc}}^2), \quad (IVC)_{ikm} \sim i.i.d. N(0, \sigma_{V^*C^*\text{Loc}}^2), \quad \varepsilon_{ijkm} \sim i.i.d. N(0, \sigma_{\text{cell}}^2)$$

The model assumptions are expressed where locations and blocks nested within locations are considered as random effects (see Chapter 18). As a consequence, all interactions involving location or blocks nested within location are also random effects. The levels of cutting at each location are the repeated measurements as cutting 1 is first, followed by cutting 2 which is followed by cutting 3. The levels of cutting cannot be randomly assigned to the layers within a location, thus providing the repeated measurements. The assumptions in this model are that the layers are independent of each other with equal variances, but because of the repeated measurements a more complex covariance structure may be more appropriate. The modeling of the covariance structure for repeated measures designs is discussed in Chapters 26 and 27. The analysis of variance table for this model is displayed in Table 5.43 where the expected mean squares are obtained from the model assumptions about the random effects (see Chapter 18).

5.7 Concluding Remarks

In this chapter, design structures involving several different sizes of experimental units were considered. These designs are also called multilevel designs and some multilevel designs are called hierarchical designs. Design types discussed included split-plot designs, strip-plot designs, nested designs, repeated measures designs, and several variations and combinations. The designs discussed in this chapter can be combined into some very complex designs. In such cases, the key to determining an appropriate model needed to describe

the data is to identify the sizes of the experimental units and to identify the treatment structure and design structure for each size of experimental unit. Almost always, a design structure for a given size of experimental unit will be one of the four basic design structures and the error sum of squares will be computed using the process for the specific basic design structure. The emphasis in this chapter was on recognizing such designs and on when and how to use them. Analyses of designs involving more than one size of experimental unit are presented in Chapters 24–30, where the assumptions concerning the models are discussed in detail. For further discussion and examples of multilevel designs, refer to Milliken (2003a, b) and Milliken et al. (1998).

5.8 Exercises

- 5.1 Find two published research papers in the literature which use a two-way or higher-order treatment structure within a multilevel design which has two or more different sizes of experimental units. For each paper, describe in detail the treatment structure, the design structure and the different sizes of experimental units used in the experiment. Comment as to the appropriateness of the design and its analysis [at least as far as the information provided by the author(s) is concerned].
- 5.2 The effects of four chemical treatments and two irrigation levels on the growth of two cultivars of wheat were studied using eight growth chambers. Each growth chamber was assigned a level of irrigation and a cultivar. There were eight pots in each growth chamber, and the two levels of chemical were randomly assigned to the pots, denoted by Chem 1 or Chem 2, for a total of four pots for each chemical within an irrigation \times cultivar combination. The randomization scheme is displayed in the following table.

Chamber Treatments				Pots in Chambers							
Chamber	Irrigation	Cultivar	Chem	Chem	Chem	Chem	Chem	Chem	Chem	Chem	Chem
1	1	1	1	2	1	2	2	1	2	1	1
2	2	1	2	2	1	1	2	2	1	1	1
3	1	1	2	2	1	1	2	1	1	1	2
4	1	2	2	2	1	2	2	1	1	1	1
5	2	1	1	2	1	1	2	2	2	1	1
6	2	2	1	2	1	2	1	2	1	2	1
7	1	2	1	1	2	2	2	2	1	1	1
8	2	2	1	2	2	1	2	1	1	1	2

- 1) Write out an analysis of variance table for this experiment.
 - 2) Identify each size of experimental unit and the corresponding design structure and treatment structure.
 - 3) Write an effects model to describe data from this experiment.
- 5.3 The effects of four chemical treatments and two irrigation levels on the growth of two cultivars of wheat were studied using four growth chambers. There were eight pots in each growth chamber, denoted by Chem 1 or Chem 2, for a total of four pots for each chem \times irr \times cult combination as displayed below, that is, the table shows the randomization scheme.

Chamber Treatments			Pots in Chambers									
Chamber	Irrigation	Cultivar	Chem	Chem	Chem	Chem	Chem	Chem	Chem	Chem	Chem	Chem
1	1	1	1	2	1	2	1	2	1	2	1	2
Chamber	Irrigation	Cultivar	Chem	Chem	Chem	Chem	Chem	Chem	Chem	Chem	Chem	Chem
2	1	2	1	2	1	2	1	2	1	2	1	2
Chamber	Irrigation	Cultivar	Chem	Chem	Chem	Chem	Chem	Chem	Chem	Chem	Chem	Chem
3	2	1	1	2	1	2	1	2	1	2	1	2
Chamber	Irrigation	Cultivar	Chem	Chem	Chem	Chem	Chem	Chem	Chem	Chem	Chem	Chem
4	2	2	1	2	1	2	1	2	1	2	1	2

- 1) Write out an analysis of variance table for this experiment. Identify the different experimental units and their corresponding design structures and treatment structures.
 - 2) What, if anything, do you see wrong with this design?
 - 3) If you were asked to help develop a design structure for this experiment, how would you advise the experimenter to carry out an experiment. Write out the analysis of variance table for your design.
 - 4) Provide an alternative design to the one in part 3 and write out the analysis of variance table for this experimental design.
 - 5) Discuss what you believe to be some of the advantages and disadvantages of the two designs in parts 3 and 4.
- 5.4 An experiment was to be conducted to study the effects of three different melt mixing procedures (M1, M2, and M3), and nine coating lay-downs (C1, C2, ..., C9), on the quality of photographic film. One of the steps in film-making requires that the film be placed in baskets while being processed in a standard way. This standard process was not being changed in any way. The baskets are large enough to hold 18 strips of film. The nine coating lay-downs were randomly assigned to two of the 18 film strips within each basket. The melt mixing procedures were randomly assigned to the baskets. The three baskets formed one complete replication of the experiment. This whole process was repeated three more times on different days, providing four complete replicates of the treatment combinations.
- 1) What are the different sizes of experimental units?
 - 2) What is the treatment structure for this experiment?
 - 3) What is the design structure for this experiment?

- 4) What is the design structure and treatment structure for each size of experimental unit?
- 5) Write out the analysis of variance table for this experiment.
- 5.5 Suppose the experiment in Exercise 5.4 is conducted in the following manner: For a basket selected at random, one of the nine coating lay-downs and two of the three melt mixing procedures are assigned at random to the 18 film strips in the basket. For example, one basket might contain the combinations, M1C1, M1C2, ..., M1C9, M2C1, M2C2, ..., M2C9. A second basket gets a different pair of the melt mixing procedures combined with all nine coating lay-downs. Thus the second basket may contain the combinations, M1C1, M1C2, ..., M1C9, M3C1, M3C2, ..., M3C9. The third basket gets the remaining pair of melt mixing procedures and all nine coating lay-downs, so the third basket would contain the combinations, M2C1, M2C2, ..., M2C9, M3C1, M3C2, ..., M3C9. This whole process would be repeated three more times on different days, providing four complete replicates.
- 1) What are the different sizes of experimental units?
 - 2) What is the treatment structure for this experiment?
 - 3) What is the design structure for this experiment?
 - 4) What is the design structure and treatment structure for each size of experimental unit?
 - 5) Write out the analysis of variance table for this experiment.
- 5.6 Discuss what you believe to be the advantages and disadvantages of the two designs used in Exercises 5.4 and 5.5. Which design would you recommend and why?
- 5.7 A baker wants to evaluate the effect of baking temperature on different formulations of bread. During one time period, she has two ovens, each of which can bake three loaves of bread at one time. The baker wants to evaluate three baking temperatures and four bread formulations. On a given day, she mixes three batches of bread dough using three of the formulations and forms two loaves from each batch. Next, one loaf from each batch is placed into the two ovens and baked at one of the three temperatures. The following table displays the formulations and temperatures used each of 12 days:

Temperatures and Formulations Used Each Day of the Experiment for Exercise 5.7

Day	Baking		Bread Formulations		
	Temperatures (°C)		A	B	D
1	160	190	A	B	D
2	160	190	A	C	D
3	160	190	A	B	C
4	160	175	A	C	D
5	160	175	B	C	D
6	175	190	A	B	D
7	175	190	A	B	C
8	175	190	B	C	D
9	160	175	A	B	D
10	160	190	B	C	D
11	175	190	A	C	D
12	160	175	A	B	C

- 1) Identify the different sizes of experimental units (draw a diagram).
- 2) What are the design and treatment structures for each of the experimental units you identified in part 1?
- 3) Write out an analysis of variance table for each size of experimental unit.
- 4) Write out a model to describe data from this experiment and write out the analysis of variance table.

6

Matrix Form of the Model

Summation notation becomes very laborious and sometimes nearly impossible to use when one works with unbalanced fixed effects models, random effects models, or mixed effects models. This problem can be solved by using a matrix form representation of the model. This chapter contains a discussion of the construction of the matrix form of the model and a description of how to use the matrices to obtain least squares estimators, to test hypotheses, to compute least squares or population marginal means, and to construct confidence intervals. The concept of estimability is discussed in Section 6.3.

6.1 Basic Notation

The matrix form of a model can be expressed as

$$\underset{n \times 1}{\mathbf{y}} = \underset{n \times p}{\mathbf{X}} \underset{p \times 1}{\boldsymbol{\beta}} + \underset{n \times 1}{\boldsymbol{\varepsilon}} \quad (6.1)$$

where \mathbf{y} denotes an $n \times 1$ vector of observations, \mathbf{X} denotes an $n \times p$ matrix of known constants, called the *design matrix*, $\boldsymbol{\beta}$ denotes a $p \times 1$ vector of unknown parameters, and $\boldsymbol{\varepsilon}$ denotes an $n \times 1$ vector of unobserved errors. The model for i th observation (the i th element of \mathbf{y}) is of the form

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_{p-1} x_{ip-1} + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (6.2)$$

The vectors and matrices used to represent model (6.2) as a matrix model (6.1) are

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p-1} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np-1} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{p-1} \end{bmatrix}, \quad \text{and} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (6.3)$$

Matrices of the type in Equation 6.3 can be used to represent many model types, including such design models as one-way models, two-way models, factorial models, and fractional factorial models as well as regression models, analysis of covariance models, random effects models, mixed effects models, split-plot models, repeated measures models, and random coefficient regression models by specifying the appropriate elements for \mathbf{X} and the appropriate assumptions on $\boldsymbol{\beta}$ and $\boldsymbol{\varepsilon}$. The following sections present some matrix models for various experimental situations.

6.1.1 Simple Linear Regression Model

The simple linear regression model can be expressed as, $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, $i = 1, 2, \dots, n$ and can be represented in matrix form as

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

The column of 1s in \mathbf{X} corresponds to the intercept of the regression model β_0 and the column of x_i corresponds to the slope of the regression model.

6.1.2 One-Way Treatment Structure Model

To represent the model for a one-way treatment structure with t treatments in a completely randomized design structure with n_i observations for the i th treatment, let the independent variables x_{ij} be defined by

$$x_{kij} = \begin{cases} 0 & \text{if the } ij\text{th observation is not from the } k\text{th treatment} \\ 1 & \text{if the } ij\text{th observation is from the } k\text{th treatment} \end{cases}$$

for $i = 1, 2, \dots, t$ and $j = 1, 2, \dots, n_i$. The variable x_{kij} is called an indicator variable in that, when it takes on the value of 1, it indicates that the observation is from treatment k . When

the value of x_{kij} is equal to 0, it indicates that the observation is *not* from treatment k . The means model can be expressed as

$$y_{ij} = \mu_1 x_{1ij} + \mu_2 x_{2ij} + \cdots + \mu_t x_{tij} + \varepsilon_{ij} \quad \text{for } i = 1, 2, \dots, t \text{ and} \\ j = 1, 2, \dots, n_i, \text{ or in matrix notation}$$

$$\begin{bmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1n_1} \\ y_{21} \\ y_{22} \\ \vdots \\ y_{2n_2} \\ \vdots \\ y_{t1} \\ \vdots \\ y_{tn_t} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_t \end{bmatrix} + \varepsilon$$

The means model is generally expressed as $y_{ij} = \mu_i + \varepsilon_{ij}$ for $i = 1, 2, \dots, t$ and $j = 1, 2, \dots, n_i$. The effects model can be expressed as

$$y_{ij} = \mu + \tau_1 x_{1ij} + \tau_2 x_{2ij} + \cdots + \tau_t x_{tij} + \varepsilon_{ij} \quad \text{for } i = 1, 2, \dots, t \text{ and} \\ j = 1, 2, \dots, n_i, \text{ or in matrix notation}$$

$$\begin{bmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1n_1} \\ y_{21} \\ y_{22} \\ \vdots \\ y_{2n_2} \\ \vdots \\ y_{t1} \\ \vdots \\ y_{tn_t} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \vdots \\ \tau_t \end{bmatrix} + \varepsilon$$

The effects model is generally expressed as $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$ for $i = 1, 2, \dots, t$ and $j = 1, 2, \dots, n_i$. The difference between the means model and the effects model is that the design

matrix for the effects model contains a column of 1s for the intercept of the model μ , while the means model does not contain that column of 1s.

6.1.3 Two-Way Treatment Structure Model

A form of the model used for a two-way treatment structure in a completely randomized design structure with t row treatments and b column treatments is

$$y_{ijk} = \mu_{ij} + \varepsilon_{ijk} \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, b, \text{ and } k = 1, 2, \dots, n_{ij} \quad (6.4)$$

The model used in Equation 6.4, called the *means model*, can be represented in matrix form as

$$\begin{bmatrix} y_{111} \\ y_{112} \\ \vdots \\ y_{11n_{11}} \\ y_{121} \\ \vdots \\ y_{12n_{12}} \\ \vdots \\ y_{1b1} \\ \vdots \\ y_{1bn_{1b}} \\ y_{211} \\ \vdots \\ y_{21n_{21}} \\ \vdots \\ y_{tb1} \\ \vdots \\ y_{tn_b} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \end{bmatrix} + \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \vdots \\ \mu_{1b} \\ \mu_{21} \\ \vdots \\ \mu_{tb} \end{bmatrix} + \varepsilon$$

The two-way effects model can be expressed as

$$y_{ijk} = \mu + \tau_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk} \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, b, \quad k = 1, 2, \dots, n_{ij} \quad (6.5)$$

The matrix form of the two-way effects model is:

$$\begin{bmatrix} y_{111} \\ y_{112} \\ \vdots \\ y_{11n_{11}} \\ y_{121} \\ \vdots \\ y_{12n_{12}} \\ \vdots \\ y_{1b1} \\ \vdots \\ y_{1bn_{1b}} \\ y_{211} \\ \vdots \\ y_{21n_{21}} \\ \vdots \\ y_{tb1} \\ \vdots \\ y_{tbn_{tb}} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots \\ 1 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots \\ 1 & 1 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots \\ 1 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & \vdots & 1 & 0 & \cdots & 0 \\ \vdots & \vdots \\ 1 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots \\ 1 & 0 & 1 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots \\ 1 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \\ \vdots & \vdots \\ 1 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \vdots \\ \tau_t \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_b \\ \gamma_{11} \\ \gamma_{12} \\ \vdots \\ \gamma_{1b} \\ \gamma_{21} \\ \vdots \\ \gamma_{tb} \end{bmatrix} + \varepsilon$$

This book emphasizes models that correspond to experimental design situations rather than purely regression situations. The next two examples demonstrate how to construct such models from the data structure.

6.1.4 Example 6.1: Means Model for Two-Way Treatment Structure

The information in Table 6.1 represents data from a two-way treatment structure in a completely randomized design structure where there are three row treatments and three

TABLE 6.1

Data for a Two-Way Treatment Structure in a CRD Structure

Row Treatment	Column Treatment		
	1	2	3
1	3, 6	9	10
2	2	5, 3	8
3	4	2	6

column treatments and one or two observations per cell. The matrix form of the *means model* for the data in Table 6.1 is:

$$\begin{bmatrix} 3 \\ 6 \\ 9 \\ 10 \\ 2 \\ 5 \\ 3 \\ 8 \\ 4 \\ 2 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \\ \mu_{21} \\ \mu_{22} \\ \mu_{23} \\ \mu_{31} \\ \mu_{32} \\ \mu_{33} \end{bmatrix} + \boldsymbol{\varepsilon}$$

The matrix form of the effects model for the data in Table 6.1 is

$$\begin{bmatrix} 3 \\ 6 \\ 9 \\ 10 \\ 2 \\ 5 \\ 3 \\ 8 \\ 4 \\ 2 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \gamma_{11} \\ \gamma_{12} \\ \gamma_{13} \\ \gamma_{21} \\ \gamma_{22} \\ \gamma_{23} \\ \gamma_{31} \\ \gamma_{32} \\ \gamma_{33} \end{bmatrix} + \boldsymbol{\varepsilon}$$

or

$$\mathbf{y} = \mathbf{j}\boldsymbol{\mu} + \mathbf{X}_1\boldsymbol{\tau} + \mathbf{X}_2\boldsymbol{\beta} + \mathbf{X}_3\boldsymbol{\gamma} + \boldsymbol{\varepsilon}$$

where \mathbf{j} is an 11×1 vector of ones corresponding to the first column of the above design matrix, \mathbf{X}_1 is an 11×3 matrix corresponding to columns 2–4, \mathbf{X}_2 is an 11×3 matrix corresponding to columns 5–7, and \mathbf{X}_3 is an 11×9 matrix corresponding to the last nine columns of the above design matrix.

The design matrices for other treatment and design structures are constructed in a similar fashion. Fortunately, most software that fits models to unbalanced data structures use the above types of representations and automatically generate the necessary columns of the design matrix when one specifies the categorical effects in the model.

6.2 Least Squares Estimation

Once the model is specified in matrix form, the next step in the analysis is to obtain the least squares estimator for the parameter vector $\boldsymbol{\beta}$. The method of least squares can be used to estimate the parameters of the model. To use this method, assume that the model can be expressed as

$$y_i = f(x_i; \boldsymbol{\beta}) + \varepsilon_i \quad \text{for } i = 1, 2, \dots, n \quad (6.6)$$

where $f(x_i; \boldsymbol{\beta})$ is a function of the vector of design variables indicated by x_i and depends on the parameter vector $\boldsymbol{\beta}$. The least squares estimator of $\boldsymbol{\beta}$ is the value of $\boldsymbol{\beta}$, usually denoted by $\hat{\boldsymbol{\beta}}$, that minimizes the sum of squares

$$SS(\boldsymbol{\beta}) = \sum_{i=1}^n [y_i - f(x_i; \boldsymbol{\beta})]^2 \quad (6.7)$$

If, in addition to assuming the model is of the form (6.6), one assumes that $\varepsilon_i \sim i.i.d. N(0, \sigma^2)$, $i = 1, 2, \dots, n$, then the least squares estimate of $\boldsymbol{\beta}$ is also a maximum likelihood estimator.

For example, the model function for the means model of a one-way treatment structure in a completely randomized design is

$$f(x_{ij}; \boldsymbol{\beta}) = \mu_i, \quad i = 1, 2, \dots, t; \quad j = 1, 2, \dots, n_i$$

The least squares estimators of the μ_i are the values, say $\hat{\mu}_1, \dots, \hat{\mu}_t$, that minimize

$$SS(\boldsymbol{\mu}) = \sum_{i=1}^t \sum_{j=1}^{n_i} (y_{ij} - \mu_i)^2$$

The model function for the means model of a two-way treatment structure in a completely randomized design is

$$f(x_{ijk}; \boldsymbol{\beta}) = \mu_{ij}, \quad i = 1, 2, \dots, t; \quad j = 1, 2, \dots, b; \quad k = 1, 2, \dots, n_{ij}$$

The least squares estimators of the μ_{ij} are the values, say $\hat{\mu}_{11}, \dots, \hat{\mu}_{tb}$ that minimize

$$SS(\boldsymbol{\mu}) = \sum_{i=1}^t \sum_{j=1}^b \sum_{k=1}^{n_{ij}} (y_{ijk} - \mu_{ij})^2$$

The model function for the effects model of a two-way treatment structure in a completely randomized design is

$$f(x_{ijk}; \boldsymbol{\beta}) = \mu + \tau_i + \beta_j + \gamma_{ij} \quad i = 1, 2, \dots, t; j = 1, 2, \dots, b; k = 1, 2, \dots, n_{ij}$$

The least squares estimators of μ , τ_i , β_j , and γ_{ij} are obtained by minimizing

$$SS(\mu, \tau_i, \beta_j, \gamma_{ij}) = \sum_{i=1}^t \sum_{j=1}^b \sum_{k=1}^{n_{ij}} (y_{ijk} - \mu - \tau_i - \beta_j - \gamma_{ij})^2$$

In general, the model can be written in a matrix form like that in Equation 6.1, and the least squares estimator of $\boldsymbol{\beta}$ is the value $\hat{\boldsymbol{\beta}}$ that minimizes the sum of squares

$$SS(\boldsymbol{\beta}) = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \quad (6.8)$$

6.2.1 Least Squares Equations

Matrix representations and calculus can be used to determine the value of $\boldsymbol{\beta}$ that minimizes the residual sum of squares. When one carries out the minimization, a set of equations are obtained which $\hat{\boldsymbol{\beta}}$ must satisfy and those equations are called the *least squares equations* or the *normal equations* of the model. The normal equations for model (6.1) are given by

$$\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y} \quad (6.9)$$

Any vector $\hat{\boldsymbol{\beta}}$ that satisfies the normal equations is a least squares estimator of $\boldsymbol{\beta}$. The least squares estimator need not be unique for some models. To help the reader become more familiar with the normal equations, the normal equations for the models discussed in Section 6.1 are provided below.

The normal equations for the one-way means model are

$$\begin{bmatrix} n_1 & 0 & \cdots & 0 \\ 0 & n_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & n_t \end{bmatrix} \begin{bmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \\ \vdots \\ \hat{\mu}_t \end{bmatrix} = \begin{bmatrix} y_{1\cdot} \\ y_{2\cdot} \\ \vdots \\ y_{t\cdot} \end{bmatrix} \quad \text{where } y_{i\cdot} = \sum_{j=1}^{n_i} y_{ij}$$

The normal equations for the one-way effects model are

$$\begin{bmatrix} n_\cdot & n_1 & n_2 & \cdots & n_t \\ n_1 & n_1 & 0 & \cdots & 0 \\ n_2 & 0 & n_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ n_t & 0 & 0 & \cdots & n_t \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\tau}_1 \\ \hat{\tau}_2 \\ \vdots \\ \hat{\tau}_t \end{bmatrix} = \begin{bmatrix} y_{\cdot\cdot} \\ y_{1\cdot} \\ y_{2\cdot} \\ \vdots \\ y_{t\cdot} \end{bmatrix} \quad \text{where } y_{\cdot\cdot} = \sum_{i=1}^t \sum_{j=1}^{n_i} y_{ij} \text{ and } n_\cdot = \sum_{i=1}^t n_i$$

The normal equations for the two-way means model using the data in Table 6.1 are

$$\begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\mu}_{11} \\ \hat{\mu}_{12} \\ \hat{\mu}_{13} \\ \hat{\mu}_{21} \\ \hat{\mu}_{22} \\ \hat{\mu}_{23} \\ \hat{\mu}_{31} \\ \hat{\mu}_{32} \\ \hat{\mu}_{33} \end{bmatrix} = \begin{bmatrix} y_{11\cdot} \\ y_{12\cdot} \\ y_{13\cdot} \\ y_{21\cdot} \\ y_{22\cdot} \\ y_{23\cdot} \\ y_{31\cdot} \\ y_{32\cdot} \\ y_{33\cdot} \end{bmatrix}$$

where

$$y_{ij\cdot} = \sum_{k=1}^{n_{ij}} y_{ijk}$$

The normal equations for the two-way effects model corresponding to the data in Table 6.1 are

$$\begin{bmatrix} 11 & 4 & 4 & 3 & 4 & 4 & 3 & 2 & 1 & 1 & 1 & 2 & 1 & 1 & 1 & 1 \\ 4 & 4 & 0 & 0 & 2 & 1 & 1 & 2 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 0 & 4 & 0 & 1 & 2 & 1 & 0 & 0 & 0 & 1 & 2 & 1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 3 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 4 & 2 & 1 & 1 & 4 & 0 & 0 & 2 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 4 & 1 & 2 & 1 & 0 & 4 & 0 & 0 & 1 & 0 & 0 & 2 & 0 & 0 & 1 & 0 \\ 3 & 1 & 1 & 1 & 0 & 0 & 3 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 2 & 2 & 0 & 0 & 2 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 2 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \gamma_{11} \\ \gamma_{12} \\ \gamma_{13} \\ \gamma_{21} \\ \gamma_{22} \\ \gamma_{23} \\ \gamma_{31} \\ \gamma_{32} \\ \gamma_{33} \end{bmatrix} = \begin{bmatrix} y_{\dots} \\ y_{1\cdot\cdot} \\ y_{2\cdot\cdot} \\ y_{3\cdot\cdot} \\ y_{\cdot 1\cdot} \\ y_{\cdot 2\cdot} \\ y_{\cdot 3\cdot} \\ y_{11\cdot} \\ y_{12\cdot} \\ y_{13\cdot} \\ y_{21\cdot} \\ y_{22\cdot} \\ y_{23\cdot} \\ y_{31\cdot} \\ y_{32\cdot} \\ y_{33\cdot} \end{bmatrix}$$

where

$$y_{\dots} = \sum_{i=1}^t \sum_{j=1}^b \sum_{k=1}^{n_{ij}} y_{ijk}, \quad y_{i\cdot\cdot} = \sum_{j=1}^b \sum_{k=1}^{n_{ij}} y_{ijk}, \quad y_{\cdot j\cdot} = \sum_{i=1}^t \sum_{k=1}^{n_{ij}} y_{ijk}, \quad \text{and} \quad y_{ij\cdot} = \sum_{k=1}^{n_{ij}} y_{ijk}$$

When the $X'X$ matrix is of full rank (Graybill, 1976), that is $X'X$ is nonsingular, then the inverse of $X'X$ exists and the least squares estimator for β (the solution for $\hat{\beta}$ in Equation 6.9) is

$$\hat{\beta} = (X'X)^{-1} X'y \quad (6.10)$$

When $X'X$ is of full rank, the least squares estimator is unique. Computing the inverse of $X'X$ is generally not an easy task. One of the most important aspects of the development of computing software is that now statisticians can invert very large matrices that would not have been attempted before computers became available. However, when there are certain patterns in $X'X$, the patterns can be exploited and the inverse can be more easily computed. For the normal equations for the one-way means model and the two-way means model $X'X$ is diagonal (all diagonal elements are nonzero and off diagonal elements are zero) and the inverse of $X'X$ is obtained by simply replacing each diagonal element by its reciprocal. Thus, the least squares estimators of the μ_i for the one-way means model are

$$\begin{bmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \\ \vdots \\ \hat{\mu}_t \end{bmatrix} = \begin{bmatrix} \frac{1}{n_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{n_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{n_t} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_t \end{bmatrix}$$

or equivalently,

$$\hat{\mu}_i = \frac{y_{i\cdot}}{n_i} = \bar{y}_{i\cdot}, \quad i = 1, 2, \dots, t$$

Similarly, the least squares estimator of μ_{ij} for the two-way means model is

$$\hat{\mu}_{ij} = \frac{y_{ij\cdot}}{n_{ij}} = \bar{y}_{ij\cdot}, \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, b$$

Unlike the normal equations in means models, the $X'X$ matrices for the effects models are *singular* and the inverses of the $X'X$ matrices do not exist. In this case there are many solutions to the normal equations (in fact an infinite number of least squares solutions). The effects models are called *overspecified* or *singular models* in that the models have more parameters than can be uniquely estimated from the data collected. Overspecified models are commonly used and there are several ways to solve their corresponding normal equations. The following discussion addresses two-way treatment structures in a completely randomized design structure, but similar techniques can be used in other factorial effects models.

Theoretically a generalized inverse can be used to solve the normal equations for $\hat{\beta}$ (Graybill, 1976), but a commonly used method for solving the normal equations of an over-specified model is to place restrictions on the parameters in the model (which in effect generates a g -inverse solution). Placing restrictions on the parameters of the model can be accomplished in many ways, two of which are considered here.

6.2.2 Sum-to-Zero Restrictions

One common technique is to require the sums of certain parameters to be equal to zero. This procedure has been used to solve normal equations from the very beginning of the analysis of variance. For the two-way effects model using the data in Table 6.1, the sum-to-zero restrictions are

$$\begin{aligned}\sum_{i=1}^3 \tau_i &= 0, \quad \sum_{j=1}^3 \beta_j = 0, \quad \sum_{i=1}^3 \gamma_{i1} = 0, \quad \sum_{i=1}^3 \gamma_{i2} = 0, \quad \sum_{i=1}^3 \gamma_{i3} = 0 \\ \sum_{j=1}^3 \gamma_{1j} &= 0, \quad \sum_{j=1}^3 \gamma_{2j} = 0, \quad \text{and} \quad \sum_{j=1}^3 \gamma_{3j} = 0\end{aligned}$$

Next, these restrictions are incorporated into the model by solving for some of the parameters in terms of others with the restrictions being taken into account, and then substituting the expressions back into the model. For example, the parameters that can be replaced are

$$\begin{aligned}\tau_3 &= -\tau_1 - \tau_2, \quad \beta_3 = -\beta_1 - \beta_2, \quad \gamma_{13} = -\gamma_{11} - \gamma_{12} \\ \gamma_{23} &= -\gamma_{21} - \gamma_{22}, \quad \gamma_{31} = -\gamma_{11} - \gamma_{21}, \quad \gamma_{32} = -\gamma_{12} - \gamma_{22} \\ \gamma_{33} &= -\gamma_{13} - \gamma_{23} = -\gamma_{31} - \gamma_{32} = \gamma_{11} + \gamma_{12} + \gamma_{21} + \gamma_{22}\end{aligned}$$

Thus, replace, τ_3 , β_3 , γ_{13} , γ_{23} , γ_{33} , γ_{31} and γ_{32} in the model to obtain a reparameterized model

$$\begin{bmatrix} y_{111} \\ y_{112} \\ y_{121} \\ y_{131} \\ y_{211} \\ y_{221} \\ y_{222} \\ y_{231} \\ y_{311} \\ y_{321} \\ y_{331} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & -1 & -1 & -1 & -1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & -1 & -1 & 0 & 0 & -1 & -1 \\ 1 & -1 & -1 & 0 & 0 & -1 & 0 & -1 & 0 \\ 1 & -1 & -1 & 0 & 1 & 0 & -1 & 0 & -1 \\ 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \mu^* \\ \tau_1^* \\ \tau_2^* \\ \beta_1^* \\ \beta_2^* \\ \gamma_{11}^* \\ \gamma_{12}^* \\ \gamma_{21}^* \\ \gamma_{22}^* \end{bmatrix} + \boldsymbol{\varepsilon}$$

which is expressed as

$$\mathbf{y} = \mathbf{X}^* \boldsymbol{\beta}^* + \boldsymbol{\varepsilon}$$

The solution to the normal equations for the data in Table 6.1 corresponding to the sum-to-zero restrictions is $\hat{\boldsymbol{\beta}}^* = (\mathbf{X}^{*\prime} \mathbf{X}^*)^{-1} \mathbf{X}^{*\prime} \mathbf{y}$. One obtains

$$\begin{aligned}\hat{\boldsymbol{\beta}}^* &= [\hat{\mu}^*, \hat{\tau}_1^*, \hat{\tau}_2^*, \hat{\beta}_1^*, \hat{\beta}_2^*, \hat{\gamma}_{11}^*, \hat{\gamma}_{12}^*, \hat{\gamma}_{21}^*, \hat{\gamma}_{22}^*] \\ &= [5.500, 2.333, -0.833, -2.000, -0.5000, -1.333, 1.667, -0.667, -0.167]\end{aligned}$$

The estimators for the remaining elements of β are obtained from the restrictions as follows:

$$\begin{aligned}\hat{\tau}_3^* &= -\hat{\tau}_1^* - \hat{\tau}_2^* = -1.500, \quad \hat{\beta}_3^* = -\hat{\beta}_1^* - \hat{\beta}_2^* = 2.500 \\ \hat{\gamma}_{13}^* &= -\hat{\gamma}_{11}^* - \hat{\gamma}_{12}^* = -0.333, \quad \hat{\gamma}_{23}^* = -\hat{\gamma}_{21}^* - \hat{\gamma}_{22}^* = 0.833 \\ \hat{\gamma}_{31}^* &= -\hat{\gamma}_{11}^* - \hat{\gamma}_{21}^* = 2.000, \quad \hat{\gamma}_{32}^* = -\hat{\gamma}_{12}^* - \hat{\gamma}_{22}^* = -1.500 \\ \hat{\gamma}_{33}^* &= \hat{\gamma}_{11}^* + \hat{\gamma}_{12}^* + \hat{\gamma}_{21}^* + \hat{\gamma}_{22}^* = -0.500\end{aligned}$$

In relation to the means model parameters, μ_{ij} , the parameters μ^* , τ_i^* , β_j^* , and γ_{ij}^* can be selected to satisfy the sum-to-zero restrictions by defining

$$\mu^* = \bar{\mu}_{..}, \quad \tau_i^* = \bar{\mu}_{i..} - \bar{\mu}_{..}, \quad \beta_j^* = \bar{\mu}_{.j} - \bar{\mu}_{..}, \quad \text{and} \quad \gamma_{ij}^* = \mu_{ij} - \bar{\mu}_{i..} - \bar{\mu}_{.j} + \bar{\mu}_{..}$$

6.2.3 Set-to-Zero Restrictions

Another reparameterization technique often used to solve the two-way effects model's normal equations uses restrictions that set the last parameter in each set equal to zero (the last parameter is selected for convenience; one could also use the first, or second, or any other). For the two-way effects model for the data in Table 6.1, the restrictions are:

$$\tau_3 = 0, \beta_3 = 0, \gamma_{13} = 0, \gamma_{23} = 0, \gamma_{33} = 0, \gamma_{31} = 0, \text{ and } \gamma_{23} = 0$$

The resulting reparameterized model obtained by incorporating the above restrictions into the two-way effects model is

$$\begin{bmatrix} y_{111} \\ y_{112} \\ y_{121} \\ y_{131} \\ y_{211} \\ y_{221} \\ y_{222} \\ y_{231} \\ y_{311} \\ y_{321} \\ y_{331} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mu^+ \\ \tau_1^+ \\ \tau_2^+ \\ \beta_1^+ \\ \beta_2^+ \\ \gamma_{11}^+ \\ \gamma_{12}^+ \\ \gamma_{21}^+ \\ \gamma_{22}^+ \end{bmatrix} + \varepsilon$$

which can be expressed as $\mathbf{y} = \mathbf{X}^+ \boldsymbol{\beta}^+ + \varepsilon$. The matrix \mathbf{X}^+ is obtained from the full design matrix of the two-way effects model, \mathbf{X} , by deleting the columns corresponding to τ_3 , β_3 , γ_{31} , γ_{32} , γ_{33} , γ_{13} , and γ_{23} . The process of obtaining the design matrix for the reparameterized model corresponding to the set-to-zero restrictions is much simpler than obtaining the design matrix for the sum-to-zero restrictions. The solution to the normal equations corresponding to the set-to-zero restrictions is

$$\hat{\boldsymbol{\beta}}^+ = (\mathbf{X}^{+\prime} \mathbf{X}^+)^{-1} \mathbf{X}^{+\prime} \mathbf{y}$$

One obtains

$$\begin{aligned}\hat{\beta}^{+*} &= [\hat{\mu}^+, \hat{\tau}_1^+, \hat{\tau}_2^+, \hat{\beta}_1^+, \hat{\beta}_2^+, \hat{\gamma}_{11}^+, \hat{\gamma}_{12}^+, \hat{\gamma}_{21}^+, \hat{\gamma}_{22}^+] \\ &= [6.0, 4.0, 2.0, -2.0, -4.0, -3.5, 3.0, -4.0, 0.0]\end{aligned}$$

The estimates of the remaining parameters are zero since they were specified by the set-to-zero restrictions, that is,

$$\hat{\tau}_3^+ = \hat{\beta}_3^+ = \hat{\gamma}_{31}^+ = \hat{\gamma}_{32}^+ = \hat{\gamma}_{33}^+ = \hat{\gamma}_{13}^+ = \hat{\gamma}_{23}^+ = 0$$

To relate the set-to-zero restrictions to the mean model parameters, μ_{ij} , define

$$\begin{aligned}\mu^+ &= \mu_{tb}, \quad \tau_i^+ = \mu_{ib} - \mu_{tb} \quad \beta_j^+ = \mu_{tj} - \mu_{tb} \quad \text{and} \quad \gamma_{ij}^+ = \mu_{ij} - \mu_{tj} - \mu_{ib} + \mu_{tb}\end{aligned}$$

Thus there are several possible solutions to the normal equations when $X'X$ is not of full rank (that is, when $X'X$ is singular). This occurs because the model is overparameterized; that is, there are more parameters in the model (16 in the case of the two-way effects model) than can be uniquely estimated from the available data (there are nine cells of data, so one can estimate at most nine parameters). The number of parameters that can be estimated uniquely might be called the number of *essential* parameters. To cope with the overparameterized model and nonunique least squares solutions, the concept of estimability must be considered, which is the topic of Section 6.3. The next example presents the two possible solutions for the effects model for a one-way treatment structure.

6.2.4 Example 6.2: A One-Way Treatment Structure

This is an example of a one-way treatment structure with four treatments in a completely randomized design structure. The data are shown in Table 6.2. The X^* matrix is constructed by reparameterizing the model by using the sum-to-zero restrictions; that is, it is assumed that $\tau_1^* + \tau_2^* + \tau_3^* + \tau_4^* = 0$.

TABLE 6.2

Data for One-Way Treatment Structure for Means and Effects Models in Section 6.2.4

Treatment 1	Treatment 2	Treatment 3	Treatment 4
2.2	2.4	1.8	1.9
2.4	2.6	1.7	2.0
2.5	3.0	1.8	2.3
2.3	3.1	1.6	2.1
2.0	2.5	1.4	1.9
1.9		1.9	2.0
1.9			2.4
1.9			1.5

The resulting design matrix is:

$$\mathbf{X}^* = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{bmatrix}$$

The normal equations for the sum-to-zero restriction model are

$$\begin{bmatrix} 26 & -1 & -3 & -2 \\ -1 & 15 & 8 & 8 \\ -3 & 8 & 13 & 8 \\ -2 & 8 & 8 & 14 \end{bmatrix} \begin{bmatrix} \hat{\mu}^* \\ \hat{\tau}_1^* \\ \hat{\tau}_2^* \\ \hat{\tau}_3^* \end{bmatrix} = \begin{bmatrix} 55.1 \\ -0.9 \\ -2.5 \\ -5.9 \end{bmatrix}$$

The corresponding least squares solution to the sum-to-zero restriction normal equations is

$$\hat{\beta}^* = \begin{bmatrix} \hat{\mu}^* \\ \hat{\tau}_1^* \\ \hat{\tau}_2^* \\ \hat{\tau}_3^* \end{bmatrix} = \begin{bmatrix} 2.1510 \\ 0.0204 \\ 0.5690 \\ -0.4510 \end{bmatrix}$$

and $\hat{\tau}_4^* = -\hat{\tau}_1^* - \hat{\tau}_2^* - \hat{\tau}_3^* = -0.1384$. In terms of the means model parameters, the sum-to-zero restriction parameters can be expressed as

$$\mu^* = \bar{\mu}, \quad \tau_1^* = \mu_1 - \bar{\mu}, \quad \tau_2^* = \mu_2 - \bar{\mu}, \quad \tau_3^* = \mu_3 - \bar{\mu}, \quad \tau_4^* = \mu_4 - \bar{\mu}.$$

The set-to-zero restriction design matrix, X^+ , with $\tau_4^+ = 0$, is constructed from X^* by replacing the “-1” values with “0” values. The normal equations for the set-to-zero restriction model are

$$\begin{bmatrix} 26 & 7 & 5 & 6 \\ 7 & 7 & 0 & 0 \\ 5 & 0 & 5 & 0 \\ 6 & 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} \hat{\mu}^+ \\ \hat{\tau}_1^+ \\ \hat{\tau}_2^+ \\ \hat{\tau}_3^+ \end{bmatrix} = \begin{bmatrix} 55.1 \\ 15.2 \\ 13.6 \\ 10.2 \end{bmatrix}$$

The least squares solution for the set-to-zero restriction model is

$$\hat{\beta}^+ = \begin{bmatrix} \hat{\mu}^+ \\ \hat{\tau}_1^+ \\ \hat{\tau}_2^+ \\ \hat{\tau}_3^+ \end{bmatrix} = \begin{bmatrix} 2.0125 \\ 0.1589 \\ 0.7075 \\ -0.3125 \end{bmatrix}$$

and $\hat{\tau}_4^+ = 0$. In terms of the means model parameters, the set-to-zero restriction parameters can be expressed as

$$\mu^+ = \mu_4, \quad \tau_1^+ = \mu_1 - \mu_4, \quad \tau_2^+ = \mu_2 - \mu_4, \quad \tau_3^+ = \mu_3 - \mu_4, \quad \tau_4^+ = \mu_4 - \mu_4 = 0$$

The last parameter that needs to be estimated is the population variance σ^2 . An estimate of σ^2 based on the least squares solution for β is

$$\begin{aligned} \hat{\sigma}^2 &= \frac{1}{n-r}(y - X\hat{\beta})'(y - X\hat{\beta}) \\ &= \frac{1}{n-r} \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - \cdots - \hat{\beta}_{p-1} x_{ip-1})^2 \end{aligned} \quad (6.11)$$

where $r = \text{rank}(X)$.

If the errors are assumed to be independently distributed with the first four moments equal to the first four moments of a normal distribution, then $\hat{\sigma}^2$ is the best quadratic unbiased estimate of σ^2 . If the errors are also normally distributed, then $\hat{\sigma}^2$ is the best unbiased estimator of σ^2 and the sampling distribution of $(n-r)\hat{\sigma}^2/\sigma^2$ is a central chi-square distribution with $n-r$ degrees of freedom.

6.3 Estimability and Connected Designs

When an overspecified model or a less than full rank model is used for an experimental situation, there are many different least squares solutions (in fact there are an infinite number of

solutions). If two researchers analyzed the above two data sets, one using the sum-to-zero restriction and the other the set-to-zero restriction, they might appear to obtain different conclusions. For the two-way example in the last section, $\hat{\tau}_2^* = -0.833$ and $\hat{\tau}_2^+ = 2.000$; thus one researcher might say that τ_2 is most likely to be negative while the other might say that τ_2 is likely to be positive—and both statements are incorrect.

6.3.1 Estimable Functions

Since both researchers are analyzing the same data set, it seems that they should consider only parameters or functions of the parameters that have identical estimates for both reparameterized models. Such functions of the parameters are called *estimable functions of the parameters*.

Definition 6.3.1: A parameter or function of the parameters $f(\boldsymbol{\beta})$ is *estimable* if and only if the estimate of the parameter or function of parameters is invariant with respect to the choice of a least squares solution; that is, the value of the estimate is the same regardless of which solution to the normal equations is used.

If two researchers obtain an estimate of an estimable function of the parameters, they both will obtain the same value even if they have two different least squares solutions. Therefore, they will make the same decisions about estimable functions of the parameters. For matrix models, linear estimable functions of $\boldsymbol{\beta}$ take on the form of linear combinations of the parameter vector such as $\mathbf{a}'\boldsymbol{\beta}$ where \mathbf{a} is a $p \times 1$ vector of constants. A linear function $\mathbf{a}'\boldsymbol{\beta}$ is estimable if and only if there exists a vector \mathbf{r} such that $\mathbf{a} = \mathbf{X}'\mathbf{X}\mathbf{r}$. Each function $\mathbf{x}_i'\boldsymbol{\beta}$ is estimable where \mathbf{x}_i is the i th row of \mathbf{X} . Also, any linear combination of the $\mathbf{x}_i'\boldsymbol{\beta}$'s is an estimable function. Consider the two solutions obtained for the one-way example in Section 6.2. Because there are two different solutions for each of the parameters μ , τ_1 , τ_2 , τ_3 , and τ_4 , these parameters are considered to be nonestimable, but by computing the estimate $\mu + \tau_i$ from each method, it is seen that

$$\hat{\mu}^* + \hat{\tau}_i^* = \hat{\mu}^+ + \hat{\tau}_i^+, \quad i = 1, 2, 3, 4$$

demonstrating that $\mu + \tau_i$ is an estimable function of the parameters.

All contrasts of the τ_i , such as the differences $\tau_1 - \tau_2$, $\tau_1 - \tau_3$, $\tau_2 - \tau_3$, or $\sum_{i=1}^t c_i \tau_i$ where $\sum_{i=1}^t c_i = 0$ can be shown to be estimable functions for the one-way model.

For the two-way effects model, some estimable functions are

$$\mu + \tau_i + \beta_j + \gamma_{ij}, \quad \gamma_{ij} - \gamma_{ij'} - \gamma_{i'j} + \gamma_{i'j'}, \quad \beta_j - \beta_{j'} + \bar{\gamma}_{.j} - \bar{\gamma}_{.j'}, \quad \tau_i - \tau_{i'} + \bar{\gamma}_{i.} - \bar{\gamma}_{i'}$$

Estimable functions are discussed in more detail in Chapter 10. The important thing to remember here is the definition of an estimable function. In making inferences from a data set, one must consider only functions of the parameters that are estimable, since they are functions of the parameters with estimates that do not depend on which least squares solution is chosen. SAS®-GLM and SAS®-Mixed both check to make sure that a parameter or function of the parameters that is being requested to be estimated is in fact estimable. If the parameter or function of the parameters is not estimable, no estimate is provided.

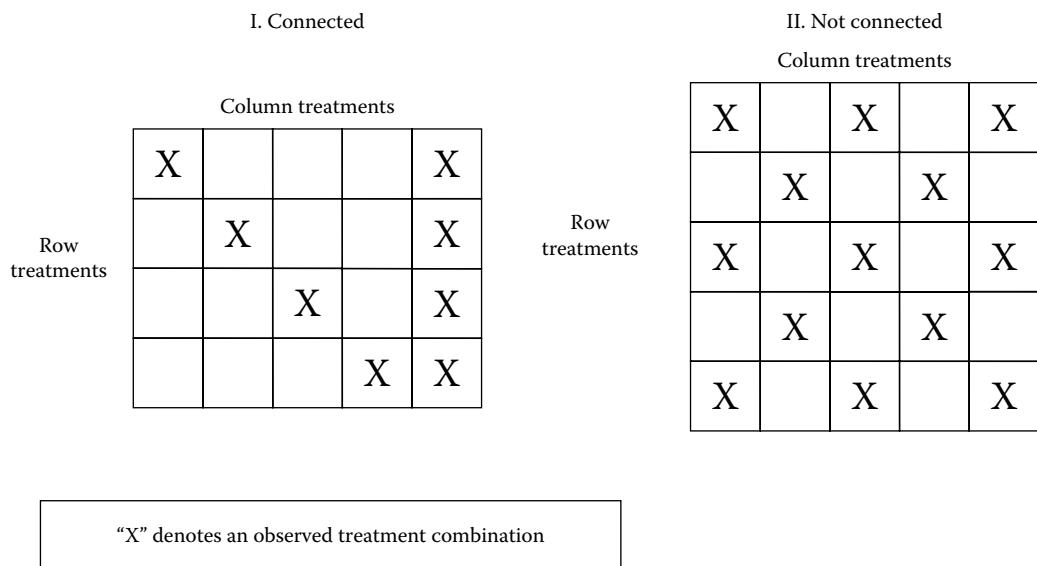


FIGURE 6.1 Connected and unconnected two-way treatment structures.

6.3.2 Connectedness

Another concept related to estimable functions is that of connectedness of a two-way treatment structure. If it can be assumed that the levels of the row and column treatments do not interact, then the treatment combination means can be modeled as

$$\mu_{ij} = \mu + \tau_i + \beta_j, \quad i = 1, 2, \dots, b, \quad j = 1, 2, \dots, t \quad (6.12)$$

A two-way treatment structure is said to be connected if and only if data occur in the two-way cells in such a way that $\beta_j - \beta_{j'}$ and $\tau_i - \tau_{i'}$ are estimable for all $(j \neq j')$ and $(i \neq i')$ for model (6.12). Arrangement I in Figure 6.1 is a connected experiment, while arrangement II is not. For example, using arrangement I,

$$\beta_1 - \beta_2 = (\mu + \tau_1 + \beta_1) - (\mu + \tau_1 + \beta_2) + (\mu + \tau_2 + \beta_5) - (\mu + \tau_2 + \beta_2)$$

a linear combination of the cell means, thus $\beta_1 - \beta_2$ is estimable. No such linear combination of the cell means for arrangement II provides $\beta_1 - \beta_2$, thus $\beta_1 - \beta_2$ is not estimable. The next section discusses the testing of hypotheses about estimable functions of the parameters.

6.4 Testing Hypotheses about Linear Model Parameters

There are several ways to develop appropriate statistics for testing hypotheses about linear functions of the parameters of a linear model. The method used here, which is expressed

in matrix notation, is equivalent to the principle of conditional error (Chapter 1) and the likelihood ratio statistic. The discussion is limited to testing hypotheses about estimable functions of the parameters. In particular, consider testing the hypothesis

$$H_0: \mathbf{H}\boldsymbol{\beta} = \mathbf{h} \text{ vs } H_a: \mathbf{H}\boldsymbol{\beta} \neq \mathbf{h} \quad (6.13)$$

where the linear combinations $\mathbf{H}\boldsymbol{\beta}$ are estimable functions of $\boldsymbol{\beta}$ and \mathbf{H} is a $q \times p$ matrix of rank q (that is, all of the rows of \mathbf{H} are linearly independent). The corresponding test statistic is

$$F_c = \frac{SSH_0/q}{\hat{\sigma}^2} \quad (6.14)$$

where $\hat{\sigma}^2$ was given by Equation 6.11 and

$$SSH_0 = (\mathbf{H}\hat{\boldsymbol{\beta}} - \mathbf{h})' [\mathbf{H}(\mathbf{X}'\mathbf{X})^{-} \mathbf{H}']^{-1} (\mathbf{H}\hat{\boldsymbol{\beta}} - \mathbf{h}) \quad (6.15)$$

which is called the *sum of squares due to deviations from the null hypothesis* [the notation “ $(\mathbf{X}'\mathbf{X})^{-}$ ” denotes a generalized inverse of the matrix $\mathbf{X}'\mathbf{X}$, Graybill, 1976]. Under the assumption that the elements of the error vector are *i.i.d.* $N(0, \sigma^2)$, F_c is distributed as an F -distribution with q and $n - r$ degrees of freedom.

The hypothesis in Equation 6.13 can always be equivalently stated in terms of a reparameterized model $\mathbf{y} = \mathbf{X}^*\boldsymbol{\beta}^* + \boldsymbol{\varepsilon}$ where $\mathbf{X}^*\mathbf{X}^*$ is nonsingular, as

$$H_0: \mathbf{H}^*\boldsymbol{\beta}^* = \mathbf{h}^* \text{ vs } H_a: \mathbf{H}^*\boldsymbol{\beta}^* \neq \mathbf{h}^*$$

Then the SSH_0 of Equation 6.15 can be computed as

$$SSH_0 = (\mathbf{H}^*\hat{\boldsymbol{\beta}}^* - \mathbf{h}^*)' [\mathbf{H}^*(\mathbf{X}^*\mathbf{X}^*)^{-1}(\mathbf{H}^{*\prime}\hat{\boldsymbol{\beta}}^* - \mathbf{h}^*)]$$

For a one-way model, testing the hypothesis

$$H_0: \tau_1 = \tau_2 = \cdots = \tau_t \text{ vs } H_a: \tau_i \neq \tau_{i'} \quad \text{for some } i \neq i'$$

in the original effects model is equivalent to testing

$$H_0: \tau_1^* = \tau_2^* = \cdots = \tau_{t-1}^* = 0 \text{ vs } H_a: \tau_i^* \neq 0 \quad \text{for some } i \leq t-1$$

in the sum-to-zero reparameterized model. The null hypothesis in terms of $\boldsymbol{\beta}^*$ for Example 6.2 is

$$H_0: \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu^* \\ \tau_1^* \\ \tau_2^* \\ \tau_3^* \end{bmatrix} = 0 \quad \text{or} \quad \mathbf{H}^*\boldsymbol{\beta}^* = 0$$

and the sum of squares due to deviations from H_0 : is

$$SSH_0 = (\hat{\tau}_1^* \quad \hat{\tau}_1^* \quad \hat{\tau}_1^*) \mathbf{Z}^{-1} \begin{bmatrix} \hat{\tau}_1^* \\ \hat{\tau}_2^* \\ \hat{\tau}_3^* \end{bmatrix}$$

where $\mathbf{Z} = \mathbf{H}^*(\mathbf{X}^{*\prime}\mathbf{X}^*)^{-1}\mathbf{H}^*$, which is the portion of $(\mathbf{X}^{*\prime}\mathbf{X}^*)^{-1}$ corresponding to the rows and columns associated with $(\tau_1^*, \tau_2^*, \tau_3^*)$.

A $(1 - \alpha)100\%$ confidence interval about an estimable function $\mathbf{a}'\boldsymbol{\beta}$ is

$$\mathbf{a}'\hat{\boldsymbol{\beta}} - [t_{\alpha/2, n-p}]S_{\mathbf{a}'\hat{\boldsymbol{\beta}}} \leq \mathbf{a}'\boldsymbol{\beta} \leq \mathbf{a}'\hat{\boldsymbol{\beta}} + [t_{\alpha/2, n-p}]S_{\mathbf{a}'\hat{\boldsymbol{\beta}}}$$

where $S_{\mathbf{a}'\hat{\boldsymbol{\beta}}}^2 = \hat{\sigma}^2 \mathbf{a}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{a}$. Simultaneous confidence intervals can be constructed about several estimable functions by using one of the multiple comparison procedures discussed in Chapter 3.

6.5 Population Marginal Means

After analyzing a cross-classified data set via an analysis of variance, the experimenter is usually interested in estimating the means of particular effects or cells. The *population marginal mean* is defined to be a linear combination of the parameters averaged over specified classes as if there were one observation in each cell (Searle et al., 1980). If every cell has at least one observation, then all population marginal means are estimable, whereas they are not necessarily estimable if some cells are empty. This definition does not depend on the sample sizes in the cells. If the data represent a proportional sampling of cells, then the experimenter might want to consider a weighted average of the cell means where the weights are given by the sample sizes (see Chapter 10). Or there could be some other weighting scheme that needs to be used to average over the various cells. An example is presented at the end of this section to demonstrate different possibilities.

For a one-way treatment structure in a completely randomized design, the population marginal mean for the i th treatment is $\mu + \tau_i = \mu_i$ and is estimated by

$$\widehat{\mu + \tau_i} = \hat{\mu} + \hat{\tau}_i$$

where $\hat{\mu}$ and $\hat{\tau}_i$ are obtained from any solution to the normal equations. These estimated values are called *estimated population marginal means*.

For a two-way treatment structure in a completely randomized design, the population marginal mean for the (i, j) th cell is $\mu_{ij} = \mu + \tau_i + \beta_j + \gamma_{ij}$. The population marginal mean for row i is the average of the μ_{ij} in that row, or

$$\bar{\mu}_{i \cdot} = \sum_{j=1}^b \frac{\mu_{ij}}{b} = \mu + \tau_i + \bar{\beta}_{\cdot} + \bar{\gamma}_{i \cdot}$$

The population marginal mean for column j is the average of the μ_{ij} in that column, or

$$\bar{\mu}_{\cdot j} = \sum_{i=1}^t \frac{\mu_{ij}}{t} = \mu + \bar{\tau}_j + \beta_j + \bar{\gamma}_j,$$

The estimates of the population marginal means are

$$\hat{\mu}_{ij} = \hat{\mu} + \hat{\tau}_i + \hat{\beta}_j + \hat{\gamma}_{ij},$$

$$\hat{\mu}_{\cdot j} = \hat{\mu} + \hat{\tau}_i + \sum_{j=1}^b \frac{\hat{\beta}_j}{b} + \sum_{j=1}^b \frac{\hat{\gamma}_{ij}}{b},$$

and

$$\hat{\mu}_{\cdot j} = \hat{\mu} + \sum_{i=1}^t \frac{\hat{\tau}_i}{t} + \hat{\beta}_j + \sum_{i=1}^t \frac{\hat{\gamma}_{ij}}{t}, \text{ respectively.}$$

These estimates are unique for any of the possible least squares solutions

$$\hat{\mu}, \hat{\tau}_1, \hat{\tau}_2, \dots, \hat{\tau}_t, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_b, \hat{\gamma}_{11}, \hat{\gamma}_{12}, \dots, \hat{\gamma}_{tb}$$

The estimate of the population marginal mean $\bar{\mu}_1$, for the two-way example in Table 6.1, computed from the sum-to-zero restricted model, is

$$\begin{aligned}\hat{\mu}_1 &= \hat{\mu}^* + \hat{\tau}_1^* + \frac{\hat{\beta}_1^* + \hat{\beta}_2^* + \hat{\beta}_3^*}{3} + \frac{\hat{\gamma}_{11}^* + \hat{\gamma}_{12}^* + \hat{\gamma}_{13}^*}{3} \\ &= 5.500 + 2.333 + \frac{-2.000 - 0.500 + 2.500}{3} + \frac{-1.333 + 1.667 - 0.333}{3} \\ &= 7.833\end{aligned}$$

One obtains the same value for $\hat{\mu}_1$ when the set-to-zero least squares solution is used as expected, since $\bar{\mu}_1$ is estimable for this example.

When there are no empty cells, then all population marginal means are estimable. If there are empty cells, then any population marginal mean involving one or more of the missing cells is not estimable. For example, if the (2, 2) cell is missing in a 2×2 treatment structure, there is no information about μ_{22} and hence μ_{22} is not estimable. The population marginal mean for column 2 is

$$\bar{\mu}_{\cdot 2} = \frac{\mu_{12} + \mu_{22}}{2}$$

Because $\bar{\mu}_{\cdot 2}$ depends on μ_{22} , it follows that $\bar{\mu}_{\cdot 2}$ is not estimable ($\bar{\mu}_{\cdot 2}$ is not estimable either).

Any population marginal mean that is estimable can be expressed as a linear combination of the elements of the parameter vector of a reparameterized model; that is, the population marginal mean can be expressed as $\mathbf{a}'\hat{\beta}^*$ for a proper choice of \mathbf{a} . The variance of the estimated population marginal mean is

$$\text{Var}(\mathbf{a}'\hat{\beta}^*) = \sigma^2 \mathbf{a}'(\mathbf{X}^*\mathbf{X})^{-1}\mathbf{a}$$

and the estimated standard error of $\mathbf{a}'\hat{\beta}^*$ is

$$\widehat{s.e.}(\mathbf{a}'\hat{\beta}^*) = \hat{\sigma}\sqrt{\mathbf{a}'(\mathbf{X}^{*'}\mathbf{X}^*)^{-1}\mathbf{a}}$$

For the one-way model with the set-to-zero restriction, the estimated population marginal means are

$$\widehat{\mu + \tau_i} = \hat{\mu}^* + \hat{\tau}_i^*, \quad i = 1, 2, \dots, t-1$$

and

$$\widehat{\mu + \tau_t} = \hat{\mu}^*$$

The variance of these estimated population marginal means are

$$\text{Var}(\widehat{\mu + \tau_i}) = \text{Var}(\hat{\mu}^*) + 2\text{Cov}(\hat{\mu}^*, \hat{\tau}_i^*) + \text{Var}(\hat{\tau}_i^*) \quad i = 1, 2, \dots, t-1$$

and

$$\text{Var}(\widehat{\mu + \tau_t}) = \text{Var}(\hat{\mu}^*)$$

The data in Table 6.3 are grade point averages (GPA) of a sample of students from a school where the students were classified by year in school and sex. The two-way model

$$y_{ijk} = \mu + \tau_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk} \quad i = 1, 2, 3, 4, \quad j = 1, 2, \quad k = 1, 2, \dots, n_{ij}$$

is used to describe the data where the y_{ijk} are the GPA values, τ_i is the effect for the i th year, β_j is the effect of the j th sex, and γ_{ij} is the interaction effect. The set-to-zero least squares solution for the parameters is displayed in Table 6.4. The estimated standard error is zero for those parameters that are set to zero. The estimate of the cell mean for year 1 and female is

$$\hat{\mu}_{1f} = \hat{\mu} + \hat{\tau}_1 + \hat{\beta}_f + \hat{\gamma}_{1f} = 3.333 - 0.133 - 0.033 + 0.205 = 3.372$$

and the remaining cell means can be computed similarly, and they are displayed in Table 6.5. Estimates of the marginal year and sex means are provided in Tables 6.6 and 6.7. The

TABLE 6.3

Grade Point Average Data for Two-Way Treatment Structure

Freshmen		Sophomores		Juniors		Seniors	
Female	Male	Female	Male	Female	Male	Female	Male
3.3	2.3	3.3	3.1	3.2	2.9	3.7	3.4
3.5	3.2	3.8	3.9	3.5	3.3	2.9	3.8
3.8	3.4	2.3	2.8	3.9	3.6	—	2.8
3.7	3.8	3.0	3.4	3.3	—	—	—
2.8	3.3	3.5	—	—	—	—	—
3.3	—	—	—	—	—	—	—
3.2	—	—	—	—	—	—	—

TABLE 6.4

Least Squares Solution for GPA Data from Two-Way Treatment Structure

Parameter	Estimate	Standard Error
Intercept	3.333	0.264
Year 1	-0.133	0.334
Year 2	-0.033	0.349
Year 3	-0.067	0.373
Year 4	0.000	—
Sex f	-0.033	0.417
Sex m	0.000	—
Year × sex 1 f	0.205	0.498
Year × sex 1 m	0.000	—
Year × sex 2 f	-0.087	0.518
Year × sex 2 m	0.000	—
Year × sex 3 f	0.242	0.544
Year × sex 3 m	0.000	—
Year × sex 4 f	0.000	—
Year × sex 4 m	0.000	—

TABLE 6.5

Cell Means with Estimated Standard Errors

Year	Sex	Cell Size	Cell Mean	Standard Error
1	f	7	3.371	0.173
1	m	5	3.200	0.204
2	f	5	3.180	0.204
2	m	4	3.300	0.229
3	f	4	3.475	0.229
3	m	3	3.267	0.264
4	f	2	3.300	0.323
4	m	3	3.333	0.264

TABLE 6.6

Raw, Least Squares, and Weighted Means for the Years Averaged over Levels of Sex

Year	Raw Mean	LS Mean	Weighted
1	3.300	3.286	3.257
2	3.233	3.240	3.245
3	3.386	3.371	3.357
4	3.320	3.317	3.319

TABLE 6.7

Raw, Least Squares, and Weighted Means for the Females and Males Averaged over Levels of Year

Sex	Raw Mean	LS Mean	Weighted
f	3.333	3.332	3.327
m	3.267	3.275	3.261

raw means are weighted averages of the respective cell means using the numbers of observations in the cells as weights. For example, the raw mean or unadjusted mean for year 1 is computed as

$$\hat{\mu}_{1\cdot} = \frac{7 * 3.371 + 5 * 3.200}{7 + 5} = 3.300$$

The least squares means for the year and sex effects are the unweighted averages of the respective cell means, and the least squares mean for year 1 is

$$\hat{\mu}_{1\cdot} = \frac{3.371 + 3.200}{2} = 3.286$$

When making statements about the marginal year or marginal sex effects one may be interested in using either the raw means (weighting by observed cell size) or the least squares means that weight each cell mean equally (as if there were equal cell sizes). For designed experiments, the least squares means are likely to be the means of interest because you most likely designed the experiment with equal numbers of observations per cell, and so providing estimates of the marginal means as if the cell sizes were equal is a reasonable solution, even though some of the data may be missing. But if the data are from an observational study, the unweighted means may not be the marginal means of interest. If the data were from a simple random sample from the population, then the cell sizes may reflect the proportional membership structure of the population. If that is the case, the raw means or means weighted by sample sizes are the marginal means of interest. But if the sample sizes in the cells are not representative of the population structure, then neither the least squares means nor the raw means are of interest. This type of phenomenon occurs when some segments of the population are over- or undersampled by design or by chance. When the population structure is known, that is, the proportion of the population in each cell is known, the marginal means of interest are obtained by using the known population proportions as weights. For example, suppose the proportions of males and females in each of the years of study are as given in Table 6.8. The estimated marginal mean for year 1 using the weights is

$$\hat{\mu}_{1\cdot} = \frac{11 * 3.371 + 22 * 3.200}{33} = 3.257$$

The estimated marginal mean for females averaged over years using the weights is

$$\hat{\mu}_{.f} = \frac{11 * 3.371 + 12 * 3.180 + 10 * 3.475 + 8 * 3.300}{41} = 3.327$$

The estimated marginal mean for males averaged over years using the weights is

$$\hat{\mu}_{.m} = \frac{22 * 3.200 + 14 * 3.300 + 13 * 3.267 + 10 * 3.333}{59} = 3.261$$

It is not a problem to compute several types of adjusted means and then select those that you like, but it is necessary to specify the proportions in the population structure as correctly

TABLE 6.8

Population Distribution of Students to Classes for the Grade Point Average Study

	Freshmen	Sophomores	Juniors	Seniors	Total
Females	11%	12%	10%	8%	41%
Males	22%	14%	13%	10%	59%
Total	33%	26%	23%	18%	

TABLE 6.9

SAS®-GLM Code with Estimate Statements to Provide Estimates of the Population Marginal Means for Levels of Sex and Year Using the Weights in Table 6.8

```

proc glm data=ex6_4; class year sex;
model gpa=year sex year*sex/solution;
lsmeans year|sex/stderr;
means year|sex;
***assume the following is the population structure
***sex year=1 2 3 4 sum
***female 11 12 10 8 41
***male 22 14 13 10 59
*** sum 33 26 23 18;
estimate 'female pop' intercept 41 sex 41 0 year 11 12 10 8
year*sex 11 0 12 0 10 0 8 0/divisor=41;
estimate 'male pop' intercept 59 sex 0 59 year 22 14 13 10
year*sex 0 22 0 14 0 13 0 10/divisor=59;
estimate 'year 1 pop' intercept 33 sex 11 22 year 33
year*sex 11 22 0 0 0 0 0 0/divisor=33;
estimate 'year 2 pop' intercept 26 sex 12 14 year 0 26 0 0
year*sex 0 0 12 14 0 0 0 0/divisor=26;
estimate 'year 3 pop' intercept 23 sex 10 13 year 0 0 23 0
year*sex 0 0 0 10 13 0 0 0/divisor=23;
estimate 'year 4 pop' intercept 18 sex 8 10 year 0 0 0 18
year*sex 0 0 0 0 0 8 10 /divisor=18;

```

as possible and then compute the estimated marginal means using those weights. This can be accomplished by software packages that allow the use of an “estimate” statement where the estimate and the estimated standard error of the desired marginal mean are provided. Table 6.9 contains code to use the SAS-GLM procedure to provide the estimates of the marginal means using the proportions in Table 6.8 as weights.

6.6 Concluding Remarks

This chapter, required only for those interested in a theoretical background, introduced least squares estimation procedures and discussed the important concept of estimability. Also discussed were the definition and estimation of population marginal means. This chapter provides general formulas for those who want to develop statistical software for their own use.

6.7 Exercises

- 6.1 For the following data set fit the model $y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$, $i = 1, 2, 3$ and $j = 1, 2, 3$.
- 1) Obtain the estimates of the parameters that satisfy the set-to-zero solution.
 - 2) Obtain the estimates of the parameters that satisfy the sum-to-zero solution.
 - 3) Use the two solutions in parts 1) and 2) to verify that $\mu + \tau_2 + \beta_3$ is likely to be estimable.
 - 4) Use the two solutions in parts 1) and 2) to verify that $\tau_1 - \tau_2$ is likely to be estimable.
 - 5) Use the two solutions in parts 1) and 2) to verify that $\beta_1 - \beta_2$ is likely to be estimable.
 - 6) Use the two solutions in parts 1) and 2) to show that the row treatment and column treatment marginal means are estimable.
 - 7) Discuss the estimability of μ , β_1 , and τ_3 .

	Column Treatment 1	Column Treatment 2	Column Treatment 3
Row treatment 1	15	22	18
Row treatment 2	19	24	24
Row treatment 3	21	27	23

- 6.2 1) Show that arrangement I in the following table is connected.
 2) Show that arrangement II in the following table is not connected.

Arrangement I		Arrangement II	
X	X	X	X
X			X X
X		X	X
X X		X	X

"X" denotes that the treatment combination was observed, rows are row treatments, and columns are column treatments.

- 6.3 For the following data set fit the two-way effects model

$$y_{ijk} = \mu + \tau_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk} \quad i = 1, 2, 3, \quad j = 1, 2, 3, \quad k = 0, 1, \text{ or } 2$$

by obtaining a solution to the normal equations.

- 1) Obtain estimates of the row treatment marginal means, column treatment marginal means, and the cell marginal means.
- 2) Use the estimates of the cell marginal means from part 1) to verify that the row treatment marginal means are computed as if there is only one observation per cell.
- 3) Use the proportions in the following population structure table to obtain row treatment and column treatment weighted marginal means.
- 4) Discuss the difference between the results from parts 2) and 3) for row 2 and column 3.
- 5) Obtain an estimate of the difference between row 1 and row 3 and construct a 95% confidence interval about the difference between the two row means. Use information for both the least squares means approach and the weighted means approach.

Data for Exercise 6.3

	Column Treatment 1	Column Treatment 2	Column Treatment 3
Row treatment 1	15, 13	22, 19	18, 20
Row treatment 2	19	24, 26	
Row treatment 3	21, 22	27	23, 23

Population Structure Proportions for Exercise 6.3

	Column Treatment 1	Column Treatment 2	Column Treatment 3
Row treatment 1	15%	10%	5%
Row treatment 2	20%	15%	0%
Row treatment 3	10%	15%	10%

7

Balanced Two-Way Treatment Structures

Chapters 4 and 5 discussed how to analyze the design structure of an experiment. This chapter will discuss methods of analyzing the treatment structure of an experiment. In particular, this chapter will emphasize the analysis of a two-way cross-classified treatment structure. Suppose there are two sets of treatments T_1, T_2, \dots, T_t and B_1, B_2, \dots, B_b . Each one of the T -treatments is to be combined with each one of the B -treatments and assigned to an experimental unit. For convenience, assume that the experimental units are assigned to the treatment combinations completely at random. Alternatively, suppose a survey is taken randomly from a large set of experimental units and that experimental units in the sample can be assigned to categories according to the values of T_1, T_2, \dots, T_t and B_1, B_2, \dots, B_b . In either of these two situations, a total of bt populations are sampled in a cross-classified treatment structure.

Analyzing the treatment structure and analyzing the design structure of an experiment are usually performed independently except for split-plot and repeated measures designs and variations of these considered in later chapters of the book. Thus, it makes little difference whether the experimental units are grouped into complete blocks, balanced incomplete blocks, Latin squares, or some other grouping as the analysis of the treatment structure is similar for most standard designs. Since bt populations are sampled, there are $bt - 1$ degrees of freedom in the sum of squares for testing the hypothesis of equal treatment means, i.e. the hypothesis that $\mu_{11} = \mu_{12} = \dots = \mu_{tb}$. This chapter considers different partitions of the treatment sum of squares to test different kinds of hypotheses that are usually of interest to experimenters.

7.1 Model Definition and Assumptions

7.1.1 Means Model

The *means model* is defined by

$$Y_{ijk} = \mu_{ij} + \varepsilon_{ijk}, \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, b, \quad k = 1, 2, \dots, n \quad (7.1)$$

where it is assumed that $\varepsilon_{ijk} \sim i.i.d. N(0, \sigma^2)$, $i = 1, 2, \dots, t$, $j = 1, 2, \dots, b$, $k = 1, 2, \dots, n$ and μ_{ij} is the response expected when treatment combination (T_i, B_j) is assigned to a randomly selected experimental unit.

7.1.2 Effects Model

Another common model in experiments having a two-way treatment structure is known as the *effects model*, which is defined by Equation 7.1 with μ_{ij} replaced by

$$\mu_{ij} = \mu + \tau_i + \beta_j + \gamma_{ij}, \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, b \quad (7.2)$$

Philosophically, the effects model is intuitively appealing because it might be motivated by assuming that μ represents some overall mean effect, τ_i represents the effect on the mean as a result of assigning T_i to the experimental unit, β_j represents the effect of assigning B_j to the experimental unit, and γ_{ij} represents any additional effects, either positive or negative, that might result from using both T_i and B_j at the same time on the experimental unit. However, the effects model presents some estimability problems, as pointed out in Chapter 6. Both models are discussed below.

7.2 Parameter Estimation

In most cases, the estimate of σ^2 is derived solely from the analysis of the design structure of an experiment. However, in some instances the experimenter may be able to make certain assumptions about the μ_{ij} or the treatment effects that might provide some additional information about σ^2 . For example, if the experimenter knows that $\mu_{11} = \mu_{12}$, then the samples from the populations with means μ_{11} and μ_{12} could be combined to provide an additional degree of freedom for estimating σ^2 . In this case, the corresponding single degree of freedom sum of squares is given by $n(\bar{y}_{11} - \bar{y}_{12})^2/2$.

Most experimenters would not tend to believe that $\mu_{11} = \mu_{12}$, but there are other beliefs that might be appropriate. One such belief is that the effects of the two sets of treatments are additive; that is, the two sets of treatments do not interact. If this were true then additional information would be available for estimating σ^2 .

All of the discussion provided in Chapters 1–3 for the one-way treatment structure apply to the two-way treatment structure if one considers the bt combinations as bt different treatments; that is, $[\mu_{11}, \mu_{12}, \dots, \mu_{1b}, \dots, \mu_{t1}, \mu_{t2}, \dots, \mu_{tb}] = [\mu_1, \mu_2, \dots, \mu_b]$, say.

Consequently, the best estimates of the parameters in the means model are

$$\hat{\mu}_{ij} = \frac{1}{n} \sum_{k=1}^n y_{ijk} = \bar{y}_{ij\cdot}, \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, b$$

and

$$\hat{\sigma}^2 = \frac{1}{N-bt} \sum_{ijk} (y_{ijk} - \bar{y}_{ij\cdot})^2$$

where $N = nbt$. The sampling distribution of $\hat{\mu}_{ij}$ is $N(\mu_{ij}, \sigma^2/n)$ for $i = 1, 2, \dots, t, j = 1, 2, \dots, b$ and the sampling distribution of $(N - bt)\hat{\sigma}^2/\sigma^2$ is $\chi^2(N - bt)$. In addition, $\hat{\mu}_{11}, \hat{\mu}_{12}, \dots, \hat{\mu}_{tb}$ and $\hat{\sigma}^2$ have independent probability distributions.

Most often an experimenter will want to answer the following questions:

- 1) How do the T -treatments affect the response?
- 2) How do the B -treatments affect the response?

7.3 Interactions and Their Importance

In order to give good answers to the above questions, the experimenter must first determine whether these two sets of treatments interact. The interaction hypothesis can be stated in several equivalent ways. Some of these are given below.

- 1) $H_0: \mu_{ij} - \mu_{i'j'} = \mu_{rj} - \mu_{r'j'},$ for all $i \neq i'$ and $j \neq j',$
- 2) $H_0: \mu_{ij} - \mu_{rj} = \mu_{i'j'} - \mu_{r'j'},$ for all $i \neq i'$ and $j \neq j',$
- 3) $H_0: \mu_{ij} - \mu_{rj} - \mu_{i'j'} + \mu_{r'j'} = 0$ for all $i \neq i'$ and $j \neq j',$
- 4) $H_0: \mu_{ij} = \mu + \tau_i + \beta_j$ for all i and j for some set of parameters $\mu, \tau_1, \tau_2, \dots, \tau_t, \beta_1, \beta_2, \dots, \beta_b.$

Each of 1–4 implies that there is no interaction between the two sets of treatment effects. The interpretation of 1 is that the difference between any pair of B -treatments is the same regardless of which T -treatment they are combined with. Similarly, the interpretation of 2 is that the difference between any pair of T -treatments is the same regardless of which B -treatment they are combined with. Both 1 and 2 are algebraically equivalent to 3, which is often referred to as the set of all possible two by two table differences. The interpretation of 4 is that the effects of the two sets of treatments are additive. If any of 1–4 is true, then it is stated that there is no interaction between the T -treatments and B -treatments.

If the two sets of treatments do not interact, then the effects of each set of treatments can best be compared after averaging over the effects of the second set of treatments. Such a comparison is best in the sense that averaging provides more power for comparing the effects of two or more treatments, or equivalently, averaging gives the shortest possible confidence intervals on effect differences. If the two sets of treatments interact, then differences between the effects of one set of treatments depend on the level of the second treatment set with which they are combined, and the analysis of the experiment is slightly more complex.

7.4 Main Effects

If the experimenter concludes that the two sets of treatments do not interact, then hypotheses about the main effects can be tested. These hypotheses can be written as:

$$H_{01}: \bar{\mu}_{1.} = \bar{\mu}_{2.} = \cdots = \bar{\mu}_{t.} \quad \text{and} \quad H_{02}: \bar{\mu}_{.1} = \bar{\mu}_{.2} = \cdots = \bar{\mu}_{.b}$$

Even if there is interaction in the experiment, the above two hypotheses can still be tested. However, the interpretations of the results of the tests in these two situations will be quite different.

7.5 Computer Analyses

Most statistical analysis packages automatically give the tests for the two main effects hypotheses and the interaction hypothesis described in the preceding two sections provided that one specifies a model of the form

$$y = T B T \times B$$

Most also have an option that allows the user to compare the main effect means to one another by using one or more of the multiple comparison procedures discussed in Chapter 3. Most packages also allow the user to specify and test contrasts of the user's own choosing.

If it is determined that the two sets of treatments interact, then the experimenter may want to compare the differences between all pairs of the bt treatment combinations. This can be done by hand if such comparisons cannot be made by the statistical package being used. Alternatively, if the statistical analysis package does not allow multiple comparisons of the $T \times B$ cell means, it can often be tricked into doing so. To do this, one can include a new identification variable in the data file so that the new variable takes on bt different values, one for each of the bt treatment combinations. This new variable can be used to reanalyze the data as a one-way treatment structure, thus yielding multiple comparisons on the two-way cell means. One may also be interested in adjusting for carrying out multiple tests as described in Chapter 3.

In the next chapter, a case study is considered that illustrates the concepts discussed in this chapter.

7.6 Concluding Remarks

There are three basic preliminary hypotheses that are often tested when the treatments are arranged in a two-way treatment structure. The most important of these is the interaction hypothesis. If there is no interaction, then the main effects of each of the treatments can best be compared by averaging over the levels of the other treatment. If there is interaction, then the experimenter must be careful to determine whether it makes sense to average over the levels of the second treatment when comparing the effects of the first treatment. More often than not, it does not make sense.

7.7 Exercises

- 7.1 The following data are from a two-way treatment structure in a completely randomized design structure with three replications per treatment combination.

- 1) Estimate the parameters of the means model, $y_{ijk} = \mu_{ij} + \varepsilon_{ijk}$, $i = 1, 2, 3$ and $j = 1, 2, 3, 4$, and $k = 1, 2, 3$.
- 2) Estimate the parameters of the effects model, $y_{ijk} = \mu + \tau_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}$, $i = 1, 2, 3, 4$ and $j = 1, 2, 3$ and $k = 1, 2, 3$.
- 3) Use contrast statements with the means model to provide the sum of squares for row treatments, sum of squares for column treatments, and sum of squares due to interaction.
- 4) Use contrast statements with the effects model to provide the sum of squares for row treatments, sum of squares for column treatments, and sum of squares due to interaction.

	Column Treatment 1	Column Treatment 2	Column Treatment 3	Column Treatment 4
Row treatment 1	78, 74, 75	85, 86, 86	80, 82, 79	89, 87, 87
Row treatment 2	79, 82, 82	89, 87, 86	87, 83, 84	81, 83, 80
Row treatment 3	75, 72, 74	80, 83, 85	76, 79, 78	91, 93, 92

8

Case Study: Complete Analyses of Balanced Two-Way Experiments

In the preceding chapter, it was assumed that, when the T -treatments and the B -treatments interact in an experiment, the experimenter will want to compare the effects of the T -treatments at each possibility for one of the B -treatments, or vice versa. In many instances, interaction does not occur everywhere in the experiment—often just one or two treatment combinations are responsible for the interaction. In other cases, one possibility for one of the treatments may interact with the possibilities for the second treatment, while all other possibilities of the first treatment do not interact with any possibilities of the second.

In order to conduct a more complete analysis of data with interaction, it is helpful to determine where the interaction occurs in the data. For example, if it is known that all of the interaction in an experiment is caused by only one of the possibilities (often a control) of the T -treatments, then all of the remaining possibilities of the T -treatments could still be compared after averaging over all of the possibilities of the B -treatments. This results in more powerful tests for comparing the possibilities of the T -treatments that do not interact with the B -treatments.

8.1 Contrasts of Main Effect Means

Very often the possibilities of the T -treatments and the possibilities of the B -treatments suggest main-effect contrasts that would be particularly interesting to an experimenter. Such main effect contrasts suggest special types of interaction contrasts that should also be interesting for the experimenter. Furthermore, such interaction contrasts are often easy to interpret. Next main effect contrasts and orthogonal main effect contrasts are defined.

Definition 8.1: A linear combination of the $\bar{\mu}_{i..}$, $\sum_{i=1}^t c_i \bar{\mu}_{i..}$, is called a contrast in the T main effect means if $\sum_{i=1}^t c_i = 0$. Likewise, a linear combination of the $\bar{\mu}_{.j..}$, $\sum_{j=1}^b d_j \bar{\mu}_{.j..}$, is called a contrast in the B main effect means if $\sum_{j=1}^b d_j = 0$.

Definition 8.2: Two contrasts, $\sum_{i=1}^t c_i \bar{\mu}_{i\cdot}$ and $\sum_{i=1}^t c'_i \bar{\mu}_{i\cdot}$, are called orthogonal contrasts in the T main effect means if $\sum_{i=1}^t c_i c'_i = 0$. Similarly, two contrasts, $\sum_{j=1}^b d_j \bar{\mu}_{\cdot j}$ and $\sum_{j=1}^b d'_j \bar{\mu}_{\cdot j}$ are called orthogonal contrasts in the B main effect means if $\sum_{j=1}^b d_j d'_j = 0$.

Now suppose that

$$S_T = \left\{ \sum_{i=1}^t c_{i1} \bar{\mu}_{i\cdot}, \sum_{i=1}^t c_{i2} \bar{\mu}_{i\cdot}, \dots, \sum_{i=1}^t c_{it-1} \bar{\mu}_{i\cdot} \right\}$$

is a set of $t - 1$ orthogonal contrasts in the T main effect means and that

$$S_B = \left\{ \sum_{j=1}^b d_{j1} \bar{\mu}_{\cdot j}, \sum_{j=1}^b d_{j2} \bar{\mu}_{\cdot j}, \dots, \sum_{j=1}^b d_{jb-1} \bar{\mu}_{\cdot j} \right\}$$

is a set of $b - 1$ orthogonal contrasts in the B main effect means. Each of the sets S_T and S_B suggests a partitioning of the two main effect sums of squares. That is,

$$S_T^* = \left\{ Q_p^2 = \frac{nb(\sum_i c_{ip} \bar{y}_{i\cdot})^2}{\sum_i c_{ip}^2}, \quad p = 1, 2, \dots, t - 1 \right\} \quad (8.1)$$

defines a partitioning of the sum of squares for T and

$$S_B^* = \left\{ Q_q^2 = \frac{nt(\sum_j d_{jq} \bar{y}_{\cdot j})^2}{\sum_j d_{jq}^2}, \quad q = 1, 2, \dots, b - 1 \right\} \quad (8.2)$$

defines a partitioning of the sum of squares for B . That is, each Q_p^2 in S_T^* is a single-degree-of-freedom sum of squares used for testing whether the corresponding contrast in the T main effect means is equal to zero. Furthermore, the sum of the $t - 1$ single-degree-of-freedom contrasts in S_T^* is equal to the sum of squares for testing $H_{01}: \bar{\mu}_{1\cdot} = \bar{\mu}_{2\cdot} = \dots = \bar{\mu}_{t\cdot}$. A similar situation exists for the elements of S_B^* .

One should not overemphasize the desirability of obtaining orthogonal partitions of the basic sums of squares. Orthogonal partitions are nice from a mathematical point of view, but may not be all that nice from a practical point of view. Quite often a well-chosen set of orthogonal contrasts will enable an experimenter to interpret his/her data wisely, clearly, and completely. However, the experimenter should really consider any and all contrasts that are meaningful and should not be overly concerned about whether the selected contrasts are orthogonal or not.

8.2 Contrasts of Interaction Effects

This section begins with the definition of an interaction contrast in a two-way experiment.

Definition 8.3: A linear combination of the μ_{ij} , $\sum_i \sum_j \omega_{ij} \mu_{ij}$ is called an interaction contrast if $\sum_i \omega_{ij} = 0$ for $j = 1, 2, \dots, b$ and $\sum_j \omega_{ij} = 0$ for $i = 1, 2, \dots, t$.

Contrasts in the main effects of a two-way experiment give rise to special types of interaction contrasts. Suppose that $\sum_{i=1}^t c_i \bar{\mu}_i$ is a contrast in the T main effect means and that $\sum_{j=1}^b d_j \bar{\mu}_j$ is a contrast in the B main effect means. Then $\sum_i \sum_j c_i d_j \mu_{ij}$ is an interaction contrast. That is, if one takes $\omega_{ij} = c_i d_j$ for all i and j , then one gets an interaction contrast. Two interaction contrasts, $\sum_i \sum_j \omega_{ij} \mu_{ij}$ and $\sum_i \sum_j \omega'_{ij} \mu_{ij}$, are called orthogonal interaction contrasts if $\sum_i \sum_j \omega_{ij} \omega'_{ij} = 0$. Orthogonal contrasts in the two sets of main effects give rise to orthogonal contrasts in the interaction effects. Suppose that $\sum_{i=1}^t c_i \bar{\mu}_i$ and $\sum_{i=1}^t c'_i \bar{\mu}_i$ are two contrasts in the T main effect means and suppose that $\sum_{j=1}^b d_j \bar{\mu}_j$ and $\sum_{j=1}^b d'_j \bar{\mu}_j$ are two contrasts in the B main effect means, then $\sum_i \sum_j c_i d_j \mu_{ij}$ and $\sum_i \sum_j c'_i d'_j \mu_{ij}$ are orthogonal interaction contrasts if either $\sum_{i=1}^t c_i c'_i = 0$ or $\sum_{j=1}^b d_j d'_j = 0$; that is, if at least one of the pairs of main effect contrasts is an orthogonal pair of contrasts, the two interaction contrasts will be orthogonal to one another. Next suppose that S_T^* and S_B^* are as defined in Section 8.1, and let $S_{T \times B}^*$ be defined by

$$S_{T \times B}^* = \left\{ Q_{pq}^2 = \frac{n(\sum_i \sum_j c_{ip} d_{jq} \bar{y}_{ij})^2}{\sum_i c_i^2 \sum_j d_j^2}, \quad p = 1, 2, \dots, t-1; q = 1, 2, \dots, b-1 \right\} \quad (8.3)$$

$S_{T \times B}^*$ defines a partitioning of the sum of squares for interaction. That is, $\sum_p \sum_q Q_{pq}^2 = T \times B$ interaction sum of squares, and all of the $(t-1)(b-1)$ single-degree-of-freedom sums of squares in Equation 8.3 have independent probability distributions. In the next section, an example is discussed that illustrates the ideas described in the preceding section.

8.3 Paint–Paving Example

Consider the experiment in Table 8.1, which gives the means of three independent replications of each of the paint \times paving treatment combinations. This experiment was conducted to compare the lifetimes, measured in weeks, of two colors of paint manufactured by two different companies on three types of paving surfaces. The error sum of squares for this experiment was 455.04 with 24 degrees of freedom so that $\hat{\sigma}^2 = 18.96$. The usual analysis of variance table for this experiment is given in Table 8.2.

TABLE 8.1

Paint–Paving Cell Means

Paint	Asphalt I	Asphalt II	Concrete	Mean
Yellow I	15	17	32	21.333
Yellow II	27	30	20	25.667
White I	30	28	29	29.000
White II	34	35	36	35.000
Mean	26.5	27.5	29.25	27.750

TABLE 8.2

Analysis of Variance Table for Paint–Paving Experiment

Source of Variation	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Total	35	2039.79			
Paint	3	896.75	298.92	15.75	<0.001
Paving	2	46.50	23.25	1.25	n.s.
Paint × paving	6	641.50	106.42	5.64	<0.001
Error	24	455.04	18.96		

The possible values for each of the paint and paving treatments give rise to two sets of orthogonal contrasts in the main effect means that might be of interest. These are given in Table 8.3. These two sets of orthogonal contrasts in main effects suggest six orthogonal contrasts in the interaction effects. These are given in Table 8.4.

A more complete analysis of these data using the partitioning suggested in Tables 8.3 and 8.4 is given in Table 8.5. Next the details of the computations necessary to obtain the sum of squares for three selected contrasts from Table 8.5 are given.

TABLE 8.3

Main Effect Contrasts for Paint–Paving Experiment

Comparison Contrast	Hypothesis
Paints	
Yellow I vs yellow II	$\bar{\mu}_{1.} - \bar{\mu}_{2.} = 0$
White I vs white II	$\bar{\mu}_{3.} - \bar{\mu}_{4.} = 0$
Yellow vs white	$\bar{\mu}_{1.} + \bar{\mu}_{2.} - \bar{\mu}_{3.} - \bar{\mu}_{4.} = 0$
Pavings	
Asphalt I vs asphalt II	$\bar{\mu}_{.1} - \bar{\mu}_{.2} = 0$
Asphalt vs concrete	$\bar{\mu}_{.1} + \bar{\mu}_{.2} - 2\bar{\mu}_{.3} = 0$

TABLE 8.4

Interaction Hypotheses for Paint–Paving Data

Comparison Contrast	Hypothesis
Yellow × asphalt	$\mu_{11} - \mu_{12} - \mu_{21} + \mu_{22} = 0$
White × asphalt	$\mu_{31} - \mu_{32} - \mu_{41} + \mu_{42} = 0$
Color × asphalt	$\mu_{11} - \mu_{12} + \mu_{21} - \mu_{22} - \mu_{31} + \mu_{32} - \mu_{41} + \mu_{42} = 0$
Yellow × type	$\mu_{11} + \mu_{12} - 2\mu_{13} - \mu_{21} - \mu_{22} + 2\mu_{23} = 0$
White × type	$\mu_{31} + \mu_{32} - 2\mu_{33} - \mu_{41} - \mu_{42} + 2\mu_{43} = 0$
Color × type	$\mu_{11} + \mu_{12} - 2\mu_{13} + \mu_{21} + \mu_{22} - 2\mu_{23} - \mu_{31} - \mu_{32} + 2\mu_{33} - \mu_{41} - \mu_{42} + 2\mu_{43} = 0$

Note: “Type” refers to asphalt vs concrete.

TABLE 8.5

Analysis of Variance Table for Paint–Paving Experiment Including Single-Degree-of-Freedom Tests

Source of Variation	<i>df</i>	SS	MS	F	<i>p</i>
Total	35	2039.79			
Paint	3	896.92	298.97	15.77	<0.0001
Yellow	1	84.5	84.5	4.46	<0.05
White	1	162.0	162.0	8.54	<0.01
Color	1	650.25	650.25	34.30	<0.0001
Paving	2	46.5	23.25	1.23	n.s.
Asphalt	1	6.0	6.0	0.32	n.s.
Type	1	40.5	40.5	2.14	n.s.
Paint × paving	6	641.5	106.92	5.64	<0.001
Yellow × asphalt	1	0.75	0.75	0.04	n.s.
White × asphalt	1	6.75	6.75	0.36	n.s.
Color × asphalt	1	13.5	13.5	0.71	n.s.
Yellow × type	1	600.25	600.25	31.66	<0.0001
White × type	1	2.25	2.25	0.12	n.s.
Color × type	1	18.0	18.0	0.95	n.s.
Error	24	455.04	18.96		

The single degree of freedom sum of squares for comparing the two white paint means is, from Equation 8.1,

$$Q_1^2 = \frac{3 \cdot 3[(1)29.0 + (-1)35.0]^2}{(1)^2 + (-1)^2} = 162.0$$

The single degree of freedom sum of squares for comparing asphalt with concrete is, from Equation 8.2

$$Q_2^2 = \frac{3 \cdot 4[(1)(26.5) + (1)(27.5) + (-2)(29.5)]^2}{1^2 + 1^2 + (-2)^2} = 40.5$$

and the single degree of freedom sum of squares for comparing the white × type interaction is, from Equation 8.3,

$$Q_3^2 = \frac{3[(1)(30) + (1)(28) + (-2)(29) + (-1)(34) + (-1)(35) + (2)(36)]^2}{(1)^2 + (1)^2 + (-2)^2 + (-1)^2 + (-1)^2 + (2)^2} = 2.25$$

From examining the analysis in Table 8.5, one can make the following conclusions:

- 1) All of the interaction in the experiment is caused by the two yellow paints acting differently on the two types of surfaces, since this interaction contrast is the only single degree of freedom sum of squares for interaction that is significant.

- 2) Because we now know where the interaction exists in the data, we can make the following observations:
- Since there is no interaction between asphalt and paint, the two asphalts can be compared after averaging across all paints. The value of the F -statistic for this comparison is $F = 0.32$; thus, there is no significant difference between asphalts I and II.
 - Since there is no interaction between the white paints and the three pavings, the two white paints can be compared after averaging across all pavings. The value of the F -statistic for this comparison is $F = 8.54$, which indicates that white paint II is significantly different from white paint I. From Table 8.1, we see that white paint II lasts longer.
 - Although the statistic for comparing yellow I vs yellow II is significant ($F = 4.45$), one must be careful when making an interpretation because of the significant interaction between the brands of yellow paint and the type of paving.
 - Even though the comparison for asphalt vs concrete is not significant ($F = 2.14$), one must again be careful when making an interpretation because of the significant interaction between the brands of paint and the types of paving.
- 3) To complete the analysis of these data, we should yet examine:
- Yellow I vs yellow II on asphalt (that is, test $\mu_{11} + \mu_{12} - \mu_{21} - \mu_{22} = 0$).
 - Yellow I vs yellow II on concrete.
 - Concrete vs asphalt for yellow I.
 - Concrete vs asphalt for yellow II.
 - The three pavings for white paints.

The results are given in Table 8.6.

Examination of the results in Table 8.6 and the means in Table 8.1 reveals that 1) yellow II is significantly better than yellow I on asphalt, but 2) yellow I is significantly better than yellow II on concrete; 3) yellow I lasts significantly longer on concrete than on asphalt; 4) yellow II lasts significantly longer on asphalt than on concrete; and finally, 5) the white paints last about the same length of time on all three pavings.

All of the results obtained for our analysis of this example can be obtained using many statistical computing packages such as SAS® and SPSS using their contrast options. Some of these programs may require that the 12 treatment combinations be considered as a one-way treatment structure.

TABLE 8.6
Tests of Hypotheses in Conclusion 3

Source of Variation	df	SS	MS	F	p
a	1	468.75	468.75	24.72	<0.0001
b	1	216.00	216.00	11.39	<0.005
c	1	512.00	512.00	27.00	<0.0001
d	1	144.50	144.50	7.62	<0.02
e	2	3.00	1.50	0.08	n.s.

8.4 Analyzing Quantitative Treatment Factors

In this section it is assumed that the levels of both factors of an experiment are quantitative. In this case, one can define contrasts that measure curvilinear trends in each set of main effect treatment means. Trends of interest are often linear, quadratic, cubic, and so on. The corresponding orthogonal contrasts that partition the main effect sums of squares into effects that measure linear, quadratic, cubic, and so on, trends can then be used to construct orthogonal contrasts in the interaction effects. The resulting contrasts are called $\text{Lin } T \times \text{Lin } B$ (linear effect of T by linear effect of B), $\text{Lin } T \times \text{Quad } B$, and so on. For a 3×4 experiment where both treatments have equally spaced levels (for example, 5, 10, 15 for factor T and 2, 4, 6, 8 for factor B), the $\text{Lin } T$ contrast is defined by $-\bar{\mu}_{1\cdot} + \bar{\mu}_{3\cdot} = 0$ and the $\text{Lin } B$ contrast is identified by $-3\bar{\mu}_{1\cdot} - \bar{\mu}_{2\cdot} + \bar{\mu}_{3\cdot} + 3\bar{\mu}_{4\cdot} = 0$.

Note that for the $\text{Lin } T$ contrast, $c_1 = -1$, $c_2 = 0$, and $c_3 = 1$ in $\sum_{i=1}^3 c_i \bar{\mu}_{i\cdot} = 0$ and in the $\text{Lin } B$ contrast, $d_1 = -3$, $d_2 = -1$, $d_3 = 1$, and $d_4 = 3$ in $\sum_{j=1}^4 d_j \bar{\mu}_{\cdot j} = 0$. Thus the $\text{Lin } T \times \text{Lin } B$ interaction contrast is defined by $3\mu_{11} + \mu_{12} - \mu_{13} - 3\mu_{14} - 3\mu_{31} - \mu_{32} + \mu_{33} + 3\mu_{34} = 0$. The values of the c_i and the d_j used to define these kinds of orthogonal contrasts in the main effect means can be found in a table of orthogonal polynomial coefficients. See Beyer (1968). These are reproduced in Table 8.7 for a 3×4 experiment.

Let x_1, x_2, \dots, x_t represent the levels of factor T and let z_1, z_2, \dots, z_b represent the levels of factor B . There always exist parameters α_{kh} , $k = 0, 1, 2, \dots, t-1$, $h = 0, 1, 2, \dots, b-1$, such that the cell mean parameters can be represented as a polynomial function of x_i and z_j . That is, there exist α_{kh} such that

$$\mu_{ij} = \sum_{k=0}^{t-1} \sum_{h=0}^{b-1} \alpha_{kh} x_i^k z_j^h \quad (8.4)$$

Expanding Equation 8.4, one gets

$$\begin{aligned} \mu_{ij} = & \alpha_{00} + \alpha_{10} x_i + \alpha_{20} x_i^2 + \dots + \alpha_{t-10} x_i^{t-1} + \alpha_{01} z_j + \alpha_{02} z_j^2 + \dots + \alpha_{0b-1} z_j^{b-1} \\ & + \alpha_{11} x_i z_j + \alpha_{12} x_i z_j^2 + \dots + \alpha_{t-1b-1} x_i^{t-1} z_j^{b-1} \end{aligned}$$

Table 8.8 gives the expected values of main effect and interaction contrasts for a 3×4 experiment in terms of the α_{kh} in Equation 8.4. In constructing the table, it is assumed that the three levels of the x were coded to -1 , 0 , and 1 and that the four levels of the z were

TABLE 8.7

Orthogonal Contrast Coefficients for a 3×4 Experiment

Contrast	Coefficients			
T	c_1	c_2	c_3	
Lin T	-1	0	1	
Quad T	-1	2	-1	
B	d_1	d_2	d_3	d_4
Lin B	-3	-1	1	3
Quad B	1	-1	-1	1
Cubic B	-3	1	-1	3

TABLE 8.8Expected Values of Orthogonal Polynomials in a 3×4 Experiment

Comparison Contrast	Hypothesis
Lin T	$2\alpha_{10} + 10\alpha_{12} = 0$
Quad T	$2\alpha_{20} + 10\alpha_{22} = 0$
Lin B	$60\alpha_{01} + 40\alpha_{21} + 492\alpha_{03} + 328\alpha_{23} = 0$
Quad B	$48\alpha_{02} + 32\alpha_{22} = 0$
Cubic B	$48\alpha_{03} + 32\alpha_{23} = 0$
Lin $T \times$ Lin B	$40\alpha_{11} + 328\alpha_{13} = 0$
Lin $T \times$ Quad B	$32\alpha_{12} = 0$
Lin $T \times$ Cubic B	$96\alpha_{13} = 0$
Quad $T \times$ Lin B	$40\alpha_{21} + 328\alpha_{23} = 0$
Quad $T \times$ Quad B	$32\alpha_{22} = 0$
Quad $T \times$ Cubic B	$96\alpha_{23} = 0$

coded to -3 , -1 , 1 , and 3 . The purpose of providing Table 8.8 is to point out the hypotheses that are actually being tested when contrasts of this type are being investigated. For example, the Lin $T \times$ Lin B contrast tests the hypothesis that $40\alpha_{11} + 328\alpha_{13} = 0$. Thus, if this effect is significant, it could be because $\alpha_{11} \neq 0$ or $\alpha_{13} \neq 0$ and not only because $\alpha_{11} \neq 0$ (as many data analysts might believe).

If one is going to examine orthogonal polynomials, it is recommended that one looks at the coefficients of the highest degree term first and consider the remaining terms in descending order of degree. Once a higher-order term is determined to be in the model, then all terms whose two components both have degrees lower than that of the significant term should also be included in the model whether significant or not. For example, if one decides that x^2z^2 should be in the model, then the model should also include xz^2 , z^2 , x^2z , x^2 , xz , x , and z should also be included in the model. The reason is that orthogonal polynomials always refer to coded values of the quantitative variables. Thus, if α_{22} is nonzero, it really implies that

$$\left(\frac{x - h_1}{c_1} \right)^2 \left(\frac{z - h_2}{c_2} \right)^2$$

belongs in the model where

$$\left(\frac{x - h_1}{c_1} \right) \text{ and } \left(\frac{z - h_2}{c_2} \right)$$

are the coded values of x and z . Expansion of

$$\left(\frac{x - h_1}{c_1} \right)^2 \left(\frac{z - h_2}{c_2} \right)^2$$

demonstrates that the terms xz^2 , z^2 , x^2z , x^2 , xz , x , and z are also in the model, even though other lower degree orthogonal polynomials may not have been significant.

8.5 Multiple Comparisons

Any of the multiple comparison procedures discussed in Chapter 3 can be used with only some very minor adjustments for making multiple comparisons on the main effects of a two-factor experiment. The adjustments require that the n and the n_i be replaced by the total number of observations that were averaged to estimate the main effect means being compared. In this chapter, the sample sizes are nt for the B main effect means and nb for the T main effect means. Our recommendations for multiple comparisons on main effect means are the same as those given in Section 3.2. The only procedures given in Chapter 3 that are easily generalized to contrasts in the interaction effects are the LSD procedure, Bonferroni's method, the multivariate t method, the simulation method, and Scheffé's procedure. We have found that Scheffé's procedure is not very satisfactory because the required critical point is much too large and the procedure is much too conservative. Our recommendations for multiple comparisons of interaction contrasts are as follows:

- 1) Conduct an F -test for interaction.
 - 2) If the F -statistic is significant, make any planned comparisons by using the LSD procedure (or equivalently, the contrast procedure given in the preceding section). For data snooping and unplanned comparisons, use the procedure given by Johnson (1976), which is not discussed here.
 - 3) If the F -test for interaction is not significant, the experimenter should still examine any individual interaction contrasts that she had planned to consider but by using the multivariate t method or Bonferroni's method. The multivariate t method is used whenever the selected contrasts are linearly independent; otherwise, Bonferroni's method should be used.
-

8.6 Concluding Remarks

In this chapter, we introduced, by giving examples, methods for obtaining a maximum amount of information from an experiment. Included were methods for discovering where interaction occurs in an experiment. Knowing where interaction occurs in an experiment is valuable in determining the best answers to questions that may be raised. The techniques introduced in this chapter should help experimenters do a better job of analyzing their experiments. The analysis of quantitative treatment factors was also considered, including how to determine what kinds of trends might be related to the different levels of the treatment factors.

8.7 Exercises

- 8.1 An experiment was conducted in an RCB design structure (with days as blocks) to aid in developing a product that can be used as a substrate for making ribbons.

The treatment structure is a two-way with one factor consisting of three different base (*B*) polymers (mylar, nylon, and polyethylene). The second factor consisted of five different additives (*A*) that could be included to enhance the formulation. The additives are denoted by: *c1*, *c2*, *c3*, *c4*, and *c5*. The variable of primary interest is the tensile strength of the resulting ribbon. The data are given below:

B	A	Day	TS	B	A	Day	TS	B	A	Day	TS
Mylar	<i>c1</i>	1	9.2	Nylon	<i>c1</i>	1	8.2	Peth	<i>c1</i>	1	9.2
Mylar	<i>c2</i>	1	8.7	Nylon	<i>c2</i>	1	7.7	Peth	<i>c2</i>	1	13.4
Mylar	<i>c3</i>	1	9.1	Nylon	<i>c3</i>	1	11.4	Peth	<i>c3</i>	1	9.7
Mylar	<i>c4</i>	1	12.4	Nylon	<i>c4</i>	1	8.1	Peth	<i>c4</i>	1	9.1
Mylar	<i>c5</i>	1	10.5	Nylon	<i>c5</i>	1	9.5	Peth	<i>c5</i>	1	8.5
Mylar	<i>c1</i>	2	8.2	Nylon	<i>c1</i>	2	7.2	Peth	<i>c1</i>	2	8.4
Mylar	<i>c2</i>	2	8.7	Nylon	<i>c2</i>	2	7.7	Peth	<i>c2</i>	2	12.5
Mylar	<i>c3</i>	2	8.8	Nylon	<i>c3</i>	2	10.5	Peth	<i>c3</i>	2	9.1
Mylar	<i>c4</i>	2	11.5	Nylon	<i>c4</i>	2	7.8	Peth	<i>c4</i>	2	9.1
Mylar	<i>c5</i>	2	10.6	Nylon	<i>c5</i>	2	9.6	Peth	<i>c5</i>	2	8.9
Mylar	<i>c1</i>	3	8.4	Nylon	<i>c1</i>	3	7.4	Peth	<i>c1</i>	3	8.2
Mylar	<i>c2</i>	3	8.3	Nylon	<i>c2</i>	3	7.8	Peth	<i>c2</i>	3	8.5
Mylar	<i>c3</i>	3	8.7	Nylon	<i>c3</i>	3	7.3	Peth	<i>c3</i>	3	8.8
Mylar	<i>c4</i>	3	8.5	Nylon	<i>c4</i>	3	7.7	Peth	<i>c4</i>	3	8.1
Mylar	<i>c5</i>	3	8.8	Nylon	<i>c5</i>	3	7.1	Peth	<i>c5</i>	3	8.3

Obtain an analysis of the data and write a report that summarizes what you believe to be all of the information in the data.

- 8.2 An experiment was conducted to study the relationships between three factors as to their effect on the ability of a tire tread made with the combinations to increase friction. A RCB design structure was used. Three levels of carbon (5, 7, and 9) were combined with combinations of two types of rubber gum, type *A* at four levels (0.1, 0.3, 0.5, and 0.7) and type *B* at two levels (0.2 and 0.8). Use orthogonal polynomials to investigate the effects of the three factors on the friction index. Obtain an analysis of the data and write a report that summarizes what you believe to be all of the information in the data. The data follow.

Block	Type A	Type B	C	Friction	Block	Type A	Type B	C	Friction
1	0.3	0.2	5	14	1	0.3	0.2	7	18
1	0.3	0.2	9	20	1	0.3	0.8	5	15
1	0.3	0.8	7	21	1	0.3	0.8	9	17
1	0.5	0.2	5	15	1	0.5	0.2	7	22
1	0.5	0.2	9	21	1	0.5	0.8	5	17
1	0.5	0.8	7	23	1	0.5	0.8	9	15
1	0.7	0.2	5	17	1	0.7	0.2	7	25
1	0.7	0.2	9	27	1	0.7	0.8	5	19
1	0.7	0.8	7	27	1	0.7	0.8	9	18

Continued

Block	Type A	Type B	C	Friction	Block	Type A	Type B	C	Friction
1	0.9	0.2	5	25	1	0.9	0.2	7	35
1	0.9	0.2	9	35	1	0.9	0.8	5	25
1	0.9	0.8	7	35	1	0.9	0.8	9	25
2	0.3	0.2	5	17	2	0.3	0.2	7	23
2	0.3	0.2	9	24	2	0.3	0.8	5	15
2	0.3	0.8	7	21	2	0.3	0.8	9	17
2	0.5	0.2	5	19	2	0.5	0.2	7	27
2	0.5	0.2	9	25	2	0.5	0.8	5	22
2	0.5	0.8	7	28	2	0.5	0.8	9	18
2	0.7	0.2	5	22	2	0.7	0.2	7	29
2	0.7	0.2	9	30	2	0.7	0.8	5	24
2	0.7	0.8	7	32	2	0.7	0.8	9	23
2	0.9	0.2	5	29	2	0.9	0.2	7	37
2	0.9	0.2	9	38	2	0.9	0.8	5	30
2	0.9	0.8	7	38	2	0.9	0.8	9	32

8.3 Consider Exercise 7.1 in Chapter 7.

- 1) Use orthogonal contrasts of the row treatments to show that the sum of squares for row treatments can be expressed as the sum of two single degree of freedom sums of squares.
- 2) Use orthogonal contrasts of the column treatments to show that the sum of squares for column treatments can be expressed as the sum of three single degree of freedom sums of squares.
- 3) Use the orthogonal contrasts in parts 1 and 2 to construct orthogonal interaction contrasts to show that the sum of squares due to interaction can be expressed as the sum of the six single degree of freedom sums of squares for your interaction contrasts.

9

Using the Means Model to Analyze Balanced Two-Way Treatment Structures with Unequal Subclass Numbers

Chapters 7 and 8 considered the equal sample size case, where each treatment combination is observed an equal number of times. Chapters 13–15 consider cases where some treatment combinations are missing, but this chapter, as well as Chapters 10–12, assumes that every treatment combination is observed and at least one combination is observed more than once.

9.1 Model Definitions and Assumptions

As in Section 7.1.1, let μ_{ij} be the expected response when possibility i of treatment T and possibility j of treatment B are both applied to the same experimental unit. This chapter assumes that the observed response, Y_{ijk} , can be modeled by

$$Y_{ijk} = \mu_{ij} + \varepsilon_{ijk}, \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, b, \quad k = 1, 2, \dots, n_{ij} \quad (9.1)$$

where under the ideal conditions,

$$\varepsilon_{ijk} \sim i.i.d. N(0, \sigma^2), \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, b, \quad k = 1, 2, \dots, n_{ij}$$

and $n_{ij} > 0$ for every i and j .

9.2 Parameter Estimation

Everything discussed for the one-way treatment structure in Chapters 1–3 applies to the two-way treatment structure as well, if one considers the bt treatment combinations as bt

different treatments. For unbalanced data problems that is often the best and simplest way to analyze the data. The best estimates of the parameters in the means model are

$$\hat{\mu}_{ij} = \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} y_{ijk} = \bar{y}_{ij\cdot}, \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, b \quad (9.2)$$

and

$$\hat{\sigma}^2 = \frac{1}{N - bt} \sum_{ijk} (y_{ijk} - \bar{y}_{ij\cdot})^2 \quad (9.3)$$

where $N = n\cdot$. The sampling distributions of the $\hat{\mu}_{ij}$ and $\hat{\sigma}^2$ are

$$\hat{\mu}_{ij} \sim N(\mu_{ij}, \sigma^2/n_{ij}) \quad \text{for } i = 1, 2, \dots, t, \quad j = 1, 2, \dots, b$$

and

$$(N - bt) \hat{\sigma}^2 / \sigma^2 \sim \chi^2(N - bt)$$

In addition, all of the $\hat{\mu}_{ij}$ and $\hat{\sigma}^2$ are independently distributed as before.

The experimenter will usually want to answer the same questions when the data are unbalanced as when they are balanced. Recall that those questions were:

- 1) Do the two sets of treatments interact?
- 2) How do the T treatments affect the response?
- 3) How do the B treatments affect the response?

These questions can be stated as hypotheses in terms of the parameters of the means model, and the corresponding hypotheses are:

$$\begin{aligned} H_{T \times B}: \mu_{ij} - \mu_{i'j} - \mu_{ij'} + \mu_{i'j'} &= 0 \quad \text{for all } i \neq i' \text{ and } j \neq j' \\ H_T: \bar{\mu}_{1\cdot} = \bar{\mu}_{2\cdot} = \cdots = \bar{\mu}_{t\cdot} \\ H_B: \bar{\mu}_{\cdot 1} = \bar{\mu}_{\cdot 2} = \cdots = \bar{\mu}_{\cdot b} \end{aligned}$$

Testing the above hypotheses should be considered as a first step in analyzing any two-way experiment. There will usually be well-defined contrasts that directly address other questions of interest to the researcher. The hypotheses $H_{T \times B}$, H_T , and H_B are often tested to help choose an appropriate multiple comparison procedure for addressing other questions that may be of interest.

As in Equation 7.2, there always exist parameters μ , τ_i , β_j , and γ_{ij} , $i = 1, 2, \dots, t$, $j = 1, 2, \dots, b$ such that μ_{ij} can be expressed in an effects model as

$$\mu_{ij} = \mu + \tau_i + \beta_j + \gamma_{ij}, \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, b$$

Many experimenters prefer to look at a representation of the treatment combination means using the effects model. This may be because many statisticians have encouraged experimenters to consider such models. As a result, much of the existing computer software leads data analysts towards using effects model representations. Both types of models are considered in this book. The means model is considered in this chapter and the effects

model is considered in Chapter 10. Sections 1.5 and 1.6 introduced different procedures for calculating test statistics. For the one-way case, all those methods gave rise to the same test statistics. This is, in fact, always the case for *well-balanced* two-way experiments as well. However, it is not always the case with unbalanced data sets. *Well-balanced* means that there is an equal number of observations for each treatment combination. The matrix procedure is used in this chapter to obtain test statistics, and model fitting procedures are used in Chapter 10.

9.3 Testing Whether All Means Are Equal

Consider the data in Table 9.1. The data are from a small two-way treatment structure experiment conducted in a completely randomized design structure. To begin, the two-way cell means and the marginal means for the data in Table 9.1 are calculated. Table 9.2 gives these means where a *row marginal mean* is defined as the mean of the cell means in a given row, and a *column marginal mean* is defined as the mean of the cell means in a given column.

One thing that should be noticed is that in unbalanced two-way experiments there are two different ways that one can compute T means and B means. Table 9.1 gives means that are computed by taking row (column) totals and dividing by the number of observations in the row (column). Table 9.2 gives means that are computed in two steps. First cell means are computed for each $T \times B$ combination, and then means of the cell means are computed for each row and column. The two methods generally give different answers unless the

TABLE 9.1
An Unbalanced Two-Way Experiment

	B_1	B_2	B_3	T Totals	T Means
T_1	19	24	22		
	20	26	25	182	22.750
	21	—	25		
Cell totals	60	50	72		
T_2	25	21	31		
	27	24	32	217	27.125
	—	24	33		
Cell totals	52	69	96		
B totals	112	119	168	399	
B means	22.4	23.8	28.0		24.9375

TABLE 9.2
Cell Means and Marginal Means for the Data in Table 9.1

	B_1	B_2	B_3	T Marginal Means
T_1	20	25	24	23
T_2	26	23	32	27
B marginal means	23	24	28	25

experiment is well balanced. For example, the T_1 mean in Table 9.1 is 22.75 while the T_1 mean in Table 9.2 is 23. Consequently, one question that must be addressed is which of these two methods should be used when calculating treatment main effect means. Which method should be used is discussed in Section 9.5.

The experimental error variance for the data in Table 9.1 is

$$\hat{\sigma}^2 = \frac{1}{N-bt} \sum_{ijk} (y_{ijk} - \bar{y}_{ij.})^2 = 20/10 = 2 \quad \text{with } N-bt = 16-6 = 10 \text{ degrees of freedom}$$

Next consider the experiment as a one-way treatment structure with six treatments and test that all six treatment combination means are equal to one another. That is, consider testing

$$H_0: \mu_{11} = \mu_{12} = \mu_{13} = \mu_{21} = \mu_{22} = \mu_{23}$$

Using Equation 1.8, one gets

$$SSH_0 = \frac{60^2}{3} + \frac{50^2}{2} + \frac{72^2}{3} + \frac{52^2}{2} + \frac{69^2}{3} + \frac{96^2}{3} - \frac{399^2}{16} = 238.9375$$

which is based on 5 degrees of freedom. Thus the F -statistic for testing H_0 is

$$F_c = \frac{238.9375/5}{2} = 23.89$$

which is significant at the $\hat{\alpha}=0.00003$ level. Consequently, H_0 would be rejected, and thus there are significant differences between the means of the six different treatment combinations.

9.4 Interaction and Main Effect Hypotheses

In the previous section, it was determined that there are significant differences among the six treatment combination means. Now it is necessary to see where differences occur. As a first step, consider whether there is significant $T \times B$ interaction for the data given in Table 9.1 and test

$$H_{T \times B}: \mu_{ij} - \mu_{i'j} - \mu_{ij'} + \mu_{i'j'} = 0 \quad \text{for all } i \neq i' \text{ and } j \neq j'$$

This can be accomplished by utilizing the matrix procedure for developing test statistics that was discussed in Section 1.4. The hypothesis $H_{T \times B}$ will be true if and only if

$$\mu_{11} - \mu_{12} - \mu_{21} + \mu_{22} = 0$$

and

$$\mu_{11} - \mu_{13} - \mu_{21} + \mu_{23} = 0$$

In turn, these statements are true if and only if $C\mu = \mathbf{0}$ where

$$C = \begin{bmatrix} 1 & -1 & 0 & -1 & 1 & 0 \\ 1 & 0 & -1 & -1 & 0 & 1 \end{bmatrix} \text{ and } \mu' = [\mu_{11} \mu_{12} \mu_{13} \mu_{21} \mu_{22} \mu_{23}]$$

Then from Equation 1.11,

$$SSH_{T \times B} = [C\hat{\mu}]' [CDC']^{-1} [C\hat{\mu}] = [-8 \ 2] \begin{bmatrix} \frac{10}{6} & \frac{5}{6} \\ \frac{5}{6} & \frac{9}{6} \end{bmatrix}^{-1} \begin{bmatrix} -8 \\ 2 \end{bmatrix}$$

since $D = \text{Diag}[1/3, 1/2, 1/3, 1/2, 1/3, 1/3]$. Thus $SSH_{T \times B} = 776(6/65) = 71.631$ and it is based on 2 degrees of freedom. The corresponding F -statistic is $F_c = (71.631/2)/2 = 17.91$ which is based on 2 and 10 degrees of freedom and is significant at the $\alpha = 0.0005$ level. Other C matrices could be used to test the hypothesis of no interaction, but all such matrices will produce the same test statistic. The reader can verify that this is true.

Next consider testing the equality of the expected row marginal means. This is being done for illustration purposes even though such a test may not be appropriate here because of the significant $T \times B$ interaction. The appropriate hypothesis is $H_T: \bar{\mu}_1 = \bar{\mu}_2$. Note that H_T is true if and only if $C\mu = \mathbf{0}$ where $C = [1 \ 1 \ 1 \ -1 \ -1 \ -1]$. Using Equation 1.11 one gets

$$SSH_T = [C\hat{\mu}]' [CDC']^{-1} [C\hat{\mu}] = [-12][14/6]^{-1}[-12] = 61.714$$

which is based on 1 degree of freedom, and the corresponding F -statistic is $F_c = (61.714/1)/2 = 30.857$, which is significant at the $\alpha = 0.00024$ level. One could also take $C = [\frac{1}{3} \ \frac{1}{3} \ \frac{1}{3} \ -\frac{1}{3} \ -\frac{1}{3} \ -\frac{1}{3}]$. The reader should verify that this second choice for C leads to the same F -test statistic as does the C used above.

Finally, consider testing the equality of the expected column marginal means by testing $H_B: \bar{\mu}_{.1} = \bar{\mu}_{.2} = \bar{\mu}_{.3}$. Note that H_B is true if and only if $C = \mathbf{0}$ where

$$C = \begin{bmatrix} 1 & -1 & 0 & 1 & -1 & 0 \\ 1 & 0 & -1 & 1 & 0 & -1 \end{bmatrix}$$

Using Equation 1.11, one gets

$$\begin{aligned} SSH_B &= [C\hat{\mu}]' [CDC']^{-1} [C\hat{\mu}] = [-2 \ -10] \begin{bmatrix} \frac{10}{6} & \frac{5}{6} \\ \frac{5}{6} & \frac{9}{6} \end{bmatrix}^{-1} \begin{bmatrix} -2 \\ -10 \end{bmatrix} \\ &= [-2 \ -10] \frac{6}{65} \begin{bmatrix} 9 & -5 \\ -5 & 10 \end{bmatrix} \begin{bmatrix} -2 \\ -10 \end{bmatrix} = 77.169 \end{aligned}$$

which is based on 2 degrees of freedom. Hence, the corresponding F -statistic is $F_c = 19.29$ which is significant at the $\alpha = 0.0037$ level.

TABLE 9.3

Analysis of Variance Table for Data in Table 9.1

Source of Variation	df	SS	MS	F	p
Total	15	258.938			
$\mu_{11} = \mu_{12} = \dots = \mu_{23}$	5	238.938	47.79	23.89	0.00003
T	1	61.714	61.71	30.86	0.00024
B	2	77.169	38.58	19.29	0.00037
$T \times B$	2	71.631	35.81	17.91	0.0005
Error	10	20.00	2.00		

The above tests are summarized in the analysis of variance table given in Table 9.3. Note that there are some differences between the results in this analysis and those obtained for the balanced case, discussed in Chapter 7.

1) For balanced data, it is always true that

$$SS_T + SS_B + SS_{T \times B} = SS_{\mu_{11} = \mu_{12} = \dots = \mu_{23}}$$

This is generally not true for unbalanced data.

2) For the balanced case, SS_T , SS_B , and $SS_{T \times B}$ are statistically independent; this is generally not true for unbalanced data.

One does not need to be overly concerned about the fact that the sums of squares for T , B , and $T \times B$ do not add up to equal the sum of squares for testing that all means are equal or that the sums of squares for T , B , and $T \times B$ are not statistically independent, and these points are only being made here because they are true and not because they are issues with which one must deal.

There are other sums of squares that are often associated with analyses of two-way treatment structures. Two popular sets of these are examined in Chapter 10.

9.5 Population Marginal Means

Often the experimenter is interested in making comparisons about and between the possibilities of each main effect. In the balanced case, one may compare $\bar{\mu}_1$, $\bar{\mu}_2$, ..., $\bar{\mu}_t$ with one another. As mentioned earlier, these means are called the population marginal means for both the balanced and unbalanced case. The best estimate of $\bar{\mu}_i$ is

$$\hat{\bar{\mu}}_i = \frac{1}{b} \sum_{j=1}^b \hat{\mu}_{ij} = \bar{\hat{\mu}}_i, \quad i = 1, 2, \dots, t \quad (9.4)$$

The estimated standard error of $\hat{\bar{\mu}}_i$ is

$$\widehat{s.e.}(\hat{\bar{\mu}}_i) = \frac{\hat{\sigma}}{b} \sqrt{\sum_{j=1}^b \frac{1}{n_{ij}}}, \quad i = 1, 2, \dots, t \quad (9.5)$$

The best estimate of $\bar{\mu}_{\cdot j}$ is

$$\hat{\mu}_{\cdot j} = \frac{1}{t} \sum_{i=1}^t \hat{\mu}_{ij} = \bar{\mu}_{\cdot j}, \quad j = 1, 2, \dots, b \quad (9.6)$$

The estimated standard error of $\hat{\mu}_{\cdot j}$ is

$$\widehat{s.e.}(\hat{\mu}_{\cdot j}) = \frac{\hat{\sigma}}{t} \sqrt{\sum_{i=1}^t \frac{1}{n_{ij}}}, \quad j = 1, 2, \dots, b \quad (9.7)$$

It should be noted that, in unbalanced data problems, it is generally the case that $\hat{\mu}_{\cdot i}$ will be different from $\bar{y}_{i\cdot}$ and that $\hat{\mu}_{\cdot j}$ will be different from $\bar{y}_{\cdot j}$. In the example, $\bar{y}_{1\cdot} = 22.75$ and $\hat{\mu}_{1\cdot} = 23$.

The estimators $\hat{\mu}_{\cdot i}$ and $\hat{\mu}_{\cdot j}$ are unbiased estimates of $\bar{\mu}_{\cdot i}$ and $\bar{\mu}_{\cdot j}$, respectively. It can be noted that the estimators $\bar{y}_{i\cdot}$ and $\bar{y}_{\cdot j}$ are unbiased estimates of

$$\tilde{\mu}_{i\cdot} = \left(\sum_{j=1}^b n_{ij} \mu_{ij} \right) / n_{i\cdot}$$

and

$$\tilde{\mu}_{\cdot j} = \left(\sum_{i=1}^t n_{ij} \mu_{ij} \right) / n_{\cdot j}$$

respectively. That is, $\bar{y}_{i\cdot}$ provides an unbiased estimator of a weighted average of the cell mean parameters in the i th row using the sample sizes within the row cells as weights. Likewise, $\bar{y}_{\cdot j}$ provides an unbiased estimator of a weighted average of the cell mean parameters in the j th column using the sample sizes within the column cells as weights. When using computing packages to analyze data, it is extremely important to determine whether estimates of the main effect means are calculated as $\hat{\mu}_{\cdot i}$ and $\hat{\mu}_{\cdot j}$ or as $\bar{y}_{i\cdot}$ and $\bar{y}_{\cdot j}$. In most designed experiments, one will want to use the estimators $\hat{\mu}_{\cdot i}$ and $\hat{\mu}_{\cdot j}$.

For the data in Table 9.1, the estimated population marginal means are shown in Table 9.2 as $\hat{\mu}_{1\cdot} = 23$, and $\hat{\mu}_{2\cdot} = 27$, and $\hat{\mu}_{\cdot 1} = 23$, and $\hat{\mu}_{\cdot 2} = 27$, and $\hat{\mu}_{\cdot 3} = 28$. The estimated standard errors of these estimates are

$$\widehat{s.e.}(\hat{\mu}_{1\cdot}) = \frac{\sqrt{2}}{3} \sqrt{\frac{1}{3} + \frac{1}{2} + \frac{1}{3}} = 0.51$$

$$\widehat{s.e.}(\hat{\mu}_{2\cdot}) = \frac{\sqrt{2}}{3} \sqrt{\frac{1}{2} + \frac{1}{3} + \frac{1}{3}} = 0.51$$

$$\widehat{s.e.}(\hat{\mu}_{\cdot 1}) = \frac{\sqrt{2}}{2} \sqrt{\frac{1}{3} + \frac{1}{2}} = 0.65$$

$$\widehat{s.e.}(\hat{\mu}_{\cdot 2}) = \frac{\sqrt{2}}{2} \sqrt{\frac{1}{2} + \frac{1}{3}} = 0.65$$

and

$$\widehat{s.e.}(\hat{\mu}_{.3}) = \frac{\sqrt{2}}{2} \sqrt{\frac{1}{3} + \frac{1}{3}} = 0.58$$

To make inferences about linear combinations of the population marginal means, say $\sum_i c_i \bar{\mu}_i$, or $\sum_j d_j \bar{\mu}_{.j}$, it can be shown that

$$\frac{\sum_i c_i \hat{\mu}_i - \sum_i c_i \bar{\mu}_i}{\hat{\sigma} \sqrt{\sum_i c_i^2 \left(\sum_j \frac{1}{n_{ij}} \right)}} \sim t(v) \quad (9.8)$$

and that

$$\frac{\sum_j d_j \hat{\mu}_{.j} - \sum_j d_j \bar{\mu}_{.j}}{\hat{\sigma} \sqrt{\sum_j d_j^2 \left(\sum_i \frac{1}{n_{ij}} \right)}} \sim t(v) \quad (9.9)$$

The formulas in Equations 9.8 and 9.9 can be obtained as special cases of Equation 1.4. For example, the test statistic that tests $H_T: \bar{\mu}_{.1} = \bar{\mu}_{.2}$ using a t -test is

$$t_c = \frac{\hat{\mu}_{.1} - \hat{\mu}_{.2}}{\hat{\sigma} \sqrt{\sum_j \frac{1}{n_{1j}} + \sum_j \frac{1}{n_{2j}}}} = \frac{23 - 27}{\frac{\sqrt{2}}{3} \sqrt{\left(\frac{1}{3} + \frac{1}{2} + \frac{1}{3} \right) + \left(\frac{1}{2} + \frac{1}{3} + \frac{1}{3} \right)}} = \frac{-4}{\frac{1.414}{3} \sqrt{\frac{7}{3}}} = \frac{-4}{0.72} = -5.55$$

which is significant at the $\alpha = 0.00024$ level. A 95% confidence interval for $\bar{\mu}_{.1} - \bar{\mu}_{.2}$ is

$$\begin{aligned} \hat{\mu}_{.1} - \hat{\mu}_{.2} &\mp t_{\frac{\alpha}{2}, v} \frac{\hat{\sigma}}{t} \sqrt{\sum_i \frac{1}{n_{i1}} + \sum_j \frac{1}{n_{i2}}} = 23 - 24 \mp t_{.025, v} \cdot \frac{1.414}{2} \sqrt{\left(\frac{1}{3} + \frac{1}{2} \right) + \left(\frac{1}{2} + \frac{1}{3} \right)} \\ &= -1 \mp (2.228)(0.91) = -1 \mp 2.03 \end{aligned}$$

9.6 Simultaneous Inferences and Multiple Comparisons

There are few good procedures available for making multiple comparisons in two-way experiments in which there are unequal numbers of observations per treatment combination. If one wants to compare all pairs of two-way cell means, then any of the techniques discussed in Chapter 3 can be used simply by considering the two-way experiment as a

one-way treatment structure experiment. In this case the reader should see the recommendations given in Section 3.2. If one wishes to make multiple comparisons on the population marginal means, it is recommended that one use t -tests based on Equations 9.8 and 9.9. Use the given significance levels if the corresponding F -test for comparing the corresponding marginal means is significant. If the F -test is not significant, then it is still recommended that one use these t -tests. However, in this case, one should use Bonferroni's method and claim two population marginal means to be significantly different only when the calculated significance level is less than α/p where α is the selected experimentwise error rate and p is the number of comparisons that are being considered prior to collecting the data. If it is determined that there is interaction in the data, then one would likely want to compare the effects of one of the treatments for each possibility of the other treatment. That is, one would likely want to compare the cell means within each row to one another and the compare the cell means within each column to one another. There are $bt(t + 1)/2 + bt(b + 1)/2$ such pairwise comparisons. If the F -test comparing all means is significant, then one can use the actual significance levels for all such pairwise comparisons. This use is the equivalent of a Fisher's LSD procedure. If the F -test comparing all means is not significant, then one would take a Bonferroni approach. For data snooping and unplanned comparisons, one should use Scheffé's procedure. The test statistics described in Section 9.4 can be obtained automatically with many statistical computing packages. Since these packages employ the effects model, the interested reader should see Section 10.7.

9.7 Concluding Remarks

This chapter is the first of seven considering the analysis of two-way treatment structures with unequal subclass numbers. The analyses presented in this chapter were obtained by using the means model. An important assumption made was that all treatment combinations were observed at least once. Procedures for testing main effect and interaction hypotheses were obtained as special cases of the general techniques introduced in Chapter 1. Population marginal means were defined, and procedures for making inferences on the population marginal means were given. In Chapter 10, similar kinds of questions are answered by utilizing the effects model; however, the authors hope to make all readers equally comfortable with both the means model and the effects model.

9.8 Exercises

- 9.1 Consider an experiment which was conducted to study the effectiveness of four types of drugs on three different diseases. The experiment was conducted in a completely randomized design with six patients assigned to each drug \times disease combination initially. Data on the variable to be analyzed (change in disease score after two weeks) is missing on 13 of the 72 patients. The data are given below.
- 1) Analyze this data with SAS-GLM using a means model. Include the following options on the MODEL statement: E and SOLUTION.

- 2) Use LSMEANS to perform a multiple comparison procedure that has an $\alpha=0.05$ experimentwise error rate when comparing all 12 treatment combinations with one another.
- 3) Using CONTRAST and/or ESTIMATE options determine the observed significance levels for the following hypotheses:
 - a) $(\bar{\mu}_{1.} + \bar{\mu}_{2.})/2 = \bar{\mu}_{3.} = \bar{\mu}_{4.}$
 - b) $\mu_{11} = \mu_{12} = \mu_{13}$
 - c) $(\bar{\mu}_{1.} + \bar{\mu}_{2.})/2 = \bar{\mu}_{3.}$
 - d) $\mu_{12} = \mu_{22} = \mu_{32} = \mu_{42}$
 - e) $\bar{\mu}_{3.} - \bar{\mu}_{4.} = 0$
 - f) $\mu_{12} - \mu_{13} - \mu_{32} + \mu_{33} = 0$
 - g) $\mu_{22} - \mu_{32} = 0$
- 4) Construct 95% confidence intervals for the contrasts in e–g above.
- 5) If there is significant interaction in these data, see if you can determine which treatment combinations are responsible for the interaction.

	Disease 1	Disease 2	Disease 3
Drug 1	42, 44, 36, 13, 19, 22	33, 26, 33, 21	31, -3, 25, 24
Drug 2	28, 23, 34, 42, 13	34, 33, 31, 36	3, 26, 28, 32, 4, 16
Drug 3	1, 24, 9, 22, -2, 15	21, 1, 9, 3	11, 9, 7, 1, -6
Drug 4	1, 29, 19	22, 7, 25, 5, 12	27, 12, -5, 16, 15, 12

10

Using the Effects Model to Analyze Balanced Two-Way Treatment Structures with Unequal Subclass Numbers

Chapter 9 contained a discussion of the analysis of two-way treatment structures having unequal subclass numbers using the means model. This chapter considers using an effects model in the same situation. All questions that can be answered by using the effects model can also be answered by using the means model, and vice versa. The effects model is being discussed because it is often an important tool when using statistical computing packages to analyze two-way treatment structures as statistical software is often programmed to automatically produce test statistics for main effects and interaction effects as well as estimates of marginal means, two-way means and their estimated standard errors.

10.1 Model Definition

The effects model corresponding to the means model (9.1) is defined by

$$y_{ijk} = \mu + \tau_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}, \quad i = 1, 2, \dots, t; \quad j = 1, 2, \dots, b; \quad k = 1, 2, \dots, n_{ij} \quad (10.1)$$

where $\varepsilon_{ijk} \sim i.i.d. N(0, \sigma^2)$.

10.2 Parameter Estimates and Type I Analysis

Other sums of squares are often associated with analyses of two-way treatment structures having unequal subclass numbers besides those introduced in Chapter 9. Two of these are examined in this chapter. The first set of sums of squares involves fitting the two-way effects model to the observed data in a sequential manner using generalizations of

the model comparison method described in Section 1.6. One sequence of steps often used is the following:

- Step 1. Fit $y_{ijk} = \mu + \varepsilon_{ijk}$ and denote its residual sum of squares by RSS_1 .
- Step 2. Fit $y_{ijk} = \mu + \tau_i + \varepsilon_{ijk}$ and denote its residual sum of squares by RSS_2 .
- Step 3. Fit $y_{ijk} = \mu + \tau_i + \beta_j + \varepsilon_{ijk}$ and denote its residual sum of squares by RSS_3 .
- Step 4. Fit $y_{ijk} = \mu + \tau_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}$ and denote its residual sum of squares by RSS_4 .

The quantity RSS_i is the residual sum of squares after fitting the model in the i th step, $i = 1, 2, 3, 4$. The difference between RSS_1 and RSS_2 , denoted by $R(\tau|\mu)$, is called the reduction due to τ adjusted for μ ; that is, $R(\tau|\mu) = RSS_1 - RSS_2$. This reduction gives the amount by which one can reduce the residual sum of squares of the model in Step 1 by considering a model with τ_i included as well. The larger the value of $R(\tau|\mu)$, the more important it is to have τ_i in the model. Thus, $R(\tau|\mu)$ is a measure of the effect of the different possibilities for treatment T . The quantity $R(\beta|\mu, \tau) = RSS_2 - RSS_3$ is called the reduction due to β adjusted for both μ and τ . It gives the additional amount by which one can reduce the residual sum of squares of the model in step 2 by also including β_j in the model. $R(\beta|\mu, \tau)$ is a measure of the effect of different possibilities for treatment B above and beyond the effect of treatment T . Finally, the quantity $R(\gamma|\mu, \tau, \beta) = RSS_3 - RSS_4$ is called the reduction due to γ adjusted for μ , τ , and β . It gives the additional amount by which one can reduce the residual sum of squares of the model in step 3 by adding interaction parameters, the γ_{ij} , to the model. Clearly, $R(\gamma|\mu, \tau, \beta)$ is a measure of interaction, since the model in step 3 is an additive model that holds if and only if there is no interaction.

An analysis of variance table corresponding to this sequential analysis is given in Table 10.1. This analysis is called a type I analysis. The sums of squares in the last four lines of Table 10.1 are statistically independent, and the ratios of the T , B , and $T \times B$ mean squares to the error mean square all have noncentral F -distributions. It is quite interesting and informative to determine exactly the hypothesis that each of the F -statistics in Table 10.1 is testing. The hypotheses that are being tested are given in Section 10.4.

To illustrate, each of the four models required for the type I analysis are fit to the data in Table 9.1. An understanding of Chapter 6 is required to follow the computations made here. However, such an understanding is not necessary for those readers willing to let statistical computing packages perform the required computations. Those readers who are interested in the details may consider the next few pages. The model given in step 1 is

TABLE 10.1

Analysis of Variance Table for a Sequential Analysis (Type I Analysis)

Source of Variation	df	SS	MS	F
Total	$N - 1$	RSS_1		
T	$t - 1$	$R(\tau \mu)$	$R(\tau \mu)/(t - 1)$	$\frac{TMS}{\hat{\sigma}^2}$
B	$b - 1$	$R(\beta \mu, \tau)$	$R(\beta \mu, \tau)/(b - 1)$	$\frac{BMS}{\hat{\sigma}^2}$
$T \times B$	$(t - 1)(b - 1)$	$R(\gamma \mu, \tau, \beta)$	$R(\gamma \mu, \tau, \beta)/[(b - 1)(t - 1)]$	$\frac{(T \times B)MS}{\hat{\sigma}^2}$
Error	$N - tb$	RSS_4	$\hat{\sigma}^2$	

$y_{ijk} = \mu + \varepsilon_{ijk}$. The best estimate of μ in this model is $\hat{\mu} = \bar{y}_{...} = 24.9375$, the average of all of the observations, and the residual sum of squares is

$$RSS_1 = \sum_{i,j,k} (y_{ijk} - \hat{\mu})^2 = \sum_{i,j,k} y_{ijk}^2 - n_{..} \bar{y}_{...}^2 = 10209 - 16(24.9375)^2 = 258.9375$$

and this residual sum of squares is based on $n_{..} - 1 = 16 - 1 = 15$ degrees of freedom.

The normal equations for the model defined in step 2 are:

$$\begin{bmatrix} 16 & 8 & 8 \\ 8 & 8 & 0 \\ 8 & 0 & 8 \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\tau}_1 \\ \hat{\tau}_2 \end{bmatrix} = \begin{bmatrix} 399 \\ 182 \\ 217 \end{bmatrix}$$

One possible solution to these equations is obtained by using the set-to-zero restrictions discussed in Section 6.2, which yields the solution $\hat{\tau}_2 = 0$, $\hat{\tau}_1 = -4.375$, and $\hat{\mu} = 27.125$ (recall from Chapter 6 that a unique solution does not exist). The residual sum of squares is

$$\begin{aligned} RSS_2 &= \sum_{i,j,k} (y_{ijk} - \hat{\mu} - \hat{\tau}_i)^2 = \sum_{i,j,k} y_{ijk}^2 - \hat{\mu} \cdot y_{...} - \hat{\tau}_1 \cdot y_{1..} - \hat{\tau}_2 \cdot y_{2..} \\ &= 10209 - 27.125 \cdot 399 - (-4.375) \cdot 182 - 0 \cdot 217 \\ &= 182.375 \end{aligned}$$

which is based on $16 - 2 = 14$ degrees of freedom. Thus, $R(\tau|\mu) = 258.9375 - 182.375 = 76.5625$ and is based on $15 - 14 = 1$ degree of freedom.

The normal equations for the model in step 3 are:

$$\begin{bmatrix} 16 & 8 & 8 & 5 & 5 & 6 \\ 8 & 8 & 0 & 3 & 2 & 3 \\ 8 & 0 & 8 & 2 & 3 & 3 \\ 5 & 3 & 2 & 5 & 0 & 0 \\ 5 & 2 & 3 & 0 & 5 & 0 \\ 6 & 3 & 3 & 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\tau}_1 \\ \hat{\tau}_2 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} 399 \\ 182 \\ 217 \\ 112 \\ 119 \\ 168 \end{bmatrix}$$

To obtain a solution, one can let $\hat{\tau}_2 = 0$ and $\hat{\beta}_3 = 0$ (see Chapter 6). Then this system of equations can be reduced to an equivalent system by deleting the rows and columns that correspond to τ_2 and β_3 . The system then reduces to:

$$\begin{bmatrix} 16 & 8 & 5 & 5 \\ 8 & 8 & 3 & 2 \\ 5 & 3 & 5 & 0 \\ 5 & 2 & 0 & 5 \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\tau}_1 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} 399 \\ 182 \\ 112 \\ 119 \end{bmatrix}$$

The solution to this reduced system is $\hat{\mu} = 30.154$, $\hat{\tau}_1 = -4.308$, $\hat{\beta}_1 = -5.169$, $\hat{\beta}_2 = -4.631$. The residual sum of squares for this model is

$$\begin{aligned} RSS_3 &= \sum_{i,j,k} y_{ijk}^2 - \hat{\mu} \cdot y_{...} - \hat{\tau}_1 \cdot y_{1..} - \hat{\beta}_1 \cdot y_{.1..} - \hat{\beta}_2 \cdot y_{.2..} \\ &= 10209 - 30.154 \cdot 399 - (-4.308) \cdot 182 - (-5.169) \cdot 112 - (-4.631) \cdot 119 \\ &= 91.631 \end{aligned}$$

and this residual sum of squares is based on $16 - 4 = 12$ degrees of freedom. Thus $R(\beta|\mu, \tau) = RSS_2 - RSS_3 = 182.375 - 91.631 = 90.744$ which is based on $14 - 12 = 2$ degrees of freedom.

The normal equations for the model in step 4 are

$$\left[\begin{array}{ccccccccc|c} 16 & 8 & 8 & 5 & 5 & 6 & 3 & 2 & 3 & 2 & 3 & 3 \\ 8 & 8 & 0 & 3 & 2 & 3 & 3 & 2 & 3 & 0 & 0 & 0 \\ 8 & 0 & 8 & 2 & 3 & 3 & 0 & 0 & 0 & 2 & 3 & 3 \\ 5 & 3 & 2 & 5 & 0 & 0 & 3 & 0 & 0 & 2 & 0 & 0 \\ 5 & 2 & 3 & 0 & 5 & 0 & 0 & 2 & 0 & 0 & 3 & 0 \\ 6 & 3 & 3 & 0 & 0 & 6 & 0 & 0 & 3 & 0 & 0 & 3 \\ 3 & 3 & 0 & 3 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 & 2 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 3 & 3 & 0 & 0 & 0 & 3 & 0 & 0 & 3 & 0 & 0 & 0 \\ 2 & 0 & 2 & 2 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 3 & 0 & 3 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 3 & 0 & 3 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 3 \end{array} \right] \left[\begin{array}{c} \hat{\mu} \\ \hat{\tau}_1 \\ \hat{\tau}_2 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \hat{\gamma}_{11} \\ \hat{\gamma}_{12} \\ \hat{\gamma}_{13} \\ \hat{\gamma}_{21} \\ \hat{\gamma}_{22} \\ \hat{\gamma}_{23} \end{array} \right] = \left[\begin{array}{c} 399 \\ 182 \\ 217 \\ 112 \\ 119 \\ 168 \\ 60 \\ 50 \\ 72 \\ 52 \\ 69 \\ 96 \end{array} \right]$$

To obtain a solution, let $\hat{\tau}_2 = 0$, $\hat{\beta}_3 = 0$, $\hat{\gamma}_{13} = 0$, $\hat{\gamma}_{21} = 0$, $\hat{\gamma}_{22} = 0$, $\hat{\gamma}_{23} = 0$ (see Chapter 6). Using the reduction technique, the system reduces to

$$\left[\begin{array}{cccccc|c} 16 & 8 & 5 & 5 & 3 & 2 & \hat{\mu} \\ 8 & 8 & 3 & 2 & 3 & 2 & \hat{\tau}_1 \\ 5 & 3 & 5 & 0 & 3 & 0 & \hat{\beta}_1 \\ 5 & 2 & 0 & 5 & 0 & 2 & \hat{\beta}_2 \\ 3 & 3 & 3 & 0 & 3 & 0 & \hat{\gamma}_{11} \\ 2 & 2 & 0 & 2 & 0 & 2 & \hat{\gamma}_{12} \end{array} \right] \left[\begin{array}{c} \hat{\mu} \\ \hat{\tau}_1 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\gamma}_{11} \\ \hat{\gamma}_{12} \end{array} \right] = \left[\begin{array}{c} 399 \\ 182 \\ 112 \\ 119 \\ 60 \\ 50 \end{array} \right]$$

The solution to this reduced system is $\hat{\mu} = 32$, $\hat{\tau}_1 = -8$, $\hat{\beta}_1 = -6$, $\hat{\beta}_2 = -9$, $\hat{\gamma}_{11} = 2$, $\hat{\gamma}_{12} = 10$. Thus $\hat{\sigma}^2 = 2$, since the residual sum of squares of the full model is $RSS_4 = 10,209 - 10,189 = 20$ and is based on 10 degrees of freedom. Also, $R(\gamma|\mu, \tau, \beta) = RSS_3 - RSS_4 = 91.631 - 20 = 71.631$ and it is based on $12 - 10 = 2$ degrees of freedom. All of the preceding results can be summarized in an analysis of variance table like the one given in Table 10.2.

TABLE 10.2

Type I Analysis of Variance Table

Source of Variation	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	α
Total	15	258.9375			
<i>T</i>	1	76.5625	76.5625	38.28	0.0001
<i>B</i>	2	90.744	45.372	22.69	0.0002
<i>T</i> \times <i>B</i>	2	71.631	35.815	17.91	0.0005
Error	10	20.000	2.0		

The sum of squares as well as the test statistic for interaction in the type I analysis is the same as that obtained using the means model and the matrix procedure in Chapter 9; however, the two procedures give different sums of squares and test statistics for both of the *T* and *B* main effects. The data in a two-way treatment structure could also be analyzed by fitting μ first, then β , then τ , and finally γ . The only new sums of squares required that are not already given in Table 10.2 are $R(\beta|\mu)$ and $R(\tau|\mu, \beta)$. For unbalanced data cases, the corresponding *F*-tests for the *T* and *B* main effects will usually be different from those obtained by the fitting the effects in the order given in Table 10.1 and/or Table 10.2.

It should be recalled from Chapter 6 that the parameter estimates, $\hat{\mu} = 32$, $\hat{\tau}_1 = -8$, $\hat{\beta}_1 = -6$, $\hat{\beta}_2 = -9$, $\hat{\gamma}_{11} = 2$, $\hat{\gamma}_{12} = 10$, are not unbiased estimates of the corresponding parameters, μ , τ_1 , β_1 , β_2 , γ_{11} , and γ_{12} . Indeed, these individual parameters are not estimable. Under the set-to-zero restrictions used to solve the normal equations, it is possible to show that

$$\begin{aligned}\hat{\mu} &\text{ is an unbiased estimate of } \mu + \tau_2 + \beta_3 + \gamma_{23} \\ \hat{\tau}_1 &\text{ is an unbiased estimate of } \tau_1 - \tau_2 + \gamma_{13} - \gamma_{23} \\ \hat{\beta}_1 &\text{ is an unbiased estimate of } \beta_1 - \beta_3 + \gamma_{21} - \gamma_{23} \\ \hat{\beta}_2 &\text{ is an unbiased estimate of } \beta_2 - \beta_3 + \gamma_{22} - \gamma_{23} \\ \hat{\gamma}_{11} &\text{ is an unbiased estimate of } \gamma_{11} - \gamma_{13} - \gamma_{21} + \gamma_{23}\end{aligned}\tag{10.2}$$

and

$$\hat{\gamma}_{12} \text{ is an unbiased estimate of } \gamma_{12} - \gamma_{13} - \gamma_{22} + \gamma_{23}.$$

More about estimable functions and their estimates can be found in the next section.

10.3 Using Estimable Functions in SAS

The SAS®-GLM procedure has an option that can be used to identify estimable functions of the model parameters. (Readers who do not use SAS may skip this section.) Since it is easiest to describe estimable functions by using an example, consider again the data in Table 9.1. A SAS-GLM analysis of this data can be obtained by using the following statements:

```
PROC GLM;
CLASSES T B;
MODEL y=T B T*B/<selected options>;
```

TABLE 10.3

General Form of Estimable Functions

Effect	Coefficients	
Intercept		L1
T	1	L2
T	2	$L1 - L2$
B	1	L4
B	2	L5
B	3	$L1 - L4 - L5$
$T \times B$	1 1	L7
$T \times B$	1 2	L8
$T \times B$	1 3	$L2 - L7 - L8$
$T \times B$	2 1	$L4 - L7$
$T \times B$	2 2	$L5 - L8$
$T \times B$	2 3	$L1 - L2 - L4 - L5 + L7 + L8$

Many options can be used with the MODEL statement. One of the most important of these is the E option. This option asks SAS-GLM to print a general form of the estimable functions of the model parameters. Recall from Chapter 6 that all linear functions of the parameters in a design model are not necessarily estimable. Using the E option, SAS-GLM prints information that one can use to determine those linear combinations of the parameters that are estimable and those that are not estimable.

The general form of an estimable function given by SAS-GLM is shown in Table 10.3. It means that a linear function of the model parameters $\ell' \beta$ where $\ell' = [\ell_1, \ell_2, \dots, \ell_{12}]$ and $\beta' = [\mu, \tau_1, \tau_2, \beta_1, \beta_2, \beta_3, \gamma_{11}, \gamma_{12}, \gamma_{13}, \gamma_{21}, \gamma_{22}, \gamma_{23}]$ is estimable if and only if there exist constants L1, L2, L4, L5, L7, and L8 such that

$$\begin{aligned} \ell' \beta = & (L1)\mu + (L2)\tau_1 + (L1 - L2)\tau_2 + (L4)\beta_1 + (L5)\beta_2 + (L1 - L4 - L5)\beta_3 + (L7)\gamma_{11} \\ & + (L8)\gamma_{12} + (L2 - L7 - L8)\gamma_{13} + (L4 - L7)\gamma_{21} + (L5 - L8)\gamma_{22} \\ & + (L1 - L2 - L4 - L5 + L7 + L8)\gamma_{23} \end{aligned}$$

For example, from the general form of estimable functions, we can see:

- 1) That μ is not estimable, since in order for μ to be estimable, one would at least need to have $L1 = 1$, $L2 = 0$, and $L1 - L2 = 0$ all at the same time, and these three equations cannot all be true at the same time.
- 2) That τ_1 is not estimable, since in order for τ_1 to be estimable, one would at least need to have $L1 = 0$, $L2 = 1$, and $L1 - L2 = 0$ all at the same time which cannot be true.
- 3) That $\tau_1 - \tau_2$ is not estimable, since in order for $\tau_1 - \tau_2$ to be estimable, one would at least need to have $L1 = 0$, $L2 = 1$, $L1 - L2 = -1$, $L4 = 0$, $L5 = 0$, $L1 - L4 - L5 = 0$, $L7 = 0$, $L8 = 0$, and $L2 - L7 - L8 = 0$ all at the same time. However, $L2 = 1$, $L7 = 0$, $L8 = 0$, and $L2 - L7 - L8 = 0$ cannot all be true at the same time.

It is clear that there are many functions of the model parameters that are not estimable. Are there any interesting functions of the model parameter that are estimable? From Chapter 6, one knows that there are estimable functions of the model parameters. In fact, a basis set of estimable functions of the model parameters is defined to be a set of linearly independent

TABLE 10.4

A Basis Set of Estimable Functions

<i>L1</i>	<i>L2</i>	<i>L4</i>	<i>L5</i>	<i>L7</i>	<i>L8</i>	Estimable Function
1	0	0	0	0	0	$\mu + \tau_2 + \beta_3 + \gamma_{23}$
0	1	0	0	0	0	$\tau_1 - \tau_2 + \gamma_{13} - \gamma_{23}$
0	0	1	0	0	0	$\beta_1 - \beta_3 + \gamma_{21} - \gamma_{23}$
0	0	0	1	0	0	$\beta_2 - \beta_3 + \gamma_{22} - \gamma_{23}$
0	0	0	0	1	0	$\gamma_{11} - \gamma_{13} - \gamma_{21} + \gamma_{23}$
0	0	0	0	0	1	$\gamma_{12} - \gamma_{13} - \gamma_{22} + \gamma_{23}$

functions that are estimable and so that every other estimable function can be written as a linear combination of the estimable functions in this basis set. In the general form of the estimable functions, one can see that one is free to choose six of the *L*, namely, *L1*, *L2*, *L4*, *L5*, *L7*, and *L8*. Thus there are six linearly independent estimable functions in a basis set. One basis set that can be easily obtained is to successively let one of the six *L* be equal to 1 and let all of the remaining *L* be equal to 0. For example, taking *L1* = 1 and *L2* = *L4* = *L5* = *L7* = *L8* = 0, the general form simplifies to $\mu + \tau_2 + \beta_3 + \gamma_{23}$, giving one linear function of the model parameters that is estimable. Taking *L2* = 1 and the other *L* equal to 0, the general form simplifies to $\tau_1 - \tau_2 + \gamma_{13} - \gamma_{23}$, giving a second linear function of the model parameters that is estimable. Continuing in this manner produces the basis set of estimable functions given in Table 10.4.

Note that the number of linear functions in the basis set is equal to six, which is equal to the rank of the $X'X$ matrix as discussed in Chapter 6 and also equal to the number of treatment combinations in this 2×3 experiment. Also note that the estimable functions in Table 10.4 are exactly the same as the functions that are being estimated by the solution to the normal equations that satisfies the set-to-zero restrictions as shown in Equation 10.2. This is not a coincidence; it is always true when using SAS-GLM to analyze data that one can find the functions being estimated by the set-to-zero solution to the normal equations by simply letting each of the *L* in the general form that one is free to choose be equal to one and letting the other *L* equal to zero.

Basis sets of estimable functions are not unique, and another basis set of estimable functions is given in Table 10.5, along with the values of *L1*, *L2*, *L4*, *L5*, *L7*, and *L8* that produced this basis set.

When one uses the SOLUTION option on the SAS-GLM model statement, the computer prints out least squares estimates of the model parameters using the set-to-zero restrictions.

TABLE 10.5

Another Basis Set of Estimable Functions

<i>L1</i>	<i>L2</i>	<i>L4</i>	<i>L5</i>	<i>L7</i>	<i>L8</i>	Estimable Function
1	1/2	1/3	1/3	1/6	1/6	$\mu + \bar{\tau}_. + \bar{\beta}_. + \bar{\gamma}_..$
0	1	0	0	1/3	1/3	$\tau_1 - \tau_2 + \bar{\gamma}_{1.} - \bar{\gamma}_{2.}$
0	0	1	0	1/2	0	$\beta_1 - \beta_3 + \bar{\gamma}_{1.} - \bar{\gamma}_{3.}$
0	0	0	1	0	1/2	$\beta_2 - \beta_3 + \bar{\gamma}_{2.} - \bar{\gamma}_{3.}$
0	0	0	0	1	0	$\gamma_{11} - \gamma_{13} - \gamma_{21} + \gamma_{23}$
0	0	0	0	0	1	$\gamma_{12} - \gamma_{13} - \gamma_{22} + \gamma_{23}$

TABLE 10.6

Results Obtained with SOLUTION Option in SAS-GLM

Parameter		Estimate	Standard Error	t-Value	Pr > t
Intercept		32.00000000 B	0.81649658	39.19	<0.0001
T	1	-8.00000000 B	1.15470054	-6.93	<0.0001
T	2	0.00000000 B	—	—	—
B	1	-6.00000000 B	1.29099445	-4.65	0.0009
B	2	-9.00000000 B	1.15470054	-7.79	<0.0001
B	3	0.00000000 B	—	—	—
T × B	1 1	2.00000000 B	1.73205081	1.15	0.2751
T × B	1 2	10.00000000 B	1.73205081	5.77	0.0002
T × B	1 3	0.00000000 B	—	—	—
T × B	2 1	0.00000000 B	—	—	—
T × B	2 2	0.00000000 B	—	—	—
T × B	2 3	0.00000000 B	—	—	—

Note: The $X'X$ matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

The results of the SOLUTION option are shown in Table 10.6. Note that the results of the solution option agree with the set-to-zero solution given in Section 10.2 which is repeated here for convenience:

$$\hat{\mu} = 32, \hat{\tau}_1 = 8, \hat{\beta}_1 = -6, \hat{\beta}_2 = -9, \hat{\gamma}_{11} = 2, \hat{\gamma}_{12} = 10$$

However, as pointed out before, these least squares estimates do not estimate their respective parameters. In fact, as shown previously, the individual parameters are not estimable. SAS-GLM indicates this by putting the letter B next to each of the least squares estimates in Table 10.6. The functions of the parameters that these estimators are really unbiased estimates of are those given in Table 10.4. That is, $\hat{\mu} = 32$ is the best unbiased estimate of $\mu + \tau_2 + \beta_3 + \gamma_{23}$, and $\hat{\tau}_1 = -8$ is the best unbiased estimate of $\tau_1 - \tau_2 + \gamma_{13} - \gamma_{23}$, $\hat{\tau}_2 = 0$ is estimating zero (which it does a good job of doing, too), $\hat{\beta}_1 = -6$ is the best unbiased estimate of $\beta_1 - \beta_3 + \gamma_{11} - \gamma_{23}$, and so on. The standard errors printed in Table 10.6 are the actual estimated standard errors of the estimators. That is, $\widehat{s.e.}(\hat{\mu}) = 0.8165$, $\widehat{s.e.}(\hat{\tau}_1) = 1.1547$, and so on. The t-test statistics in Table 10.6 test that the corresponding function being estimated is equal to zero. For example $t = 39.19$, which corresponds to $\hat{\mu}$, tests $H_0: \mu + \tau_2 + \beta_3 + \gamma_{23} = 0$. Such tests are usually not very interesting. By using CONTRAST or ESTIMATE statements in the SAS-GLM procedure, one can make inferences about any estimable linear combination of the parameters that one chooses to consider. The linear combination need not be a contrast, but it must be estimable. Fortunately, SAS-GLM always checks whether or not the specified linear combination is an estimable function. If it is, the ESTIMATE statement gives its best unbiased estimate, its estimated standard error, and a t-statistic and p-value that tests whether the parameter function being estimated is equal to zero or not. The CONTRAST statement allows the simultaneous testing of several estimable functions, and gives a F-statistic along with its p-value, as described in Chapter 1.

The proper use of the ESTIMATE statement for our example requires the following form:

```
ESTIMATE 'label' INTERCEPT c1 T c2 c3 B c4 c5 c6 T*B c7 c8 c9 c10 c11 c12;
```

If all of the coefficients of a particular effect are zero, that effect and its coefficients do not need to be included in the ESTIMATE statement. To use the CONTRAST statement, one needs only to replace the word ESTIMATE with CONTRAST in the above form. For example, to obtain the best estimates of the estimable functions given in Table 10.5, one would use:

1. ESTIMATE 'OVERALL MEAN' INTERCEPT 1 T -5 -5 B .33333 .33333 .33333
T*B .16667 .16667 .16667 .16667 .16667 .16667;
 2. ESTIMATE 'TI -T2' T 1 -1 T*B .33333 .33333 .33333 -.33333 -.33333
-.33333;
 3. ESTIMATE 'B1 -B3' B 1 0 -1 T*B .5 0 -.5 .5 0 -.5;
 4. ESTIMATE 'B2 -B3' B 0 1 -1 T*B 0 .5 -.5 0 .5 -.5;
 5. ESTIMATE 'INTI' T*B 1 0 -1 -1 0 1;
 6. ESTIMATE 'INT2' T*B 0 1 -1 0 -1 1;
-

10.4 Types I–IV Hypotheses

Many readers of this book may already be aware that the SAS-GLM procedure gives users the option of selecting one of four types of sums of squares for testing hypotheses. This section is mainly concerned with defining and interpreting the corresponding four types of hypotheses that are tested. The data in Table 9.1 will be used to illustrate these hypotheses. As stated in Section 10.2, the type I sums of squares are obtained by fitting the two-way effects model in a sequential fashion. The sum of squares obtained at each step, which is a measure of the importance of the particular term being added at that step, is the amount that the residual sum of squares can be reduced by including that term in the model. The type II analysis is also obtained by utilizing the model comparison technique. The sums of squares corresponding to each effect are adjusted for every other effect in the model that is at the same or a lower level. Hence, the type II sum of squares corresponding to the T -effect is $R(\tau|\mu, \beta)$ and the type II sum of squares corresponding to the B -effect is $R(\beta|\mu, \tau)$. Readers who are slightly confused should see Section 16.3, because the definitions of the type I and type II analyses for a three-way treatment structure will help clarify the differences between the type I sum of squares and the type II sum of squares. Table 10.7 shows the type I and type II sums of squares for two-way effects models and Table 10.8 gives the type II sums of squares and test statistics for the data in Table 9.1.

We advise experimenters to think in terms of the parameters in the means model rather than the parameters in the effects model. The main advantage of using effects models to model two-way experiments is that many of the interesting hypotheses are tested automatically; this makes it easy for the experimenter to do. The disadvantage of the effects

TABLE 10.7

Definitions of Types I and II Sums of Squares

Source of Variation	df	Type I SS	Type II SS
T	$t - 1$	$R(\tau \mu)$	$R(\tau \mu, \beta)$
B	$b - 1$	$R(\beta \mu, \tau)$	$R(\beta \mu, \tau)$
$T \times B$	$(t - 1)(b - 1)$	$R(\gamma \mu, \tau, \beta)$	$R(\gamma \mu, \tau, \beta)$

TABLE 10.8

Type II Analysis of Variance Table

Source of Variation	df	SS	MS	F	α
Total	15	258.9375			
T	1	72.369	72.369	36.18	0.0001
B	2	90.744	45.372	22.69	0.0002
$T \times B$	2	71.631	35.815	17.91	0.0005
Error	10	20.000	2.0		

model is that it is difficult for experimenters to understand exactly what is being tested by the type I and type II sums of squares. The main advantage of the means model is that one is able to understand exactly what is being tested and/or estimated. The disadvantage of the means model is that the experimenter has to write her own ESTIMATE and CONTRAST statements to get the test statistics of interest. We want experimenters to be knowledgeable about both the effects model and the means model; the effects model can be used to automatically get many statistics of interest and the means model to identify what is being estimated or tested by one of statistics of interest.

What hypotheses are being tested by the type I and type II sums of squares? The hypotheses being tested by the type I sums of squares are called type I hypotheses. Table 10.9 gives the hypotheses tested by a type I analysis of the data in Table 9.1 for model (10.1) in terms of the parameters in a means model. Later we shall show how to determine these hypotheses from the SAS-GLM procedure. Clearly, an experimenter would rarely be interested in the hypotheses that correspond to T and B in Table 10.9, and as the sample sizes change in one or more of the cells in Table 9.1, the hypotheses would change. The hypothesis corresponding to $T \times B$ is equivalent to testing a no-interaction hypothesis. That is, the contrasts

$$\mu_{11} - \mu_{13} - \mu_{21} + \mu_{23} \quad \text{and} \quad \mu_{12} - \mu_{13} - \mu_{22} + \mu_{23}$$

span the interaction space for this particular example. Table 10.10 gives general formulations for the Type I hypotheses for a two-way experiment expressed in terms of the means model parameters. Note that from the general form we can see that the type I hypothesis for T compares the weighted averages of row cell means using the sample sizes within each cell as weights. For example, the type I hypothesis for T using the data in Table 9.1 can be written as

$$H_0: \frac{3\mu_{11} + 2\mu_{12} + 3\mu_{13}}{8} = \frac{2\mu_{21} + 3\mu_{212} + 3\mu_{213}}{8}$$

TABLE 10.9

Hypotheses for a Type I Analysis of the Means Model for Data in Table 9.1

Source of Variation	Type I Hypotheses
T	$3\mu_{11} + 2\mu_{12} + 3\mu_{13} - 2\mu_{21} - 3\mu_{22} - 3\mu_{23} = 0$
B	$37\mu_{11} - 2\mu_{12} - 35\mu_{13} + 28\mu_{21} + 2\mu_{22} - 30\mu_{23} = 0$ and $2\mu_{11} + 28\mu_{12} - 30\mu_{13} - 2\mu_{21} + 37\mu_{22} - 35\mu_{23} = 0$
$T \times B$	$\mu_{11} - \mu_{13} - \mu_{21} + \mu_{23} = 0$ and $\mu_{12} - \mu_{13} - \mu_{22} + \mu_{23} = 0$

TABLE 10.10

General Forms of Type I Hypotheses for a Two-Way Effects Model Expressed in Terms of Means Model Parameters

Source of Variation	Type I Hypotheses
T	$\frac{1}{n_{\cdot i}} \sum_{j=1}^b n_{1j}\mu_{1j} = \frac{1}{n_{\cdot 2}} \sum_{j=1}^b n_{2j}\mu_{2j} = \dots = \frac{1}{n_{\cdot t}} \sum_{j=1}^b n_{tj}\mu_{tj}$
B	$\sum_{i=1}^t \left(n_{ij} - \frac{n_{ij}^2}{n_{\cdot j}} \right) \mu_{ij} = \sum_{\substack{i'=1 \\ i' \neq i}}^t \sum_{j=1}^b \frac{n_{ij}n_{i'j}}{n_{\cdot j}} \mu_{i'j}, \quad j = 1, 2, \dots, b$
$T \times B$	$\mu_{ij} - \mu_{i'j} - \mu_{ij'} + \mu_{i'j'} = 0 \text{ for all } i \neq i', j \neq j'$

For the data in Table 9.1, the hypotheses tested by the type II analysis in terms of the means model parameters are given in Table 10.11.

For the general model, the type II hypothesis tested by the row corresponding to the T -effect is

$$\sum_{j=1}^b \left(n_{ij} - \frac{n_{ij}^2}{n_{\cdot j}} \right) \mu_{ij} = \sum_{\substack{i'=1 \\ i' \neq i}}^t \sum_{j=1}^b \frac{n_{ij}n_{i'j}}{n_{\cdot j}} \mu_{i'j}, \quad i = 1, 2, \dots, t \quad (10.3)$$

The other two rows are the same as they were for the type I analysis.

The type I and type II hypotheses in terms of the effects model parameters in Equation 10.1 for the data in Table 9.1 are given in Tables 10.12 and 10.13.

TABLE 10.11

Hypotheses for a Type II Analysis in Terms of the Means Model for the Data in Table 9.1

Source of Variation	Type II Hypotheses
T	$4\mu_{11} + 4\mu_{12} + 5\mu_{13} - 4\mu_{21} - 4\mu_{22} - 5\mu_{23} = 0$
B	Same as for type I analysis
$T \times B$	Same as for type I analysis

TABLE 10.12

Hypotheses for a Type I Analysis in Terms of the Effects Model for the Data in Table 9.1

Source of Variation	Type II Hypotheses
T	$\tau_1 - \tau_2 + (1/8)(\beta_1 - \beta_2) + (1/8)(3\gamma_{11} + 2\gamma_{12} + 3\gamma_{13} - 2\gamma_{21} - 3\gamma_{22} - 3\gamma_{23}) = 0$
B	$\beta_1 - \beta_3 + (1/65)(37\gamma_{11} - 2\gamma_{12} - 35\gamma_{13} + 28\gamma_{21} + 2\gamma_{22} - 30\gamma_{23}) = 0$ $\beta_2 - \beta_3 + (1/65)(2\gamma_{11} + 28\gamma_{12} - 30\gamma_{13} - 2\gamma_{21} + 37\gamma_{22} - 35\gamma_{23}) = 0$
$T \times B$	$\gamma_{11} - \gamma_{13} - \gamma_{21} + \gamma_{23} = 0 \text{ and } \gamma_{12} - \gamma_{13} - \gamma_{22} + \gamma_{23} = 0$

TABLE 10.13

Hypotheses for a Type II Analysis in Terms of the Effects Model for the Data in Table 9.1

Source of Variation	Type II Hypotheses
T	$\tau_1 - \tau_2 + (1/13)(4\gamma_{11} + 4\gamma_{12} + 5\gamma_{13} - 4\gamma_{21} - 4\gamma_{22} - 5\gamma_{23}) = 0$
B	$\beta_1 - \beta_3 + (1/65)(37\gamma_{11} - 2\gamma_{12} - 35\gamma_{13} + 28\gamma_{21} + 2\gamma_{22} - 30\gamma_{23}) = 0$ $\beta_2 - \beta_3 + (1/65)(2\gamma_{11} + 28\gamma_{12} - 30\gamma_{13} - 2\gamma_{21} + 37\gamma_{22} - 35\gamma_{23}) = 0$
$T \times B$	$\gamma_{11} - \gamma_{13} - \gamma_{21} + \gamma_{23} = 0$ and $\gamma_{12} - \gamma_{13} - \gamma_{22} + \gamma_{23} = 0$

Examination of Tables 10.10–10.13 reveals that the main effect hypotheses tested by the type I and type II analyses may not be very interesting. In addition, the rejection or acceptance of these hypotheses may not be easy to interpret.

A third way to compute sums of squares for an effects model is as follows:

- 1) For the t levels of treatment T , generate $t - 1$ dummy variables, and for the b levels of treatment B , generate $b - 1$ dummy variables. These dummy variables corresponding to the τ and β are created so that in the model with the dummy variables $\tau_i = -(\tau_1 + \tau_2 + \dots + \tau_{i-1})$ and $\beta_b = -(\beta_1 + \beta_2 + \dots + \beta_{b-1})$.
- 2) The interaction between T and B is represented by the products of their corresponding dummy variables. In particular, $\gamma_{ij} = -(\gamma_{1j} + \gamma_{2j} + \dots + \gamma_{(t-1)j})$ for $j = 1, 2, \dots, b$ and $\gamma_{ib} = -(\gamma_{i1} + \gamma_{i2} + \dots + \gamma_{ib-1})$ for $i = 1, 2, \dots, t$.
- 3) The model with all of the dummy variables for the treatment variables and their interactions is fitted, and the residual sum of squares is obtained. This is equivalent to the residual sum of squares from the full-effects model.
- 4) Next, a model is fitted that contains all of the dummy variables except those corresponding to the main effect or interaction being tested. The difference between the residual sum of squares of this reduced model and that of the model in 3 is the sum of squares corresponding to that effect.

The resulting analysis is called a type III analysis. It is also known as Yates's weighted squares of means technique. When all treatment combinations are observed, the hypotheses tested by a type III analysis are the same as those tested for balanced data sets. These type III hypotheses in terms of mean model parameters are given in Table 10.14 and in terms of effect model parameters in Table 10.15. These hypotheses are usually the ones desired by experimenters, and the type III sums of squares for the data in Table 9.1 are given in Table 10.16.

TABLE 10.14

Type III Hypotheses for the Effects Model in Terms of the Means Model

Parameters	Source of Variation	Hypotheses
	T	$\bar{\mu}_{1.} = \bar{\mu}_{2.} = \dots = \bar{\mu}_{i.}$
	B	$\bar{\mu}_{.1} = \bar{\mu}_{.2} = \dots = \bar{\mu}_{.b}$
	$T \times B$	$\mu_{ij} - \mu_{i'j} - \mu_{i'j'} + \mu_{i'j'} = 0$ for all $i \neq i', j \neq j'$

TABLE 10.15

Type III Hypotheses for the Effects Model in Terms of the Effects Model Parameters

Source of Variation	Hypotheses
T	$\tau_1 + \bar{\gamma}_1 = \tau_2 + \bar{\gamma}_2 = \cdots = \tau_t + \bar{\gamma}_t$
B	$\beta_1 + \bar{\gamma}_{.1} = \beta_2 + \bar{\gamma}_{.2} = \cdots = \beta_b + \bar{\gamma}_b$
$T \times B$	$\mu_{ij} - \mu_{i'j'} - \mu_{ij'} + \mu_{i'j} = 0$ for all $i \neq i', j \neq j'$

TABLE 10.16

Type III Analysis of Variance Table

Source of Variation	df	SS	MS	F	α
Total	15	258.9375			
T	1	61.714	61.714	30.86	0.0002
B	2	71.169	38.585	19.29	0.0004
$T \times B$	2	71.631	35.815	17.91	0.0005
Error	10	20.000	2.0		

The matrix form of the model reparameterized by Yates's method for the data in Table 9.3 is

$$\begin{array}{c}
 \begin{bmatrix} 19 \\ 20 \\ 21 \\ 24 \\ 26 \\ 22 \\ 25 \\ 25 \\ 25 \\ 27 \\ 21 \\ 24 \\ 24 \\ 31 \\ 32 \\ 33 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 0 & -1 & 0 \\ 1 & -1 & 1 & 0 & -1 & 0 \\ 1 & -1 & 0 & 1 & 0 & -1 \\ 1 & -1 & 0 & 1 & 0 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \beta_1 \\ \beta_2 \\ \gamma_{11} \\ \gamma_{12} \\ \gamma_{21} \\ \gamma_{22} \\ \gamma_{23} \\ \gamma_{24} \\ \gamma_{25} \\ \gamma_{26} \\ \gamma_{27} \\ \gamma_{28} \\ \gamma_{29} \end{bmatrix} + \begin{bmatrix} \varepsilon_{111} \\ \varepsilon_{112} \\ \varepsilon_{113} \\ \varepsilon_{121} \\ \varepsilon_{122} \\ \varepsilon_{131} \\ \varepsilon_{132} \\ \varepsilon_{133} \\ \varepsilon_{211} \\ \varepsilon_{212} \\ \varepsilon_{221} \\ \varepsilon_{222} \\ \varepsilon_{223} \\ \varepsilon_{231} \\ \varepsilon_{232} \\ \varepsilon_{233} \end{bmatrix}
 \end{array}$$

Note that the fifth column (corresponding to γ_{11}) in the above design matrix is the product of the second and third column (columns corresponding to τ_1 and β_1), and the sixth column (corresponding to γ_{12}) is the product of the second and fourth columns (columns corresponding to τ_1 and β_2).

SAS-GLM introduced a fourth way of generating sums of squares corresponding to the main effects and their interactions. When all treatment combinations are observed, the hypotheses tested by this type IV analysis are the same as those tested by the type III analysis; however, when some treatment combinations are not observed, the type III and type IV analyses do not agree. We shall discuss the construction of type IV hypotheses in Chapter 14.

To conclude this section, we make the following recommendations when analyzing data in a two-way treatment structure model with no missing treatment combinations:

- 1) If the experimenter wants to compare the effects of the two treatments, she should look at hypotheses tested by a type III analysis. These hypotheses are equivalent to the hypotheses tested in the balanced or equal-subclass-numbers case.
- 2) If the experimenter is interested in building a model with which to predict the effects of particular treatment combinations, then he or she could use type I and/or type II analyses.
- 3) In sample survey experiments, the number of observations per treatment combination is often proportional to the frequency with which those combinations actually occur in the population. In this case, the experimenter may be most interested in the hypotheses based on $R(\tau|\mu, \beta)$ and $R(\beta|\mu, \tau)$, since these sums of squares test hypotheses about the weighted averages of the row means and the column means with the weights proportional to the observed sample sizes. This may require two type I analyses if one were using SAS-GLM, one with T first in the model and another with B first.

10.5 Using Types I–IV Estimable Functions in SAS-GLM

In this section, we show how one can use the information provided by the SAS-GLM to determine the hypotheses being tested by the different types of sums of squares. (Readers who do not use SAS-GLM can skip this section.) In Section 10.3 we discussed the general form of estimable functions obtained by using the E option on the MODEL statement. If the El option is chosen, SAS-GLM will print the general form of the type I estimable functions for each effect. The results of this option for the data in Table 9.1 when using an effects model are shown in Table 10.17.

From Table 10.17, we see that $\ell'\beta$, a linear combination of the parameter vector $\beta' = [\mu, \tau_1, \tau_2, \beta_1, \beta_2, \beta_3, \gamma_{11}, \gamma_{12}, \gamma_{13}, \gamma_{21}, \gamma_{22}, \gamma_{23}]$, is a type I estimable function for T if and only if there exists a constant $L2$ such that

$$\begin{aligned}\ell'\beta &= (L2)\tau_1 - (L2)\tau_2 + (0.125 \cdot L2)\beta_1 - (0.125 \cdot L2)\beta_2 \\ &\quad + (0.375 \cdot L2)\gamma_{11} + (0.25 \cdot L2)\gamma_{12} + (0.375 \cdot L2)\gamma_{13} \\ &\quad - (0.25 \cdot L2)\gamma_{21} - (0.375 \cdot L2)\gamma_{22} - (0.375 \cdot L2)\gamma_{23}\end{aligned}$$

A basis set for the type I estimable functions for T can be obtained by choosing a specific value for $L2$, say $L2 = 1$ or $L2 = 8$. We are free to choose only one of the L in the general form of an estimable function which corresponds to the 1 degree of freedom associated with the type I sum of squares for T ; all of the remaining L are determined by our choice for $L2$. For our example, $L1 = 0$, $L4 = 0.125 \cdot L2$, $L5 = -0.125 \cdot L2$, $L7 = 0.375 \cdot L2$, and $L8 = 0.25 \cdot L2$.

TABLE 10.17

Type I Estimable Functions from SAS-GLM for the Data in Table 9.1

Coefficients			
Effect	T	B	$T \times B$
Intercept	0	0	0
T 1	L2	0	0
T 2	-L2	0	0
B 1	$0.125 \times L2$	L4	0
B 2	$-0.125 \times L2$	L5	0
B 3	0	$-L4 - L5$	0
$T \times B$ 1 1	$0.375 \times L2$	$0.5692 \times L4 + 0.0308 \times L5$	L7
$T \times B$ 1 2	$0.25 \times L2$	$-0.0308 \times L4 + 0.4308 \times L5$	L8
$T \times B$ 1 3	$0.375 \times L2$	$-0.5385 \times L4 - 0.4615 \times L5$	$-L7 - L8$
$T \times B$ 2 1	$-0.25 \times L2$	$0.4308 \times L4 - 0.0308 \times L5$	$-L7$
$T \times B$ 2 2	$-0.375 \times L2$	$0.0308 \times L4 + 0.5692 \times L5$	$-L8$
$T \times B$ 2 3	$-0.375 \times L2$	$-0.4615 \times L4 - 0.5385 \times L5$	$L7 + L8$

Taking $L2 = 1$, a basis set for the type I estimable functions is

$$\{\tau_1 - \tau_2 + (1/8)(\beta_1 - \beta_2) + (1/8)(3\gamma_{11} + 2\gamma_{12} + 3\gamma_{13} - 2\gamma_{21} - 3\gamma_{22} - 3\gamma_{23})\}$$

This function of the parameters is the one that is compared to zero by a type I analysis. See Table 10.11. Another basis set can be constructed by taking $L2 = 8$. This set is given by

$$\{8\tau_1 - 8\tau_2 + \beta_1 - \beta_2 + 3\gamma_{11} + 2\gamma_{12} + 3\gamma_{13} - 2\gamma_{21} - 3\gamma_{22} - 3\gamma_{23}\}$$

Since $\mu_{ij} = \mu + \tau_i + \beta_j + \gamma_{ij}$ the hypotheses being tested in terms of the means model parameters can be determined by assigning the coefficients on the γ_{ij} in the effects model representation to the μ_{ij} in the means model representation. Thus, the type I hypothesis for T in terms of the means model parameters is

$$(3\mu_{11} + 2\mu_{12} + 3\mu_{13} - 2\mu_{21} - 3\mu_{22} - 3\mu_{23})/8 = 0 \text{ or equivalently that} \\ 3\mu_{11} + 2\mu_{12} + 3\mu_{13} - 2\mu_{21} - 3\mu_{22} - 3\mu_{23} = 0$$

which is the hypothesis for T given in Table 10.9.

If the E2 option on the MODEL statement is chosen, SAS-GLM will print the general form of the type II estimable functions for each effect. The results for the data in Table 9.1 are shown in Table 10.18.

From Table 10.18, we see that $\ell' \beta$ is a type II estimable function for B if and only if there exist constants L4 and L5 such that

$$\begin{aligned} \ell' \beta = & (L4)\beta_1 + (L5)\beta_2 + (-L4 - L5) \cdot \beta_3 + (0.5692 \cdot L4 + 0.0308 \cdot L5)\gamma_{11} \\ & + (-0.0308 \cdot L4 + 0.4308 \cdot L5)\gamma_{12} + (-0.5385 \cdot L4 - 0.4615 \cdot L5)\gamma_{13} \\ & + (0.4308 \cdot L4 - 0.0308 \cdot L5)\gamma_{21} + (0.0308 \cdot L4 + 0.5692 \cdot L5)\gamma_{22} \\ & + (-0.4615 \cdot L4 - 0.5385 \cdot L5)\gamma_{23} \end{aligned}$$

In this case we can choose values for two of the L, namely L4 and L5. Thus, there are 2 degrees of freedom corresponding to the type II sum of squares for B. With a little luck

TABLE 10.18

Type II Estimable Functions from SAS for Data in Table 9.1

Coefficients			
Effect	T	B	$T \times B$
Intercept	0	0	0
T	1	$L2$	0
T	2	$-L2$	0
B	1	0	$L4$
B	2	0	$L5$
B	3	0	$-L4 - L5$
$T \times B$	1 1	$0.3077 \times L2$	$0.5692 \times L4 + 0.0308 \times L5$
$T \times B$	1 2	$0.3077 \times L2$	$-0.0308 \times L4 + 0.4308 \times L5$
$T \times B$	1 3	$0.3846 \times L2$	$-0.5385 \times L4 - 0.4615 \times L5$
$T \times B$	2 1	$-0.3077 \times L2$	$0.4308 \times L4 - 0.0308 \times L5$
$T \times B$	2 2	$-0.3077 \times L2$	$0.0308 \times L4 + 0.5692 \times L5$
$T \times B$	2 3	$-0.3846 \times L2$	$-0.4615 \times L4 - 0.5385 \times L5$
			$L7 + L8$

or by using the general forms given in Table 10.9, one can determine that the decimal numbers given in the above expression have a lowest common denominator of 65. Choosing $L4 = 1$ and $L5 = 0$ and then $L4 = 0$ and $L5 = 1$ provides one basis set for the type II estimable functions for B . This set is

$$\{\beta_1 - \beta_3 + (1/65)(37\gamma_{11} - 2\gamma_{12} - 35\gamma_{13} + 28\gamma_{21} + 2\gamma_{22} - 30\gamma_{23}) \text{ and} \\ \beta_2 - \beta_3 + (1/65)(2\gamma_{11} + 28\gamma_{12} - 30\gamma_{13} - 2\gamma_{21} + 37\gamma_{22} - 35\gamma_{23})\}$$

in terms of the parameters in the effects model and

$$\{37\mu_{11} - 2\mu_{12} - 35\mu_{13} + 28\mu_{21} + 2\mu_{22} - 30\mu_{23} \text{ and} \\ 2\mu_{11} + 28\mu_{12} - 30\mu_{13} - 2\mu_{21} + 37\mu_{22} - 35\mu_{23}\}$$

in terms of the parameters in the means model by letting $L4 = 65$ and $L5 = 0$ for the first function and by letting $L4 = 0$ and $L5 = 65$ for the second function.

From the general form of the type I and/or type II estimable functions for $T \times B$ in Table 10.17 and/or Table 10.18, we see that we have two L for which we can choose values. A basis set for the type I estimable functions for $T \times B$ can be obtained by letting $L7 = 1$ and $L8 = 0$ and then $L7 = 0$ and $L8 = 1$. These two choices yield

$$\{\gamma_{11} - \gamma_{13} - \gamma_{21} + \gamma_{23} \text{ and } \gamma_{12} - \gamma_{13} - \gamma_{22} + \gamma_{23}\}$$

in terms of the parameters in the effects model and by

$$\{\mu_{11} - \mu_{13} - \mu_{21} + \mu_{23} \text{ and } \mu_{12} - \mu_{13} - \mu_{22} + \mu_{23}\}$$

in terms of the parameters in the means model. The reader might wish to verify that every 2×2 table differences $\mu_{ij} - \mu_{ij'} - \mu_{rj} + \mu_{rj'} = 0$ for all $i \neq i', j \neq j'$ can be obtained as some linear combination of these two functions in this basis set.

From Table 10.18, the general form of type II estimable functions for T is given by

$$(L2)\tau_1 - (L2)\tau_2 + (0.3077 \cdot L2)\gamma_{11} + (0.3077 \cdot L2)\gamma_{12} + (0.3846 \cdot L2)\gamma_{13} \\ + (-0.3077 \cdot L2)\gamma_{21} + (-0.3077 \cdot L2)\gamma_{22} + (-0.3846 \cdot L2)\gamma_{23}$$

The lowest common denominator of these decimal fractions is 13; thus, by taking $L2 = 1$, a basis set of type II estimable functions for T is

$$\{\tau_1 - \tau_2 + (1/13)(4\gamma_{11} + 4\gamma_{12} + 5\gamma_{13} - 4\gamma_{21} - 4\gamma_{22} - 5\gamma_{23})\}$$

in terms of the parameters in the effects model and by

$$\{4\mu_{11} + 4\mu_{12} + 5\mu_{13} - 4\mu_{21} - 4\mu_{22} - 5\mu_{23}\}$$

in terms of the parameters in the means model (taking $L2 = 13$).

Finally, we consider the type III and type IV estimable functions. For the data in Table 9.1, these are the same; they are shown in Table 10.19. A basis set of type III estimable functions for T in the parameters of the effects model is

$$\{\tau_1 - \tau_2 + \bar{\gamma}_{1.} - \bar{\gamma}_{2.}\}$$

in terms of the parameters of the effects model and

$$\{\bar{\mu}_{.1} - \bar{\mu}_{.2}\}$$

in terms of the parameters of the means model (taking $L2 = 1$ in both cases). In a similar manner, we see that basis sets for the type III estimable functions for B are

$$\{\beta_1 - \beta_3 + \bar{\gamma}_{.1} - \bar{\gamma}_{.3} \text{ and } \beta_2 - \beta_3 + \bar{\gamma}_{.2} - \bar{\gamma}_{.3}\}$$

TABLE 10.19

Type III Estimable Functions for the Data in Table 9.1

Effect	Coefficients			
	T	B	$T \times B$	
Intercept	0	0	0	
T	1	$L2$	0	0
T	2	$-L2$	0	0
B	1	0	$L4$	0
B	2	0	$L5$	0
B	3	0	$-L4 - L5$	0
$T \times B$	1 1	$0.3333 \times L2$	$0.5 \times L4$	$L7$
$T \times B$	1 2	$0.3333 \times L2$	$0.5 \times L5$	$L8$
$T \times B$	1 3	$0.3333 \times L2$	$-0.5 \times L4 - 0.5 \times L5$	$-L7 - L8$
$T \times B$	2 1	$-0.3333 \times L2$	$0.5 \times L4$	$-L7$
$T \times B$	2 2	$-0.3333 \times L2$	$0.5 \times L5$	$-L8$
$T \times B$	2 3	$-0.3333 \times L2$	$-0.5 \times L4 - 0.5 \times L5$	$L7 + L8$

in terms of the parameters in the effects model and

$$\{\bar{\mu}_{.1} - \bar{\mu}_{.3} \text{ and } \bar{\mu}_{.2} - \bar{\mu}_{.3}\}$$

in terms of the parameters in the means model. Note that $H_0: \bar{\mu}_{.1} - \bar{\mu}_{.3} = 0$ and $\bar{\mu}_{.2} - \bar{\mu}_{.3} = 0$ is true if and only if $H_0: \bar{\mu}_1 = \bar{\mu}_2 = \bar{\mu}_3$ is true. The reader can verify that the type IV estimable functions are identical to the type III estimable functions for the data in Table 9.1.

10.6 Population Marginal Means and Least Squares Means

The population marginal means for the two-way effects model are defined by

$$\bar{\mu}_{i.} = \mu + \tau_i + \bar{\beta}_{.} + \bar{\gamma}_{i.}, \quad i = 1, 2, \dots, t$$

for the T -treatments and

$$\bar{\mu}_{.j} = \mu + \bar{\tau}_{.} + \bar{\beta}_j + \bar{\gamma}_{.j}, \quad j = 1, 2, \dots, b$$

for the B -treatments. The best estimates of these marginal means are

$$\hat{\mu}_{i.} = \hat{\mu} + \hat{\tau}_i + \hat{\beta}_{.} + \hat{\gamma}_{i.}, \quad i = 1, 2, \dots, t \quad \text{and} \quad \hat{\mu}_{.j} = \hat{\mu} + \hat{\tau}_{.} + \hat{\beta}_j + \hat{\gamma}_{.j}, \quad j = 1, 2, \dots, b$$

respectively, and these estimates of the marginal means are often called least squares means. Their respective standard errors are given by Equations 9.5 and 9.7. To make inferences about linear combinations of the population marginal means, one can use (9.5) and (9.6). If one uses the

```
LSMEANS T B T*B/PDIFF;
```

option in the SAS-GLM procedure, one gets the best unbiased estimates of the marginal means and the two-way cell means as well as pairwise comparisons among certain subsets of these means. These are given in Table 10.20. Note that the least squares means for T and B are the same as the T and B means given in Table 9.2, and the two-way means are the same as the cell means in Table 9.2. Table 10.20 also provides the estimated standard errors of the least squares means and p -values for pairwise comparisons between the least squares means. For example, the estimated standard error of $\hat{\mu}_{.1}$ is 0.51 and this agrees with the estimated standard error that was computed in Section 9.5. Also the p -value that compares $\bar{\mu}_{.1}$ with $\bar{\mu}_{.2}$ is given in Table 10.20 as 0.0002 and this agrees with the p -value of the t -statistic calculated in Section 9.5. Similar comparisons can be made concerning other estimates, standard errors, and test statistics.

10.7 Computer Analyses

Nearly all the statistical computing packages have been developed in order to deal with effects models rather than means models. Since three major types of hypotheses can be

TABLE 10.20

Best Estimates of the Marginal Means and Cell Means

Least Squares Means

T	Y LSMean	Standard Error	$H_0: \text{LSMean} = 0$	$H_0: \text{LSMean1} = \text{LSMean2}$
			Pr > t	Pr > t
1	23.0000000	0.5091751	<0.0001	0.0002
2	27.0000000	0.5091751	<0.0001	
B	Y LSMean	Standard Error	Pr > t	LSMean Number
1	23.0000000	0.6454972	<0.0001	1
2	24.0000000	0.6454972	<0.0001	2
3	28.0000000	0.5773503	<0.0001	3

Least Squares Means for Effect B $Pr > |t|$ for $H_0: \text{LSMean}(i) = \text{LSMean}(j)$

i/j	1	2	3		
1		0.2990	0.0002		
2	0.2990		0.0010		
3	0.0002	0.0010			
T	B	Y LSMean	Standard Error	Pr > t	LSMean Number
1	1	20.0000000	0.8164966	<0.0001	1
1	2	25.0000000	1.0000000	<0.0001	2
1	3	24.0000000	0.8164966	<0.0001	3
2	1	26.0000000	1.0000000	<0.0001	4
2	2	23.0000000	0.8164966	<0.0001	5
2	3	32.0000000	0.8164966	<0.0001	6

Least Squares Means for Effect T × B $Pr > |t|$ for $H_0: \text{LSMean}(i) = \text{LSMean}(j)$

i/j	1	2	3	4	5	6
1		0.0031	0.0061	0.0009	0.0266	<0.0001
2	0.0031		0.4565	0.4956	0.1524	0.0003
3	0.0061	0.4565		0.1524	0.4068	<0.0001
4	0.0009	0.4956	0.1524		0.0425	0.0009
5	0.0266	0.1524	0.4068	0.0425		<0.0001
6	<0.0001	0.0003	<0.0001	0.0009	<0.0001	

tested for the balanced two-way treatment structure with unequal subclass numbers, one must be careful to determine which type is tested by the statistical package being used. Although the computing packages have been developed with effects models in mind, the means model can easily be implemented as well. Using the means model allows the user to specify meaningful contrasts among the treatment means. The names used in this discussion, that is, types I, II, III, and IV, correspond to the names used by SAS. SPSS uses the same notation and will produce type I, II, and III analyses. The default in SPSS is the

type III analysis. SAS-GLM produces both the type I analysis and the type III analysis by default. Readers who use other computing packages are encouraged to analyze the data in Table 9.1 and then they can compare their analyses with those given in this chapter and Chapter 9. Such a comparison should reveal the type of analysis their computing package is producing. SAS-GLM also allows means to be calculated using a

```
MEANS T B T*B/ <options>;
```

statement. However, with unbalanced data the Means statement does not give unbiased estimates of the population marginal means, instead it computes estimates of weighted averages of the row means and of the column means. That is, the above Means option gives the best unbiased estimates of the weighted means defined by

$$\tilde{\mu}_{i \cdot} = \frac{1}{n_{i \cdot}} \sum_{j=1}^b n_{ij} \mu_{ij}, \quad i = 1, 2, \dots, t \quad \text{and} \quad \tilde{\mu}_{\cdot j} = \frac{1}{n_{\cdot j}} \sum_{i=1}^t n_{ij} \mu_{ij}, \quad j = 1, 2, \dots, b$$

Probably the only case in which an experimenter may be interested in these weighted means is in sample survey type experiments (see the third recommendation at the end of Section 10.4).

For the data in Table 9.1, the best estimates of these weighted means are given by $\bar{y}_{1 \cdot} = 22.75$ and $\bar{y}_{2 \cdot} = 27.125$ for the T -weighted means and by $\bar{y}_{\cdot 1} = 22.4$, $\bar{y}_{\cdot 2} = 23.8$, and $\bar{y}_{\cdot 3} = 28.0$ for the B -weighted means.

10.8 Concluding Remarks

In this chapter, the analysis of two-way treatment structures with unequal subclass numbers is considered using the effects model under the assumption that all treatment combinations were observed. The effects model was used, even though the means model provides answers to all questions that can be raised, because much of the existing statistical computing software utilizes the effects model. Types I–III analyses and the conditions for using each analysis type were discussed. In almost all cases, the type III analysis will be preferred analysis. The type III analysis is the same as the analysis given in Chapter 9. Population marginal means were contrasted with weighted marginal means, and the appropriateness of each kind of mean was considered.

It may be of some interest to note that, if one were to fit a two-way additive model,

$$y_{ijk} = \mu + \tau_i + \beta_j + \varepsilon_{ijk}, \quad i = 1, 2, \dots, t; \quad j = 1, 2, \dots, b; \quad k = 1, 2, \dots, n_{ij}$$

to the data in Table 9.1, the types II and III sums of squares for T and B would be identical to one another and they would both be equal to the type II sum of squares for T and B in the analysis of the two-way model with the interaction term included.

10.9 Exercises

- 10.1 Consider the experiment described in Exercise 9.1
- 1) Analyze the data with SAS-GLM using an effects model. Include the following options on the MODEL statement: E, EL, E3, and SOLUTION. Also have GLM calculate the estimates of the drug main effect means, the disease main effect means, and the two-way means and have GLM do pairwise comparisons between these three sets of means.
 - 2) Write out the hypotheses being tested by the type I analysis using the notation of the means model as was done in Table 10.8, where μ_{ij} represents the expected response to drug i and disease j . Express these in terms of the effects model.
 - 3) Write out the hypotheses being tested by the type III analysis in terms of the means model.
 - 4) Write out the hypotheses being tested by the type II analysis in terms of the means model.
 - 5) Restate the following hypotheses in terms of the effects model and use CONTRAST and/or ESTIMATE statements with the effects model to determine the observed significance levels for the following hypotheses:
 - a) $(\bar{\mu}_{1.} + \bar{\mu}_{2.})/2 = \bar{\mu}_{3.} = \bar{\mu}_{4..}$
 - b) $\mu_{11} = \mu_{12} = \mu_{13},$
 - c) $(\bar{\mu}_{1.} + \bar{\mu}_{2.})/2 = \bar{\mu}_{3..}$
 - d) $\mu_{12} = \mu_{22} = \mu_{32} = \mu_{42}.$
 - e) $\bar{\mu}_{3..} - \bar{\mu}_{4..} = 0.$
 - f) $\mu_{12} - \mu_{13} - \mu_{32} + \mu_{33} = 0.$
 - g) $\mu_{22} - \mu_{32} = 0.$
- 10.2 Using the data from Exercise 8.1 where the sample sizes in each cell are equal, demonstrate that the type I, type II, type III, and type IV analyses are identical.
- 10.3 Verify that the Drug LSMEANS for 10.1 are the averages of the two-way cell means.
- 10.4 Use the method described in Section 10.4 to compute the type III analysis for the data set in Example 6.1. Use GLM to verify your results.

11

Analyzing Large Balanced Two-Way Experiments Having Unequal Subclass Numbers

In this chapter we present a method for obtaining an approximate analysis of balanced-treatment-structure experiments that have an unequal number of observations in each cell.

11.1 Feasibility Problems

We generally recommend using a general computing package such as one of those discussed in Section 10.7 to analyze unbalanced data sets in which every treatment combination is observed at least once. However, many situations arise in practice where it may not be feasible to do so, especially in developing countries where one may have several treatment factors with each factor having several different levels. In these cases, the exact procedures require several matrix inversions, and the size of the matrices to be inverted may exceed the capabilities of both the researcher and the computer that is available for use.

For example, consider an experiment with four factors, each factor occurring at five levels. To obtain an exact analysis, one will have to be able to invert several large matrices. The size of the largest matrix requiring inversion in this situation for an exact analysis is a 624×624 matrix in the unbalanced case. Obviously, such an experiment is not too unusual. In addition, some large experiments may require good initial starting values so that an iterative algorithm can converge to a reasonable solution. The methods in this chapter may allow researchers to find good initial starting values in those cases where she is having a hard time getting a program to converge.

An alternative way to analyze these types of messy data situations is called the *method of unweighted means* (Bancroft 1968). In Section 11.2 the method is described for two-way treatment structures. While one rarely encounters two-way experiments that cannot be analyzed by general computing procedures, the method of unweighted means is easily discussed and understood for the two-way experiment, and the discussion readily generalizes to larger experimental situations. Before continuing, we want to point out that it is not

the number of observations measured that makes general procedures unfeasible, but the number of treatment combinations being studied.

11.2 Method of Unweighted Means

Basically, the method of unweighted means approximates the sums of squares corresponding to each of the effects by using the observed means of the various treatment combinations. The formulas given here are the ones used for equal sample sizes; however, if the sample sizes are only slightly unequal, they still give quite accurate results. The correct sums of squares can be obtained by using the type III analysis and the formulas presented in this chapter approximate the type III sums of squares.

Consider the two-way models described in Chapters 9 and 10, where factor T had t possibilities and factor B has b possibilities. Let μ_{ij} represent the response one expects to see on a randomly selected experimental unit that receives the combination of T_i crossed with B_j . Let

$$\hat{\mu}_{ij} = \bar{y}_{ij}, \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, b$$

Also let

$$\hat{\mu}_{i\cdot} = \frac{1}{b} \sum_{j=1}^b \hat{\mu}_{ij} \quad \text{and} \quad \hat{\mu}_{\cdot j} = \frac{1}{t} \sum_{i=1}^t \hat{\mu}_{ij}$$

For testing H_{0T} : $\bar{\mu}_{1\cdot} = \bar{\mu}_{2\cdot} = \dots = \bar{\mu}_{t\cdot}$, we compute

$$SST = b \sum_{i=1}^t (\hat{\mu}_{i\cdot} - \hat{\mu}_{\cdot\cdot})^2 = \sum_{i=1}^t b \hat{\mu}_{i\cdot}^2 - bt \hat{\mu}_{\cdot\cdot}^2$$

which is based on $t - 1$ degrees of freedom. For testing H_{0B} : $\bar{\mu}_{\cdot 1} = \bar{\mu}_{\cdot 2} = \dots = \bar{\mu}_{\cdot b}$, we compute

$$SSB = t \sum_{j=1}^b (\hat{\mu}_{\cdot j} - \hat{\mu}_{\cdot\cdot})^2 = t \sum_{j=1}^b \hat{\mu}_{\cdot j}^2 - bt \hat{\mu}_{\cdot\cdot}^2$$

which is based on $b - 1$ degrees of freedom. For testing $H_{0T \times B}$: $\mu_{ij} - \mu_{i'j} - \mu_{ij'} + \mu_{i'j'} = 0$ for all $i \neq i'$ and $j \neq j'$ we compute

$$SST \times B = \sum_{i=1}^t \sum_{j=1}^b (\hat{\mu}_{ij} - \hat{\mu}_{i\cdot} - \hat{\mu}_{\cdot j} + \hat{\mu}_{\cdot\cdot})^2 = \sum_{i=1}^t \sum_{j=1}^b \hat{\mu}_{ij}^2 - b \sum_{i=1}^t \hat{\mu}_{i\cdot}^2 - t \sum_{j=1}^b \hat{\mu}_{\cdot j}^2 + bt \hat{\mu}_{\cdot\cdot}^2$$

which is based on $(t - 1)(b - 1)$ degrees of freedom. Since the above sums of squares are computed on the basis of means, the usual sum of squares for error,

$$SSE_{\text{Error}} = \sum_{i=1}^t \sum_{j=1}^b \sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}_{ij\cdot})^2$$

TABLE 11.1
Unweighted Means Analysis of Variance Table

Source of Variation	df	SS
T	$t - 1$	$b \sum_{i=1}^t (\hat{\mu}_{i\cdot} - \hat{\mu}_{..})^2$
B	$b - 1$	$t \sum_{j=1}^b (\hat{\mu}_{\cdot j} - \hat{\mu}_{..})^2$
$T \times B$	$(t - 1)(b - 1)$	$\sum_{i=1}^t \sum_{j=1}^b (\hat{\mu}_{ij} - \hat{\mu}_{i\cdot} - \hat{\mu}_{\cdot j} + \hat{\mu}_{..})^2$
Error	$N - tb$	$\frac{1}{\tilde{n}} \sum_{i=1}^t \sum_{j=1}^b \sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}_{ij\cdot})^2$

must be adjusted. This is because the $\hat{\mu}_{ij}$ have variances given by σ^2/n_{ij} rather than σ^2 , the variance of the y_{ijk} . The adjustment is made by dividing the error sum of squares by

$$\tilde{n} = \left[\frac{1}{bt} \left(\sum_{i=1}^t \sum_{j=1}^b \frac{1}{n_{ij}} \right) \right]^{-1}$$

the harmonic mean of the cell mean sample sizes.

Because the quantity \tilde{n} is the harmonic mean of the sample sizes, it is one possible average of the n_{ij} . The degrees of freedom for error are still $N - bt$. The analysis of variance table for an unweighted means analysis is given in Table 11.1. This analysis yields reasonable approximations to the F -distribution when the cell sample sizes are not too unequal. The usual recommendation is that this analysis will be acceptable if the sample sizes vary by no more than a factor of 2.

11.3 Simultaneous Inference and Multiple Comparisons

The T marginal means and the B marginal means are defined just as they were in Chapters 9 and 10. Thus the T marginal means are given by $\bar{\mu}_{i\cdot}$, $i = 1, 2, \dots, t$ and the B marginal means are given by $\bar{\mu}_{\cdot j}$, $j = 1, 2, \dots, b$. The best estimates of these marginal means are

$$\hat{\mu}_{i\cdot} = \frac{1}{b} \sum_{j=1}^b \hat{\mu}_{ij}, \quad i = 1, 2, \dots, t$$

and

$$\hat{\mu}_{\cdot j} = \frac{1}{t} \sum_{i=1}^t \hat{\mu}_{ij}, \quad j = 1, 2, \dots, b, \text{ respectively}$$

The exact estimated standard error of $\hat{\mu}_{i\cdot}$ is

$$\widehat{s.e.}(\hat{\mu}_{i\cdot}) = \frac{\hat{\sigma}}{b} \sqrt{\sum_{j=1}^b \frac{1}{n_{ij}}}$$

and the exact estimated standard error of $\hat{\mu}_{\cdot j}$ is

$$\widehat{s.e.}(\hat{\mu}_{\cdot j}) = \frac{\hat{\sigma}}{t} \sqrt{\sum_{i=1}^t \frac{1}{n_{ij}}}$$

where

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^t \sum_{j=1}^b \sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}_{ij\cdot})^2}{N - tb}} = \sqrt{ErrorMS}$$

If the cell sample sizes are not too unequal these standard errors can be approximated by $\hat{\sigma}/\sqrt{b\tilde{n}}$ and $\hat{\sigma}/\sqrt{t\tilde{n}}$, respectively. The estimated standard error of $\hat{\mu}_{i\cdot} - \hat{\mu}_{i'\cdot}$ is

$$\widehat{s.e.}(\hat{\mu}_{i\cdot} - \hat{\mu}_{i'\cdot}) = \frac{\hat{\sigma}}{b} \sqrt{\sum_{j=1}^b \frac{1}{n_{ij}} + \sum_{j=1}^b \frac{1}{n_{i'j}}}$$

which can be approximated by $\hat{\sigma}\sqrt{2}/\sqrt{b\tilde{n}}$. Similarly,

$$\widehat{s.e.}(\hat{\mu}_{\cdot j} - \hat{\mu}_{\cdot j'}) = \frac{\hat{\sigma}}{t} \sqrt{\sum_{i=1}^t \frac{1}{n_{ij}} + \sum_{i=1}^t \frac{1}{n_{ij'}}}$$

which can be approximated by $\hat{\sigma}\sqrt{2}/\sqrt{t\tilde{n}}$.

Next consider a 2×2 interaction contrast $\mu_{ij} - \mu_{i'j} - \mu_{ij'} + \mu_{i'j'}$ for $i \neq i'$ and $j \neq j'$. The best estimate of this interaction contrast is $\hat{\mu}_{ij} - \hat{\mu}_{i'j} - \hat{\mu}_{ij'} + \hat{\mu}_{i'j'}$ and its estimated standard error is given by

$$\widehat{s.e.}(\hat{\mu}_{ij} - \hat{\mu}_{i'j} - \hat{\mu}_{ij'} + \hat{\mu}_{i'j'}) = \hat{\sigma} \sqrt{\frac{1}{n_{ij}} + \frac{1}{n_{i'j}} + \frac{1}{n_{ij'}} + \frac{1}{n_{i'j'}}}$$

which can be approximated by $\hat{\sigma}\sqrt{4/\tilde{n}}$ if the cell sample sizes are not too unequal.

For each of the above functions of the cell mean parameters, test statistics are given by

$$t_c = \frac{\text{estimate}}{\text{estimated standard error}}$$

where the corresponding hypothesis is rejected if $|t_c| > t_{\alpha/2, N-tb}$. Furthermore, a $(1 - \alpha)100\%$ confidence interval is given by

$$\text{estimate} \pm t_{\alpha/2, N-tb} (\text{estimated standard error})$$

Simultaneous inference procedures described in Chapter 3 can be used and a researcher is encouraged to do so. These ideas can be generalized to more general contrasts of the cell means and/or the marginal means. All of the formulas in Chapter 8 can be approximated if one simply replaces n with \tilde{n} .

11.4 An Example of the Method of Unweighted Means

As an example, consider again the data in Table 9.1. Recall that the error sum of squares was 20 with 10 degrees of freedom, and the table of means is repeated in Table 11.2. The value of \tilde{n} is

$$\tilde{n} = \left[\frac{1}{bt} \left(\sum_{i=1}^t \sum_{j=1}^b \frac{1}{n_{ij}} \right) \right]^{-1} = \left[\frac{1}{(3)(2)} \left(\frac{1}{3} + \frac{1}{2} + \frac{1}{3} + \frac{1}{2} + \frac{1}{3} + \frac{1}{3} \right) \right]^{-1} = \left(\frac{7}{18} \right)^{-1} = 2.5714$$

Next note that

$$\begin{aligned}\sum_{i,j} \hat{\mu}_{ij}^2 &= 20^2 + 25^2 + \cdots + 28^2 = 3830 \\ \sum_{i,j} \hat{\mu}_{i\cdot}^2 &= 23^2 + 27^2 = 1258 \\ \sum_j \hat{\mu}_{\cdot j}^2 &= 23^2 + 24^2 + 28^2 = 1889\end{aligned}$$

and

$$\hat{\mu}_{..}^2 = (25)^2 = 625$$

Therefore

$$\begin{aligned}SST &= 3(1258) - 6(625) = 24 \\ SSB &= 2(1889) - 6(625) = 28\end{aligned}$$

and

$$SST \times B = 3830 - 3(1258) - 2(1889) + 6(625) = 28$$

TABLE 11.2

Cell Means and Marginal Means from Table 9.2

$\hat{\mu}_{ij}$	B_1	B_2	B_3	$\hat{\mu}_{i\cdot}$
T_1	20	25	24	23
T_2	26	23	32	27
$\hat{\mu}_{\cdot j}$	23	24	28	$25 = \hat{\mu}_{..}$

TABLE 11.3

The Unweighted Means Analysis of Variance Table for Data in Table 9.1

Source of Variation	df	SS	MS	F	p-Value
T	1	24	24	30.85	0.0002
B	2	28	14	17.99	0.0005
T × B	2	28	14	17.99	0.0005
Error(adj)	10	$\frac{20}{2.5714} = 7.7778$	0.77778		

The analysis of variance table is given in Table 11.3. Note that the F-statistics in Table 11.3 are very similar to the exact F-statistics given in Table 9.3.

11.5 Computer Analyses

Although we do not recommend it, the statistics needed for an unweighted means analysis can be obtained in Excel® and other spreadsheet programs. The unadjusted error sum of squares can be obtained most efficiently by considering the experiment as a one-way experiment and utilizing a means model. To obtain the T, B, and T × B sums of squares required for the unweighted means analysis, one can follow the steps below.

- 1) Obtain the cell means for each treatment combination.
- 2) Obtain the T marginal means, the B marginal means, and the overall mean.
- 3) Compute the variances of the cell means, the T marginal means, and the B marginal means. Denote these by S_{TB}^2 , S_B^2 , and S_T^2 , respectively.
- 4) Then

$$\begin{aligned} \text{error } SS &= \sum_{i,j,k} y_{ijk}^2 - (bt - 1)S_{TB}^2 + bt\hat{\bar{\mu}}^2 \\ SST &= b(t - 1)S_T^2 \\ SSB &= t(b - 1)S_B^2 \end{aligned} \tag{11.1}$$

and

$$SST \times B = (bt - 1)S_{TB}^2 - b(t - 1)S_T^2 - t(b - 1)S_B^2$$

If one has programs that will perform a one-way ANOVA, and a program that will perform a two-way ANOVA for balanced data, then one can get the error sum of squares from the one-way ANOVA program by treating the two-way treatment structure as a one-way treatment structure using a means model. The sums of squares for T, B, and T × B can be obtained by analyzing the cell means with the two-way ANOVA program that works for balanced data. SAS® code that will give the statistics needed to perform an unweighted means analysis is given in Table 11.4. It will be up to the reader to run this code should she wish to do so.

TABLE 11.4

SAS Code to Provide Statistics Needed for an Unweighted Means Analysis

```

DATA;
  INPUT T B Y;
  TB =10*T+B;
  CARDS;
1 1 19
1 1 20
1 1 21
1 2 24
1 2 26
1 3 22
1 3 25
1 3 25
2 1 25
2 1 27
2 2 21
2 2 24
2 2 24
2 3 31
2 3 32
2 3 33
PROC PRINT;
  TITLE 'AN UNWEIGHTED MEANS ANALYSIS OF A TWO-WAY WITH UNEQUAL SUBCLASS NUMBERS';

PROC ANOVA;
  TITLE2 'THIS ANALYSIS GIVES THE UNADJUSTED ERROR SUM OF SQUARES IN TABLE 11.3';
  CLASS TB;
  MODEL Y=TB;
  RUN;

PROC SORT; BY T B; RUN;

PROC MEANS NOPRINT; BY T B;
  VARIABLE Y;
  OUTPUT OUT=MEANS MEAN=YBAR;
  RUN;

PROC ANOVA;
  TITLE2 'THIS ANALYSIS GIVES THE T, B, AND T*B SUMS OF SQUARES GIVEN IN TABLE 11.3';
  CLASSES T B;
  MODEL YBAR=T B T*B;
  MEANS T B T*B;
  RUN;

```

11.6 Concluding Remarks

In this chapter, we introduced a method for obtaining satisfactory statistical analyses of experiments involving large numbers of different treatment combinations where each treatment combination is observed at least once. The techniques are useful for persons who may not have access to sophisticated statistical software programs.

11.7 Exercises

- 11.1 Use the data in Table 9.1 to verify the formulas given in Equation 11.1.
- 11.2 If you have SAS available to you, run the code given in Table 11.4.
- 11.3 Compute the approximate standard errors for the T marginal means and the B marginal means for the data in Table 9.1 using the formulas $\widehat{s.e.}(\hat{\bar{\mu}}_{i \cdot}) = \hat{\sigma}/\sqrt(b\tilde{n})$ and $\widehat{s.e.}(\hat{\bar{\mu}}_{\cdot j}) = \hat{\sigma}/\sqrt(t\tilde{n})$, and compare them to the exact standard errors given in Chapter 9.
- 11.4 Obtain the unweighted means analysis of variance for the drug–disease data in Exercise 9.1.
- 11.5 In Exercise 11.4, compute the estimates and their approximate estimated standard errors for each of the following.
 - 1) $\bar{\mu}_{3 \cdot} - \bar{\mu}_{4 \cdot}$.
 - 2) $\mu_{12} - \mu_{13} - \mu_{32} + \mu_{33}$
 - 3) $\mu_{22} - \mu_{32}$
 - 4) $\bar{\mu}_{\cdot 2}$.
 - 5) $\bar{\mu}_{\cdot 2}$

12

Case Study: Balanced Two-Way Treatment Structure with Unequal Subclass Numbers

In this chapter, we analyze a set of data arising from a two-way treatment structure conducted in a randomized complete block design. The experiment was intended to have been balanced, but due to unforeseen circumstances some treatment combinations were missing from some blocks. We still assume, however, that every treatment combination is observed at least once. The cases where some treatment combinations are never observed are discussed in Chapters 13–15.

12.1 Fat–Surfactant Example

A bakery scientist wanted to study the effects of combining three different fats with each of three different surfactants on the specific volume of bread loaves baked from doughs mixed from each of the nine treatment combinations. Four flours of the same type but from different sources were used as blocking factors. That is, loaves were made using all nine treatment combinations for each of the four flours. Unfortunately, one container of yeast turned out to be ineffective, and the data from the 10 loaves made with that yeast had to be removed from the analysis. Fortunately, all nine fat \times surfactant treatment combinations were observed at least once. The data are given in Table 12.1.

The data in Table 12.1 were analyzed using the SAS®-GLM procedure. Since all treatment combinations are observed at least once, and since some of the same treatment combinations are observed in each block (flour), the type III sums of squares test hypotheses that are interesting and easy to interpret. The hypotheses tested by the type III analysis would be the same as those tested if there were no missing data. Consequently, we can predict what those hypotheses are, and we do not need to include the E3 option to identify the hypotheses that are being tested. The normal equations are being solved using the sum-to-zero restrictions, so the results from a Solution option will not be interesting and consequently the option is not used in this SAS-GLM analysis. All marginal means are estimable, and their estimates have been obtained by using the LSMeans option.

TABLE 12.1

Specific Volumes from a Baking Experiment

Fat	Surfactant	Flour			
		1	2	3	4
1	1	6.7	4.3	5.7	
	2	7.1		5.9	5.6
	3		5.5	6.4	5.8
	2		5.9	7.4	7.1
	2		5.6		6.8
	3	6.4	5.1	6.2	6.3
	3	7.1	5.9		
	2	7.3	6.6	8.1	6.8
	3		7.5	9.1	

The data are analyzed using the SAS commands given in Table 12.2. The ANOVA results from the commands in Table 12.2 are given in Table 12.3. The type III F -value for the fat \times surfactant interaction is $F = 8.52$, which is significant at the $p = 0.0011$ level. Thus, surfactants should be compared within each fat level, and the fats should be compared within each surfactant level. Consequently, the fat \times surfactant least squares means are given in Table 12.4 and pairwise comparisons among these two-way least squares means are given in Table 12.5.

Figure 12.1 gives a plot of the two-way least squares means; sample means located within the same circle are not significantly different. The p -values used are those given in Table 12.5.

TABLE 12.2

SAS Analyses of Data in Table 12.1

```

DATA BREAD;
  INPUT FAT SURF F1-F4;
  LINES;
  1 1 6.7 4.3 5.7 .
  1 2 7.1 . 5.9 5.6
  1 3 . 5.5 6.4 5.8
  2 1 . 5.9 7.4 7.1
  2 2 . 5.6 . 6.8
  2 3 6.4 5.1 6.2 6.3
  3 1 7.1 5.9 . .
  3 2 7.3 6.6 8.1 6.8
  3 3 . 7.5 9.1 .
  ; RUN;
DATA BREAD;
  SET BREAD;
  DROP F1-F4;
  FLOUR=1; SPVOL=F1; OUTPUT;
  FLOUR=2; SPVOL=F2; OUTPUT;
  FLOUR=3; SPVOL=F3; OUTPUT;
  FLOUR=4; SPVOL=F4; OUTPUT;
RUN;
PROC PRINT DATA=BREAD; RUN;

```

Continued

TABLE 12.2 (continued)

```

PROC GLM DATA=BREAD;
  CLASSES FLOUR FAT SURF;
  MODEL SPVOL=FLOUR FAT | SURF;
    LSMEANS FAT|SURF/PDIFF STDERR;
    ODSOUTPUT LSMEANS=LSM;
  RUN;

SYMBOL1 V=SQUARE I=JOIN C=BLACK L=1;
SYMBOL2 V=CIRCLE I=JOIN C=BLACK L=2;
SYMBOL3 V=DIAMOND I=JOIN C=BLACK L=3;
AXIS1 ORDER=(1 TO 3 BY 1) OFFSET=(1 CM,);

PROC GPLOT; WHERE EFFECT='FAT_SURF';
  PLOT LSMEAN*FAT=SURF/HAXIS=AXIS1;
  RUN;

```

TABLE 12.3

Model ANOVA and Tests on Main Effects and Interaction

Source	df	Sum of Squares	Mean Square	F-Value	Pr > F
Model	11	22.51952891	2.04722990	12.38	<0.0001
Error	14	2.31585570	0.16541826		
Corrected total	25	24.83538462			
Type III SS					
Flour	3	8.69081097	2.89693699	17.51	<0.0001
Fat	2	10.11784983	5.05892492	30.58	<0.0001
Surfactant	2	0.99720998	0.49860499	3.01	0.0815
Fat × surfactant	4	5.63876453	1.40969113	8.52	0.0011

TABLE 12.4

Fat × Surfactant Least Squares Means

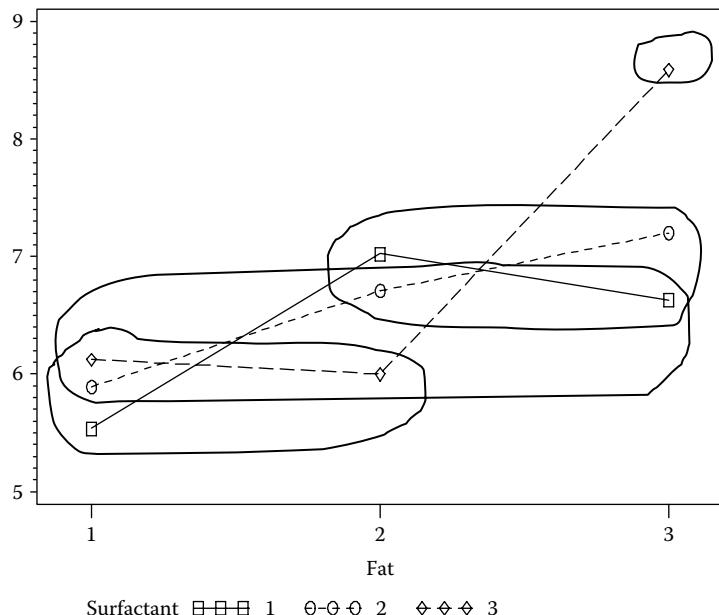
Fat	Surfactant	SPVOL LSMean	Standard Error	Pr > t	LSMean Number
1	1	5.53635388	0.24036653	<0.0001	1
1	2	5.89132489	0.23921852	<0.0001	2
1	3	6.12291175	0.24137422	<0.0001	3
2	1	7.02291175	0.24137422	<0.0001	4
2	2	6.70848186	0.30057982	<0.0001	5
2	3	6.00000000	0.20335822	<0.0001	6
3	1	6.62864505	0.30066843	<0.0001	7
3	2	7.20000000	0.20335822	<0.0001	8
3	3	8.58889843	0.30013634	<0.0001	9

TABLE 12.5Pairwise Comparisons among the Fat \times Surfactant Least Squares Means

Least Squares Means for Effect Fat \times Surfactant Pr> |t| for H₀: LSMean(i) = LSMean(j)
Dependent Variable: SPVOL

<i>i/j</i>	1	2	3	4	5	6	7	8	9
1		0.3156	0.1105	0.0007	0.0098	0.1630	0.0118	0.0001	<0.0001
2	0.3156		0.5099	0.0052	0.0546	0.7344	0.0788	0.0009	<0.0001
3	0.1105	0.5099		0.0169	0.1428	0.7028	0.2203	0.0042	<0.0001
4	0.0007	0.0052	0.0169		0.4184	0.0059	0.3341	0.5836	0.0010
5	0.0098	0.0546	0.1428	0.4184		0.0712	0.8550	0.1971	0.0006
6	0.1630	0.7344	0.7028	0.0059	0.0712		0.1053	0.0009	<0.0001
7	0.0118	0.0788	0.2203	0.3341	0.8550	0.1053		0.1378	0.0004
8	0.0001	0.0009	0.0042	0.5836	0.1971	0.0009	0.1378		0.0018
9	<0.0001	<0.0001	<0.0001	0.0010	0.0006	<0.0001	0.0004	0.0018	

SPVOL LSMean

**FIGURE 12.1** Plot of least squares means. Means located within the same circle are not significantly different.

From Figure 12.1, we can make the following observations:

- 1) The combination of fat 3 with surfactant 3 gives a response that is significantly higher than those given by all other treatment combinations.
- 2) Fat 3 generally gives a response that is significantly higher than that given by fat 1.
- 3) There is no difference in the surfactant levels when they are used with fat 1.

12.2 Concluding Remarks

In the next chapter, the case where some treatment combinations are never observed is discussed. Such cases require additional care in selecting a proper analysis. In this chapter, we considered the analysis of a balanced two-way treatment structure in a randomized complete block design when some treatment combinations are missing in some blocks. The analysis described is appropriate only when each treatment combination is observed at least once. SAS-GLM was used to obtain the analysis. The data set can be more appropriately analysed by considering sources of flow as a random effect as discussed in Chapter 22.

12.3 Exercises

- 12.1 The data in the following table are breaking strengths of beams made from combinations of types of cement and mixtures of aggregate. Four beams were made from each combination, but some of the beams were of poor quality due to the fabrication process and were not tested for strength.
- 1) Use a means model and determine if there is an interaction between the levels of cement and levels of aggregate.
 - 2) Use an effects model and determine if there is an interaction between the levels of cement and levels of aggregate.
 - 3) Carry out a complete analysis of the data.
 - 4) Exclude cement type 3 from the data set and work through parts 1–3 with the reduced data set.

Cement	Aggregate A	Aggregate B	Aggregate C	Aggregate D
Type 1	21	19	19	23
Type 1	27	19	16	24
Type 1	19	22	—	23
Type 1	—	—	—	—
Type 2	25	23	19	28
Type 2	23	20	18	27
Type 2	24	24	—	25
Type 2	—	18	—	—
Type 3	20	28	14	23
Type 3	24	—	16	25
Type 3	—	—	12	22
Type 3	—	—	—	22

13

Using the Means Model to Analyze Two-Way Treatment Structures with Missing Treatment Combinations

In this chapter and the next two chapters, we discuss the analysis of two-way treatment structures when some treatment combinations are never observed. These kinds of experimental situations often occur in practice, mostly by chance but sometimes by design. When the experimenter does have control over the experiment, extreme care should be taken to ensure that all treatment combinations are observed.

Many statistical packages contain routines that calculate test statistics for experiments with missing treatment combinations, but it is shown in this chapter that the observed values of those test statistics often have little, if any, meaning. Thus, the available statistical packages may give the experimenter a false sense of security about the analysis when, in fact, the analysis automatically provided is not generally an analysis of interest. The following sections point out some of the problems and provide methods to obtain an appropriate, meaningful analysis.

13.1 Parameter Estimation

As in Chapter 9, the effect of missing treatment combinations complicates the analysis enough that a very simple set of hypothetical data is used to aid the discussion. A realistic example is discussed in Chapter 15.

Consider the hypothetical data in Table 13.1 from a two-way treatment structure in a completely randomized design with treatments T and B each having three levels. Let μ_{ij} represent the response expected when treatments T_i and B_j are applied to a randomly selected experimental unit. A general means model for this experiment is

$$Y_{ijk} = \mu_{ij} + \varepsilon_{ijk}, \quad i = 1, 2, \dots, t; j = 1, 2, \dots, b; k = 1, 2, \dots, n_{ij} \quad (13.1)$$

TABLE 13.1

A Two-Way Experiment with Missing Treatment Combinations

	B_1	B_2	B_3
T_1	2, 4		7, 6
T_2	3	14	10, 9
T_3	6, 6	9	

TABLE 13.2

Cell Mean Parameters for Data in Table 13.1

	B_1	B_2	B_3
T_1	μ_{11}		μ_{13}
T_2	μ_{21}	μ_{22}	μ_{23}
T_3	μ_{31}	μ_{32}	

where $\varepsilon_{ijk} \sim i.i.d. N(0, \sigma^2)$. If $n_{ij} = 0$ for any i and j , then the treatment combination T_i with B_j is not observed. Table 13.2 contains the cell mean parameters for those treatment combinations that were observed at least once. Note that, even though Table 13.2 does not have cell mean parameters corresponding to the (1, 2) cell and the (3, 3) cell, it is assumed that such parameters exist. That is, we let μ_{12} and μ_{33} represent the means model parameters corresponding to the (1, 2) cell and (3, 3) cell, respectively.

Whenever treatment combinations are missing, certain hypotheses cannot be tested without making some additional assumptions about the parameters in the model. Hypotheses involving parameters corresponding to the missing cells generally cannot be tested. For example, hypotheses that involve μ_{12} and/or μ_{33} cannot be tested without making some assumptions about μ_{12} and μ_{33} . If one were able to assume that μ_{12} was equal to μ_{11} , for example, then a test of a hypothesis involving μ_{12} can be carried out since μ_{12} can be estimated by $\hat{\mu}_{11} = (2 + 4)/2 = 3$ for the data in Table 13.1. Most experimenters would not be willing to make this kind of an assumption. For the data in Table 13.1, it is not possible to estimate (or test hypotheses about) linear combinations that involve the parameters μ_{12} and μ_{33} unless one is willing to make some assumptions about these two parameters. One common assumption is that there is no interaction between the levels of T and the levels of B . This would be equivalent to assuming that

$$\mu_{12} = \mu_{11} - \mu_{21} + \mu_{22} \quad \text{and} \quad \mu_{33} = -\mu_{22} + \mu_{23} + \mu_{32}$$

as well as assuming that

$$\mu_{11} - \mu_{13} - \mu_{21} + \mu_{23} = 0 \quad \text{and} \quad \mu_{21} - \mu_{22} - \mu_{31} + \mu_{32} = 0$$

In our opinion, such assumptions should not be made without some supporting experimental evidence that the assumptions are likely to be true. All too often experimenters are willing to assume no interaction exists among the factors or treatments in any of their experiments, mainly because they do not understand how to deal with such an interaction or because they believe they are not interested in it. Neither of these is a justifiable reason

for assuming no interaction between the two sets of treatments. If interaction exists, the experimenter must deal with it and must be interested in making inferences about the interaction. In Chapter 8, we discussed methods for dealing with interaction when all treatment combinations are observed. The kinds of questions considered there can also be considered here.

As previously stated, it is not possible to make inferences about functions of parameters involving missing treatment combinations. For example, it is not possible to test $\bar{\mu}_{1.} = \bar{\mu}_{2.} = \bar{\mu}_{3.}$ or $\bar{\mu}_{.1} = \bar{\mu}_{.2} = \bar{\mu}_{.3}$ since these hypotheses involve parameters about which we have no information. Indeed, it is not possible to estimate all of the population marginal means. For the above data, one cannot estimate $\bar{\mu}_{1.}$ and $\bar{\mu}_{3.}$, nor can one estimate $\bar{\mu}_{.2}$ and $\bar{\mu}_{.3}$. However, one can estimate $\bar{\mu}_{2.}$ and $\bar{\mu}_{.1}$ since these marginal mean parameters do not involve parameters corresponding to missing cells. As one would expect, the best estimates of the parameters of model (13.1) are

$$\hat{\mu}_{ij} = \bar{y}_{ij..}, \quad i = 1, 2, \dots, t; j = 1, 2, \dots, b \text{ if } n_{ij} > 0$$

and

$$\hat{\sigma}^2 = \frac{\sum_{ijk} (y_{ijk} - \bar{y}_{ij..})^2}{N - C}$$

where $N = n_{..}$ and C = the total number of observed treatment combinations. If $n_{ij} > 0$, the sampling distribution of $\hat{\mu}_{ij}$ is $N(\mu_{ij}, \sigma^2/n_{ij})$, $i = 1, 2, \dots, t; j = 1, 2, \dots, b$ and the sampling distribution of $(N - C)\hat{\sigma}^2/\sigma^2$ is $\chi^2(N - C)$. In addition, $\hat{\mu}_{ij}$, $i = 1, 2, \dots, t; j = 1, 2, \dots, b$ and $\hat{\sigma}^2$ are statistically independent.

13.2 Hypothesis Testing and Confidence Intervals

Clearly, one method of analyzing experiments with missing treatment combinations is to use the procedures discussed in Chapter 1; in fact, this is often the best method. That is, the procedures in Chapter 1 can be used to test hypotheses about any linear combinations of the μ_{ij} corresponding to observed treatment combinations. We illustrate using the data in Table 13.1.

13.2.1 Example 13.1

Suppose we wish to obtain a 95% confidence interval for $\bar{\mu}_{2.}$ in Table 13.1. First the estimates of the means in the observed cells are $\hat{\mu}_{11} = 3$, $\hat{\mu}_{13} = 6.5$, $\hat{\mu}_{21} = 3$, $\hat{\mu}_{22} = 14$, $\hat{\mu}_{23} = 9.5$, $\hat{\mu}_{31} = 6$, $\hat{\mu}_{32} = 9$, and the estimate of σ^2 is

$$\begin{aligned} \hat{\sigma}^2 &= \frac{(2-3)^2 + (4-3)^2 + (7-6.5)^2 + (6-6.5)^2 + (3-3)^2 + (14-14)^2 \\ &\quad + (10-9.5)^2 + (9-9.5)^2 + (6-6)^2 + (6-6)^2 + (9-9)^2}{11-7} \\ &= \frac{3}{4} = 0.75 \end{aligned}$$

Also this estimate of experimental error, $\hat{\sigma}^2$, is based on 4 degrees of freedom. The best estimate of $\bar{\mu}_2$, is $\hat{\bar{\mu}}_2 = (3 + 14 + 9.5)/3 = 5.5$ and its estimated standard error is

$$\widehat{s.e.}(\hat{\bar{\mu}}_2) = \sqrt{\frac{\hat{\sigma}^2 \left(\frac{1}{n_{21}} + \frac{1}{n_{22}} + \frac{1}{n_{23}} \right)}{b}} = \sqrt{\frac{0.75 \left(\frac{1}{1} + \frac{1}{1} + \frac{1}{2} \right)}{3}} = 0.4564$$

Thus a 95% confidence interval for $\bar{\mu}_2$ is $5.5 \pm (2.776)(0.4564) = 5.5 \pm 1.267$ or $4.233 < \bar{\mu}_2 < 6.767$.

Suppose we wish to determine whether there is interaction in this two-way experiment with missing cells. Generally, the test for interaction in a 3×3 experiment would be based on 4 degrees of freedom if all nine treatment combinations were observed, but because of the two missing cells, there are only two linearly independent contrasts that measure interaction in this case. Two linearly dependent contrasts that measure interaction are

$$\mu_{11} - \mu_{13} - \mu_{21} + \mu_{23} \quad \text{and} \quad \mu_{21} - \mu_{22} - \mu_{31} + \mu_{32}$$

Thus, consider testing

$$H_{0T \times B}: \mu_{11} - \mu_{13} - \mu_{21} + \mu_{23} = 0 \quad \text{and} \quad \mu_{21} - \mu_{22} - \mu_{31} + \mu_{32} = 0$$

Using the matrix procedure introduced in Chapter 1, we can test the above hypothesis by testing $\mathbf{C}\boldsymbol{\mu} = \mathbf{0}$ where

$$\mathbf{C} = \begin{bmatrix} 1 & -1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & -1 & 1 \end{bmatrix} \quad \text{and} \quad \boldsymbol{\mu} = \begin{bmatrix} \mu_{11} \\ \mu_{13} \\ \mu_{21} \\ \mu_{22} \\ \mu_{23} \\ \mu_{31} \\ \mu_{32} \end{bmatrix}$$

The sum of squares due to $H_{0T \times B}$ is

$$SS_{T \times B} = (\mathbf{C}\hat{\boldsymbol{\mu}})'(\mathbf{C}\mathbf{D}\mathbf{C})^{-1}(\mathbf{C}\hat{\boldsymbol{\mu}})' \quad \text{where } \mathbf{D} = \text{Diag}\left(\frac{1}{2}, \frac{1}{2}, 1, 1, \frac{1}{2}, \frac{1}{2}, 1\right)$$

We get

$$SS_{T \times B} = \begin{bmatrix} 3 & -8 \end{bmatrix} \begin{bmatrix} \frac{5}{2} & -1 \\ -1 & \frac{7}{2} \end{bmatrix}^{-1} \begin{bmatrix} 3 \\ -8 \end{bmatrix} = 18.5161$$

and this sum of squares is based on 2 degrees of freedom. The corresponding F -statistic is $F = 18.5161/0.75 = 12.34$ with 2 and 4 degrees of freedom; this F is significant at the 0.0194 level.

It is not possible to test $\bar{\mu}_{1\cdot} = \bar{\mu}_{2\cdot} = \bar{\mu}_{3\cdot}$, for the data in Table 13.1 because of the two missing cells. However, it is possible to test

$$H_{0T}: (\mu_{11} + \mu_{13})/2 = (\mu_{21} + \mu_{23})/2 \quad \text{and} \quad (\mu_{21} + \mu_{22})/2 = (\mu_{31} + \mu_{32})/2$$

In a very broad sense, these are T main-effect type hypotheses as the first equation compares T_1 with T_2 after averaging over B_1 and B_3 while the second equation compares T_2 with T_3 after averaging over B_1 and B_2 . An experimenter must determine if either of these hypotheses are of interest, either individually or simultaneously; we use them to illustrate this method.

To test the hypotheses given by H_{0T} , we can take

$$C = \begin{bmatrix} 1 & 1 & -1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & -1 & -1 \end{bmatrix}$$

Again, $D = \text{Diag}\left(\frac{1}{2}, \frac{1}{2}, 1, 1, \frac{1}{2}, \frac{1}{2}, 1\right)$. Then the sum of squares due to H_{0T} is

$$SS_T = (C\hat{\mu})'(CDC)^{-1}(C\hat{\mu})' = \begin{bmatrix} -3 & 2 \end{bmatrix} \begin{bmatrix} \frac{5}{2} & -1 \\ -1 & \frac{7}{2} \end{bmatrix}^{-1} \begin{bmatrix} -3 \\ 2 \end{bmatrix} = 3.8065$$

with 2 degrees of freedom. The appropriate F -statistic is $F = (3.8065/2)/0.75 = 2.5377$ with 2 and 4 degrees of freedom, the corresponding p -value is 0.1943 and hence, H_{0T} cannot be rejected.

The reader might have noticed that CDC' was the same for both of these last two examples; this is coincidental and is not generally true. When some treatment combinations are not observed, it is often best to consider the experiment as a one-way experiment and use computing routines similar to those described in Section 1.7 to answer the important questions. However, since many statistical packages provide certain tests automatically when an effects model is used, many experimenters have preferred them. Such analyses are described in Chapter 14.

13.3 Computer Analyses

The data in Table 13.1 can be analyzed using the SAS®-GLM procedure using the statements in Table 13.3. Since the T and B main effects are not included in the Model statement and the Noint option is used as an option on the Model statement, a two-way means model is used to describe these data. The two Contrast statements that are used in the analysis correspond to the two sums of squares calculated in Section 13.2, namely the sums of squares for H_{0TxB} and H_{0T} .

A listing of the data is in Table 13.4 and the general form of the estimable functions is shown in Table 13.5. An examination of Table 13.5 reveals that every linear combination of the cell means is estimable. Also note that there are no rows in the general form of estimable functions corresponding to the two missing cells.

TABLE 13.3
SAS-GLM Commands

```

DATA;
  INPUT T B Y@@;
CARDS;
1 1      2 1 1      4
1 3      7 1 3 6
2 1      3
2 2 14
2 3 10 2 3 9
3 1      6 3 1 6
3 2      9
PROC PRINT;
TITLE 'EX. 13.1 - A TWO-WAY WITH MISSING CELLS';
RUN:
PROC GLM;
CLASSES T B;
MODEL Y=T*B/NOINT E SOLUTION;
ESTIMATE 'T2_LSM' T*B 0 0 1 1 1 0 0/DIVISOR=3;
CONTRAST 'H0_T*B' T*B 1 -1 -1 0 1 0 0, T*B 0 0 1 -1 0 -1 1;
CONTRAST 'H0_T' T*B 1 1 -1 0 -1 0 0, T*B 0 0 1 1 0 -1 -1;
LSMEANS T*B/PDIFF STDERR;
RUN;

```

TABLE 13.4
Data for Example 13.1

Observation	T	B	Y
1	1	1	2
2	1	1	4
3	1	3	7
4	1	3	6
5	2	1	3
6	2	2	14
7	2	3	10
8	2	3	9
9	3	1	6
10	3	1	6
11	3	2	9

The ANOVA table is shown in Table 13.6. The F -value of 122.10 corresponding to the model F , the type I analysis, and the type III analysis is testing the hypothesis that all μ_{ij} are equal to zero. That is, $F = 122.10$ tests

$$H_0: \mu_{11} = \mu_{13} = \mu_{21} = \mu_{22} = \mu_{23} = \mu_{31} = \mu_{32} = 0$$

Note also that the error mean square is 0.75, which agrees with the estimate of σ^2 given in Section 13.2.

TABLE 13.5

Estimable Functions

Effect	Coefficients	
$T \times B$	1 1	L_1
$T \times B$	1 3	L_2
$T \times B$	2 1	L_3
$T \times B$	2 2	L_4
$T \times B$	2 3	L_5
$T \times B$	3 1	L_6
$T \times B$	3 2	L_7

TABLE 13.6

ANOVA Table for Example 13.1

Dependent Variable: Y

Source	df	Sum of Squares	Mean Square	F-Value	Pr > F
Model	7	641.0000000	91.5714286	122.10	0.0002
Error	4	3.0000000	0.7500000		
Uncorrected total	11	644.0000000			
R-Square					
	Coefficient of Variation	Root MSE	Y Mean		
0.974771	12.53458	0.866025	6.909091		
Source					
$T \times B$	7	Type I SS 641.0000000	91.5714286	122.10	0.0002
Source					
$T \times B$	7	Type III SS 641.0000000	91.5714286	122.10	0.0002

TABLE 13.7

Results from the Contrast Statements for Example 13.1

Contrast	df	Contrast SS	Mean Square	F-Value	Pr > F
H_{0TB}	2	18.51612903	9.25806452	12.34	0.0194
H_{0T}	2	3.80645161	1.90322581	2.54	0.1943

The results from the two Contrast statements are shown in Table 13.7. Note that these agree with the calculations shown in the last section. The results of the Estimate statement are given in Table 13.8. This corresponds to the T_2 main effect mean given in the last section. Table 13.9 contains the results from the Solution option, which are the estimates and standard errors of each of the cell mean parameters. These estimates are the same as those in Table 13.10, which came from the LSMeans statement. Table 13.10, also gives comparisons between all pairs of the two-way cell means, which are produced by including the pdiff option on the LSMeans statement.

TABLE 13.8

Results from the Estimate Statements for Example 13.1

Parameter	Estimate	Standard Error	t-Value	Pr > t
T2_LSM	8.83333333	0.45643546	19.35	<0.0001

TABLE 13.9

Results from Solution Option on the Model Statement for Example 13.1

Parameter	Estimate	Standard Error	t-Value	Pr > t
T×B	1 1 3.0000000	0.61237244	4.90	0.0080
T×B	1 3 6.5000000	0.61237244	10.61	0.0004
T×B	2 1 3.0000000	0.86602540	3.46	0.0257
T×B	2 2 14.0000000	0.86602540	16.17	<0.0001
T×B	2 3 9.5000000	0.61237244	15.51	0.0001
T×B	3 1 6.0000000	0.61237244	9.80	0.0006
T×B	3 2 9.0000000	0.86602540	10.39	0.0005

TABLE 13.10

Two-Way Least Squares Means for Example 13.1

Least Squares Means

T	B	Y LSMean	Standard Error	Pr > t	LSMean Number
1	1	3.0000000	0.6123724	0.0080	1
1	3	6.5000000	0.6123724	0.0004	2
2	1	3.0000000	0.8660254	0.0257	3
2	2	14.0000000	0.8660254	<0.0001	4
2	3	9.5000000	0.6123724	0.0001	5
3	1	6.0000000	0.6123724	0.0006	6
3	2	9.0000000	0.8660254	0.0005	7

*Least Squares Means for Effect T×B*Pr > |t| for $H_0: LSMean(i) = LSMean(j)$

Dependent Variable: Y

i j	1	2	3	4	5	6	7
1		0.0156	1.0000	0.0005	0.0017	0.0257	0.0048
2	0.0156		0.0299	0.0021	0.0257	0.5946	0.0779
3	1.0000	0.0299		0.0009	0.0036	0.0474	0.0080
4	0.0005	0.0021	0.0009		0.0132	0.0017	0.0151
5	0.0017	0.0257	0.0036	0.0132		0.0156	0.6619
6	0.0257	0.5946	0.0474	0.0017	0.0156		0.0474
7	0.0048	0.0779	0.0080	0.0151	0.6619	0.0474	

Note: To ensure overall protection level, only probabilities associated with preplanned comparisons should be used.

13.4 Concluding Remarks

In this chapter we discussed some of the complications that result whenever some treatment combinations are not observed. In this chapter we used the means model (the effects model is used in Chapter 14) to describe the kinds of analyses that are possible. The important thing to remember when some treatment combinations are not observed is that some hypotheses of interest may not be testable unless some additional assumptions can be made about the parameters in the model. However, such assumptions should not be made without evidence to support them.

13.5 Exercises

- 13.1 The following data was collected from a two-way treatment structure in a completely randomized design structure. Use a means model to answer the following questions.

	B1	B2	B3
A1	19, 22, 17	29, 25	
A2		34, 34, 42, 43	26, 31, 23
A3	37		
A4	23, 26	27, 23, 31, 38	30, 14, 18

- 1) Test the equal means hypothesis.
- 2) List the contrasts that measure interaction.
- 3) Use SAS-GLM or another statistical program of your choice to find estimates and their corresponding standard errors for each of the following linear combinations of the μ_{ij} provided that the linear combination is estimable.
 - a) $\bar{\mu}_{3.} - \bar{\mu}_{4.}$
 - b) $\mu_{11} - \mu_{12}$
 - c) $\frac{\mu_{11} + \mu_{12}}{2}$
 - d) $\frac{\mu_{22} + \mu_{23}}{2}$
 - e) $\bar{\mu}_{4.}$
 - f) $\frac{\mu_{12} + \mu_{22} + \mu_{42}}{3}$
 - g) $\bar{\mu}.$
- 4) Find the F -statistic corresponding to each of the following hypotheses if the hypothesis is testable.
 - a) $\frac{\mu_{11} + \mu_{12}}{2} = \frac{\mu_{41} + \mu_{42}}{2}$

- b) $\mu_{41} = \mu_{42} = \mu_{43}$
 c) $\frac{\mu_{21} + \mu_{42}}{2} = \frac{\mu_{23} + \mu_{43}}{2}$

- 13.2 An experiment was conducted to determine the amount of a product (g) produced from 1000 g of product with a given concentration of the target compound where the reaction was carried out at different temperatures. A completely randomized design was used to assign the combinations of concentration of the compound and temperature to a set of runs where each combination was to be observed three times. Some of the runs did not produce usable data, the reaction did not occur at lower concentrations at lower temperatures, and some of the higher concentrations would have produced dangerous results at some of the higher temperatures. Thus there are unequal numbers of observations per cell and some cells could not be observed. Use the following data to:
- 1) Determine the number of linear combinations of the cell means that measure the interaction between the levels of concentration and temperature. Write out those combinations in terms of the cell means and estimate those contrasts.
 - 2) Carry out a complete analysis of the data set and make possible inferences about the levels of concentration, the levels of temperature, and their interactions.

Amount of Product Produced during the Reaction, for Exercise 13.2

	Temperature = 1	Temperature = 2	Temperature = 3	Temperature = 4	Temperature = 5
Concentration = 1				20, 22	26, 29, 29
Concentration = 2			16, 14, 18	22	25, 24, 22
Concentration = 3		12, 15, 13	20, 22, 18	24, 23	29, 30, 26
Concentration = 4	12, 15	17, 12	23, 20, 20	27, 25	
Concentration = 5	14, 16	19, 18, 21	22, 24		

14

Using the Effects Model to Analyze Two-Way Treatment Structures with Missing Treatment Combinations

In Chapter 13 we discussed the use of the means model in analyzing two-way treatment structures when some treatment combinations are not observed. In this chapter we consider the use of the effects model for analyzing the same type of situation. Using the effects model does not enable one to answer any questions that cannot be answered using the means model, and vice versa. While the means model is very simple and easy to understand, the effects model appears to be much more complex than it really is. We prefer to use the means model and are discussing the effects model here only because many statistical packages seem to recommend and encourage the use of the effects model. The effects model considered in this chapter is:

$$y_{ijk} = \mu + \tau_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}, \quad i = 1, 2, \dots, t; \quad j = 1, 2, \dots, b; \quad k = 0, 1, 2, \dots, n_{ij}$$

where $\varepsilon_{ijk} \sim i.i.d. N(0, \sigma^2)$, and $n_{ij} = 0$ implies that the (i, j) th treatment combination is not observed.

14.1 Type I and II Hypotheses

Type I and II analyses for two-way treatment structures with missing treatment combinations can be defined as for treatment structures where all combinations are observed. That is, successive models can be fit, and the resulting reductions in the residual sum of squares are determined as different effects are added to the model.

To illustrate, consider a type I analysis of the data in Table 13.1. We will only consider the results here. Readers interested in the actual model fitting results can refer to the previous edition of this book. The type I and II analyses of variance tables for these data are given in Tables 14.1 and 14.2, respectively.

TABLE 14.1

Type I Analysis of Variance Table

Source of Variation	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	α
Total	10	118.909			
<i>T</i>	2	36.159	18.080	24.11	0.0059
<i>B</i>	2	61.234	30.617	40.82	0.0022
<i>T</i> \times <i>B</i>	2	18.516	9.258	12.34	0.0194
Error	4	3.000	0.75		

TABLE 14.2

Type II Analysis of Variance Table

Source of Variation	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	α
Total	10	118.909			
<i>T</i>	2	13.784	6.892	9.19	0.0319
<i>B</i>	2	61.234	30.617	40.82	0.0022
<i>T</i> \times <i>B</i>	2	18.516	9.258	12.34	0.0194
Error	4	3.000	0.75		

In Chapter 10 it was shown that the type I and II hypotheses may not make much sense when the data are unbalanced even though all treatment combinations are observed at least once. It is perhaps obvious then that these two kinds of hypotheses would not all of a sudden make sense in the case where there are missing treatment combinations. To see that this, in fact, is true, the hypotheses that are being tested by the type I and II analyses are given in Tables 14.3 and 14.4 in terms of a means model and in Tables 14.5 and 14.6 in terms of an effects model. The entries in these tables can be determined from an SAS®-GLM analysis of the data in Table 13.1 or by using the general formulas given in Table 10.9 and Equation 10.3 as these formulas are also correct for the missing-cells problem.

Next, we discuss possible interpretations of the type I and II main-effect hypotheses. The type I and II hypotheses will generally not make much sense unless one's objective is to build a simple model for making predictions rather than to test hypotheses about the effects of the different treatment combinations. For model building, the interpretations are exactly the same as they were in Chapter 10, where we discussed the case with no missing treatment combinations. As in Chapter 10, if the numbers of observations in each cell are

TABLE 14.3

Hypotheses for a Type I Analysis for the Data in Table 13.1 in Terms of the Means Model

Source of Variation	Type I Hypotheses
<i>T</i>	$\mu_{11} + \mu_{13} - \mu_{31} - \mu_{32} = 0$ and $\mu_{21} + \mu_{22} + 2\mu_{23} - 2\mu_{31} - 2\mu_{32} = 0$
<i>B</i>	$5\mu_{11} - 5\mu_{13} + 3\mu_{21} + \mu_{22} - 4\mu_{23} + \mu_{31} - \mu_{32} = 0$ and $\mu_{11} - \mu_{13} + 2\mu_{22} - 2\mu_{23} - \mu_{31} + \mu_{32} = 0$
<i>T</i> \times <i>B</i>	$\mu_{11} - \mu_{13} - \mu_{21} + \mu_{23} = 0$ and $\mu_{21} - \mu_{22} - \mu_{31} + \mu_{32} = 0$

TABLE 14.4

Hypotheses for a Type II Analysis for the Data in Table 13.1 in Terms of the Means Model

Source of Variation	Type II Hypotheses
T	$2\mu_{11} + \mu_{13} + \mu_{22} - \mu_{23} - 2\mu_{31} - \mu_{32} = 0$ and $2\mu_{11} - 2\mu_{13} + 3\mu_{21} + 4\mu_{22} + 2\mu_{23} - 5\mu_{31} - 4\mu_{32} = 0$
B	$5\mu_{11} - 5\mu_{13} + 3\mu_{21} + \mu_{22} - 4\mu_{23} + \mu_{31} - \mu_{32} = 0$ and $\mu_{11} - \mu_{13} + 2\mu_{22} - 2\mu_{23} - \mu_{31} + \mu_{32} = 0$
$T \times B$	$\mu_{11} - \mu_{13} - \mu_{21} + \mu_{23} = 0$ and $\mu_{21} - \mu_{22} - \mu_{31} + \mu_{32} = 0$

TABLE 14.5

Hypotheses for a Type I Analysis for the Data in Table 13.1 in Terms of the Effects Model

Source of Variation	Type I Hypotheses
T	$\tau_1 - \tau_3 - 1/2\beta_2 + 1/2\beta_3 + 1/2\gamma_{11} + 1/2\gamma_{13} - 1/2\gamma_{31} - 1/2\gamma_{32} = 0$ and $\tau_2 - \tau_3 - 1/4\beta_1 - 1/4\beta_2 + 1/2\beta_3 + 1/4\gamma_{21} + 1/4\gamma_{22} + 1/2\gamma_{23} - 1/2\gamma_{31} - 1/2\gamma_{32} = 0$
B	$\beta_1 - \beta_3 + (1/9)(5\gamma_{11} - 5\gamma_{13} + 3\gamma_{21} + \gamma_{22} - 4\gamma_{23} + \gamma_{31} - \gamma_{32}) = 0$ and $\beta_2 - \beta_3 + (1/3)(\gamma_{11} - \gamma_{13} + 2\gamma_{22} - 2\gamma_{23} - \gamma_{31} + \gamma_{32}) = 0$
$T \times B$	$\gamma_{11} - \gamma_{13} - \gamma_{21} + \gamma_{23} = 0$ and $\gamma_{21} - \gamma_{22} - \gamma_{31} + \gamma_{32} = 0$

TABLE 14.6

Hypotheses for a Type II Analysis for the Data in Table 13.1 in Terms of the Effects Model

Source of Variation	Type II Hypotheses
T	$\tau_1 - \tau_3 + (1/3)(2\gamma_{11} + \gamma_{13} + \gamma_{22} - \gamma_{23} - 2\gamma_{31} - \gamma_{32}) = 0$ and $\tau_2 - \tau_3 + (1/9)(2\gamma_{11} - 2\gamma_{13} + 3\gamma_{21} + 4\gamma_{22} + 2\gamma_{23} - 5\gamma_{31} - 4\gamma_{32}) = 0$
B	$\beta_1 - \beta_3 + (1/9)(5\gamma_{11} - 5\gamma_{13} + 3\gamma_{21} + \gamma_{22} - 4\gamma_{23} + \gamma_{31} - \gamma_{32}) = 0$ and $\beta_2 - \beta_3 + (1/3)(\gamma_{11} - \gamma_{13} + 2\gamma_{22} - 2\gamma_{23} - \gamma_{31} + \gamma_{32}) = 0$
$T \times B$	$\gamma_{11} - \gamma_{13} - \gamma_{21} + \gamma_{23} = 0$ and $\gamma_{21} - \gamma_{22} - \gamma_{31} + \gamma_{32} = 0$

proportional to the actual numbers of each treatment combination existing in the population, then the experimenter might be interested in $R(\tau|\mu)$ and $R(\beta|\mu)$. Both can be obtained by conducting two type I analyses, one with T first in the model statement and the other with B first in the model statement.

14.2 Type III Hypotheses

When all treatment combinations are observed, the type III hypotheses are the same as those tested when there are equal subclass numbers. When some treatment combinations are missing, such hypotheses cannot be tested since they involve parameters about which there is no information. For the data in Table 13.1, we cannot estimate $\bar{\mu}_1$ and $\bar{\mu}_3$ since

we cannot estimate μ_{12} and μ_{33} . Likewise, we cannot estimate $\bar{\mu}_{.2}$ and $\bar{\mu}_{.3}$. Hence, it is not possible to test $\bar{\mu}_{1.} = \bar{\mu}_{2.} = \bar{\mu}_{3.}$ or $\bar{\mu}_{.1} = \bar{\mu}_{.2} = \bar{\mu}_{.3}$.

Both the type I and II hypotheses for the main effects depend on the numbers of observations in each cell. As long as there is at least one observation in a cell, then that cell mean is estimable. Thus functions of the parameters that are estimable depend only on which of the treatment combinations are observed and not on how many times they are observed.

Type III hypotheses are developed so that they do not depend on the cell sizes, but only on which cells are observed. This is consistent with the definition of type III hypotheses for two-way experiments, where all treatment combinations are observed. That is, the hypotheses $\bar{\mu}_{1.} = \bar{\mu}_{2.} = \bar{\mu}_{3.}$ and $\bar{\mu}_{.1} = \bar{\mu}_{.2} = \bar{\mu}_{.3}$ do not depend on the cell sizes. We are not going to discuss the construction of type III hypotheses for the missing data case. Even though the objectives may seem reasonable, we think that the type III hypotheses are the worst hypotheses to consider when there are missing cells because there seems to be no reasonable way to interpret them. To illustrate, a type III analysis of the data in Table 13.1 is given in Table 14.7. The hypotheses being tested by the type III analysis are given in Tables 14.8 and 14.9. Examination of Tables 14.8 and 14.9 reveals that the type III hypotheses

TABLE 14.7

Type III Analysis of Variance Table

Source of Variation	df	SS	MS	F	α
Total	10	118.909			
T	2	10.788	5.394	7.19	0.0473
B	2	68.906	34.453	45.94	0.0017
$T \times B$	2	18.516	9.258	12.34	0.0194
Error	4	3.000	0.75		

TABLE 14.8

Hypotheses for a Type III Analysis for the Data in Table 13.1 in Terms of the Means Model

Source of Variation	Type III Hypotheses
T	$2\mu_{11} + \mu_{13} + \mu_{22} - \mu_{23} - 2\mu_{31} - \mu_{32} = 0$ and $2\mu_{11} - 2\mu_{13} + 6\mu_{21} + 7\mu_{22} + 2\mu_{23} - 8\mu_{31} - 7\mu_{32} = 0$
B	$7\mu_{11} - 7\mu_{13} + 6\mu_{21} + 2\mu_{22} - 8\mu_{23} + 2\mu_{31} - 2\mu_{32} = 0$ and $\mu_{11} - \mu_{13} + 2\mu_{22} - 2\mu_{23} - \mu_{31} + \mu_{32} = 0$
$T \times B$	$\mu_{11} - \mu_{13} - \mu_{21} + \mu_{23} = 0$ and $\mu_{21} - \mu_{22} - \mu_{31} + \mu_{32} = 0$

TABLE 14.9

Hypotheses for a Type III Analysis for the Data in Table 13.1 in Terms of the Effects Model

Source of Variation	Type III Hypotheses
T	$\tau_1 - \tau_3 + (1/3)(2\gamma_{11} + \gamma_{13} + \gamma_{22} - \gamma_{23} - 2\gamma_{31} - \gamma_{32}) = 0$ and $\tau_2 - \tau_3 + (1/15)(2\gamma_{11} - 2\gamma_{13} + 6\gamma_{21} + 7\gamma_{22} + 2\gamma_{23} - 8\gamma_{31} - 7\gamma_{32}) = 0$
B	$\beta_1 - \beta_3 + (1/15)(7\gamma_{11} - 7\gamma_{13} + 6\gamma_{21} + 2\gamma_{22} - 8\gamma_{23} + 2\gamma_{31} - 2\gamma_{32}) = 0$ and $\beta_2 - \beta_3 + (1/3)(\gamma_{11} - \gamma_{13} + 2\gamma_{22} - 2\gamma_{23} - \gamma_{31} + \gamma_{32}) = 0$
$T \times B$	$\gamma_{11} - \gamma_{13} - \gamma_{21} + \gamma_{23} = 0$ and $\gamma_{21} - \gamma_{22} - \gamma_{31} + \gamma_{32} = 0$

are not meaningful except for the $T \times B$ interaction. The results in Tables 14.7–14.9 were taken from a SAS-GLM analysis of the data in Table 13.1.

14.3 Type IV Hypotheses

In the previous two sections, we attempted to show that none of the so-called main-effect hypotheses tested by the type I, type II, or type III analyses are entirely satisfactory when there are missing treatment combinations, since they rarely have any reasonable interpretations if there is a possibility of interaction between the two factors. Such hypotheses are extremely difficult to interpret because the coefficients of cell means occurring in the same row or column are rarely the same. Type IV hypotheses are constructed so that the cell mean coefficients are balanced; hence, the resulting hypotheses are interpretable.

To illustrate, let us look at all possible type IV hypotheses that are testable in the data set of Table 13.1. Table 14.10 gives the cell mean parameters for the non-missing cells. *Basically, for a two-way treatment structure, a hypothesis is defined to be a Type IV type hypothesis if it compares the levels of one treatment averaged over one or more common levels of the other treatment.* Thus, the hypotheses are expected to be marginal means hypotheses except that, when treatment combinations are missing, one cannot average across all levels of the other treatment but only across some of the levels of the other factor.

A type IV hypothesis that compares T_1 with T_2 after averaging over B_1 and B_3 is $H_0: (\mu_{11} + \mu_{13})/2 = (\mu_{21} + \mu_{23})/2$. Another type IV hypothesis that compares T_1 with T_2 is $H_0: \mu_{11} = \mu_{21}$. This latter hypothesis compares T_1 with T_2 after averaging over B_1 only. All possible type IV hypotheses for T in terms of means model parameters are given in

TABLE 14.10

Cell Mean Parameters for Data in Table 13.1

	B_1	B_2	B_3
T_1	μ_{11}	—	μ_{13}
T_2	μ_{21}	μ_{22}	μ_{23}
T_3	μ_{31}	μ_{32}	—

TABLE 14.11

All Possible Type IV Hypotheses for T for the Data in Table 13.1

$$\frac{\mu_{11} + \mu_{13}}{2} = \frac{\mu_{21} + \mu_{23}}{2}$$

$$\frac{\mu_{21} + \mu_{22}}{2} = \frac{\mu_{31} + \mu_{32}^*}{2}$$

$$\mu_{11} = \mu_{21}$$

$$\mu_{11} = \mu_{31}^*$$

$$\mu_{21} = \mu_{31}$$

$$\mu_{22} = \mu_{32}$$

$$\mu_{13} = \mu_{23}$$

* Hypotheses automatically tested by a SAS-GLM type IV analysis.

TABLE 14.12

All Possible Type IV Hypotheses for B for the Data in
Table 13.1

$\frac{\mu_{11} + \mu_{21}}{2} = \frac{\mu_{13} + \mu_{23}^*}{2}$
$\frac{\mu_{21} + \mu_{31}}{2} = \frac{\mu_{22} + \mu_{32}}{2}$
$\mu_{11} = \mu_{13}$
$\mu_{21} = \mu_{22}$
$\mu_{21} = \mu_{23}$
$\mu_{22} = \mu_{23}^*$
$\mu_{31} = \mu_{32}$

* Hypotheses automatically tested by a SAS-GLM type IV analysis.

Table 14.11. Similar results can be obtained for the type IV hypotheses for B . All possible type IV hypotheses for B are given in Table 14.12.

SAS-GLM automatically generates type IV hypotheses that can usually be interpreted, but an appropriate interpretation cannot be made without first examining the type IV estimable functions to see exactly what hypotheses SAS-GLM generated. That is, there is no unique interpretation appropriate for all data sets as there is in the case when there are no missing treatment combinations. In fact, relabeling the treatments before doing the analysis may result in different type IV hypotheses being generated, and hence there will be different sums of squares and F -values in the type IV analysis. Thus, the type IV analysis obtained is not a unique characteristic of the data and depends on how the treatments are labeled. Obviously, this is not very desirable, but it is unavoidable. SAS-GLM indicates this situation has occurred by placing an asterisk on the printed degrees of freedom and noting that "Other Type IV Testable Hypotheses exist which may yield different SS." The type IV hypothesis for T that was automatically tested by an SAS-GLM analysis of the data in Table 13.1 is equivalent to simultaneously testing and

$$\mu_{11} = \mu_{31} \quad \text{and} \quad \frac{\mu_{21} + \mu_{22}}{2} = \frac{\mu_{31} + \mu_{32}}{2}$$

Thus, the type IV hypothesis for T simultaneously compares the effect of T_1 and T_3 at level 1 of B and the effect of T_2 and T_3 averaged over levels 1 and 2 of B . These two functions are identified with the asterisks in Table 14.11. We note that level 3 of B is not involved at all in this particular set. The type IV analysis of variance table obtained from SAS-GLM is given in Table 14.13.

Type IV hypotheses not tested by SAS-GLM, but perhaps just as interesting to the experimenter as those that are automatically tested, are given in Table 14.11. In order to test such interesting type IV hypotheses, one can (and should) use Estimate or Contrast options. For example, to test all of the type IV hypotheses in Table 14.11 when using the effects model, we would use the following Estimate statements in a SAS-GLM analysis:

```
ESTIMATE 'T1 VS T2 AVE OVER B1 AND B3' T 1 -1 0 T*B .5 .5 -.5 0 -.5 0 0;
ESTIMATE 'T2 VS T3 AVE OVER B1 AND B2' T 0 1 -1 T*B 0 0 .5 .5 0 -.5 -.5;
ESTIMATE 'T1 VS T2 AT B1' T 1 -1 0 T*B 1 0 -1 0 0 0 0;
ESTIMATE 'T1 VS T3 AT B1' T 1 0 -1 T*B 1 0 0 0 0 -1 0;
ESTIMATE 'T2 VS T3 AT B1' T 0 1 -1 T*B 0 0 1 0 0 -1 0;
```

TABLE 14.13
Type IV Analysis of Variance Table

Source of Variation	df	SS	MS	F	α
Total	10	118.909			
T	2*	12.769	6.385	8.51	0.0362
B	2*	70.179	35.089	46.79	0.0017
$T \times B$	2	18.516	9.258	12.34	0.0194
Error	4	3.000	0.75		

* Other type IV testable hypotheses exist which may yield different SS.

TABLE 14.14
Results of ESTIMATE Options

Parameter	Estimate	Standard Error	t-Value	Pr > t
T1 vs T2 ave over B1 and B3	-1.50000000	0.68465320	-2.19	0.0936
T2 vs T3 ave over B1 and B2	1.00000000	0.81009259	1.23	0.2846
T1 vs T2 at B1	-0.00000000	1.06066017	-0.00	1.0000
T1 vs T3 at B1	-3.00000000	0.86602540	-3.46	0.0257
T2 vs T3 at B1	-3.00000000	1.06066017	-2.83	0.0474
T2 vs T3 at B2	5.00000000	1.22474487	4.08	0.0151
T1 vs T2 at B3	-3.00000000	0.86602540	-3.46	0.0257

```
ESTIMATE 'T2 VS T3 AT B2' T 0 1 -1 T*B 0 0 0 1 0 0 -1;
ESTIMATE 'T1 VS T2 AT B3' T 1 -1 0 T*B 0 1 0 0 -1 0 0;
```

The results of the above Estimate statements are shown in Table 14.14.

If the reader does not have SAS-GLM available, all of the hypotheses in Tables 14.11 and 14.12 can be tested by using the means model and procedures similar to those given in Section 13.2.

14.4 Population Marginal Means and Least Squares Means

Population marginal means and least squares means are defined here in the same way that they were defined in Sections 9.5 and 10.6. However, if a particular treatment is not observed with all possibilities for the other treatment factor, then the corresponding population marginal mean is not estimable. In this case, the table of two-way cell means (for example, the $\bar{\mu}_{ij}$) can be used to compare each observed treatment combination with all other observed treatment combinations. If a data set is quite sparse, few if any population marginal means will be estimable.

For the data in Table 13.1, $\bar{\mu}_{2\cdot}$ and $\bar{\mu}_{\cdot 1}$ are the only population marginal means that are estimable. Their best estimates are $\hat{\bar{\mu}}_{2\cdot} = 8.333$ and $\hat{\bar{\mu}}_{\cdot 1} = 4.00$, respectively. In general, the best estimate of $\sum_{i,j} c_{ij} \bar{\mu}_{ij}$ is $\sum_{i,j} c_{ij} \hat{\bar{\mu}}_{ij}$, and its estimated standard error is $\hat{\sigma} \sqrt{[\sum_{i,j} (c_{ij}^2/n_{ij})]}$ where the

sums are taken over all nonempty cells. A $(1 - \alpha)100\%$ confidence interval for $\sum_{i,j} c_{ij}\mu_{ij}$ is given by

$$\sum_{i,j} c_{ij}\hat{\mu}_{ij} \pm t_{\alpha/2,v} \hat{\sigma} \sqrt{\frac{c_{ij}^2}{n_{ij}}}$$

A t -statistic with v degrees of freedom for testing $\sum_{i,j} c_{ij}\mu_{ij} = 0$ is given by

$$t = \frac{\sum_{i,j} c_{ij}\hat{\mu}_{ij}}{\hat{\sigma} \sqrt{\sum_{i,j} \frac{c_{ij}^2}{n_{ij}}}}$$

In both instances, v is the degrees of freedom corresponding to $\hat{\sigma}^2$, the error mean square.

For researchers wishing to make multiple comparisons, we recommend using the observed p -values given by the above t -tests whenever the F -value for comparing all treatment combinations is significant. If this F -value is not significant, then one should use Bonferroni's method on all comparisons of interest. That is, declare that linear combinations are significantly different from zero if the p -value obtained is less than α/p where p is the total number of planned comparisons. For data snooping and unplanned comparisons, one should use a Scheffé procedure.

14.5 Computer Analyses

The reader should use his or her own statistical package to analyze the examples given in this chapter and in Chapters 15 and 17. Comparing the results of the analyses so obtained with those given in this book will give the reader valuable insight into the kinds of hypotheses tested by the packages she is accustomed to using. We know of no package that handles the analysis of data with missing treatment combinations adequately or completely. Several do a good job with unbalanced data provided that there are no missing treatment combinations.

Anyone who does many statistical analyses on data with missing treatment combinations should learn how to use a package that allows a specified set of hypotheses to be tested. Then, and only then, can one be sure that the hypotheses tested are reasonable, meaningful, and interpretable. SAS-GLM and SAS-Mixed allow users to specify their own hypotheses.

14.6 Concluding Remarks

In summary, an acceptable analysis of data with missing treatment combinations requires a great deal of thought. An experimenter or statistician cannot simply run a computer

program on the data and then select numbers from that program to report in a paper. Unfortunately, this has been done and is being done by uninformed experimenters and data analysts. We hope that anyone who has studied this chapter will never do so again. Those willing to exert the necessary effort to analyze their data correctly are advised to use the means model discussed in Chapter 13.

A more realistic example is discussed in Chapter 15.

14.7 Exercises

- 14.1 The data in Exercise 13.1 were collected from a two-way treatment structure in a completely randomized design structure. Use an effects model to answer the following questions.
- 1) List all possible type IV hypotheses for A .
 - 2) List all possible type IV hypotheses for B .
 - 3) Find the type IV sums of squares for both A and B that are automatically computed by SAS-GLM.
 - 4) Give the hypotheses being tested by the analysis in part 3.
 - 5) Compute LSMEANS for A , for B , and for $A \times B$; perform all possible multiple comparisons among the two-way $A \times B$ means.
 - 6) Compute another type IV sum of squares for both A (with 3 degrees of freedom) and B (with 2 degrees of freedom) that are different from the type IV sum of squares automatically given by SAS-GLM in part 1.
 - 7) Compare the analyses from Exercise 13.1 and this one, that is, compare the means model analysis to the effects model analysis.

15

Case Study: Two-Way Treatment Structure with Missing Treatment Combinations

15.1 Case Study

In Chapters 13 and 14 we discussed the analysis of two-way treatment structures in a completely randomized design structure when there are missing treatment combinations. In this chapter we illustrate how to analyze a two-way treatment structure in a randomized complete block design when some treatment combinations are not observed. Consider the data in Table 15.1, which is obtained from the experiment described in Chapter 12, but with a few treatment combinations not being observed in any of the blocks. Figure 15.1 shows the treatment combinations observed at least once. Any hypothesis that involves treatment combinations (fat 1, surfactant 3) or (fat 2, surfactant 2) cannot be tested unless additional assumptions are made. In this discussion, let FS_{ij} represent the response expected when fat i and surfactant j are assigned to a randomly selected experimental unit.

To get the error sum of squares for this experiment, we can fit either an effects model or a means model in a randomized block design structure using any of the available statistical packages. The model to fit an effects model is

```
MODEL SPVOL = BLK FAT SURF FAT*SURF;
```

And the model to fit a means model for the two-way Fat \times Surfactant combinations is

```
MODEL SPVOL = BLK FAT*SURF / NOINT;
```

After fitting either of these models, one finds that the error sum of squares is equal to 2.0941, with 11 degrees of freedom. Thus

$$\hat{\sigma}^2 = 2.0941/11 = 0.1904$$

If all treatment combinations had been observed, there would have been 4 degrees of freedom for interaction hypotheses. Since two treatment combinations are never observed,

TABLE 15.1

Specific Volumes from the Baking Experiment in Chapter 12

Fat	Surfactant	Flour			
		1	2	3	4
1	1	6.7	4.3	5.7	
	2	7.1		5.9	5.6
	3				
2	1		5.9	7.4	7.1
	2				
	3	6.4	5.1	6.2	6.3
3	1	7.1	5.9		
	2	7.3	6.6	8.1	6.8
	3		7.5	9.1	

only two degrees of freedom remain for interaction hypotheses. Two independent contrasts in the interaction space are $FS_{11} - FS_{12} - FS_{31} + FS_{32}$ and $FS_{21} - FS_{23} - FS_{31} + FS_{33}$. An SAS®-GLM type IV analysis tests these two contrasts equal to zero simultaneously. The value of the F -statistic for testing the two contrasts equal to zero simultaneously is

$$F = (5.4002/2)/0.1904 = 14.18$$

with 2 and 11 degrees of freedom.

All possible type IV hypotheses for fat are given in Table 15.2. The hypothesis in 4 was automatically tested by the SAS-GLM type IV analysis for fat. Hypotheses 1–3 can be tested with Contrast statements should one want to do so, and hypotheses 2 and 3 can also be examined using Estimate statements. The results of the statistical tests performed are also shown in Table 15.2.

All possible type IV hypotheses for surfactant are specified in Table 15.3. The last equality in hypothesis 3 and hypothesis 5 were automatically tested by an SAS-GLM

		Surfactant		
		1	2	3
Fat	1	x	x	—
	2	x	—	x
	3	x	x	x

FIGURE 15.1 Treatment combination observed in baking experiment.**TABLE 15.2**

Type IV Hypotheses for Fat

	Hypothesis	df	F	p-Value
1	$FS_{11} = FS_{21} = FS_{31}$	2	8.69	0.006
2	$FS_{12} = FS_{32}$	1	15.18	0.003
3	$FS_{23} = FS_{33}$	1	43.89	0.000
4	$FS_{11} + FS_{12} = FS_{31} + FS_{32}$	2	13.25	0.001
	$FS_{21} + FS_{23} = FS_{31} + FS_{32}$			

TABLE 15.3
Type IV Hypotheses for Surfactant

	Hypothesis	df	F	p-Value
1	$FS_{11} = FS_{12}$	1	0.85	0.376
2	$FS_{21} = FS_{23}$	1	9.11	0.012
3	$FS_{31} = FS_{32} = FS_{33}$	2	9.95	0.003
4	$FS_{11} + FS_{31} = FS_{12} + FS_{32}$	1	2.64	0.132
5	$FS_{21} + FS_{31} = FS_{23} + FS_{33}$	1	2.79	0.123

type IV analysis for surfactant. The F -value given was $F = 6.34$ with 2 and 11 degrees of freedom. All five hypotheses can be tested using Contrast statements, and all but hypothesis 3 can be tested with an Estimate statement.

Since there is significant interaction in these data, it is probably best to compare all observed treatment combinations by examining the least squares means of the seven treatment combinations observed. The least squares means and the p -values for comparing them to one another using pairwise t -tests are given in Table 15.4. We noted that, when the

TABLE 15.4
Least Squares Means, t -Statistics and p -Values for Pairwise Comparisons

Least Squares Means				
Fat	Surfactant	SPVOL LSMean	LSMean Number	
1	1	5.54120221		1
1	2	5.88331748		2
2	1	7.02225604		3
2	3	6.00000000		4
3	1	6.64163972		5
3	2	7.20000000		6
3	3	8.59518737		7

Least Squares Means for Effect Fat \times Surfactant t for $H_0: LSMean(i) = LSMean(j)/Pr > |t|$
Dependent Variable: SPVOL

<i>i/j</i>	1	2	3	4	5	6	7
1	-0.92322 0.3757	-3.98937 0.0021	-1.35133 0.2037	-2.71158 0.0202	-4.88578 0.0005	-7.50717 0.0001	
2	0.923221 0.3757	-3.09314 0.0102	-0.34525 0.7364	-1.79001 0.1010	-3.89591 0.0025	-6.39489 0.0001	
3	3.989367 0.0021	3.093142 0.0102	3.019 0.0117	0.894121 0.3904	-0.52493 0.6101	-3.85221 0.0027	
4	1.351332 0.2037	0.34525 0.7364	-3.019 0.0117	-1.6363 0.1300	-3.88951 0.0025	-6.62522 0.0001	
5	2.711581 0.0202	1.790011 0.1010	-0.89412 0.3904	1.636298 0.1300	-1.42392 0.1822	-4.26043 0.0013	
6	4.885782 0.0005	3.895912 0.0025	0.524926 0.6101	3.889509 0.0025	1.42392 0.1822	-3.56176 0.0045	
7	7.507171 <0.0001	6.394888 <0.0001	3.852207 0.0027	6.625224 <0.0001	4.260426 0.0013	3.561758 0.0045	

Note: To ensure overall protection level, only probabilities associated with preplanned comparisons should be used.

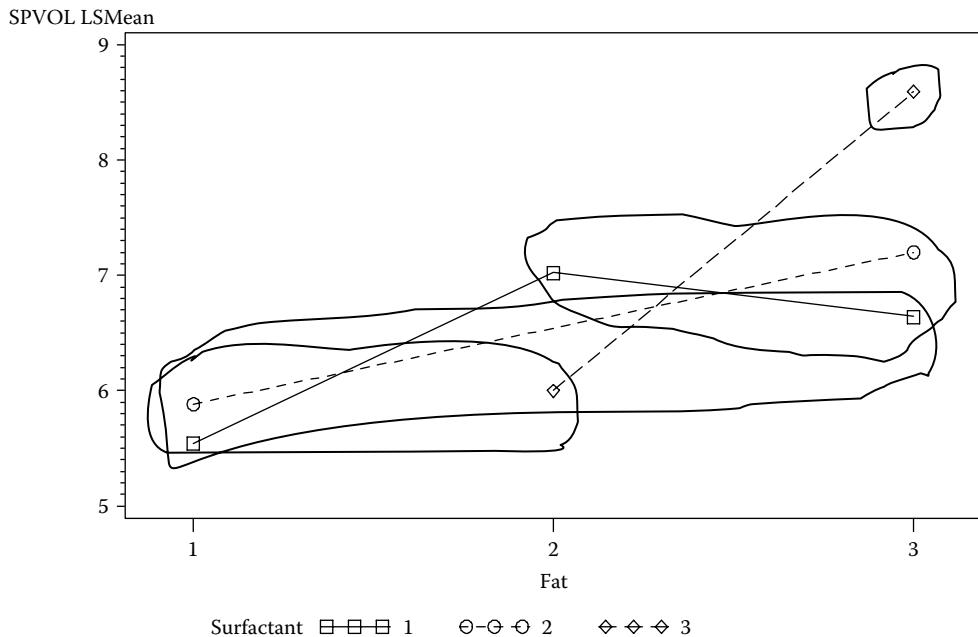


FIGURE 15.2 Fat \times surfactant least squares means. Means located within the same circle are not significantly different.

design structure is completely randomized, the best estimates of the population cell means are the means of the observations in each cell. (This is not true in a randomized block design.) The best estimates can easily be obtained with a computing package or by using the methods of Chapter 6. From such estimates, one can construct Figure 15.2, where means that are not significantly different have been enclosed in the same circle.

15.2 Concluding Remarks

This chapter illustrated the analysis of a two-way treatment structure experiment in a randomized complete block design when some treatment combinations are not observed.

15.3 Exercise

- 15.1 A study was conducted to investigate the effects of three different levels of exercise on the systolic blood pressure of female humans in five age groups. Subjects were selected from three different gyms where gym is considered to be a blocking factor. The subjects were randomly assigned to an exercise level within a gym and their systolic blood pressure was evaluated after six months on the program. Four subjects were selected from each age group at the beginning of the

study, but several dropped out before the end of the six months, thus providing an unbalanced data set. The systolic blood pressure data are given in the following table. The age groups are 1 = (age < 25), 2 = (25 ≤ age < 35), 3 = (35 ≤ age < 45), 4 = (45 ≤ age < 55), and 5 = (age ≥ 55). The exercise levels are 1 = (low intensity for 30 min), 2 = (medium intensity for 45 min), and 3 = (high intensity for 60 min), where each person exercises 3 days a week.

Exercise Level	Age 1	Age 2	Age 3	Age 4	Age 5
<i>Gym = 1</i>					
1	—	—	132	134	138
1	—	—	—	—	137
1	—	—	—	—	129
2	—	118	121	142	—
2	—	129	123	140	—
2	—	115	—	131	—
3	110	109	126	—	—
3	116	117	120	—	—
3	113	—	125	—	—
3	107	—	108	—	—
<i>Gym = 2</i>					
1	—	—	116	—	136
1	—	—	128	—	—
1	—	—	134	—	—
2	—	114	118	124	—
2	—	108	118	125	—
2	—	121	115	132	—
2	—	—	118	135	—
3	—	98	118	—	—
3	—	119	123	—	—
3	—	118	—	—	—
3	—	110	—	—	—
<i>Gym = 3</i>					
1	—	—	137	139	142
1	—	—	141	148	147
1	—	—	144	129	141
1	—	—	133	137	—
2	—	120	120	137	—
2	—	120	122	143	—
2	—	129	117	127	—
3	118	117	138	—	—
3	110	123	—	—	—

- 1) Determine the number of linearly independent interaction contrasts between age and exercise intensity, and write out one set.
- 2) Use both the means model and the effects model to test the set of interaction contrasts that you gave in part 1.

- 3) Determine all of the type IV estimable functions for age main effects.
- 4) Determine all of the type IV estimable functions for the exercise intensity main effects.
- 5) Carry out a complete analysis using the effects model (include tests of hypotheses, confidence intervals and multiple comparisons).
- 6) Carry out a complete analysis using the means model (include tests of hypotheses, confidence intervals and multiple comparisons).

16

Analyzing Three-Way and Higher-Order Treatment Structures

In Chapters 7–15, we discussed the analysis of two-way treatment structures. The methods and results given in those nine chapters can be generalized to more complex treatment structures; such analyses become only slightly more complicated as the complexity of the treatment structure increases. We illustrate the method of generalization by specifically addressing the analysis of three-way treatment structures.

In Section 16.1 we give a general strategy to follow when analyzing higher-order treatment structures. Section 16.2 discusses the analysis of balanced and unbalanced treatment structures. The discussion of unbalanced experiments includes the case where each treatment combination is observed at least once and the case where some treatment combinations are missing.

16.1 General Strategy

Suppose that treatments T_i , B_j , and C_k are applied simultaneously to the same experimental unit. Let μ_{ijk} represent the expected response to the treatment combination (T_i, B_j, C_k) for $i = 1, 2, \dots, t$; $j = 1, 2, \dots, b$; and $k = 1, 2, \dots, c$. There is no three-way interaction among these treatment combinations provided that

$$(\mu_{ijk} - \mu_{i'jk} - \mu_{ij'k} + \mu_{i'j'k}) - (\mu_{ijk'} - \mu_{i'jk'} - \mu_{ij'k'} + \mu_{i'j'k'}) = 0 \quad \text{for all } i, i', j, j', k, \text{ and } k'$$

This implies that the $T \times B$ interaction at level k of factor C is the same as the $T \times B$ interaction at level k' of factor C for all values of k and k' . Similarly, the $T \times C$ interaction is the

same at all levels of factor B , and the $B \times C$ interaction is the same at all levels of factor T . Equivalent expressions of the no-interaction statements are:

- 1) $\mu_{ijk} - \bar{\mu}_{ij\cdot} - \bar{\mu}_{i\cdot k} - \bar{\mu}_{\cdot jk} + \bar{\mu}_{i\cdot\cdot} + \bar{\mu}_{\cdot j\cdot} + \bar{\mu}_{\cdot\cdot k} - \bar{\mu}_{\dots} = 0$ for all i, j , and k .
- 2) There exist parameters $\mu, \tau_1, \tau_2, \dots, \tau_b, \beta_1, \beta_2, \dots, \beta_{bc}, \xi_1, \xi_2, \dots, \xi_c, \gamma_{11}, \gamma_{12}, \dots, \gamma_{tb}, \eta_{11}, \eta_{12}, \dots, \eta_{tc}, \theta_{11}, \theta_{12}, \dots, \theta_{bc}$ such that

$$\mu_{ijk} = \mu + \tau_i + \beta_j + \xi_k + \gamma_{ij} + \eta_{ik} + \theta_{jk} \quad \text{for all } i, j, \text{ and } k$$

That is, the μ_{ijk} can be described by main effects and two-factor interaction effects. When analyzing three-way treatment structures, the first and most important step is to determine whether there is a three-factor interaction, even though the experimenter may not be

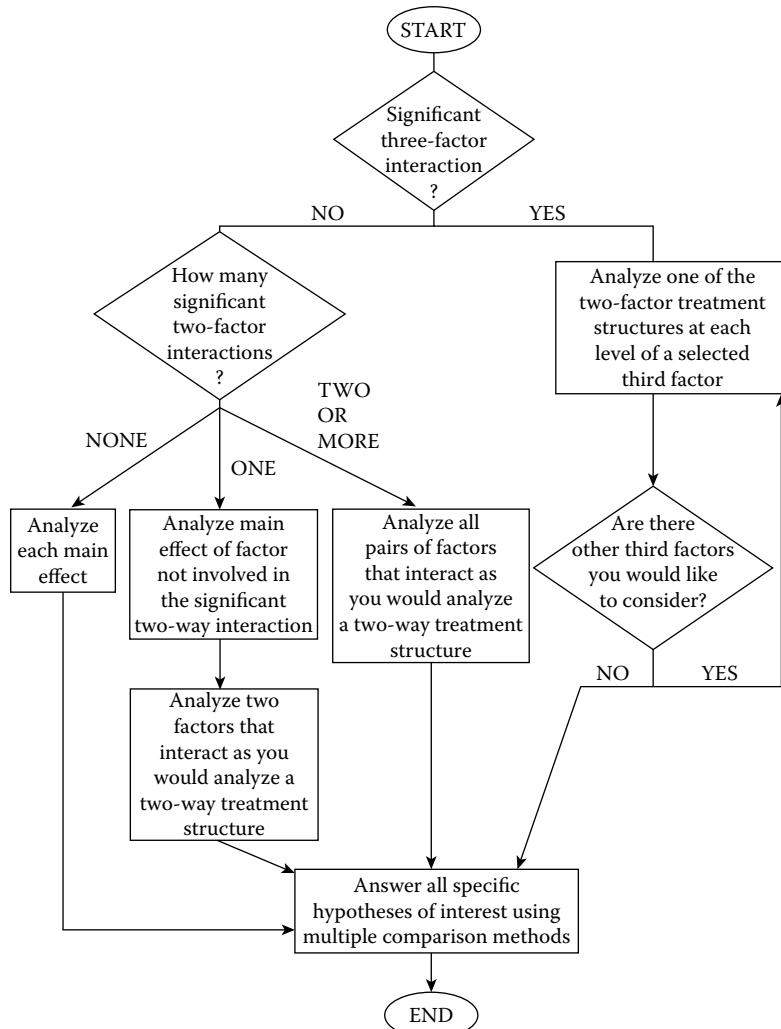


FIGURE 16.1 Strategy for analyzing three-factor experiments.

interested in it. If there is no three-factor interaction, then the second step is to determine whether there are any two-factor interactions. If there are also no two-factor interactions, then each of the main effects can be analyzed. If a three-factor interaction exists, the experimenter should analyze the two-way treatment structures of two of the treatment factors at each level of a selected third treatment factor, usually the factor of least interest. Obviously, these two-way analyses could be done by letting each treatment be the selected third one. The types of analyses that can be obtained with statistical computing packages are similar to those available for two-way treatment structures.

Figure 16.1 presents a general strategy for analyzing three-way treatment structures. This strategy can also be applied to four-way and higher-order treatment structures. If all treatment combinations are observed an equal number of times, the resulting data can be analyzed by using many different kinds of statistical software.

16.2 Balanced and Unbalanced Experiments

If all treatment combinations are observed, but observed an unequal number of times, then type III analyses can be used. If every treatment combination is observed at least once, all main-effect and interaction hypotheses can still be tested, and the questions answered are the same as those that would be answered with complete balance everywhere.

If some treatment combinations are missing, then, as was the case in Chapter 13, no hypothesis involving the missing treatment combinations can be tested. The experimenter should specify her own type IV hypotheses between treatments of interest. Such hypotheses can be tested by using the matrix procedure described in Chapter 1 or by contrast statements available in many statistical computing packages.

The example in Chapter 17 demonstrates some of the steps that may be required in order to obtain a complete analysis of data with missing treatment combinations.

16.3 Type I and II Analyses

In Chapter 10, type I and II analyses were described for a two-way treatment structure experiment. Both of these analyses produce sums of squares for each of the effects using the model comparison procedure. The type I analysis fits models sequentially and each effect is adjusted for all of the other effects that preceded it in the model. In the type II analysis, the sum of squares for each effect is adjusted for all other effects that are at the same or lower level. Consider the three-way model given by

$$y_{ijk\ell} = \mu + T_i + B_j + (TB)_{ij} + C_k + (TC)_{ik} + (BC)_{jk} + (TBC)_{ijk} + \varepsilon_{ijk\ell}$$

for $i = 1, 2, \dots, t$, $j = 1, 2, \dots, b$; $k = 1, 2, \dots, c$; and $\ell = 1, 2, \dots, n_{ijk}$

Suppose that $n_{ijk} > 0$ for all i, j , and k . That is, each of the three-way cells is observed at least once. Table 16.1 shows the type I and II sums of square along with their corresponding degrees of freedom using the reduction notation described in Chapter 10.

TABLE 16.1

Type I and II Sums of Squares for a Three-Way Experiment with at Least One Observation in Each Cell

Source of Variation	df	Type I SS	Type II SS
T	$t - 1$	$R(T \mu)$	$R(T \mu, B, C)$
B	$b - 1$	$R(B \mu, T)$	$R(B \mu, T, C)$
$T \times B$	$(t - 1)(b - 1)$	$R(T \times B \mu, T, B)$	$R(T \times B \mu, T, B, C, T \times C, B \times C)$
C	$c - 1$	$R(C \mu, T, B, T \times B)$	$R(C \mu, T, B)$
$T \times C$	$(t - 1)(c - 1)$	$R(T \times C \mu, T, B, T \times B, C)$	$R(T \times C \mu, T, B, T \times B, C, B \times C)$
$B \times C$	$(b - 1)(c - 1)$	$R(B \times C \mu, T, B, T \times B, C, T \times C)$	$R(B \times C \mu, T, B, T \times B, C, T \times C, B \times C)$
$T \times B \times C$	$(t - 1)(b - 1)(c - 1)$	$R(T \times B \times C \mu, T, B, T \times B, C, T \times C, B \times C)$	$R(T \times B \times C \mu, T, B, T \times B, C, T \times C, B \times C)$

16.4 Concluding Remarks

This chapter discussed the analysis of three-way and higher-way treatment structure experiments. A flow chart was given that provides a general strategy for analyzing such experiments.

It is important to examine the highest-order interaction effects first. Many experimenters avoid considering higher-order interactions because they are often not quite sure how to deal with these interactions. Such temptations should be avoided. Even though experience has shown us that very-high-order interactions are seldom significant, they must be dealt with whenever they are. The techniques discussed in Chapter 8 can be generalized to three-way and higher-way treatment structures and may help determine which treatment combinations are causing the interaction. The cause of the interaction may be the most important information identified in the study.

16.5 Exercises

- 16.1 Consider a study with four factors, denoted by A, B, C , and D , each at two levels for a total of 16 treatment combinations. The means model that can be used to describe the resulting data is

$$y_{ijklm} = \mu_{ijkl} + \varepsilon_{ijklm}, \quad i = 1, 2, \quad j = 1, 2, \quad k = 1, 2, \quad l = 1, 2, \quad m = 1, 2, \dots, n_{ijkl}, \quad \text{where } n_{ijkl} > 0.$$

Specify the contrasts of the μ_{ijkl} that measure the following:

- 1) The main effects of A, B, C , and D .
- 2) All possible two-way interactions.
- 3) All possible three-way interactions.
- 4) The four-way interaction.

- 16.2 Use the following data to compute the contrasts, estimated standard errors and 95% confidence intervals for the contrasts in Exercise 16.1.

Data for Exercise 16.2

$A = 0$		$B = 0$		$A = 0$		$B = 1$		$A = 1$		$B = 0$		$A = 1$		$B = 1$	
$D = 0$		$D = 1$		$D = 0$		$D = 1$		$D = 0$		$D = 1$		$D = 0$		$D = 1$	
$C = 0$	6, 8	9		8, 5		11		7		13, 10		10, 11		14	
$C = 1$	10		12, 10, 11		12, 10		12, 13		8, 9, 11		12, 16		14, 12		18, 16, 18

17

Case Study: Three-Way Treatment Structure with Many Missing Treatment Combinations

In this chapter we show a detailed analysis of a three-way treatment structure when many treatment combinations are missing.

17.1 Nutrition Scores Example

A home economist conducted a sample survey experiment to study how much lower-socioeconomic-level mothers knew about nutrition and to judge the effect of a training program designed to increase their knowledge of nutrition. A test was administered to the mothers both before and after the training program, and the changes in their test scores were measured. These changes are reported in Table 17.1. The mothers tested were classified according to three factors: age, race, and whether they were receiving food stamps.

17.2 An SAS-GLM Analysis

Table 17.2 gives the type IV analysis of variance table for the data in Table 17.1 obtained from SAS®-GLM using the SAS commands:

```
PROC GLM;
CLASSES GROUP AGE RACE;
MODEL GAIN=GROUP|AGE|RACE/SOLUTION E4;
LSMEANS GROUP|AGE|RACE/PDIFF STDERR;
RUN;
```

Table 17.2 reveals that the estimate of σ^2 for the data in Table 17.1 is $\hat{\sigma}^2 = 2627.4724/92 = 28.56$ with 92 degrees of freedom. Table 17.2 also shows that there are zero degrees of freedom

TABLE 17.1

Changes in Scores on a Nutrition Test between Post-Training and Pretraining

Group

Age Classification	Food Stamps			No Food Stamps		
	Black	Hispanic	White	Black	Hispanic	White
1	4, 4		-8, 9			5, -2, -10
2			5, 0, 10, 3, 3, 7, 7, 4	-4, -2, 0, 0, 5, -6, 2		7, 2, -13, 2, 3, 3, -4, -5
3	1, 5, 15, 9	0	4, 5, 0, 5, 2, 8, 1, -2, 6, 6, 4, -5, 6, 3, 7, 4, 5, 12, 3, 8, 3, 8, 13, 4, 7, 9, 3, 12, 11, 4, 12	3, -14, -14, -1, 3, 1	-1, 6	-20, 6, 9, -5, 3, -1, 3, 0, 4, -3, 2, 3, -5, 2, -1, -1, 6, -8, 0, 2
4	-3		-6, -5, 5, 8, 5, 6, 7, 6, 2, 7, 5	0		

for the three-factor interaction hypothesis, which indicates that there are no contrasts in these data that can be used for estimating a three-factor interaction. This does not imply that there is not a three-factor interaction among the factors group, age, and race, only that there are no testable hypotheses in the three-factor interaction effects.

The type IV *F*-values seem to suggest that there are no significant differences in any of the main effects and two-way interactions. This seems a bit strange, particularly since a visual examination of the data in Table 17.1 shows a large number of negative values for GAIN in the group that did not receive food stamps while the values for GAIN in the group that did receive food stamps are mainly positive. Thus, one would expect that there would be an effect due to food stamps, at least in these two subgroups. Also note that the *F*-value that

TABLE 17.2

Type IV Analysis of Variance Tables

*The GLM Procedure**Dependent Variable: Gain*

Source	df	Sum of Squares	Mean Square	F-Value	Pr > F
Model	14	1068.546279	76.324734	2.67	0.0026
Error	92	2627.472413	28.559483		
Corrected total	106	3696.018692			
<i>R-Square</i>		<i>Coefficient of Variation</i>		<i>Gain Mean</i>	
0.289107		233.3957		5.344107	2.289720
Type IV SS					
Group	1 ^a	75.7378078	75.7378078	2.65	0.1068
Age	3 ^a	41.5257840	13.8419280	0.48	0.6938
Group × age	3	91.5762463	30.5254154	1.07	0.3663
Race	2 ^a	11.6770165	5.8385082	0.20	0.8155
Group × race	2	113.7034419	56.8517209	1.99	0.1424
Age × race	3	87.3013862	29.1004621	1.02	0.3880
Group × age × race	0	0.0000000			

^a Other type IV testable hypotheses exist which may yield different SS.

compares the 15 cell means corresponding to the cells that have data is $F = 2.67$ with 14 and 92 degrees of freedom. Its corresponding p -value is 0.0026, which also seems to suggest that there are significant differences among these 15 treatment combination means. Why then, is the test for the group main effect ($F = 2.65$, $p = 0.1068$) not significant? The answer, of course, may lie in the hypothesis that is actually being tested by the type IV F -value for GROUP.

Table 17.3 shows the hypotheses being tested by the SAS-GLM Type IV analysis. An examination of Table 17.3 reveals that the type IV hypothesis for group compares the food stamp group to the no food stamp group averaging over the six cell means that they have in common, namely the cells corresponding to (age = 1, white), (age = 2, white), (age = 3, black), (age = 3, Hispanic), (age = 3, white), and (age = 4, black). This seems like a very reasonable hypothesis as it averages across the maximum number of similar categories that one can average over. To further explore why this test is not significant, let us look at the standard error of the corresponding type IV contrast that would compare these two groups. The estimated standard error is

$$\begin{aligned}\widehat{s.e.} \text{ (type IV group contrast)} &= \hat{\sigma} \sqrt{\sum_i \sum_j \sum_k \left(\frac{c_{ijk}^2}{n_{ijk}} \right)} \\ &= 5.344 \sqrt{\left(\frac{1}{2} + \frac{1}{8} + \frac{1}{4} + \frac{1}{1} + \frac{1}{31} + \frac{1}{1} \right) + \left(\frac{1}{3} + \frac{1}{8} + \frac{1}{6} + \frac{1}{2} + \frac{1}{20} + \frac{1}{1} \right)} \\ &= 12.047\end{aligned}$$

Notice that the size of the standard error depends more on the small sample sizes within the cells than it does on the large sample sizes. Since the type IV contrast for groups

TABLE 17.3

Type IV Hypotheses Tested by SAS-GLM

Source of Variation	Hypothesis
Group	$\mu_{N1W} + \mu_{N2W} + \mu_{N3B} + \mu_{N3H} + \mu_{N3W} + \mu_{N4B} = \mu_{Y1W} + \mu_{Y2W} + \mu_{Y3B} + \mu_{Y3H} + \mu_{Y3W} + \mu_{Y4B}$
Age	$\mu_{Y1B} + \mu_{Y1W} = \mu_{Y4B} + \mu_{Y4W}, \mu_{N2B} + \mu_{Y2W} = \mu_{N4B} + \mu_{N4W}$, and $\mu_{N3B} + \mu_{Y3B} + \mu_{Y3W} = \mu_{N4B} + \mu_{Y4B} + \mu_{Y4W}$
Group \times age	$(\mu_{N1W} - \mu_{N3W} - \mu_{Y1W} + \mu_{Y3W}) + (\mu_{N1B} - \mu_{N4B} - \mu_{Y1B} + \mu_{N3B}) = 0$, $(\mu_{N2W} - \mu_{N3W} - \mu_{Y2W} + \mu_{Y3W}) + (\mu_{N3B} - \mu_{N4B} - \mu_{Y3B} + \mu_{Y4B}) = 0$, and $\mu_{N3B} - \mu_{N4B} - \mu_{Y3B} + \mu_{Y4B} = 0$
Race	$\mu_{N2B} + \mu_{N3B} + \mu_{Y1B} + \mu_{Y3B} + \mu_{Y4B} = \mu_{N2W} + \mu_{N3W} + \mu_{Y1W} + \mu_{Y3W} + \mu_{Y4W}$, and $\mu_{N3H} + \mu_{Y3H} = \mu_{N3W} + \mu_{Y3W}$
Group \times race	$\mu_{N3B} - \mu_{N3W} - \mu_{Y3B} + \mu_{Y3W} = 0$, and $\mu_{N3H} - \mu_{N3W} - \mu_{Y3H} + \mu_{Y3W} = 0$
Age \times race	$\mu_{Y1B} - \mu_{Y1W} - \mu_{Y4B} + \mu_{Y3W} = 0$, $\mu_{N2B} - \mu_{N2W} - \mu_{N3B} + \mu_{N3W} = 0$, and $\mu_{Y3B} - \mu_{Y3W} - \mu_{Y4B} + \mu_{Y4W} = 0$
Group \times age \times race	None

involves three cells which have only one observation in them, the observed standard error must be greater than

$$5.344 \sqrt{\frac{1}{1} + \frac{1}{1} + \frac{1}{1}} = 9.256$$

regardless of the numbers of observations in the other cells. This illustrates another aspect of data analysis that data analysts must be aware of. Certain tests may not have much power associated with them if the corresponding hypotheses involve cells that have small sample sizes. This could have been true even if all cells were observed. That is, one may have low power even when one has a balanced treatment structure if some of the cell sample sizes are small.

It seems likely that few, if any, of the hypotheses automatically tested by the type IV analysis of SAS-GLM will be of particular interest to the experimenter for these data. We do not even consider the type I-III hypotheses, since they usually make little sense in cases where there are missing cells.

In this kind of a messy experiment, the safest and easiest way to obtain useful information is to look at the three-way least squares means, and pairwise comparisons between them. Table 17.4 gives the three-way least squares means and their estimated standard errors, and Table 17.5 gives *p*-values corresponding to pairwise comparisons among the least squares means.

Suppose we wish to test the type IV group hypothesis that $\mu_{Y3W} = \mu_{N3W}$. That is, how do the two cell means that have the largest sample sizes compare to one another? These two cells correspond to LSMeans 6 and 13 in Table 17.5, and the corresponding *p*-value that compares these two means is *p* = 0.0004. Thus, there is a highly significant difference due to food stamps for the age = 3 and race = white subgroup of mothers. Next, consider age = 2,

TABLE 17.4
Three-Way Least Squares Means

Group	Age	Race	Gain LSMean	Standard Error	Pr > <i>t</i>	LSMean Number
N	1	W	-2.33333333	3.08542178	0.4514	1
N	2	B	-0.71428571	2.01988270	0.7244	2
N	2	W	-0.62500000	1.88942725	0.7416	3
N	3	B	-3.66666667	2.18172267	0.0962	4
N	3	H	2.50000000	3.77885451	0.5099	5
N	3	W	-0.20000000	1.19497872	0.8674	6
N	4	B	0.00000000	5.34410729	1.0000	7
Y	1	B	4.00000000	3.77885451	0.2926	8
Y	1	W	0.50000000	3.77885451	0.8950	9
Y	2	W	4.87500000	1.88942725	0.0115	10
Y	3	B	7.50000000	2.67205365	0.0061	11
Y	3	H	0.00000000	5.34410729	1.0000	12
Y	3	W	5.41935484	0.95983000	<0.0001	13
Y	4	B	-3.00000000	5.34410729	0.5759	14
Y	4	W	3.63636364	1.61130898	0.0264	15

TABLE 17.5
Pairwise Comparisons between All Least Squares Means

		Least Squares Means for Effect Group \times Age \times Race														
		$Pr > t \text{ for } H_0: L\bar{S}Mean(i) = L\bar{S}Mean(j)$														
		Dependent Variable: Gain														
<i>ij</i>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
1	0.6617	0.6379	0.7250	0.3244	0.5207	0.7062	0.1975	0.5628	0.0493	0.0180	0.7062	0.0184	0.9142	0.0897		
2	0.6617	0.9743	0.3233	0.4551	0.8270	0.9008	0.2741	0.7775	0.0462	0.0161	0.9008	0.0073	0.6900	0.0956		
3	0.6379	0.9743	0.2947	0.4614	0.8496	0.9124	0.2765	0.7906	0.0424	0.0149	0.9124	0.0054	0.6762	0.0895		
4	0.7250	0.3233	0.2947	0.1610	0.1668	0.5269	0.0822	0.3421	0.0039	0.0017	0.5269	0.0002	0.9083	0.0084		
5	0.3244	0.4551	0.4614	0.1610	0.4974	0.7034	0.7796	0.7091	0.5754	0.2828	0.7034	0.4559	0.4029	0.7827		
6	0.5207	0.8270	0.8496	0.1668	0.4974	0.9709	0.2920	0.8602	0.0255	0.0100	0.9709	0.0004	0.6104	0.0589		
7	0.7062	0.9008	0.9124	0.5269	0.7034	0.9709	0.5426	0.9393	0.3920	0.2126	1.0000	0.3208	0.6923	0.5164		
8	0.1975	0.2741	0.2765	0.0822	0.7796	0.2920	0.5426	0.5141	0.8364	0.4514	0.5426	0.7167	0.2876	0.9297		
9	0.5628	0.7775	0.7906	0.3421	0.7091	0.8602	0.9393	0.5141	0.3031	0.1338	0.9393	0.2102	0.5941	0.4471		
10	0.0493	0.0462	0.0424	0.0039	0.5754	0.0255	0.3920	0.8364	0.3031	0.4246	0.3920	0.7979	0.1681	0.6191		
11	0.0180	0.0161	0.0149	0.0017	0.2828	0.0100	0.2126	0.4514	0.1338	0.4246	0.2126	0.4655	0.0822	0.2188		
12	0.7062	0.9008	0.9124	0.5269	0.7034	0.9709	1.0000	0.5426	0.9393	0.3920	0.2126	0.3208	0.6923	0.5164		
13	0.0184	0.0073	0.0054	0.0002	0.4559	0.0004	0.3208	0.7167	0.2102	0.7979	0.4655	0.3208	0.1244	0.3443		
14	0.9142	0.6900	0.6762	0.9083	0.4029	0.6104	0.6923	0.2876	0.5941	0.1681	0.0822	0.6923	0.1244	0.2375		
15	0.0897	0.0956	0.0895	0.0084	0.7827	0.0589	0.5164	0.9297	0.4471	0.6191	0.2188	0.5164	0.3443	0.2375		

whites. The corresponding cell means are LSMeans 3 and 10, and the corresponding p -value is $p = 0.0424$, which also indicates a significant difference due to food stamps. Finally, consider the groups corresponding to age = 3, blacks. These two means correspond to LSMeans 4 and 11, and the corresponding p -value is $p = 0.0017$. Thus, it seems clear that if one only considers cells that have reasonable sample sizes, then we find that there are significant differences due to food stamps.

17.3 A Complete Analysis

Since we cannot test for three-factor interaction and hence do not know whether there is a three-factor interaction, we next examine two-way analyses at each level of a third treatment factor. Let us suppose that the experimenter is most interested in the effects of, or differences between the race \times group combinations. We thus examine these two-way combinations at each level of the age factor, as shown in Table 17.6.

For age = 1 in Table 17.6, we observe the following: 1) It is not possible to test for race \times group interaction in this age group, since no contrast exists that measures a two-factor interaction; 2) the only type IV hypothesis comparing groups that can be tested is $\mu_{Y1W} = \mu_{N1W}$; 3) the only type IV hypothesis in the age = 1 group that concerns races that can be tested is $\mu_{Y1B} = \mu_{Y1W}$; and 4) the hypothesis $\mu_{Y1B} = \mu_{N1W}$ can also be tested in the age = 1 group, but this hypothesis is probably of secondary interest since it involves different levels in both race and group. From Table 17.5, the p -value for testing $\mu_{Y1W} = \mu_{N1W}$ is 0.5628. The p -values corresponding to the hypotheses in 3 and 4 are 0.5141 and 0.1975, respectively.

Tables 17.7 and 17.8 give the testable hypotheses, the p -values of their respective test statistics, and a (subjective) importance rating for the age groups, age = 2 and age = 4, respectively.

Finally, we examine the age = 3 group. In this group all race \times group combinations are observed; hence, it is possible to test for race \times group interaction. If this interaction is not

TABLE 17.6

Observed Race \times Age Combinations for Each Age Level for the Data in Table 17.1

	Group	Race		
		Black	Hispanic	White
Age = 1	No			X
	Yes	X		X
Age = 2	No	X		X
	Yes			X
Age = 3	No	X	X	X
	Yes	X	X	X
Age = 4	No	X		
	Yes	X		X

Note: "X" indicates that the cell was observed at least once in the experiment.

TABLE 17.7

Testable Hypotheses for Age = 2 for Data in Table 17.1

Hypothesis	p-Value	Importance
$\mu_{N2W} = \mu_{N2B}$	0.974	Primary
$\mu_{Y2W} = \mu_{N2W}$	0.042	Primary
$\mu_{Y2W} = \mu_{N2B}$	0.046	Secondary

Note: The p-values are from the Table 17.5.

TABLE 17.8

Testable Hypotheses for Age = 4 for Data in Table 17.1

Hypothesis	p-Value	Importance
$\mu_{Y4B} = \mu_{N4B}$	0.692	Primary
$\mu_{Y4BW} = \mu_{Y4W}$	0.238	Primary
$\mu_{Y4W} = \mu_{N4B}$	0.516	Secondary

Note: The p-values are from the Table 17.5.

significant, we can examine the main-effect means for both race and group in the age = 3 subgroup of mothers. These three hypotheses are specified by

$$H_{01}: \mu_{N3B} - \mu_{N3W} - \mu_{Y3B} + \mu_{Y3W} = 0, \text{ and}$$

$$\mu_{N3H} - \frac{1}{2}(\mu_{N3B} + \mu_{N3W}) - \mu_{Y3H} + \frac{1}{2}(\mu_{Y3B} + \mu_{Y3W}) = 0$$

$$H_{02}: \bar{\mu}_{Y3.} = \bar{\mu}_{N3.}$$

and

$$H_{03}: \bar{\mu}_{.3B} = \bar{\mu}_{.3H} = \bar{\mu}_{.3W}, \text{ respectively}$$

The test statistic for H_{01} was actually given by the original SAS-GLM analysis shown in Table 17.2, while the tests of the other two hypotheses were not. Hypothesis H_{01} is equivalent to the one tested by the race \times group type IV F-value, which can be seen by examining the estimable functions for race \times group in Table 17.3. The tests for these three hypotheses can be obtained by using the Contrast statements in SAS-GLM, as illustrated below, or by hand. In this case, it is probably easier to do the testing by hand. When using SAS-GLM, it is also easier to use a Means model. With this model, only the coefficients corresponding to the group \times age \times race effect need to be entered. The SAS-GLM Contrast statements needed for an effects model are shown in Table 17.9.

The test statistics for the above three hypotheses can be obtained from SAS-GLM by including the following Contrast statements with the SAS commands given in Section 17.2. The required statements are given in Table 17.9. The results from these options are given in Table 17.10.

If the experimenter is also interested in the effects of the race \times age combinations, these combinations can be analyzed at each value of the group factor. A similar situation exists if the experimenter wanted to examine the effects of the group \times age combinations for

TABLE 17.9Contrast Statements for Testing H_{01} – H_{03}

For H_{01} :	CONTRAST 'H01' GROUP 0 0 AGE 0 0 0 0 GROUP*AGE 0 0 0 0 0 0 0 0 RACE 0 0 0 GROUP*RACE 1 0 -1 -1 0 1 AGE*RACE 0 0 0 0 0 0 0 0 0 0 0 0 0 GROUP*AGE*RACE 0 0 0 1 0 -1 0 0 0 0 -1 0 1 0 0, GROUP 0 0 AGE 0 0 0 0 GROUP*AGE 0 0 0 0 0 0 0 0 RACE 0 0 0 GROUP*RACE -.5 1 -.5 .5 -1 .5 AGE*RACE 0 0 0 0 0 0 0 0 0 0 0 0 0 GROUP*AGE*RACE 0 0 0 -.5 1 -.5 0 0 0 0 .5 -1 .5 0 0;
For H_{02} :	CONTRAST 'H02' GROUP 3 -3 AGE 0 0 0 0 GROUP*AGE 0 0 3 0 0 0 -3 0 GROUP*RACE 1 1 1 -1 -1 AGE*RACE 0 0 0 0 0 0 0 0 GROUP*AGE*RACE 0 0 0 1 1 1 0 0 0 0 -1 -1 -1 0 0;
For H_{03} :	CONTRAST 'H03' GROUP 0 0 AGE 0 0 0 0 GROUP*AGE 0 0 0 0 0 0 0 0 RACE 1 -1 0 GROUP*RACE .5 -.5 0 .5 -.5 0 AGE*RACE 0 0 0 0 1 -1 0 0 0 GROUP*AGE*RACE 0 0 0 .5 -.5 0 0 0 0 0 .5 -.5 0 0 0, GROUP 0 0 AGE 0 0 0 0 GROUP*AGE 0 0 0 0 0 0 0 0 RACE 1 0 -1 GROUP*RACE .5 0 -.5 .5 0 -.5 AGE*RACE 0 0 0 0 1 0 -1 0 0 GROUP*AGE*RACE 0 0 0 .5 0 -.5 0 0 0 0 .5 0 -.5 0 0;

TABLE 17.10

Results from the SAS Contrast Statements

*The GLM Procedure**Dependent Variable: Gain*

Contrast	df	Contrast SS	Mean Square	F-Value	Pr > F
H_{01}	2	113.7034419	56.8517209	1.99	0.1424
H_{02}	1	102.1000978	102.1000978	3.57	0.0618
H_{03}	2	7.8054511	3.9027256	0.14	0.8724

each race. Both of these analyses can be done either by hand, by using Contrast statements, or by using three-way least squares means when possible, as illustrated at the beginning of this section for the different levels of the age factor.

The analysis of higher-order cross-classified treatment structures can be carried out in ways similar to those illustrated in this chapter.

17.4 Concluding Remarks

This chapter presented the analysis of a three-way treatment structure having a large number of missing treatment combinations. An SAS-GLM analysis was obtained and interpreted. Questions not answered by the SAS-GLM analysis were also raised, and techniques for answering these questions were illustrated.

17.5 Exercises

- 17.1 The following data was collected from a three-way treatment structure in a completely randomized design structure. Use a means model to answer the following questions.

Level 1 of C	B1	B2	B3
A1	21, 24	23, 23, 27	35
A2		18, 16	28, 23, 25
A3	37, 37		34
Level 2 of C	B1	B2	B3
A1	32		18, 21
A2	27, 29, 26	35	
A3	37, 32	34, 36, 42	16, 13, 15
Level 3 of C	B1	B2	B3
A1	26, 22	23, 25, 20	22
A2	30, 33		40
A3		27, 30, 33	36, 38

- 1) Use the means model to provide a complete analysis of the data.
 - 2) Determine all of the type IV hypotheses for A .
 - 3) Determine all of the type IV hypotheses for $A \times B$.
 - 4) Determine a set of linearly independent contrasts that measure three-way interaction.
- 17.2 The data for Exercise 17.1 was collected from a three-way treatment structure in a completely randomized design structure.
- 1) Use the effects model to provide a complete analysis of the data.
 - 2) Compare the analyses from the effects model with that of the means model.
- 17.3 The following data are from a four-way treatment structure in a completely randomized design structure.

Data for Exercise 17.3

	$A = 0 B = 0$		$A = 0 B = 1$		$A = 1 B = 0$		$A = 1 B = 1$	
	$D = 0$	$D = 1$						
C = 0	7, 8		9	12, 11	12, 10	11, 16	13	18, 15
C = 1		10		11				16, 18
C = 2	9	10, 13		14	12, 14	14, 17	13, 15, 16	
C = 3	10, 11		13		15		18, 17	

- 1) Determine the numbers of degrees of freedom for each two-way, three-way and four-way interactions.
- 2) Use a means model to carry out an analysis of this data by determining the important factors as well as making the necessary comparisons. Be sure to use multiple comparisons when needed.
- 3) Use a effects model to carry out an analysis of this data by determining the important factors as well as making the necessary comparisons. Be sure to use multiple comparisons when needed.

18

Random Effects Models and Variance Components

Models with more than one random component are applied to several situations, including random effects and mixed effects models where some or all of the factors in the treatment structure are random or where there are several sizes of experimental units as in split-plot and repeated measures designs. The parameters of interest for such models include the variances associated with the distributions of the random components (usually called variance components). It is important to be able to identify the random components of a model and be able to utilize them in the analysis of the model. When carrying out an analysis of variance for a given model, the expected values of the mean squares (which are functions of the variance components) are needed in order to construct proper test statistics and determine standard errors for comparisons of fixed effect parameters. It is also important to be able to obtain estimates of the variance components and test hypotheses and construct confidence intervals about functions of the variance components. The discussion of random effects models and methods of analyzing them is divided into four chapters. This chapter defines the random effects model and describes a general procedure for computing expectations of sums of squares. The procedure can easily be used by computer software to evaluate the expectations of sums of squares. The problem of estimation is discussed in Chapter 19, methods for testing hypotheses and constructing confidence intervals are presented in Chapter 20, and a detailed analysis of an example is presented in Chapter 21.

18.1 Introduction

The philosophy behind the use of random effects models is quite different from that behind the use of the fixed-effects models (discussed in the previous chapters) in both the sampling scheme and the parameters of interest. Before these differences are discussed, the definitions of a random effect and a fixed effect are given.

Definition 18.1: A factor is a *random effect* factor if its levels consist of a random sample of levels from a population of possible levels.

Definition 18.2: A factor is a *fixed effect* factor if its levels are selected by a nonrandom process or if its levels consist of the entire population of possible levels.

Thus, in order to determine whether a factor is a fixed effect or a random effect, one needs to know how the experimenter selected the levels of that factor. If all possible levels of the factor or a set of selected levels of the factor are included in the experiment, the factor is considered as a fixed effect. If some form of randomization is used to select the levels included in the experiment, then the factor is a random effect.

Rule: The levels of a factor are fixed until proven random.

To establish that the levels of a factor are random, the population or conceptual population of possible levels must be described and the method of randomly selecting the levels of the factor must be specified. Inferences are to be made to the population of levels, so if that population is not describable, the inferences may not be meaningful.

For example, suppose a plant breeder wants to study a characteristic (say, yield) of wheat varieties. There are many possible wheat varieties (a population of varieties), but if he wants to study a certain set of varieties, then he would select just those varieties for his experiment. In this case, the factor “variety” is called a fixed effect, since the levels of varieties are chosen or fixed. However, if the plant breeder is interested in how a characteristic is distributed among the varieties in the population, then he is not interested in which set of varieties is included in the experiment. In this case, the plant breeder can randomly select the varieties to be included in the experiment from the population of varieties. Therefore the factor variety in this second experiment is a random effect. When constructing a model to describe a given experimental situation, it must be stated whether a factor is a random or fixed effect. The models considered in the previous chapters were constructed under the assumption that all factors in the treatment structure were fixed effects and there was one size of experimental unit. However, the idea of a random effect was alluded to when blocking was introduced in Chapter 4, where it was assumed that the factor “blocks” was a random effect. Further, a basic assumption is that those factors associated with the design structure of a model are random effects. Some factors can have both a set of levels that are fixed effects and a set that are random effects (Njuho and Milliken, 2005), but that topic is not elaborated on here. Three basic types of models can be constructed, depending on the number of sizes of experimental units in the design structure and assumptions about the factors in the treatment structure. These types of models are defined below.

Definition 18.3: A model is called a fixed or fixed effects model if all of the factors in the treatment structure are fixed effects and there is only one size of experimental unit in the study, no blocking, and the variances are all equal.

Definition 18.4: A model is called a random or random effects model if all of the factors in the treatment structure are random effects (all of the factors in the design structure are already assumed to be random effects).

Definition 18.5: A model is called a mixed or mixed effects model if some of the factors in the treatment structure are fixed effects and some are random effects, or if all of the factors in the treatment structure are fixed effects and there is more than one size of experimental unit in the design structure (or there are at least two variance components in the model).

The models discussed in this chapter are all random effects models. The discussion of mixed models is presented in Chapters 22 and 23. The following example is presented to help motivate the application and analysis of random effects models.

18.1.1 Example 18.1: Random Effects Nested Treatment Structure

A consumer group studied the variation in coffee prices in U.S. cities with populations of at least 20,000. Three factors that the group wished to investigate were states, cities within states, and stores within cities within states. The treatment structure or sampling design is a three-way, two-level, nested system involving states, cities, and stores, where cities are nested within states and stores are nested within cities. The sampling procedure was to select r states at random ($r < 50$) from the population of all possible states. Next, randomly select t_i cities from the C_i cities in the i th state ($t_i < C_i$) with populations of at least 20,000. Finally, randomly select n_{ij} stores ($n_{ij} < S_{ij}$) from the S_{ij} stores in the j th city in the i th state and determine the price of a particular grade of coffee at each randomly selected store. A model that can be used to describe the variability in the coffee prices is

$$y_{ijk} = \mu + s_i + c_{j(i)} + a_{k(j)} \quad i = 1, 2, \dots, r, \quad j = 1, 2, \dots, t_i, \quad k = 1, 2, \dots, n_{ij}$$

where μ denotes the average price of coffee in the United States, s_i denotes the effect of the i th randomly selected state, $c_{j(i)}$ denotes the effect of the j th randomly selected city from the i th state, and $a_{k(j)}$ denotes the effect of the k th randomly selected store from the j th city in the i th state. The assumptions are that

- 1) The s_i are distributed *i.i.d.* $N(0, \sigma_{\text{State}}^2)$.
- 2) The $c_{j(i)}$ are distributed *i.i.d.* $N(0, \sigma_{\text{City}}^2)$.
- 3) The $a_{k(j)}$ are distributed *i.i.d.* $N(0, \sigma_{\text{Store}}^2)$.

It is also assumed that all of the random effects are distributed independently of one another. The parameters in this random effects model are μ , σ_{State}^2 , σ_{City}^2 , and σ_{Store}^2 . The terms s_i , $c_{j(i)}$, and $a_{k(j)}$ are random variables and are not parameters in the model. Most applications are only interested in the estimation of the parameters and are not interested in predicting the values of the random variables. But, when predictions of the random variables are of interest, estimated best linear unbiased predictors (EBLUP) of the random effects can be obtained (Littell et al., 2006; Milliken and Johnson, 2002).

The variance of a coffee price can be expressed through the parameters of the model as

$$\text{Var}(y_{ijk}) = \sigma_{\text{Price}}^2 = \sigma_{\text{State}}^2 + \sigma_{\text{City}}^2 + \sigma_{\text{Store}}^2$$

The covariance between two coffee prices from stores within the same city (and state) is

$$\text{Cov}(y_{ij1}, y_{ij2}) = \sigma_{y_{ij1}y_{ij2}} = \sigma_{\text{State}}^2 + \sigma_{\text{City}}^2$$

The correlation between these two prices is

$$\rho_{y_{ij1}y_{ij2}} = \frac{\sigma_{\text{State}}^2 + \sigma_{\text{City}}^2}{\sigma_{\text{State}}^2 + \sigma_{\text{City}}^2 + \sigma_{\text{Store}}^2}$$

The covariance between two coffee prices from stores from different cities within the state is $\text{Cov}(y_{i11}, y_{i22}) = \sigma_{y_{i11}y_{i22}} = \sigma_{\text{State}}^2$. The correlation between two prices from two different cities within the same state is

$$\rho_{y_{i11}y_{i22}} = \frac{\sigma_{\text{State}}^2}{\sigma_{\text{State}}^2 + \sigma_{\text{City}}^2 + \sigma_{\text{Store}}^2}$$

Thus, when a model includes random effects, a correlation structure similar to that for the coffee prices is imposed on the resulting data. It is of interest to be able to construct models and then identify the sources of variation as well as the correlations among the various groupings of observations. The following examples are used to demonstrate the construction of models with one or more random effects and to evaluate the resulting implied covariance or correlation structure among the observations in the data set.

18.2 General Random Effects Model in Matrix Notation

In order to describe methods used to evaluate the expectations of sums of squares, it is necessary to have some general notation to describe the general random effects model. This section presents a matrix representation of a general random effects model, which is used in later sections to demonstrate general methods for computing expected mean squares. To help visualize the general random effects model and its expression in terms of matrices, a random effects model for a one-way treatment structure in a completely randomized design structure is examined.

18.2.1 Example 18.2: One-Way Random Effects Model

A model describing a one-way random effects treatment structure in a completely randomized design structure is

$$y_{ij} = \mu + u_i + \varepsilon_{ij}, \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, n_i \quad (18.1)$$

where μ is the population mean of the response, u_i denotes the effect of i th randomly selected treatment and is assumed to be distributed *i.i.d.* $N(0, \sigma_u^2)$, and ε_{ij} denotes the random error of the j th observation of the i th treatment where it is assumed the ε_{ij} are distributed *i.i.d.* $N(0, \sigma_\varepsilon^2)$. It is also assumed that u_i and ε_{ij} are independent random variables. These assumptions allow the variances and covariances of the observations to be evaluated. The variance of an observation is

$$\begin{aligned} \text{Var}(y_{ij}) &= \sigma_y^2 = \text{Var}(\mu + u_i + \varepsilon_{ij}) = \text{Var}(u_i) + \text{Var}(\varepsilon_{ij}) \\ &= \sigma_u^2 + \sigma_\varepsilon^2 \end{aligned}$$

There are two components in the variance of y_{ij} which include the variance of the population of treatments or levels of u and the variance of the experimental units, hence the name variance components or components of variance. The covariance of two observations obtained from the same possibility for u_i is

$$\begin{aligned} \text{Cov}(y_{ij}, y_{ij'}) &= \text{Cov}(\mu + u_i + \varepsilon_{ij}, \mu + u_i + \varepsilon_{ij'}) \\ &= \text{Cov}(u_i, u_i) = \text{Var}(u_i) = \sigma_u^2 \end{aligned}$$

The covariance between two observations obtained from different i values is zero. Hence, observations obtained from the same i are correlated and this correlation is called the intraclass correlation and is defined as

$$\rho = \frac{\text{Cov}(y_{ij}, y_{ij'})}{\sqrt{\text{Var}(y_{ij})\text{Var}(y_{ij'})}} = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

Observations from different i values are uncorrelated. The model in Equation 18.1 can be described in matrix notation as

$$\mathbf{y} = \mathbf{j}\mu + \mathbf{Z}_1\mathbf{u} + \boldsymbol{\varepsilon} \quad (18.2)$$

where \mathbf{j} is an $N \times 1$ vector of ones ($N = \sum_{i=1}^t n_i$), \mathbf{Z}_1 is an $N \times t$ design matrix, \mathbf{u} is the $t \times 1$ vector random variable assumed to be distributed as the multivariate normal distribution $N_t(\mathbf{0}, \sigma_u^2 \mathbf{I}_t)$, and $\boldsymbol{\varepsilon}$ is the $N \times 1$ vector random variable assumed to be distributed $N_N(\mathbf{0}, \sigma_e^2 \mathbf{I}_N)$. The covariance matrix of the vector of observations, \mathbf{y} , is

$$\begin{aligned} \boldsymbol{\Sigma} &= \text{Var}(\mathbf{y}) = \text{Var}(\mathbf{j}\mu + \mathbf{Z}_1\mathbf{u} + \boldsymbol{\varepsilon}) \\ &= \mathbf{Z}_1 \text{Var}(\mathbf{u}) \mathbf{Z}_1' + \text{Var}(\boldsymbol{\varepsilon}) \\ &= \sigma_u^2 \mathbf{Z}_1 \mathbf{Z}_1' + \sigma_e^2 \mathbf{I}_N \end{aligned}$$

The variances of the y_{ij}' s are the diagonal elements of $\boldsymbol{\Sigma}$ and the covariances between pairs of y_{ij} are off-diagonal elements of $\boldsymbol{\Sigma}$.

Equivalently, the model can be written as

$$\begin{bmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1n_1} \\ y_{21} \\ y_{22} \\ \vdots \\ y_{2n_2} \\ \vdots \\ y_{t1} \\ y_{t2} \\ \vdots \\ y_{tn_t} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \mu + \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \vdots \\ \varepsilon_{1n_1} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \vdots \\ \varepsilon_{2n_2} \\ \vdots \\ \varepsilon_{t1} \\ \varepsilon_{t2} \\ \vdots \\ \varepsilon_{tn_t} \end{bmatrix}$$

The general random effects model will have r random components representing the main effects and interactions for the random effect factors of the treatment structure and for those factors used to describe the design structure as well as possible interactions between components of the design and treatment structures used to describe the necessary error terms. The general random effects model includes an overall mean parameter denoted by μ

as well as $\boldsymbol{\varepsilon}$, the vector representing the residual or smallest size of experimental unit errors. The general random effects model can be expressed in matrix notation as

$$\mathbf{y} = j\mu + \mathbf{Z}_1\mathbf{u}_1 + \mathbf{Z}_2\mathbf{u}_2 + \cdots + \mathbf{Z}_r\mathbf{u}_r + \boldsymbol{\varepsilon} \quad (18.3)$$

where the \mathbf{u}_s , $s = 1, 2, \dots, r$ denote random effects and $\boldsymbol{\varepsilon}$ denotes the residual error and where all of these random variables are assumed to be distributed independently with assumed marginal distributions

$$\mathbf{u}_1 \sim N(0, \sigma_1^2 \mathbf{I}_{t_1}), \quad \mathbf{u}_2 \sim N(0, \sigma_2^2 \mathbf{I}_{t_2}), \dots, \quad \mathbf{u}_r \sim N(0, \sigma_r^2 \mathbf{I}_{t_r}), \quad \text{and} \quad \boldsymbol{\varepsilon} \sim N(0, \sigma_{\varepsilon}^2 \mathbf{I}_N)$$

N is the total number of observations in the data vector, \mathbf{y} , and the \mathbf{Z}_i are the $N \times t_i$ design matrices corresponding to the i th random effect vector. This general random effects model can be expressed as

$$\mathbf{y} = j_N\mu + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon} \quad \text{where } \mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_r],$$

$$\mathbf{u} = \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_r \end{bmatrix} \text{ and } \text{Var}(\mathbf{u}) = \begin{bmatrix} \sigma_1^2 \mathbf{I}_{t_1} & 0 & 0 & 0 \\ 0 & \sigma_2^2 \mathbf{I}_{t_2} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \sigma_r^2 \mathbf{I}_{t_r} \end{bmatrix}$$

Consequently,

$$\text{Var}(\mathbf{y}) = \mathbf{Z}' \text{Var}(\mathbf{u}) \mathbf{Z} + \sigma_{\varepsilon}^2 \mathbf{I}_N$$

The covariance matrix of the data vector \mathbf{y} is

$$\begin{aligned} \boldsymbol{\Sigma} = \text{Var}(\mathbf{y}) &= \text{Var}(j\mu + \mathbf{Z}_1\mathbf{u}_1 + \mathbf{Z}_2\mathbf{u}_2 + \cdots + \mathbf{Z}_r\mathbf{u}_r + \boldsymbol{\varepsilon}) \\ &= \mathbf{Z}_1 \text{Var}(\mathbf{u}_1) \mathbf{Z}_1' + \mathbf{Z}_2 \text{Var}(\mathbf{u}_2) \mathbf{Z}_2' + \cdots + \mathbf{Z}_r \text{Var}(\mathbf{u}_r) \mathbf{Z}_r' + \text{Var}(\boldsymbol{\varepsilon}) \\ &= \sigma_1^2 \mathbf{Z}_1 \mathbf{Z}_1' + \sigma_2^2 \mathbf{Z}_2 \mathbf{Z}_2' + \cdots + \sigma_r^2 \mathbf{Z}_r \mathbf{Z}_r' + \sigma_{\varepsilon}^2 \mathbf{I}_N \end{aligned}$$

This general form of the random effects model can be used in many situations to help identify the sources of variation in the data collection system. One method of looking at these sources of variation is to compute a set of sums of squares corresponding to each random effect and then determine the functions of the parameters each sum of square estimates. The functions of the variance components being estimated by each effect sum of squares can be evaluated by obtaining the expected values of the effects sums of squares. In the next section, this matrix form of the general random effects model is used to describe the method for evaluating the expectations of sums of squares involving the observations.

18.3 Computing Expected Mean Squares

The expected values of the sums of squares from an analysis of variance of a random effects model involve the variance components. For a given model, at least two methods can be used to evaluate the expected mean squares (remember that a mean square is a sum

of squares divided by its degrees of freedom). The first method is to algebraically evaluate the expected values by using the model assumptions and the second method is to evaluate the expected values by means of a computer algorithm. The algebraic method is presented by applying it to the sum of squares obtained from the analysis of a model with a one-way random effects treatment structure. The computer algorithm method is discussed in general terms and demonstrated by more complex examples.

18.3.1 Algebraic Method

There are two variance components in the one-way random effects model in Equation 18.1, thus two sums of squares are used in the analysis to describe the variability in the response. The two sums of squares that are usually computed are the sum of squares within levels of the random effect, designated Q_1 , and the sum of squares between the levels of the random effect, designated Q_2 . For the one-way random effects model, these sums of squares are given by

$$Q_1 = \sum_{i=1}^t \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i\cdot})^2 = \sum_{i=1}^t \sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^t n_i \bar{y}_{i\cdot}^2$$

and

$$Q_2 = \sum_{i=1}^t n_i (\bar{y}_{i\cdot} - \bar{y}_{..})^2 = \sum_{i=1}^t n_i \bar{y}_{i\cdot}^2 - n_{..} \bar{y}_{..}^2$$

In terms of the random variables of model (18.1), the quantities in Q_1 and Q_2 are expressed as

$$\left. \begin{aligned} y_{ij} &= \mu + u_i + \varepsilon_{ij} \\ \bar{y}_{i\cdot} &= \mu + u_i + \bar{\varepsilon}_{i\cdot} \quad \text{and} \\ \bar{y}_{..} &= \mu + \tilde{u}_\cdot + \bar{\varepsilon}_{..} \quad \text{where } \tilde{u}_\cdot = \frac{1}{N} \sum_{i=1}^t n_i u_i \end{aligned} \right\} \quad (18.4)$$

Substituting the terms from Equation 18.4 into Q_1 , the expression becomes

$$Q_1 = \sum_{i=1}^t \sum_{j=1}^{n_i} [(\mu + u_i + \varepsilon_{ij}) - (\mu + u_i + \bar{\varepsilon}_{i\cdot})]^2 = \sum_{i=1}^t \sum_{j=1}^{n_i} (\varepsilon_{ij} - \bar{\varepsilon}_{i\cdot})^2$$

The expectation of Q_1 can be evaluated using properties of the distribution of ε as

$$\begin{aligned} E(Q_1) &= \sum_{i=1}^t \sum_{j=1}^{n_i} E(\varepsilon_{ij} - \bar{\varepsilon}_{i\cdot})^2 = \sum_{i=1}^t \sum_{j=1}^{n_i} [E(\varepsilon_{ij}^2) + E(\bar{\varepsilon}_{i\cdot}^2) - 2E(\varepsilon_{ij}\bar{\varepsilon}_{i\cdot})] \quad (\text{by squaring}) \\ &= \sum_{i=1}^t \sum_{j=1}^{n_i} \left(\sigma_\varepsilon^2 + \frac{\sigma_\varepsilon^2}{n_i} - 2 \frac{\sigma_\varepsilon^2}{n_i} \right) \quad \text{using } E(\varepsilon_{ij}^2) = \sigma_\varepsilon^2 \text{ and } E(\bar{\varepsilon}_{i\cdot}^2) = \frac{\sigma_\varepsilon^2}{n_i} \\ &= \sum_{i=1}^t \sum_{j=1}^{n_i} \frac{(n_i - 1)\sigma_\varepsilon^2}{n_i} = \sigma_\varepsilon^2 \sum_{i=1}^t (n_i - 1) = (N - t)\sigma_\varepsilon^2 \end{aligned}$$

Using this expression for the expectation for the sum of squares, the expected mean square is

$$E(\text{mean square of } Q_1) = E\left(\frac{Q_1}{N-t}\right) = \sigma_\varepsilon^2$$

Substituting the expressions in Equation 18.4 into the equation defining Q_2 provides

$$\begin{aligned} Q_2 &= \sum_{i=1}^t n_i (\mu + u_i + \varepsilon_{ij} - \mu - \tilde{u}_{..} + \bar{\varepsilon}_{i..})^2 \\ &= \sum_{i=1}^t n_i [(u_i - \tilde{u}_{..}) + (\varepsilon_{ij} - \bar{\varepsilon}_{i..})]^2 \end{aligned}$$

The expectation of Q_2 is

$$\begin{aligned} E(Q_2) &= \sum_{i=1}^t n_i [E(u_i - \tilde{u}_{..})^2 + E(\bar{\varepsilon}_{i..} - \bar{\varepsilon}_{..})^2] \\ &= \sum_{i=1}^t n_i [E(u_i)^2 + E(\tilde{u}_{..})^2 - 2E(u_i \tilde{u}_{..}) + E(\bar{\varepsilon}_{i..})^2 + E(\bar{\varepsilon}_{..})^2 - 2E(\bar{\varepsilon}_{i..} \bar{\varepsilon}_{..})] \end{aligned}$$

To simplify this, the expectation of Q_2 is evaluated in two parts. The first part evaluates the expectation involving the ε . The distributions associated with the means $\bar{\varepsilon}_{i..}$ and $\bar{\varepsilon}_{..}$ are

$$\bar{\varepsilon}_{i..} \sim N\left(0, \frac{\sigma_\varepsilon^2}{n_i}\right) \quad \text{and} \quad \bar{\varepsilon}_{..} \sim N\left(0, \frac{\sigma_\varepsilon^2}{N}\right)$$

Then the expectation of the part of Q_2 involving the ε is

$$\begin{aligned} &\sum_{i=1}^t n_i [E(\bar{\varepsilon}_{i..})^2 + E(\bar{\varepsilon}_{..})^2 - 2E(\bar{\varepsilon}_{i..} \bar{\varepsilon}_{..})] \\ &= \sum_{i=1}^t n_i \left\{ \sigma_\varepsilon^2 \left(\frac{1}{n_i} \right) + \frac{\sigma_\varepsilon^2}{N} - 2E\left[\bar{\varepsilon}_{i..} \frac{\sum_{i'=1}^t n_{i'} \bar{\varepsilon}_{i'..}}{N} \right] \right\} \\ &= \sum_{i=1}^t n_i \left\{ \sigma_\varepsilon^2 \left(\frac{1}{n_i} \right) + \frac{\sigma_\varepsilon^2}{N} - 2 \frac{n_i}{N} E(\bar{\varepsilon}_{i..})^2 \right\} \\ &= \sum_{i=1}^t n_i \left\{ \sigma_\varepsilon^2 \left(\frac{1}{n_i} \right) + \frac{\sigma_\varepsilon^2}{N} - 2 \frac{n_i}{N} \left(\frac{\sigma_\varepsilon^2}{n_i} \right) \right\} \\ &= \sum_{i=1}^t n_i \left(\frac{1}{n_i} - \frac{1}{N} \right) \sigma_\varepsilon^2 \\ &= \sum_{i=1}^t \left(1 - \frac{n_i}{N} \right) \sigma_\varepsilon^2 = (t-1) \sigma_\varepsilon^2 \end{aligned}$$

To evaluate the part of the expectation of Q_2 involving the u_i , let $\tilde{u}_i = \sum_{i=1}^t (n_i u_i / N)$. Since the u_i are independent with mean equal to zero,

$$\text{Var}(\tilde{u}_i) = E(\tilde{u}_i)^2 = \sum_{i=1}^t \left(\frac{n_i^2}{N^2} \right) \sigma_u^2 = \left(\frac{1}{N^2} \right) \sum_{i=1}^t (n_i^2) \sigma_u^2$$

The covariance between u_i and \tilde{u}_i is

$$\text{Cov}(\tilde{u}_i, u_i) = E(\tilde{u}_i u_i) = \left(\frac{n_i}{N} \right) E(u_i^2) = \left(\frac{n_i}{N} \right) \sigma_u^2$$

since $E(u_i u_{i'}) = 0$ for $i \neq i'$. Putting these pieces together, the part of the expectation of Q_2 involving the u_i is

$$\begin{aligned} & \sum_{i=1}^t n_i [E(u_i^2) + E(\tilde{u}_i^2) - 2E(u_i \tilde{u}_i)] \\ &= \sum_{i=1}^t n_i \sigma_u^2 + \sum_{i=1}^t \frac{n_i \sigma_u^2}{N^2} \left(\sum_{i=1}^t n_i^2 \right) - 2 \sum_{i=1}^t \left(\frac{n_i^2}{N} \right) \sigma_u^2 \\ &= \left(\sum_{i=1}^t n_i \sigma_u^2 - \frac{\sum_{i=1}^t n_i^2}{N} \sigma_u^2 \right) = \left(N - \frac{\sum_{i=1}^t n_i^2}{N} \right) \sigma_u^2 \end{aligned}$$

Finally, putting the two parts together, the expectation of Q_2 is

$$E(Q_2) = (t-1)\sigma_e^2 + \left(N - \frac{\sum_{i=1}^t n_i^2}{N} \right) \sigma_u^2$$

and the expected mean square for Q_2 is

$$E\left(\frac{Q_2}{t-1}\right) = \sigma_e^2 + \left(\frac{1}{t-1} \right) \left(N - \frac{\sum_{i=1}^t n_i^2}{N} \right) \sigma_u^2$$

An analysis of variance table with sources, degrees of freedom, sums of squares and expected means squares for the one-way random effects model is given in Table 18.1.

18.3.2 Computing Using Hartley's Method of Synthesis

Hartley (1967) described a technique for evaluating expectations of mean squares for a specific design and sum of squares that he called the method of synthesis. In order to describe and later apply the technique, the expectation of a sum of squares computed from the general random effects model (18.3) is expressed in matrix notation. Any sum of squares that is computed using an equation or is extracted via computer software can always be represented as a quadratic form in the data as

$$Q = \mathbf{y}' \mathbf{A} \mathbf{y} \tag{18.5}$$

TABLE 18.1

Analysis of Variance Table for One-Way Random Effects Model

Source	df	Sum of Squares	Expected Mean Square
Between treatments or populations	$t - 1$	$Q_2 = \sum_{i=1}^t n_i (\bar{y}_{i\cdot} - \bar{y}_{..})^2$	$\sigma_e^2 + \left(\frac{1}{t-1} \right) \left(N - \frac{\sum_{i=1}^t n_i^2}{N} \right) \sigma_u^2$
Within error	$N - t$	$Q_1 = \sum_{i=1}^t \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i\cdot})^2$	σ_e^2

where \mathbf{y} is the vector of observations and A is an appropriately chosen symmetric matrix of constants called the matrix of the quadratic form (Graybill, 1976). As an example, the sample variance of a vector of n observations is

$$s^2 = \sum_{i=1}^n \frac{(y_i - \bar{y}_i)^2}{n-1} = \mathbf{y}' \left[\frac{1}{n-1} \left(\mathbf{I}_n - \frac{1}{n} \mathbf{J}_n \right) \right] \mathbf{y}$$

where \mathbf{I}_n is an $n \times n$ identity matrix, \mathbf{J}_n is an $n \times n$ matrix of ones, and the matrix of the quadratic form $\mathbf{y}'A\mathbf{y}$ is

$$A = \frac{1}{n-1} \left(\mathbf{I}_n - \frac{1}{n} \mathbf{J}_n \right)$$

For different models, certain choices of A always exist that yield the desired sums of squares, but fortunately, as will be seen shortly, it is not necessary to know the elements in A , nor even to know how to determine the elements of matrix A . You only need to know that A exists.

For the general random effects model (18.3) and its corresponding covariance matrix Σ , the expectation of a quadratic form (Graybill, 1976) is

$$E(\mathbf{y}'A\mathbf{y}) = \text{Tr}[A\Sigma] + \frac{1}{2} \mu^2 j'_n A j_n \quad (18.6)$$

where $\text{Tr}[B] = \sum_{i=1}^n b_{ii}$ and b_{ii} , $i = 1, 2, \dots, n$, denote the diagonal elements of the square matrix B . The sums of squares in the analysis of variance are constructed such that $\mu^2 j'_n A j_n = 0$. Thus the expectations of the sums of squares do not depend on $\mu^2 j'_n A j_n$ and are given by

$$E(\mathbf{y}'A\mathbf{y}) = \text{Tr}[A\Sigma] \quad (18.7)$$

The covariance matrix of the general random effects model in Equation 18.3 is

$$\Sigma = \sigma_1^2 \mathbf{Z}_1 \mathbf{Z}'_1 + \sigma_2^2 \mathbf{Z}_2 \mathbf{Z}'_2 + \dots + \sigma_r^2 \mathbf{Z}_r \mathbf{Z}'_r + \sigma_e^2 \mathbf{I}_N$$

Thus, the expectation of the quadratic form $\mathbf{y}'A\mathbf{y}$ is

$$\begin{aligned} E(\mathbf{y}'A\mathbf{y}) &= \text{Tr}[A\Sigma] = \text{Tr}[A(\sigma_1^2 \mathbf{Z}_1 \mathbf{Z}'_1 + \sigma_2^2 \mathbf{Z}_2 \mathbf{Z}'_2 + \dots + \sigma_r^2 \mathbf{Z}_r \mathbf{Z}'_r + \sigma_e^2 \mathbf{I}_N)] \\ &= \sigma_1^2 \text{Tr}[A \mathbf{Z}_1 \mathbf{Z}'_1] + \sigma_2^2 \text{Tr}[A \mathbf{Z}_2 \mathbf{Z}'_2] + \dots + \sigma_r^2 \text{Tr}[A \mathbf{Z}_r \mathbf{Z}'_r] + \sigma_e^2 \text{Tr}[A] \end{aligned}$$

Therefore, the coefficient of σ_e^2 is $\text{Tr}[A]$ and this is equal to the number of degrees of freedom associated with the sum of squares $\mathbf{y}'\mathbf{A}\mathbf{y}$. The coefficient of σ_s^2 is $\text{Tr}[\mathbf{A}\mathbf{Z}_s\mathbf{Z}'_s]$ for $s = 1, 2, \dots, r$.

One property of the trace operator is that $\text{Tr}[\mathbf{A}\mathbf{Z}_s\mathbf{Z}'_s] = \text{Tr}[\mathbf{Z}'_s\mathbf{A}\mathbf{Z}_s]$ where $\text{Tr}[\mathbf{Z}'_s\mathbf{A}\mathbf{Z}_s]$ is the sum of the diagonal elements of $\mathbf{Z}'_s\mathbf{A}\mathbf{Z}_s$ or $\text{Tr}[\mathbf{Z}'_s\mathbf{A}\mathbf{Z}_s] = \sum_{j=1}^{t_s} z'_{sj} \mathbf{A} z_{sj}$ since there are t_s columns in \mathbf{Z}_s . But $z'_{sj} \mathbf{A} z_{sj}$ is the same sum of squares as $\mathbf{y}'\mathbf{A}\mathbf{y}$ except that the column vector \mathbf{z}_{sj} is used as data in place of \mathbf{y} . Hence, if you have an equation or a computer program that calculates $\mathbf{y}'\mathbf{A}\mathbf{y}$ it can also be used to calculate $z'_{sj} \mathbf{A} z_{sj}$. Thus, the coefficient of σ_s^2 in the expectation of $\mathbf{y}'\mathbf{A}\mathbf{y}$ is the sum $z'_{s1} \mathbf{A} z_{s1} + z'_{s2} \mathbf{A} z_{s2} + \dots + z'_{st_s} \mathbf{A} z_{st_s}$.

If the elements of \mathbf{A} are known, then the above sums of squares can be evaluated explicitly. If \mathbf{A} is not known, which is likely since a computer code is probably used to calculate $\mathbf{y}'\mathbf{A}\mathbf{y}$, each $z'_{sj} \mathbf{A} z_{sj}$ can be computed by having the computer calculate the sum of squares where the column \mathbf{z}_{sj} is used as the data vector (instead of \mathbf{y}). Thus, the sum of squares must be computed for each column of the matrix of design matrices from all of the random effects in the model, $[\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_r]$ as if it were data, and then the expectation of the sum of squares $\mathbf{y}'\mathbf{A}\mathbf{y}$ is evaluated as

$$E(\mathbf{y}'\mathbf{A}\mathbf{y}) = v\sigma_e^2 + \left(\sum_{q=1}^{t_1} z'_{1q} \mathbf{A} z_{1q} \right) \sigma_1^2 + \left(\sum_{q=1}^{t_2} z'_{2q} \mathbf{A} z_{2q} \right) \sigma_2^2 + \dots + \left(\sum_{q=1}^{t_r} z'_{rq} \mathbf{A} z_{rq} \right) \sigma_r^2$$

where v is number of the degrees of freedom associated with $\mathbf{y}'\mathbf{A}\mathbf{y}$.

To help demonstrate the idea of synthesis, the expectations of the sums of squares, Q_1 and Q_2 are recomputed for the one-way random effects model. First, a specific model with $t = 4$ and $n_i = 4$ for each i is used to show how to compute the expectations and then the expectations are computed for the general one-way random effects model.

The matrix form for a model to describe the yield of four randomly selected varieties of wheat in a completely randomized design structure with four replications is

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \\ y_{21} \\ y_{22} \\ y_{23} \\ y_{24} \\ y_{31} \\ y_{32} \\ y_{33} \\ y_{34} \\ y_{41} \\ y_{42} \\ y_{43} \\ y_{44} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_{11} \\ \boldsymbol{\varepsilon}_{12} \\ \boldsymbol{\varepsilon}_{13} \\ \boldsymbol{\varepsilon}_{14} \\ \boldsymbol{\varepsilon}_{21} \\ \boldsymbol{\varepsilon}_{22} \\ \boldsymbol{\varepsilon}_{23} \\ \boldsymbol{\varepsilon}_{24} \\ \boldsymbol{\varepsilon}_{31} \\ \boldsymbol{\varepsilon}_{32} \\ \boldsymbol{\varepsilon}_{33} \\ \boldsymbol{\varepsilon}_{34} \\ \boldsymbol{\varepsilon}_{41} \\ \boldsymbol{\varepsilon}_{42} \\ \boldsymbol{\varepsilon}_{43} \\ \boldsymbol{\varepsilon}_{44} \end{bmatrix}$$

$$= \mu + \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3 \\ \mathbf{u}_4 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_{11} \\ \boldsymbol{\varepsilon}_{12} \\ \boldsymbol{\varepsilon}_{13} \\ \boldsymbol{\varepsilon}_{14} \\ \boldsymbol{\varepsilon}_{21} \\ \boldsymbol{\varepsilon}_{22} \\ \boldsymbol{\varepsilon}_{23} \\ \boldsymbol{\varepsilon}_{24} \\ \boldsymbol{\varepsilon}_{31} \\ \boldsymbol{\varepsilon}_{32} \\ \boldsymbol{\varepsilon}_{33} \\ \boldsymbol{\varepsilon}_{34} \\ \boldsymbol{\varepsilon}_{41} \\ \boldsymbol{\varepsilon}_{42} \\ \boldsymbol{\varepsilon}_{43} \\ \boldsymbol{\varepsilon}_{44} \end{bmatrix}$$

or $\mathbf{y} = j_{16}\mu + [\mathbf{z}_1 \ \mathbf{z}_2 \ \mathbf{z}_3 \ \mathbf{z}_4]\mathbf{u} + \boldsymbol{\varepsilon}$.

The expectation of Q_1 , the within sum of squares, is

$$E(Q_1) = E(\mathbf{y}' \mathbf{A}_w \mathbf{y}) = \sigma_u^2 \left[\sum_{j=1}^4 z_j' \mathbf{A}_w z_j \right] + 12\sigma_e^2$$

where 12 is the degrees of freedom associated with Q_1 and \mathbf{A}_w denotes the matrix of the quadratic form for the within sum of squares Q_1 . To obtain the coefficient of σ_u^2 , compute the within sum of squares using z_1 as data, compute Q_1 using z_2 as data, compute Q_1 by using z_3 as data and compute Q_1 using z_4 as data. The within sum of squares for column z_1

$$Q_1(z_1) = \sum_{i=1}^4 \sum_{j=1}^4 z_{1ij}^2 - 4 \sum_{i=1}^4 \bar{z}_{1i..}^2 = 4 - 4(1) = 0$$

Likewise, the values of $Q_1(z_2)$, $Q_1(z_3)$, and $Q_1(z_4)$, are also zero, which implies that the coefficient of σ_u^2 in $E(Q_1)$ is zero, thus, $E(Q_1) = 12\sigma_e^2$.

The expectation of Q_2 , the between sum of squares, is

$$E(Q_2) = E(\mathbf{y}' \mathbf{A}_B \mathbf{y}) = \sigma_u^2 \left[\sum_{j=1}^4 z_j' \mathbf{A}_B z_j \right] + 3\sigma_e^2$$

where

$$Q_2 = \sum_{i=1}^t n_i (\bar{y}_{i..} - \bar{y}_{..})^2 = \sum_{i=1}^t n_i \bar{y}_{i..}^2 - n_{..} \bar{y}_{..}^2$$

and 3 is the number of degrees of freedom associated with Q_2 ; \mathbf{A}_B denotes the matrix of the quadratic form for the between sum of squares Q_2 . To compute the coefficient of σ_u^2 , compute the between sum of squares or Q_2 using each column z_1 , z_2 , z_3 and z_4 and add these four values together. The value of the between sum of squares computed by using z_1 as the data is

$$Q_2(z_1) = 4 \sum_{i=1}^4 \bar{z}_{1i..}^2 - 16 \bar{z}_{..}^2 = 4(1^2 + 0^2 + 0^2 + 0^2) - 16(0.25)^2 = 3$$

The values of $Q_2(z_2)$, $Q_2(z_3)$, and $Q_2(z_4)$, are also equal to 3; thus, the coefficient of σ_u^2 $3 + 3 + 3 + 3 = 12$. Using these values, the expression for $E(Q_2)$ is

$$E(Q_2) = 12\sigma_u^2 + 3\sigma_e^2$$

Next, Hartley's method of synthesis is used to compute the expectations of the between and within sums of squares for the general one-way random effects model in Equation 18.1. The matrix form of the model is

$$\begin{bmatrix} y_{11} \\ \vdots \\ y_{1n_1} \\ y_{21} \\ \vdots \\ y_{2n_2} \\ \vdots \\ y_{t1} \\ \vdots \\ y_{tn_t} \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} \mu + \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_t \end{bmatrix} + \boldsymbol{\varepsilon}$$

or

$$\mathbf{y} = \mathbf{j}_{n_t} \mu + [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_t] \mathbf{u} + \boldsymbol{\varepsilon}$$

The within or error sum of squares is

$$Q_1 = \sum_{i=1}^t \sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^t n_i \bar{y}_{i*}^2 = \mathbf{y}' \mathbf{A}_w \mathbf{y}$$

and its expectation is

$$E(Q_1) = \sigma_u^2 \sum_{i=1}^t \mathbf{z}_i' \mathbf{A}_w \mathbf{z}_i + (N-t) \sigma_e^2$$

where $N = \sum_{i=1}^t n_i$ and \mathbf{A}_w is the matrix of the quadratic form for the within sum of squares. The value of Q_1 when using the first column, \mathbf{z}_1 , as data is

$$Q_1(\mathbf{z}_1) = \sum_{i=1}^t \sum_{j=1}^{n_i} z_{1ij}^2 - \sum_{i=1}^t n_i \bar{z}_{1i*}^2 = n_1 - [n_1(1) + n_2(0) + \cdots + n_t(0)] = 0$$

Likewise, the values of $Q_2(\mathbf{z}_2), Q_2(\mathbf{z}_3), \dots, Q_2(\mathbf{z}_t)$, are also zero; thus, the coefficient of σ_u^2 in $E(Q_1)$ is zero, implying that $E(Q_1) = (N-t)\sigma_e^2$ and where there are $(N-t)$ degrees of freedom associated with Q_1 .

The between sum of squares is

$$Q_2 = \sum_{i=1}^t n_i \bar{y}_{i..}^2 - n_{..} \bar{y}_{..}^2 = \mathbf{y}' \mathbf{A}_B \mathbf{y}$$

and its expectation is

$$E(Q_2) = \sigma_u^2 \sum_{i=1}^t \mathbf{z}_i' \mathbf{A}_B \mathbf{z}_i + (t-1)\sigma_e^2$$

where \mathbf{A}_B is the matrix of the quadratic form for the between sum of squares. The value of Q_2 using the first column, \mathbf{z}_1 , as data is

$$Q_2(\mathbf{z}_1) = \sum_{i=1}^t n_i \bar{z}_{1i..}^2 - N \bar{z}_{1..}^2$$

For column \mathbf{z}_1 , $\bar{z}_{11..} = 1$, $\bar{z}_{12..} = \bar{z}_{13..} = \dots = \bar{z}_{1t..} = 0$ and $\bar{z}_{1..} = n_1/N$. Thus,

$$Q_2(\mathbf{z}_1) = n_1(1)^2 + n_2(0)^2 + \dots + n_t(0)^2 - N(n_1/N)^2 = \frac{n_1 - n_1^2}{N}$$

Similarly, the values of Q_2 using the other columns of \mathbf{Z} as data are

$$Q_2(\mathbf{z}_2) = \frac{n_2 - n_2^2}{N}, \quad Q_2(\mathbf{z}_3) = \frac{n_3 - n_3^2}{N}, \dots, \quad Q_2(\mathbf{z}_t) = \frac{n_t - n_t^2}{N}$$

Combining these results, the expectation of the between sum of squares is

$$E(Q_2) = (t-1)\sigma_e^2 + \left[\sum_{i=1}^t Q_2(\mathbf{z}_i) \right] \sigma_u^2 = (t-1)\sigma_e^2 + \left[N - \frac{\sum_{i=1}^t n_i^2}{N} \right] \sigma_u^2$$

The expectations of Q_1 and Q_2 obtained via Hartley's method of synthesis are equivalent to those obtained using the algebraic technique.

Next, Hartley's method of synthesis is used to evaluate the expectation of sums of squares for a model used to describe a two-way random effects treatment structure with interaction in a completely randomized design structure. In this case, both the row treatments and column treatments are random effects. Data from this experiment can be modeled by

$$y_{ijk} = \mu + a_i + b_j + c_{ij} + \varepsilon_{ijk}, \quad i = 1, 2, \dots, s, j = 1, 2, \dots, t, k = 1, 2, \dots, n_{ij}$$

where the a_i , $i = 1, 2, \dots, s$, denote the random effects corresponding to the rows with distributions that are *i.i.d.* $N(0, \sigma_a^2)$; the b_j , $j = 1, 2, \dots, t$, denote the random effects corresponding to the columns with distributions that are *i.i.d.* $N(0, \sigma_b^2)$; the c_{ij} denote the random effects of the row-column combinations with distributions that are *i.i.d.* $N(0, \sigma_c^2)$, and the ε_{ijk} denote the experimental unit errors with distributions that are *i.i.d.* $N(0, \sigma_e^2)$. The schematic in Figure 18.1 represents data from an unbalanced two-way treatment structure in a completely randomized design structure where there are two or three observations per cell.

		Column treatments		
Row treatments	1	1	2	3
		$y_{111} y_{112} y_{113}$	$y_{121} y_{122}$	$y_{131} y_{132}$
	2	$y_{211} y_{212}$	$y_{221} y_{222} y_{223}$	$y_{231} y_{232}$

FIGURE 18.1 Example of two-way random effects treatment structure.

The matrix form for this two-way random effects treatment structure in a completely randomized design structure for the data structure in Figure 18.1 is

$$\begin{bmatrix} y_{111} \\ y_{112} \\ y_{113} \\ y_{121} \\ y_{122} \\ y_{131} \\ y_{132} \\ y_{211} \\ y_{212} \\ y_{221} \\ y_{222} \\ y_{223} \\ y_{231} \\ y_{232} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \mu + \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_{11} \\ c_{12} \\ c_{13} \\ c_{21} \\ c_{22} \\ c_{23} \end{bmatrix} + \boldsymbol{\varepsilon}$$

or $\mathbf{y} = \mathbf{j}_{14} \mu + \mathbf{Z}_1 \mathbf{a} + \mathbf{Z}_2 \mathbf{b} + \mathbf{Z}_3 \mathbf{c} + \boldsymbol{\varepsilon}$ where the parameters of the model are μ , σ_a^2 , σ_b^2 , σ_c^2 , and σ_e^2 .

Sums of squares for this unbalanced two-way random effects model can be computed in several different ways (see Chapters 9 and 10). In the analysis for this model, there are four variance components and thus one needs four different sums of squares. To demonstrate the method of synthesis, four sums of squares have been selected that correspond to the balanced case of SSROWS, SSCOLUMNS, SSINTERACTION, and SSERROR but have been modified for the unequal sample sizes (which correspond to Henderson's type I sums of squares, discussed later in this section). The four sums of squares are

$$Q_1 = \sum_{i=1}^s \frac{y_{i..}^2}{n_{i..}} - \frac{y_{...}^2}{n...} \quad (\text{SSROWS})$$

$$Q_2 = \sum_{j=1}^t \frac{y_{.j}^2}{n_{.j}} - \frac{y_{...}^2}{n...} \quad (\text{SSCOLUMNS})$$

$$Q_3 = \sum_{i=1}^s \sum_{j=1}^t \frac{y_{ij.}^2}{n_{ij.}} - Q_1 - Q_2 + \frac{y_{...}^2}{n...} \quad (\text{SSINTERACTION})$$

and

$$Q_4 = \sum_{i=1}^s \sum_{j=1}^t \sum_{k=1}^{n_{ij.}} (y_{ijk} - \bar{y}_{ij.})^2 \quad (\text{SSERROR})$$

For the data structure in Figure 18.1,

$$Q_2 = \frac{y_{\cdot 1}^2}{5} + \frac{y_{\cdot 2}^2}{5} + \frac{y_{\cdot 3}^2}{4} - \frac{y_{\cdot \cdot}^2}{14}$$

and $E(Q_2)$ has the form $E(Q_2) = k_1\sigma_a^2 + k_2\sigma_b^2 + k_3\sigma_c^2 + 2\sigma_e^2$ for some values of k_1 , k_2 , and k_3 where 2 is the number of degrees of freedom associated with Q_2 . The next step is to use Hartley's method of synthesis to determine k_1 , k_2 , and k_3 . To determine the value of k_1 , compute Q_2 for each of the two columns of Z_1 . The value of Q_2 using the first column of Z_1 as data is

$$Q_2(z_{11}) = \frac{3^2}{5} + \frac{2^2}{5} + \frac{2^2}{4} - \frac{7^2}{14} = 0.1$$

and the value of Q_2 using the second column of Z_1 as data is

$$Q_2(z_{12}) = \frac{2^2}{5} + \frac{3^2}{5} + \frac{2^2}{4} - \frac{5^2}{14} = 0.1$$

Thus, $k_1 = Q_2(z_{11}) + Q_2(z_{12}) = 0.1 + 0.1 = 0.2$. To determine the value of k_2 , compute Q_2 for each column of Z_2 as

$$Q_2(z_{21}) = \frac{5^2}{5} + \frac{0^2}{5} + \frac{0^2}{4} - \frac{5^2}{14} = 3.214$$

$$Q_2(z_{22}) = \frac{0^2}{5} + \frac{5^2}{5} + \frac{0^2}{4} - \frac{5^2}{14} = 3.214$$

and

$$Q_2(z_{23}) = \frac{0^2}{5} + \frac{0^2}{5} + \frac{4^2}{4} - \frac{4^2}{14} = 2.857$$

The value of k_2 is $k_2 = Q_2(z_{21}) + Q_2(z_{22}) + Q_2(z_{23}) = 3.214 + 3.214 + 2.857 = 9.285$.

The value of k_3 is obtained by computing Q_2 for each column of Z_3 . The values of Q_2 are

$$Q_2(z_{311}) = \frac{3^2}{5} + \frac{0^2}{5} + \frac{0^2}{4} - \frac{3^2}{14} = 1.157$$

$$Q_2(z_{312}) = \frac{0^2}{5} + \frac{2^2}{5} + \frac{0^2}{4} - \frac{2^2}{14} = 0.514$$

$$Q_2(z_{313}) = \frac{0^2}{5} + \frac{0^2}{5} + \frac{2^2}{4} - \frac{2^2}{14} = 0.714$$

$$Q_2(z_{321}) = \frac{2^2}{5} + \frac{0^2}{5} + \frac{0^2}{4} - \frac{2^2}{14} = 0.514$$

$$Q_2(z_{322}) = \frac{0^2}{5} + \frac{3^2}{5} + \frac{0^2}{4} - \frac{3^2}{14} = 1.157$$

and

$$Q_2(z_{323}) = \frac{0^2}{5} + \frac{0^2}{5} + \frac{2^2}{4} - \frac{2^2}{14} = 0.714$$

Thus, the value of k_3 is computed as $k_3 = 1.157 + 0.514 + 0.714 + 0.514 + 1.157 + 0.714 = 4.770$. Using the above values for k_1 , k_2 , and k_3 , the expectation of the SSROWS is

$$E(Q_2) = 0.200\sigma_a^2 + 9.285\sigma_b^2 + 4.770\sigma_c^2 + 2\sigma_e^2$$

Similarly, applying Hartley's method of synthesis, the expectations of Q_1 and Q_3 are determined to be

$$E(Q_1) = 7.000\sigma_a^2 + 0.143\sigma_b^2 + 2.429\sigma_c^2 + (1)\sigma_e^2$$

and

$$E(Q_3) = 4.371\sigma_c^2 + 2\sigma_e^2$$

In general, the expectation of the SSERROR is equal to the number of degrees of freedom associated with SSERROR times σ_e^2 , which in this case provides

$$E(Q_4) = E(\text{SSERROR}) = 8\sigma_e^2$$

There are various ways to compute sums of squares for unbalanced treatment structures, including the type I-IV sums of squares of SAS®-GLM, and type I-III from SAS-MIXED, as well as four methods due to Henderson. The computations of the expectations of the type I-III from SAS-GLM are shown below as well as a demonstration of how to use SAS to provide the computations of the expectations of sums of squares. Table 18.2 contains the SAS code to generate the columns of Z_1 (denoted by a_1 and a_2), Z_2 (denoted by b_1 , b_2 , and b_3), and Z_3 (denoted by c_{11} , c_{12} , c_{13} , c_{21} , c_{22} , and c_{23}) and compute the three types of

TABLE 18.2

SAS Code to Evaluate the Type I-III Sums of Squares for Each Column of $[Z1, Z2, Z3]$ and y for Data Structure in Figure 18.1 and Data in Table 19.2

```

data ex_18; input row col y @@;
a1=(row=1); a2=(row=2); b1=(col=1); b2=(col=2); b3=(col=3);
c11=a1*b1;c12=a1*b2;c13=a1*b3;c21=a2*b1;c22=a2*b2;c23=a2*b3;
datalines;
1 1 10 1 1 12 1 1 11 1 2 13 1 2 15 1 3 21 1 3 19
2 1 16 2 1 18 2 2 13 2 2 19 2 2 14 2 3 11 2 3 13
proc glm data=ex_18;
class row col;
model y a1 a2 b1 b2 b3 c11 c12 c13 c21 c22 c23=row col row*col / ss1
ss2 ss3 e1 e2 e3;
random row col row*col;

```

TABLE 18.3

Type I–III Sums of Squares for Each Column of [Z1, Z2, Z3] and y for Data Structure in Figure 18.1 and Data in Table 19.2

SAS Sums of Squares

Effect	Type I			Type II			Type III		
	Row	Column	Interaction	Row	Column	Interaction	Row	Column	Interaction
a1	3.50000	0.0000	0.000	3.40000	0.0000	0.000	3.37500	0.00000	0.000
a2	3.50000	0.0000	0.000	3.40000	0.0000	0.000	3.37500	0.00000	0.000
b1	0.07143	3.1429	0.000	0.00000	3.1429	0.000	0.00000	3.10588	0.000
b2	0.07143	3.1429	0.000	0.00000	3.1429	0.000	0.00000	3.10588	0.000
b3	0.00000	2.8571	0.000	0.00000	2.8571	0.000	0.00000	2.82353	0.000
c11	0.64286	0.9378	0.776	0.42353	0.9378	0.776	0.37500	0.77647	0.776
c12	0.28571	0.6521	0.776	0.42353	0.6521	0.776	0.37500	0.77647	0.776
c13	0.28571	0.7227	0.706	0.29412	0.7227	0.706	0.37500	0.70588	0.706
c21	0.28571	0.6521	0.776	0.42353	0.6521	0.776	0.37500	0.77647	0.776
c22	0.64286	0.9378	0.776	0.42353	0.9378	0.776	0.37500	0.77647	0.776
c23	0.28571	0.7227	0.706	0.29412	0.7227	0.706	0.37500	0.70588	0.706
Y	0.64286	14.7597	109.145	0.18824	14.7597	109.145	0.16667	8.90980	109.145

sums of squares for each column using SAS-GLM. The sums of squares for each of the columns are displayed in Table 18.3. The expectation of the type I sum of squares due to rows is computed as

$$\begin{aligned} E[\text{SSROW(I)}] &= (1)\sigma_e^2 + (0.6429 + 0.2857 + 0.2857 + 0.6429 + 0.2857)\sigma_c^2 \\ &\quad + (0.0714 + 0.0714 + 0.000)\sigma_b^2 + (3.500 + 3.500)\sigma_a^2 \\ &= \sigma_e^2 + 2.4284\sigma_c^2 + 0.1428\sigma_b^2 + 7\sigma_a^2 \end{aligned}$$

The sums of the sums of squares for the respective effects are displayed in Table 18.4. Thus, Table 18.5 contains the expectations of the remaining sums of squares that were computed using the summary displayed in Table 18.4. The coefficients can be computed

TABLE 18.4

Sums of the Type I–III Sums of Squares for Each Column of [Z1, Z2, Z3] for Each of the Effects Using Data from Table 18.3

SAS Sums of Squares

Effect	Type I			Type II			Type III		
	Row	Column	Interaction	Row	Column	Interaction	Row	Column	Interaction
A Columns	7.00000	0.0000	0.000	6.80000	0.0000	0.000	6.75000	0.00000	0.000
B Columns	0.14286	9.1429	0.000	0.00000	9.1429	0.000	0.00000	9.03529	0.000
C Columns	2.42857	4.6252	4.518	2.28235	4.6252	4.518	2.25000	4.51765	4.518

TABLE 18.5

Expected Values of the SAS Type I–III Sums of Squares,
Computed Using Hartley's Method of Synthesis

$E[\text{SSROWS(I)}] =$	$\sigma_e^2 + 7.0\sigma_a^2 + 0.1429\sigma_b^2 + 2.4286\sigma_c^2$
$E[\text{SSCOLUMNS(I)}] =$	$2\sigma_e^2 + 9.1429\sigma_b^2 + 4.6252\sigma_c^2$
$E[\text{SSINTERACTION (I)}] =$	$2\sigma_e^2 + 4.518\sigma_c^2$
$E[\text{SSERROR(I)}] =$	$8\sigma_e^2$
$E[\text{SSROWS (II)}] =$	$\sigma_e^2 + 6.8\sigma_a^2 + 2.2834\sigma_c^2$
$E[\text{SSCOLUMNS (II)}] =$	$2\sigma_e^2 + 9.1429\sigma_b^2 + 4.6252\sigma_c^2$
$E[\text{SSINTERACTION (II)}] =$	$2\sigma_e^2 + 4.518\sigma_c^2$
$E[\text{SSERROR (II)}] =$	$8\sigma_e^2$
$E[\text{SSROWS (III)}] =$	$\sigma_e^2 + 6.75\sigma_a^2 + 2.25\sigma_c^2$
$E[\text{SSCOLUMNS (III)}] =$	$2\sigma_e^2 + 9.0353\sigma_b^2 + 4.5177\sigma_c^2$
$E[\text{SSINTERACTION (III)}] =$	$2\sigma_e^2 + 4.518\sigma_c^2$
$E[\text{SSERROR(III)}] =$	$8\sigma_e^2$

for each expected sum of squares where the coefficient of σ_e^2 is the number of degrees of freedom associated with the respective sum of squares.

There are several alternative methods for computing sums of squares (with names attached) in the statistical literature that are discussed next. The discussion contains indications as to when the techniques can be applied to mixed effects models as well as random effects models. Henderson (1953) introduced four methods of computing sums of squares, called Henderson's methods I, II, III, and IV (also see Searle, 1987; Henderson, 1984). The following discussion considers Henderson's methods I and III. The analysis of variance method, or Henderson's method I, is appropriate only for random effects models and is a technique that consists of computing sums of squares analogous to those computed for balanced data sets except that they are altered to account for unequal numbers of observations per treatment combination or unbalanced data sets. Henderson's method I sum of squares was used above without justification, but that justification follows. The two-way classification random effects model is used to demonstrate this method.

A model that can be used to describe the balanced two-way data set is

$$y_{ijk} = \mu + a_i + b_j + c_{ij} + \varepsilon_{ijk} \quad i = 1, 2, \dots, s, \quad j = 1, 2, \dots, t, \quad k = 1, 2, \dots, n$$

where

$$a_i \sim i.i.d. N(0, \sigma_a^2), \quad b_j \sim i.i.d. N(0, \sigma_b^2), \quad c_{ij} \sim i.i.d. N(0, \sigma_c^2), \quad \text{and} \quad \varepsilon_{ijk} \sim i.i.d. N(0, \sigma_e^2)$$

The sums of squares used in an analysis of variance are

$$\begin{aligned} SSA &= nt \sum_{i=1}^s (\bar{y}_{i..} - \bar{y}_{...})^2 = \sum_{i=1}^s \frac{\bar{y}_{i..}^2}{nt} - \frac{\bar{y}_{...}^2}{nst}, \\ SSB &= ns \sum_{j=1}^t (\bar{y}_{.j.} - \bar{y}_{...})^2 = \sum_{j=1}^t \frac{\bar{y}_{.j.}^2}{ns} - \frac{\bar{y}_{...}^2}{nst}, \end{aligned}$$

and

$$\begin{aligned} SSAB &= n \sum_{i=1}^s \sum_{j=1}^t (\bar{y}_{ij\cdot} - \bar{y}_{i..} - \bar{y}_{.j\cdot} + \bar{y}_{...})^2 \\ &= \sum_{i=1}^s \sum_{j=1}^t \frac{y_{ij\cdot}^2}{n} - \sum_{i=1}^s \frac{y_{i..}^2}{nt} - \sum_{j=1}^t \frac{y_{.j\cdot}^2}{ns} + \frac{y_{...}^2}{nst} \end{aligned}$$

To convert these sums of squares from equal sample sizes to unequal sample sizes, replace the products nst with $n_{..}$, ns with n_j and nt with n_i . Thus, the sums of squares for unequal sample sizes become

$$\begin{aligned} SSA &= \sum_{i=1}^s \frac{y_{i..}^2}{n_i} - \frac{y_{...}^2}{n_{..}}, \\ SSB &= \sum_{j=1}^t \frac{y_{.j\cdot}^2}{n_j} - \frac{y_{...}^2}{n_{..}}, \end{aligned}$$

and

$$SSAB = \sum_{i=1}^s \sum_{j=1}^t \frac{y_{ij\cdot}^2}{n_{ij}} - \sum_{i=1}^s \frac{y_{i..}^2}{n_i} - \sum_{j=1}^t \frac{y_{.j\cdot}^2}{n_j} + \frac{y_{...}^2}{n_{..}}$$

These are the sums of squares used at the beginning of the discussion of the data structure in Figure 18.1.

The fitting-constants method, or Henderson's method III, involves fitting various linear models to the data and then computing the corresponding sums of squares. The method uses what is called the reduction in the sums of squares due to fitting the full model and those due to fitting various submodels (see Chapter 10). To set the notation, consider the model

$$\mathbf{y} = \mathbf{X}_1 \mathbf{b}_1 + \mathbf{X}_2 \mathbf{b}_2 + \mathbf{X}_3 \mathbf{b}_3 + \boldsymbol{\varepsilon}$$

The reduction in the total sum of squares due to fitting the full model is

$$R(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3) = \mathbf{y}'\mathbf{y} - \text{SSERROR}(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$$

where $\text{SSERROR}(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ is the residual sum of squares after fitting the full model. The reduction due to fitting \mathbf{b}_1 and \mathbf{b}_2 is $R(\mathbf{b}_1, \mathbf{b}_2) = \mathbf{y}'\mathbf{y} - \text{SSERROR}(\mathbf{b}_1, \mathbf{b}_2)$ where $\text{SSERROR}(\mathbf{b}_1, \mathbf{b}_2)$ is the residual sum of squares for the model $\mathbf{y} = \mathbf{X}_1 \mathbf{b}_1 + \mathbf{X}_2 \mathbf{b}_2 + \boldsymbol{\varepsilon}$.

The reduction due to \mathbf{b}_3 after fitting \mathbf{b}_1 and \mathbf{b}_2 is denoted by $R(\mathbf{b}_3|\mathbf{b}_1, \mathbf{b}_2)$ and is given by $R(\mathbf{b}_3|\mathbf{b}_1, \mathbf{b}_2) = R(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3) - R(\mathbf{b}_1, \mathbf{b}_2) = \text{SSERROR}(\mathbf{b}_1, \mathbf{b}_2) - \text{SSERROR}(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$. Likewise, the reduction due to \mathbf{b}_1 after fitting \mathbf{b}_2 is $R(\mathbf{b}_1|\mathbf{b}_2) = R(\mathbf{b}_1, \mathbf{b}_2) - R(\mathbf{b}_2) = \text{SSERROR}(\mathbf{b}_2) - \text{SSERROR}(\mathbf{b}_1, \mathbf{b}_2)$. Finally, the reduction due to \mathbf{b}_1 and \mathbf{b}_2 after fitting \mathbf{b}_3 is $R(\mathbf{b}_1, \mathbf{b}_2|\mathbf{b}_3) = R(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3) - R(\mathbf{b}_3) = \text{SSERROR}(\mathbf{b}_3) - \text{SSERROR}(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$.

One of the advantages of using this technique is that $E[R(\mathbf{b}_1, \mathbf{b}_2|\mathbf{b}_3)]$ does not depend on \mathbf{b}_3 if it is a fixed effect or σ_3^2 if it is a random effect unless \mathbf{b}_3 denotes an interaction between \mathbf{b}_1 and \mathbf{b}_2 or an interaction with \mathbf{b}_1 or \mathbf{b}_2 .

For the model

$$\mathbf{y} = j\mu + X_1\mathbf{b} + X_2\mathbf{t} + X_3\mathbf{g} + \boldsymbol{\epsilon}$$

where $\mathbf{b} \sim N(0, \sigma_b^2)$, $\mathbf{t} \sim N(0, \sigma_t^2)$, $\mathbf{g} \sim N(0, \sigma_g^2)$, and $\boldsymbol{\epsilon} \sim N(0, \sigma_\epsilon^2)$, one set of possible sums of squares (which are type I sums of squares from SAS-GLM and SAS-MIXED) are

$$\begin{aligned} R(\mathbf{b} | \mu) &= R(\mu, \mathbf{b}) - R(\mu) \\ R(\mathbf{t} | \mu, \mathbf{b}) &= R(\mu, \mathbf{b}, \mathbf{t}) - R(\mu, \mathbf{b}) \\ R(\mathbf{g} | \mu, \mathbf{b}, \mathbf{t}) &= R(\mu, \mathbf{b}, \mathbf{t}, \mathbf{g}) - R(\mu, \mathbf{b}, \mathbf{t}) \\ \text{SSERROR}(\mu, \mathbf{b}, \mathbf{t}, \mathbf{g}) &= \mathbf{y}'\mathbf{y} - R(\mu, \mathbf{b}, \mathbf{t}, \mathbf{g}) \end{aligned}$$

The expectations of these sums of squares have the forms

$$\begin{aligned} E[\text{SSERROR}(\mu, \mathbf{b}, \mathbf{t}, \mathbf{g})] &= (n - p)\sigma_\epsilon^2 \\ E[R(\mathbf{g} | \mu, \mathbf{b}, \mathbf{t})] &= k_1\sigma_\epsilon^2 + k_2\sigma_g^2 \\ E[R(\mathbf{t} | \mu, \mathbf{b})] &= k_3\sigma_\epsilon^2 + k_4\sigma_g^2 + k_5\sigma_t^2, \end{aligned}$$

and

$$E[R(\mathbf{b} | \mu)] = k_6\sigma_\epsilon^2 + k_7\sigma_g^2 + k_8\sigma_t^2 + k_9\sigma_b^2$$

The type I sums of squares given by $R(\mathbf{t} | \mu)$, $R(\mathbf{b} | \mu, \mathbf{t})$, $R(\mathbf{g} | \mu, \mathbf{b}, \mathbf{t})$ and $\text{SSERROR}(\mu, \mathbf{b}, \mathbf{t}, \mathbf{g})$ could also be used, and the method of fitting constants can be utilized for both random effects and mixed effects models as long as the fixed effects are fit before the random effects.

The method of synthesis can be used to evaluate the expectations of sums of squares for any model involving random effects and/or multiple error terms. A set of sums of squares and their expectations (or mean squares) can be used to estimate the variance components, to develop tests of hypotheses, and construct confidence intervals about individual variance components and/or functions of the variance components. Methods of estimation are discussed in Chapter 19, and inference techniques are presented in Chapter 20. Many statistical software programs automatically use Hartley's synthesis to compute expected mean squares.

18.4 Concluding Remarks

In this chapter, the concepts of random and fixed effects were defined as well as the concept of a random effects model. The random effects model is expressed in matrix form in order to describe methods for computing the expected mean squares. An unbalanced one-way treatment structure in a completely randomized design structure was used to demonstrate the algebraic method and Hartley's synthesis method of computing expected mean squares. Different methods for computing sums of squares were described, and an unbalanced two-way treatment structure in a completely randomized design structure was used to demonstrate the computations of the respective expected mean squares. SAS code is presented to demonstrate Hartley's method of synthesis.

18.5 Exercises

- 18.1 The model for a one-way random effects treatment structure in a randomized complete block design structure is $y_{ij} = \mu + b_i + a_j + \varepsilon_{ij}$, $i = 1, 2, \dots, n_{blk}$ and $j = 1, 2, \dots, t$ with assumptions $b_i \sim N(0, \sigma_{\text{Block}}^2)$, $a_j \sim N(0, \sigma_{\text{trt}}^2)$, and $\varepsilon_{ij} \sim N(0, \sigma_e^2)$. Use the algebraic method to evaluate the expectations of the three sums of squares

$$SS(\text{blocks}) = t \sum_{i=1}^{n_{blk}} (\bar{y}_{i..} - \bar{y}_{...})^2, \quad SS(\text{treatments}) = n_{blk} \sum_{j=1}^t (\bar{y}_{.j} - \bar{y}_{...})^2$$

and

$$SS(\text{error}) = \sum_{i=1}^{n_{blk}} \sum_{j=1}^t (y_{ij} - \bar{y}_{i..} - \bar{y}_{.j} + \bar{y}_{...})^2$$

- 18.2 For the one-way random effects treatment structure in a randomized complete block design structure model in Exercise 18.1, use the $n_{blk} = 5$ and $t = 4$ and evaluate the expectations of the three sums of squares using Hartley's method of synthesis.

- 18.3 The usual split-plot design (Section 5.2) can be expressed as

$$y_{ijk} = \mu_{ik} + b_j + w_{ij} + \varepsilon_{ijk}, \quad i = 1, 2, \dots, a, \quad j = 1, 2, \dots, b, \quad \text{and} \quad k = 1, 2, \dots, t$$

where $b_j \sim N(0, \sigma_{blk}^2)$, $w_{ij} \sim N(0, \sigma_{\text{whole-plot}}^2)$, and $\varepsilon_{ijk} \sim N(0, \sigma_e^2)$.

Let $a = 3$, $b = 4$, and $t = 2$, then use Hartley's method of synthesis to evaluate the expected values of the three sums of squares

$$\begin{aligned} SS(\text{blocks}) &= at \sum_{j=1}^b (\bar{y}_{.j} - \bar{y}_{...})^2 \\ SS(\text{whole-plot error}) &= t \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij} - \bar{y}_{.j} - \bar{y}_{i..} + \bar{y}_{...})^2 \end{aligned}$$

and

$$SS(\text{subplot error}) = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^t (y_{ijk} - \bar{y}_{i..} - \bar{y}_{.j} + \bar{y}_{i..})^2$$

For extra credit, use the algebraic method to evaluate the expectations of these sums of squares for general a , b , and t .

- 18.4 Use the assumptions for the model in Section 18.3 and evaluate the expectations of the type I sums of squares given by $R(t|\mu)$, $R(b|\mu, t)$, and $R(g|\mu, b, t)$ $\text{SSERROR}(\mu, b, t, g)$.

19

Methods for Estimating Variance Components

There are several ways to estimate variance components for the general random effects model. Some of the procedures yield the same estimators when the design is balanced (equal sample sizes per cell and no missing cells) and different estimators when the design is not balanced. The four techniques discussed in this chapter are the method of moments, maximum likelihood, restricted, or residual maximum likelihood (REML), and MIVQUE. The method of moments produces unbiased estimates, maximum likelihood and REML estimators are consistent and have the usual large-sample-size properties of maximum likelihood estimates, and the MIVQUE method produces estimates having minimum variance within the class of quadratic unbiased estimates. When the design is balanced and the solutions for the variance components are all positive, the method of moments, REML, and MIVQUE estimators are identical. When the design is unbalanced, method-of-moments estimates are easiest to compute, while the other three methods require iterative algorithms. On the other hand, the maximum likelihood, REML, and MIVQUE methods provide estimators with better properties than does the method of moments. REML is generally the preferred method of estimating the variance components.

19.1 Method of Moments

The method of moments has been used to obtain estimates of variance components since Eisenhart (1947) gave the name MODEL II to the random effects model. Many researchers worked on the method of moments over the next 20 years and derived estimators and developed methods for testing hypotheses, and methods to construct confidence intervals about variance components (see Searle, 1987; Graybill, 1976; Henderson 1984; Searle et al. 1992; Burdick and Graybill, 1992 for good lists of references). In this section, a generalized version of the method of moments estimation process is discussed.

The general random effects model can be written as

$$\mathbf{y} = \mathbf{j}_N \boldsymbol{\mu} + \mathbf{Z}_1 \mathbf{u}_1 + \mathbf{Z}_2 \mathbf{u}_2 + \cdots + \mathbf{Z}_r \mathbf{u}_r + \boldsymbol{\epsilon}$$

where

$$\begin{aligned} E(\boldsymbol{u}_i) &= \mathbf{0}, \quad i = 1, 2, \dots, r \\ \text{Var}(\boldsymbol{u}_i) &= \sigma_i^2 \mathbf{I}_{l_i}, \quad i = 1, 2, \dots, r \\ E(\boldsymbol{\varepsilon}) &= \mathbf{0}, \quad \text{Var}(\boldsymbol{\varepsilon}) = \sigma_{\varepsilon}^2 \mathbf{I}_N \end{aligned} \tag{19.1}$$

and $\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_r$, and $\boldsymbol{\varepsilon}$ are independent random vectors.

The method of moments technique for estimating the variance components of the general random effects model of Equation 19.1 involves the following steps:

- 1) Compute as many sums of squares and their corresponding mean squares as there are variance components in the model.
- 2) Evaluate the expectation of each mean of square in terms of the variance components; these expectations must not involve μ (or any other fixed effect parameters) and each variance component must be included in the expectation of at least one of the mean squares.
- 3) Equate the expectation of the mean squares to the observed values of the mean squares, thus generating a system of linear equations in the variance components (replace the variance component parameters with variance component solutions in the set of equations).
- 4) Solve the resulting system of equations to obtain an estimate of each of the variance components.

One problem with the method of moments solution is that some of the estimates of the variance component can have negative values. When the solution for a variance component is negative, the estimator of the variance component is set to zero (keeping the estimator in the parameter space).

When the random effects $\boldsymbol{u}_i, i = 1, 2, \dots, r$, and $\boldsymbol{\varepsilon}$ of model (19.1) are jointly independent and normally distributed and when the sums of squares are distributed independently of one another, the resulting estimators of the variance components are minimum variance unbiased. If $\boldsymbol{u}_i, i = 1, 2, \dots, r$, and $\boldsymbol{\varepsilon}$ have the same first four moments as those of a normal distribution, the estimators are minimum variance quadratic unbiased (Graybill, 1976, p. 632). The method of moments technique does not require an assumption of normality in order to obtain estimators. The only known property these estimators possess without the assumption of normality or the assumption that the distributions of the random vectors have the same first four moments of a normal distribution is that they are unbiased. However, the process of setting the estimator equal to zero when the solution is negative implies that the estimators are no longer unbiased.

The key to the method of moments is determining how to compute sums of squares and then evaluating the expectations of the resulting mean squares. These topics were discussed in Chapter 18.

If the model has $r + 1$ variance components, $\sigma_{\varepsilon}^2, \sigma_1^2, \sigma_2^2, \dots, \sigma_r^2$, then $r + 1$ sums of squares (or mean squares) are required. Let $Q_0 = \mathbf{y}' \mathbf{A}_0 \mathbf{y}$, $Q_1 = \mathbf{y}' \mathbf{A}_1 \mathbf{y}$, ..., $Q_r = \mathbf{y}' \mathbf{A}_r \mathbf{y}$, denote the sums of squares with respective expectations given by

$$E(Q_i) = b_{i0} \sigma_{\varepsilon}^2 + b_{i1} \sigma_1^2 + b_{i2} \sigma_2^2 + \dots + b_{ir} \sigma_r^2, \quad i = 0, 1, 2, \dots, r$$

Equate each sum of squares to its expectation, inserting a tilde (\sim) over the variances to denote a solution as

$$Q_i = b_{i0} \tilde{\sigma}_\epsilon^2 + b_{i1} \tilde{\sigma}_1^2 + b_{i2} \tilde{\sigma}_2^2 + \cdots + b_{ir} \tilde{\sigma}_r^2, \quad i = 0, 1, 2, \dots, r$$

or, in matrix notation as,

$$\begin{bmatrix} Q_0 \\ Q_1 \\ Q_2 \\ \vdots \\ Q_r \end{bmatrix} = \begin{bmatrix} b_{00} & b_{01} & b_{02} & \cdots & b_{0r} \\ b_{10} & b_{11} & b_{12} & \cdots & b_{1r} \\ b_{20} & b_{21} & b_{22} & \cdots & b_{2r} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{r0} & b_{r1} & b_{r2} & \cdots & b_{rr} \end{bmatrix} \begin{bmatrix} \tilde{\sigma}_\epsilon^2 \\ \tilde{\sigma}_1^2 \\ \tilde{\sigma}_2^2 \\ \vdots \\ \tilde{\sigma}_r^2 \end{bmatrix}$$

or $Q = B\tilde{\sigma}^2$.

If the rank of the matrix B is $r + 1$, then all of the variance components are estimable. If the rank of B is less than $r + 1$, then not all of the variance components are estimable and only some linear combinations of the variance components are estimable. Assuming that the rank of B is $r + 1$, the solution to the system of equations is

$$\tilde{\sigma}^2 = B^{-1}Q = CQ \text{ (say)}$$

The solution is obtained without restricting the values to the parameter space; that is, some of the solutions may be negative. The solution for $\tilde{\sigma}_i^2$ is denoted by $\hat{\sigma}_i^2$ and the estimate is denoted by $\hat{\sigma}_i^2$ where

$$\hat{\sigma}_i^2 = \begin{cases} \tilde{\sigma}_i^2 & \text{if } \tilde{\sigma}_i^2 > 0 \\ 0 & \text{if } \tilde{\sigma}_i^2 \leq 0 \end{cases} \quad i = 0, 1, 2, \dots, r$$

In many models and methods for computing sums of squares, the B matrix is triangular and thus the solution can be obtained without inverting B . Each solution is a linear combination of the observed sums of squares, $Q_0, Q_1, Q_2, \dots, Q_r$ as

$$\hat{\sigma}_i^2 = c_{i0}Q_0 + c_{i1}Q_1 + c_{i2}Q_2 + \cdots + c_{ir}Q_r \quad i = 0, 1, 2, \dots, r$$

where $c'_i = [c_{i0} \quad c_{i1} \quad c_{i2} \quad \cdots \quad c_{ir}]$ is the i th row of $C = B^{-1}$. The variance of $\hat{\sigma}_i^2$ is

$$\text{Var}(\hat{\sigma}_i^2) = \text{Var}(c_{i0}Q_0 + c_{i1}Q_1 + c_{i2}Q_2 + \cdots + c_{ir}Q_r)$$

When the Q_i , $i = 0, 1, \dots, r$, are uncorrelated, the variance of $\hat{\sigma}_i^2$ is

$$\text{Var}(\hat{\sigma}_i^2) = c_{i0}^2 \text{Var}(Q_0) + c_{i1}^2 \text{Var}(Q_1) + c_{i2}^2 \text{Var}(Q_2) + \cdots + c_{ir}^2 \text{Var}(Q_r)$$

A summary of method-of-moment estimators and their variances for several models is presented in Searle (1971, Chapter 11). For most balanced models (assuming the moments of the distributions of the random variables correspond to the first four moments of a normal distribution), the method of moments estimators are uniformly minimum variance

unbiased estimators of the variance components (Graybill, 1976). Thus, for nearly balanced models, the method-of-moments estimators should have fairly good properties.

19.1.1 Applications. Example 19.1: Unbalanced One-Way Model

The unbalanced one-way random effects model of Example 18.2 is

$$y_{ij} = \mu + u_i + \varepsilon_{ij} \quad i = 1, 2, \dots, t \text{ and } j = 1, 2, \dots, n_i$$

where the u_i are uncorrelated with mean 0 and variance σ_u^2 , the ε_{ij} are uncorrelated with mean 0 and variance σ_ε^2 , and the u_i and the ε_{ij} are uncorrelated. Two sums of squares that can be used are the sum of squares within, Q_0 or SSW and the sums of squares between, Q_1 or SSB where

$$Q_0 = \sum_{i=1}^t \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i\cdot})^2 = \sum_{i=1}^t \sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^t n_i \bar{y}_{i\cdot}^2 = SSW$$

and

$$Q_1 = \sum_{i=1}^t n_i (\bar{y}_{i\cdot} - \bar{y}_{..})^2 = \sum_{i=1}^t n_i \bar{y}_{i\cdot}^2 - \left(\sum_{i=1}^t n_i \right) \bar{y}_{..}^2 = SSB$$

The expectations of Q_0 and Q_1 were evaluated in Chapter 18 as

$$E(Q_0) = (N - t) \sigma_\varepsilon^2$$

and

$$E(Q_1) = (t-1) \sigma_\varepsilon^2 + \left(N - \frac{\sum_{i=1}^t n_i^2}{N} \right) \sigma_u^2$$

where $N = \sum_{i=1}^t n_i$

The equations obtained by equating the sums of squares to their expectations are

$$Q_0 = (N - t) \tilde{\sigma}_\varepsilon^2$$

$$Q_1 = (t-1) \tilde{\sigma}_\varepsilon^2 + \left(N - \frac{\sum_{i=1}^t n_i^2}{N} \right) \tilde{\sigma}_u^2$$

or, in matrix notation,

$$\begin{bmatrix} Q_0 \\ Q_1 \end{bmatrix} = \begin{bmatrix} N-t & 0 \\ (t-1) & N - \frac{\sum_{i=1}^t n_i^2}{N} \end{bmatrix} \begin{bmatrix} \tilde{\sigma}_\varepsilon^2 \\ \tilde{\sigma}_u^2 \end{bmatrix}$$

The set of equations can also be generated by equating the observed mean squares to their expectations as the mean squares are obtained by dividing each equation by its corresponding degrees of freedom. The solutions are the same. The resulting system of equations to be solved involving the mean squares is

$$\frac{Q_0}{N-t} = \tilde{\sigma}_\varepsilon^2$$

and

$$\frac{Q_1}{t-1} = \tilde{\sigma}_\varepsilon^2 + \frac{\left(N - \frac{\sum_{i=1}^t n_i^2}{N} \right)}{t-1} \tilde{\sigma}_1^2$$

The solution to this system of equations is

$$\tilde{\sigma}_\varepsilon^2 = \frac{Q_0}{N-t}$$

and

$$\tilde{\sigma}_1^2 = \frac{Q_1 - (t-1)\tilde{\sigma}_\varepsilon^2}{N - \frac{\sum_{i=1}^t n_i^2}{N}}$$

The method moments estimators are

$$\hat{\sigma}_\varepsilon^2 = \tilde{\sigma}_\varepsilon^2$$

and

$$\hat{\sigma}_1^2 = \begin{cases} \tilde{\sigma}_1^2 & \text{if } \tilde{\sigma}_1^2 > 0 \\ 0 & \text{if } \tilde{\sigma}_1^2 \leq 0 \end{cases}$$

19.1.2 Example 19.2: Wheat Varieties in a One-Way Random Effects Model

An experimenter randomly selected four varieties of wheat from a population of varieties of wheat and conducted an experiment to evaluate damage caused by insects on the wheat plants just prior to heading. The design structure was a completely randomized design with four replications or plots per variety (the plot is the experimental unit). Because of environmental conditions, some of the plots were destroyed (flooded out by excess rain). A day just before the wheat plants started to head, the experimenter randomly selected 20 plants from each plot and rated the amount of insect damage done to each plant using a scale from 0 to 10 where 0 indicates no damage and 10 indicates

severe damage. Thus, the response measured on each plot is the mean of the ratings from the 20 plants. The data are shown in Table 19.1. The computations necessary to compute the sums of squares, Q_0 and Q_1 , are given in Table 19.1 along with the resulting sums of squares, mean squares, the expected mean squares, the system of equations, the resulting solution, and the estimates of the variance components. The information obtained from the estimates of the variance components is that the plot-to-plot variance within a variety is about 0.056, while the variance of the population of varieties is about 0.067. The variance of a randomly selected plot planted to a randomly selected variety is the sum of the two variance components or

$$\hat{\sigma}_{\text{Damage}}^2 = \hat{\sigma}_e^2 + \hat{\sigma}_{\text{Var}}^2 = 0.056 + 0.067 = 0.123$$

TABLE 19.1

Data and Computations for Insect Damage on Wheat Varieties of Example 19.2

<i>Variety</i>	A	B	C	D
3.90		3.60	4.15	3.35
4.05		4.20	4.60	3.80
4.25		4.05	4.15	
		3.85	4.40	
$\bar{y}_{..} = \frac{52.35}{13} = 4.0269$		$Q_0 = 0.50792$		
$\sum_{i=1}^4 \sum_{j=1}^{n_i} y_{ij}^2 = 212.1275$		$Q_1 = 0.81016$		
$\sum_{i=1}^4 n_i \bar{y}_i^2 = 211.61958$				

Expected Mean Squares

$$E\left(\frac{Q_0}{9}\right) = \sigma_e^2 \quad E\left(\frac{Q_1}{3}\right) = \sigma_e^2 + 3.1795 \sigma_{\text{Var}}^2$$

System of Equations

$$\frac{Q_0}{9} = 0.05644 = \tilde{\sigma}_e^2$$

$$\frac{Q_1}{3} = 0.27005 = \tilde{\sigma}_e^2 + 3.1795 \tilde{\sigma}_{\text{Var}}^2$$

Solution to System of Equations

$$\tilde{\sigma}_e^2 = 0.05644$$

$$\tilde{\sigma}_{\text{Var}}^2 = 0.06719$$

Estimates of the Variance Components

$$\hat{\sigma}_e^2 = 0.05644$$

$$\hat{\sigma}_{\text{Var}}^2 = 0.06719$$

The estimate of the intraclass correlation is

$$\hat{\rho} = \frac{\hat{\sigma}_{\text{Var}}^2}{\hat{\sigma}_\epsilon^2 + \hat{\sigma}_{\text{Var}}^2} = \frac{0.067}{0.123} = 0.545$$

When experiments are done with treatment structures that are other than one-way or other than a completely randomized design structure, there is no universally accepted technique for obtaining sums of squares from which to derive estimates of the variance components. The methods presented in Chapter 18 for computing sums of squares are used to estimate the variance components for a two-way random effects model.

19.1.3 Example 19.3: Data for Two-Way Design in Table 18.2

The data in Table 19.2 are observations for Figure 18.1, where the expectations of several types of sums of squares were evaluated via synthesis. The values Q_0 , Q_1 , Q_2 , and Q_3 correspond to Henderson's method I sums of squares are

$$Q_0 = \text{SSError} = 30.6666$$

$$Q_1 = \text{SSA} = \sum_{i=1}^2 \frac{y_{i..}^2}{n_{i..}} - \frac{y_{...}^2}{n_{...}} = 0.6428$$

$$Q_2 = \text{SSB} = \sum_{j=1}^3 \frac{y_{.j}^2}{n_{.j}} - \frac{y_{...}^2}{n_{...}} = 15.2143$$

$$Q_3 = \text{SSAB} = \sum_{i=1}^2 \sum_{j=1}^3 \frac{y_{ij}^2}{n_{ij}} - \sum_{i=1}^2 \frac{y_{i..}^2}{n_{i..}} - \sum_{j=1}^3 \frac{y_{.j}^2}{n_{.j}} + \frac{y_{...}^2}{n_{...}} = 108.6905$$

The expectations of these sums of squares are:

$$E(Q_0) = 8\sigma_\epsilon^2$$

$$E(Q_1) = \sigma_\epsilon^2 + 0.1429\sigma_b^2 + 2.4286\sigma_c^2 + 7.0\sigma_a^2$$

$$E(Q_2) = 2\sigma_\epsilon^2 + 9.286\sigma_b^2 + 4.77\sigma_c^2 + 0.20\sigma_a^2$$

$$E(Q_3) = 2\sigma_\epsilon^2 + 6.37\sigma_c^2$$

TABLE 19.2

Data for a Two-Way Random Effects Treatment Structure for Example 19.3

Row Treatments	Column Treatments		
	1	2	3
1	10	13	21
	12	15	19
	11		
2	16	13	11
	18	19	13
		14	

Equate the values of the sums of squares to their respective expected values to provide following system of equations

$$\begin{aligned} 30.6666 &= 8\tilde{\sigma}_e^2 \\ 0.6428 &= \tilde{\sigma}_e^2 + 0.1429\tilde{\sigma}_b^2 + 2.4286\tilde{\sigma}_c^2 + 7.0\tilde{\sigma}_a^2 \\ 15.2143 &= 2\tilde{\sigma}_e^2 + 9.286\tilde{\sigma}_b^2 + 4.77\tilde{\sigma}_c^2 + 0.20\tilde{\sigma}_a^2 \\ 108.6905 &= 2\tilde{\sigma}_e^2 + 6.37\tilde{\sigma}_c^2 \end{aligned}$$

The solution to the system of equations

$$\begin{aligned} \tilde{\sigma}_e^2 &= 3.83325 \\ \tilde{\sigma}_c^2 &= 15.8593 \\ \tilde{\sigma}_b^2 &= -10.8817 \\ \tilde{\sigma}_a^2 &= -8.2523 \end{aligned}$$

provides unbiased estimates of the variance components. Because some of the above values are negative, the final method of moments estimates of the variance components are taken to be

$$\begin{aligned} \hat{\sigma}_e^2 &= 3.83325 \\ \hat{\sigma}_c^2 &= 15.8593 \\ \hat{\sigma}_b^2 &= 0.00 \\ \hat{\sigma}_a^2 &= 0.00 \end{aligned}$$

The results from Example 19.3 point out one of the problems often encountered when using the method of moments technique: It can yield negative solutions for the variance components, which are not admissible as estimators. Listed below are the method of moments solutions and the resulting estimates for Example 19.3 obtained by solving the systems of equations generated by the method of fitting constants or Henderson's method III sums of squares (SAS® type I) and the SAS type III sums of squares (the expectations of these sums of squares were evaluated by synthesis and are given in Table 18.4).

The type I solutions are

$$\begin{aligned} \tilde{\sigma}_e^2 &= 3.8333 \\ \tilde{\sigma}_c^2 &= \frac{109.1451 - 2(3.8333)}{4.5178} = 22.4620 \\ \tilde{\sigma}_b^2 &= \frac{14.5797 - 2(3.8333) - 4.6252(22.4620)}{9.1429} = -10.5877 \\ \tilde{\sigma}_a^2 &= \frac{0.6429 - 3.833 - 0.1428(-10.71322) - 2.4284(22.4620)}{7.00} = -8.0329 \end{aligned}$$

and the type I estimates are

$$\begin{aligned}\hat{\sigma}_e^2 &= 3.8333 \\ \hat{\sigma}_c^2 &= 22.4620 \\ \hat{\sigma}_b^2 &= 0 \\ \hat{\sigma}_a^2 &= 0\end{aligned}$$

The type III solutions are

$$\begin{aligned}\tilde{\sigma}_e^2 &= 3.8333 \\ \tilde{\sigma}_c^2 &= \frac{109.1451 - 2(3.8333)}{4.5178} = 22.4620 \\ \tilde{\sigma}_b^2 &= \frac{8.9098 - 2(3.8333) - 4.5178(22.4620)}{9.0353} = -11.1080 \\ \tilde{\sigma}_a^2 &= \frac{0.677 - 3.833 - 2.2500(22.4620)}{6.75} = -8.0305\end{aligned}$$

and the type III solutions are

$$\begin{aligned}\hat{\sigma}_e^2 &= 3.8333 \\ \hat{\sigma}_c^2 &= 22.4620 \\ \hat{\sigma}_b^2 &= 0 \\ \hat{\sigma}_a^2 &= 0\end{aligned}$$

When the solution for a variance component is negative, the standard process is to set the corresponding estimate to zero. To demonstrate the consequences of this process, the expected mean squares for the mean square within and mean square between for a one-way random effects model can be expressed as

$$\begin{aligned}E(MSWithin) &= \sigma_e^2 \\ E(MSBetween) &= \sigma_e^2 + c\sigma_u^2\end{aligned}$$

The solution for σ_u^2 is

$$\tilde{\sigma}_u^2 = \frac{MSBetween - MSWithin}{c}$$

Under the normality assumption, $MSWithin$ and $MSBetween$ are independent random variables. If $\sigma_u^2 = 0$, then both expected mean squares are equal to σ_e^2 , thus, the probability

the solution for σ_u^2 is negative is approximately 0.50, depending on the degrees of freedom. If the numerator and denominator degrees of freedom are equal then

$$P(MSBetween < MSWithin \mid \sigma_u^2 = 0) = 0.50$$

As σ_u^2 gets larger, the probability of getting a negative solution decreases. Thus, it is reasonable to set $\hat{\sigma}_u^2$ equal to zero when the solution is negative (see Searle et al. 1992 for a more detailed discussion). Confidence intervals and tests of hypotheses constructed from methods of moments estimators are discussed in Chapter 20.

19.2 Maximum Likelihood Estimators

In statistics, the most common technique for estimating parameters of a distribution is the method of maximum likelihood. The process uses the assumed distribution of the observations and constructs a likelihood function which is a function of the data and the unknown model parameters. The maximum likelihood estimators are those values of the parameters from the parameter space that maximize the value of the likelihood function. In practice, the \log_e of the likelihood function is maximized. Equivalently, one can find the values of the parameters in the parameter space that minimize $-2(\log_e)$ of the likelihood function. The parameter space for the general random effects model (19.1) is

$$\{-\infty < \mu < +\infty, \quad 0 < \sigma_i^2 < +\infty, \quad i = 1, 2, \dots, k; \quad 0 < \sigma_\varepsilon^2 < \infty\}$$

For the general random effects model 19.1, the distribution of the vector of observations is

$$\mathbf{y} \sim N(j_n \boldsymbol{\mu}, \sigma_\varepsilon^2 \mathbf{I}_n + \sigma_1^2 \mathbf{Z}_1 \mathbf{Z}'_1 + \sigma_2^2 \mathbf{Z}_2 \mathbf{Z}'_2 + \dots + \sigma_k^2 \mathbf{Z}_k \mathbf{Z}'_k) \quad \text{or} \quad \mathbf{y} \sim N(j_n \boldsymbol{\mu}, \boldsymbol{\Sigma})$$

The likelihood function of the observations is

$$L(\boldsymbol{\mu}, \sigma_\varepsilon^2, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2 \mid \mathbf{y}) = (2\pi)^{-n/2} |\boldsymbol{\Sigma}|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{y} - j_n \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{y} - j_n \boldsymbol{\mu})\right]$$

and $-2\log_e$ of the likelihood function is

$$\begin{aligned} \ell(\boldsymbol{\mu}, \sigma_\varepsilon^2, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2 \mid \mathbf{y}) &= -2 \log_e[L(\boldsymbol{\mu}, \sigma_\varepsilon^2, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2 \mid \mathbf{y})] \\ &= n \log_e(2\pi) + \log_e(|\boldsymbol{\Sigma}|) + (\mathbf{y} - j_n \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{y} - j_n \boldsymbol{\mu}) \end{aligned}$$

The process of minimizing $\ell(\boldsymbol{\mu}, \sigma_\varepsilon^2, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2 \mid \mathbf{y})$ over the parameter space generally requires an iterative procedure utilizing likelihood equations generated by taking either the first derivatives or the first and second derivatives of $\ell(\boldsymbol{\mu}, \sigma_\varepsilon^2, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2 \mid \mathbf{y})$ with respect to each of the parameters of the model. When the data are from a balanced design

(equal n and no missing cells), the set of likelihood equations generated by equating the first derivatives of the likelihood function to zero can often be solved explicitly. The solutions obtained are not restricted to the parameter space since some of the values of the solutions for the variance components can have negative values. For some balanced models, it can be shown that the maximum likelihood estimate of σ_i^2 is $\hat{\sigma}_i^2 = 0$ when the solution for σ_i^2 from the likelihood equations is negative (Searle, 1971). For unbalanced designs, an iterative technique is required where the estimation process should restrict the estimates of the variance components to belong to the parameter space.

19.2.1 Example 19.4: Maximum Likelihood Solution for Balanced One-Way Model

The balanced one-way random effects model can be expressed as

$$y_{ij} = \mu + u_i + \varepsilon_{ij}, \quad i = 1, 2, \dots, t \text{ and } j = 1, 2, \dots, n$$

where $u_i \sim i.i.d. N(0, \sigma_u^2)$, $\varepsilon_{ij} \sim i.i.d. N(0, \sigma_\varepsilon^2)$, and the u_i and the ε_{ij} are independent, or $\mathbf{y} \sim N[(\mathbf{j}_n \otimes \mathbf{j}_t)\mu, \boldsymbol{\Sigma}]$ where $\boldsymbol{\Sigma} = \sigma_u^2 \mathbf{J}_n \otimes \mathbf{I}_t + \sigma_\varepsilon^2 \mathbf{I}_n \otimes \mathbf{I}_t$.

The notation $A \otimes B$ denotes the Kronecker or direct product of the two matrices, A and B (Graybill, 1976). The covariance matrix $\boldsymbol{\Sigma}$ can be expressed as

$$\boldsymbol{\Sigma} = \sigma_\varepsilon^2 \left[\left(\mathbf{I}_n - \frac{1}{n} \mathbf{J}_n \right) \otimes \mathbf{I}_t \right] + (\sigma_\varepsilon^2 + n\sigma_u^2) \left[\left(\frac{1}{n} \mathbf{J}_n \right) \otimes \mathbf{I}_t \right]$$

where σ_ε^2 and $(\sigma_\varepsilon^2 + n\sigma_u^2)$ are the characteristic roots of $\boldsymbol{\Sigma}$ and $[(\mathbf{I}_n - (1/n)\mathbf{J}_n) \otimes \mathbf{I}_t]$ and $[(1/n)\mathbf{J}_n] \otimes \mathbf{I}_t$ are orthogonal idempotent matrices.

The inverse of the covariance matrix is

$$\boldsymbol{\Sigma}^{-1} = \frac{1}{\sigma_\varepsilon^2} \left[\left(\mathbf{I}_n - \frac{1}{n} \mathbf{J}_n \right) \otimes \mathbf{I}_t \right] + \frac{1}{\sigma_\varepsilon^2 + n\sigma_u^2} \left[\left(\frac{1}{n} \mathbf{J}_n \right) \otimes \mathbf{I}_t \right]$$

Using this representation of the inverse of the covariance matrix, the following can be obtained

$$|\boldsymbol{\Sigma}| = (\sigma_\varepsilon^2)^{t(n-1)} (\sigma_\varepsilon^2 + n\sigma_u^2)^t$$

or

$$\log_e |\boldsymbol{\Sigma}| = t(n-1) \log_e (\sigma_\varepsilon^2) + t \log_e (\sigma_\varepsilon^2 + n\sigma_u^2)$$

and

$$(\mathbf{y} - \mathbf{j}_{nt}\mu)' \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{j}_{nt}\mu) = \frac{nt(\bar{y}_{..} - \mu)^2}{\sigma_\varepsilon^2 + n\sigma_u^2} + \frac{SSE}{\sigma_\varepsilon^2} + \frac{SSU}{\sigma_\varepsilon^2 + n\sigma_u^2}$$

where

$$SSE = \sum_{i=1}^t \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2 \quad \text{and} \quad SSU = n \sum_{i=1}^t (\bar{y}_{i.} - \bar{y}_{..})^2$$

Using these expressions, $-2 \log_e(L(\mu, \sigma_u^2, \sigma_\epsilon^2 | \mathbf{y})$ can be written as

$$\begin{aligned} \ell(\mu, \sigma_u^2, \sigma_\epsilon^2 | \mathbf{y}) &= tn \log_e(2\pi) + t(n-1) \log_e(\sigma_\epsilon^2) + t \log_e(\sigma_\epsilon^2 + n\sigma_u^2) \\ &\quad + \frac{nt(\bar{y}_{..} - \mu)^2}{\sigma_\epsilon^2 + n\sigma_u^2} + \frac{SSE}{\sigma_\epsilon^2} + \frac{SSU}{\sigma_\epsilon^2 + n\sigma_u^2} \end{aligned}$$

The likelihood equations are obtained by differentiating $\ell(\mu, \sigma_u^2, \sigma_\epsilon^2 | \mathbf{y})$ with respect to the three parameters, μ , σ_u^2 and σ_ϵ^2 and then setting the derivatives equal zero. The derivatives evaluated at the solution for the parameters when set equal to zero are

$$\begin{aligned} \frac{\partial \ell(\mu, \sigma_u^2, \sigma_\epsilon^2 | \mathbf{y})}{\partial \mu} &= \frac{-2nt(\bar{y}_{..} - \tilde{\mu})}{\sigma_\epsilon^2 + n\sigma_u^2} = 0 \\ \frac{\partial \ell(\mu, \sigma_u^2, \sigma_\epsilon^2 | \mathbf{y})}{\partial \sigma_\epsilon^2} &= \frac{t(n-1)}{\tilde{\sigma}_\epsilon^2} + \frac{t}{\tilde{\sigma}_\epsilon^2 + n\tilde{\sigma}_u^2} - \frac{nt(\bar{y}_{..} - \tilde{\mu})^2}{(\tilde{\sigma}_\epsilon^2 + n\tilde{\sigma}_u^2)^2} - \frac{SSE}{(\tilde{\sigma}_\epsilon^2)^2} - \frac{SSU}{(\tilde{\sigma}_\epsilon^2 + n\tilde{\sigma}_u^2)^2} = 0 \\ \frac{\partial \ell(\mu, \sigma_u^2, \sigma_\epsilon^2 | \mathbf{y})}{\partial \sigma_u^2} &= \frac{nt}{\tilde{\sigma}_\epsilon^2 + n\tilde{\sigma}_u^2} - \frac{n^2t(\bar{y}_{..} - \tilde{\mu})^2}{(\tilde{\sigma}_\epsilon^2 + n\tilde{\sigma}_u^2)^2} - \frac{nSSU}{(\tilde{\sigma}_\epsilon^2 + n\tilde{\sigma}_u^2)^2} = 0 \end{aligned}$$

The solution to the maximum likelihood equations is

$$\tilde{\mu} = \bar{y}_{..}, \quad \tilde{\sigma}_\epsilon^2 = \frac{SSE}{t(n-1)} = MSEerror, \quad \text{and} \quad \tilde{\sigma}_u^2 = \frac{1}{n} \left[\frac{SSU}{t} - MSEerror \right]$$

Thus, the maximum likelihood estimates are

$$\hat{\mu} = \tilde{\mu} = \bar{y}_{..}, \quad \hat{\sigma}_\epsilon^2 = \tilde{\sigma}_\epsilon^2 = \frac{SSE}{t(n-1)} = MSEerror$$

and

$$\hat{\sigma}_u^2 = \begin{cases} \tilde{\sigma}_u^2 & \text{if } \tilde{\sigma}_u^2 \geq 0 \\ 0 & \text{if } \tilde{\sigma}_u^2 < 0 \end{cases}$$

When the estimate of σ_u^2 is zero, then the estimate of σ_ϵ^2 is recomputed by pooling SSE and SSU as well as their degrees of freedom to obtain

$$\hat{\sigma}_\epsilon^2 = \frac{SSE + SSU}{tn - 1}$$

If there happens to be a negative intraclass correlation, then the estimate of σ_ϵ^2 obtained by pooling will be an underestimate of the variance. A careful investigation of the assumptions and their appropriateness must be evaluated because if there is a degree of competition among experimental units within a level of u_i , then a negative correlation would be appropriate. In this case, the covariance matrix could be expressed as

$$\Sigma = \sigma_A^2 \rho [J_n \otimes I_t] + \sigma_A^2 (1 - \rho) [I_n \otimes I_t]$$

where $\sigma_\epsilon^2 = \sigma_A^2 (1 - \rho)$ and $\sigma_u^2 = \sigma_A^2 \rho$.

Several computational algorithms have been developed for maximizing the likelihood function, thus, providing maximum likelihood estimates of the model's parameters (values in the parameter space) (Hemmerle and Hartley, 1973; Corbeil and Searle, 1976). The large sample size variances of the maximum likelihood estimates can be obtained by inverting the matrix of second derivatives where the second derivatives are evaluated at the values of the maximum likelihood estimates. Maximum likelihood estimators and their variances have been obtained for several designed experiments and are reported in Searle (1971) and Searle et al. (1992).

When using computer software to fit these models, the experimenter should thoroughly investigate the algorithm being used and determine its properties, that is, whether it always yields meaningful estimates and whether it maximizes the likelihood function over the parameter space.

The maximum likelihood estimates of the variance components in Examples 19.2 and 19.3 were obtained by using SAS-Mixed. The maximum likelihood estimates for σ_ϵ^2 and σ_u^2 in Example 19.2 are $\hat{\sigma}_u^2 = 0.05749$ and $\hat{\sigma}_\epsilon^2 = 0.04855$, as displayed in Table 19.3. The maximum likelihood estimates for the parameters of the model in Example 19.3 are $\hat{\sigma}_\epsilon^2 = 3.8428$, $\hat{\sigma}_c^2 = 7.40708$, $\hat{\sigma}_a^2 = 0$, and $\hat{\sigma}_e^2 = 0$, as displayed in Table 19.4. The ML algorithm in SAS-Mixed does the maximization over the parameter space, as is shown by the values of $\hat{\sigma}_a^2$ and $\hat{\sigma}_e^2$ being set equal to zero.

TABLE 19.3

Proc Mixed Code to Compute Maximum Likelihood Estimates of the Variance Components and Mean for the Data in Example 19.2

```
proc mixed method=ml data=ex19_2 covest cl;
class variety;
model damage=/solution;
random variety;
```

Covariance Parameter	Estimate	Standard Error	Lower CL	Upper CL
Variety $\hat{\sigma}_{Var}^2$	0.04855	0.05075	0.01267	2.5700
Residual $\hat{\sigma}_\epsilon^2$	0.05749	0.02760	0.02690	0.1972
Estimate of Mean	Standard Error	df	t-Value	Pr > t
$\hat{\mu} = 3.9909$	0.1297	3	30.78	<0.0001

TABLE 19.4

Proc Mixed Code to Obtain Maximum Likelihood Estimates of the Variance Components for the Two-Way Random Effects Model in Example 19.3

```
Proc Mixed data=ex19_3 covtest cl method=ML;
title2 "Using Maximum Likelihood";
class row col;
model y=/solution;
random row col row*col;
```

Covariance Parameter	Estimate	Standard Error	Lower	Upper
Row	0	—	—	—
Col	0	—	—	—
Row × col	8.5604	5.9559	3.1104	67.2317
Residual	3.5989	1.8071	1.6376	13.3052
Estimate of Mean	Standard Error	df	t Value	Pr > t
14.6910	1.3008	1	11.29	0.0562

19.3 Restricted or Residual Maximum Likelihood Estimation

Restricted or residual maximum likelihood estimates are obtained by maximizing that part of the likelihood function that does not include any fixed effects or by maximizing the likelihood function of the residuals after the fixed effects have been removed from the model. For the models in this chapter there is just one fixed effect parameter, μ . This is also equivalent to looking at the conditional distribution of a set of sums of squares given the overall sample mean. The process is accomplished by factoring the likelihood function into parts where one part involves fixed effect parameters and a second part just involves the variance components. The REML equations are obtained by differentiating $-2\log_e$ of the residual likelihood with respect to the variance components and setting them equal to zero.

For the general random effects model (19.1), the likelihood function of the observations is

$$L(\mu, \sigma_\varepsilon^2, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2 | \mathbf{y}) = (2\pi)^{-n/2} |\boldsymbol{\Sigma}|^{-1/2} e^{[-(1/2)(\mathbf{y} - \mathbf{j}_n\mu)' \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{j}_n\mu)]}$$

and the $-2 \log_e$ of the likelihood function is

$$\begin{aligned} \ell(\mu, \sigma_\varepsilon^2, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2 | \mathbf{y}) &= -2 \log_e [L(\mu, \sigma_\varepsilon^2, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2 | \mathbf{y})] \\ &= n \log_e (2\pi) + \log_e (|\boldsymbol{\Sigma}|) + (\mathbf{y} - \mathbf{j}_n\mu)' \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{j}_n\mu) \\ &= \ell(\mu, \sigma_\varepsilon^2, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2 | \bar{y}) \\ &\quad + \ell(\sigma_\varepsilon^2, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2 | SSE, SSU_1, SSU_2, \dots, SSU_k) \end{aligned}$$

where $SSE, SSU_1, SSU_2, \dots, SSU_k$ denote independent set of sums of squares that do not depend on μ and $\ell(\sigma_\varepsilon^2, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2 | SSE, SSU_1, SSU_2, \dots, SSU_k)$ is the residual likelihood function. The solution to the residual likelihood equations provides the REML estimates of the variance components.

19.3.1 Example 19.5: REML Solution for Balanced One-Way Model

Using the balanced one-way random effects model described in Example 19.4, the $-2\log_e$ of the likelihood function can be expressed as functions of the sufficient statistics $\bar{y}_{..}$, SSE and SSU as

$$\begin{aligned}\ell(\mu, \sigma_u^2, \sigma_\epsilon^2 | \mathbf{y}) &= tn \log_e(2\pi) + t(n-1) \log_e(\sigma_\epsilon^2) + t \log_e(\sigma_\epsilon^2 + n\sigma_u^2) \\ &\quad + \frac{nt(\bar{y}_{..} - \mu)^2}{\sigma_\epsilon^2 + n\sigma_u^2} + \frac{SSE}{\sigma_\epsilon^2} + \frac{SSU}{\sigma_\epsilon^2 + n\sigma_u^2} \\ &= \left[\frac{nt(\bar{y}_{..} - \mu)^2}{\sigma_\epsilon^2 + n\sigma_u^2} + \log_e(2\pi) + \log_e(\sigma_\epsilon^2 + n\sigma_u^2) \right] \\ &\quad + \left[(tn-1)\log_e(2\pi) + t(n-1)\log_e(\sigma_\epsilon^2) + (t-1)\log_e(\sigma_\epsilon^2 + n\sigma_u^2) + \frac{SSE}{\sigma_\epsilon^2} + \frac{SSU}{\sigma_\epsilon^2 + n\sigma_u^2} \right] \\ &= \ell(\mu, \sigma_u^2, \sigma_\epsilon^2 | \bar{y}_{..}) + \ell(\sigma_u^2, \sigma_\epsilon^2 | SSE, SSU)\end{aligned}$$

The residual likelihood function is $\ell(\sigma_u^2, \sigma_\epsilon^2 | SSE, SSU)$ where

$$\begin{aligned}\ell(\sigma_u^2, \sigma_\epsilon^2 | SSE, SSU) &= \left[(tn-1)\log_e(2\pi) + t(n-1)\log_e(\sigma_\epsilon^2) \right. \\ &\quad \left. + (t-1)\log_e(\sigma_\epsilon^2 + n\sigma_u^2) + \frac{SSE}{\sigma_\epsilon^2} + \frac{SSU}{\sigma_\epsilon^2 + n\sigma_u^2} \right]\end{aligned}$$

The restricted maximum likelihood equations are obtained by differentiating $\ell(\sigma_u^2, \sigma_\epsilon^2 | SSE, SSU)$ with respect to the two parameters σ_u^2 and σ_ϵ^2 and then setting the derivatives equal zero. The derivatives, evaluated at the solution for the parameters, set equal to zero are

$$\begin{aligned}\frac{\partial \ell(\sigma_u^2, \sigma_\epsilon^2 | SSE, SSU)}{\partial \sigma_\epsilon^2} &= \frac{t(n-1)}{\tilde{\sigma}_\epsilon^2} + \frac{t-1}{\tilde{\sigma}_\epsilon^2 + n\tilde{\sigma}_u^2} - \frac{SSE}{(\tilde{\sigma}_\epsilon^2)^2} - \frac{SSU}{(\tilde{\sigma}_\epsilon^2 + n\tilde{\sigma}_u^2)^2} = 0 \\ \frac{\partial \ell(\sigma_u^2, \sigma_\epsilon^2 | SSE, SSU)}{\partial \sigma_u^2} &= \frac{n(t-1)}{\tilde{\sigma}_\epsilon^2 + n\tilde{\sigma}_u^2} - \frac{nSSU}{(\tilde{\sigma}_\epsilon^2 + n\tilde{\sigma}_u^2)^2} = 0\end{aligned}$$

The solution to the residual maximum likelihood equations is

$$\tilde{\sigma}_\epsilon^2 = \frac{SSE}{t(n-1)} = MSError, \quad \text{and} \quad \tilde{\sigma}_u^2 = \frac{1}{n} \left[\frac{SSU}{t-1} - MSError \right] = \frac{1}{n} [MSU - MSError]$$

The residual maximum likelihood estimates are

$$\hat{\sigma}_\epsilon^2 = \tilde{\sigma}_\epsilon^2 = \frac{SSE}{t(n-1)} = MSError$$

and

$$\hat{\sigma}_u^2 = \begin{cases} \tilde{\sigma}_u^2 & \text{if } \tilde{\sigma}_u^2 \geq 0 \\ 0 & \text{if } \tilde{\sigma}_u^2 < 0 \end{cases}$$

When the estimate of σ_u^2 is zero, then the estimate of σ_e^2 is recomputed by pooling SSE and SSU as well as their degrees of freedom to obtain

$$\hat{\sigma}_e^2 = \frac{SSE + SSU}{tn - 1}$$

Tables 19.5 and 19.6 contain the SAS-Mixed Code and results for extracting REML estimates of the variance components for the data sets in Examples 19.2 and 19.3, respectively.

TABLE 19.5

Proc Mixed Code to Compute Restricted Maximum Likelihood Estimates of the Variance Components and Mean for the Data in Example 19.2

```
proc mixed method=reml data=ex19_2 covtest cl;
class variety;
model damage=/solution;
random variety;
```

Covariance Parameter	Estimate	Standard Error	Lower CL	Upper CL
Variety $\hat{\sigma}_{\text{Var}}^2$	0.07316	0.07802	0.01875	4.4672
Residual $\hat{\sigma}_e^2$	0.05700	0.02713	0.02681	0.1929
Estimate of Mean	Standard Error	df	t-Value	Pr > t
$\hat{\mu} = 3.9863$	0.1515	3	26.31	0.0001

TABLE 19.6

Proc Mixed Code to Obtain Restricted Maximum Likelihood Estimates of the Variance Components for the Two-Way Random Effects Model in Example 19.3

```
Proc Mixed data=ex19_3 covtest cl method=REML;
title2 "Using Maximum Likelihood";
class row col;
model y=/solution;
random row col row*col;
```

Covariance Parameter	Estimate	Standard Error	Lower	Upper
Row	0	—	—	—
Col	0	—	—	—
Row × col	9.2425	6.9923	3.1503	95.7334
Residual	3.8398	1.9228	1.7502	14.1289
Estimate of Mean	Standard Error	df	t-Value	Pr > t
$\hat{\mu} = 14.8547$	1.3503	1	11.00	0.0577

19.4 MIVQUE Method

Rao (1971) described a general procedure for obtaining minimum variance quadratic unbiased estimators (MIVQUE) of the variance components. For the general random effects model of Section 18.2, the MIVQUE of a linear combination of variances

$$\theta = C_0\sigma_\epsilon^2 + C_1\sigma_1^2 + C_2\sigma_2^2 + \dots + C_k\sigma_k^2$$

is a quadratic function of the observations which is unbiased for θ and has minimum variance within the class of quadratic unbiased estimators of θ . Thus, MIVQUE estimators of variance components possess the minimum variance property, whereas the method of moments estimators generally do not. Each individual variance component can be selected as a possible parameter to be estimated. Selecting $C_0=1$, $C_1=C_2=\dots=C_k=0$ provides $\theta=\sigma_\epsilon^2$. Other choices of C_i values will provide $\theta=\sigma_i^2$ as well as other linear combinations of the variance components.

19.4.1 Description of the Method

The estimate of θ must be a quadratic function of \mathbf{y} , thus, for some matrix A , the estimator of θ has the form $\mathbf{y}'A\mathbf{y}$. The expectation of $\mathbf{y}'A\mathbf{y}$ is

$$E(\mathbf{y}'A\mathbf{y}) = \text{tr}(\Sigma A) + \mu^2 j'_n A j_n$$

By assumption, $E(\mathbf{y}'A\mathbf{y}) = \theta$. Since the expectation does not depend on μ , A must be chosen to satisfy $\mu^2 j'_n A j_n = 0$. Under the conditions of normality, the variance of $\mathbf{y}'A\mathbf{y}$ when $\mu^2 j'_n A j_n = 0$ is

$$\text{Var}(\mathbf{y}'A\mathbf{y}) = 2 \text{tr}[\Sigma A]^2$$

Thus the MIVQUE of θ is $\mathbf{y}'A\mathbf{y}$ where A is chosen such that $\text{tr}[\Sigma A] = \theta$ and $\text{tr}[\Sigma A]^2$ is minimized over the parameter space

$$\{0 < \sigma_\epsilon^2 < \infty, 0 < \sigma_1^2 < \infty, 0 < \sigma_2^2 < \infty, \dots, 0 < \sigma_k^2 < \infty\}$$

Rao (1971) shows that the MIVQUE of $\sigma^2 = [\sigma_\epsilon^2, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2]'$ is $\hat{\sigma}^2 = S^{-1}\mathbf{f}$ where S is a $(k+1) \times (k+1)$ matrix with elements

$$s_{ii'} = \text{tr}[\mathbf{X}_i \mathbf{X}_{i'}' \mathbf{R} \mathbf{X}_{i'} \mathbf{X}_i'] \quad i, i' = 0, 1, 2, \dots, k$$

and \mathbf{f} is a $(k+1) \times 1$ vector with elements

$$f_i = \mathbf{y}' \mathbf{R} \mathbf{X}_i \mathbf{X}_i' \mathbf{R} \mathbf{y}, \quad i = 0, 1, 2, \dots, k$$

and

$$\mathbf{R} = \boldsymbol{\Sigma}^{-1} [\mathbf{I}_n - \mathbf{j}_n (\mathbf{j}'_n \boldsymbol{\Sigma}^{-1} \mathbf{j}_n)^{-1} \mathbf{j}'_n] \boldsymbol{\Sigma}^{-1}$$

The solution for $\hat{\sigma}^2$ depends on the elements of $\boldsymbol{\Sigma}$ that are functions of the unknown variance components. In order to compute the MIVQUE of σ^2 , some constants must be substituted into $\boldsymbol{\Sigma}$ for $\sigma_\varepsilon^2, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$. For that set of constants, the estimate of σ^2 is MIVQUE (and is a quadratic function of \mathbf{y}). In order for $\hat{\sigma}^2$ to be MIVQUE of σ^2 , the elements of \mathbf{R} must not depend on the data vector. Some software uses 1 as the value of the residual variance and 0 as the values of each of the other variances and covariances.

Usually it is best to substitute values for $\sigma_\varepsilon^2, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$ into $\boldsymbol{\Sigma}$ that are close to the true values of the parameters. One possible procedure is to obtain values from other experiments. The process of using the fixed values (not dependent on the current data) as starting values for σ^2 in a non iterative process or using zero iterations provides a solution that is called the MIVQUE0 solution where 0 denotes that no iterations have been performed. Swallow and Monahan (1984) used the method of moments estimates as initial values for the variances and used MIVQUE-A to describe the resulting estimators of the variance components. Another method is to use an iterative procedure (Brown, 1976) by using some initial values of the variance components, say $\sigma_{\varepsilon 0}^2, \sigma_{10}^2, \sigma_{20}^2, \dots, \sigma_{k0}^2$, to start the process. Use those initial values to evaluate $\boldsymbol{\Sigma}$ and obtain $\hat{\sigma}_{(0)}^2$. Here $\hat{\sigma}_{(0)}^2$ depends on the values chosen for σ_0^2 . Then use $\hat{\sigma}_{(0)}^2$ to evaluate $\boldsymbol{\Sigma}$ to obtain the second iteration estimate, $\hat{\sigma}_{(1)}^2$. Continue the iteration process until there is very little change from one iteration to the next. The resulting iterative MIVQUE values are no longer quadratic functions of \mathbf{y} , since the elements of $\boldsymbol{\Sigma}$ are functions of \mathbf{y} . The final estimator of σ_i^2 , say at step $m + 1$, can be called MIVQUE given the previous values of $\hat{\sigma}_{(m)}^2$. For balanced models, Swallow and Searle (1978) have shown that the equations simplify so that an explicit solution can be obtained. The solutions are identical to those provided by the method of moments. When there are unequal sample sizes and/or empty cells, the iterative procedure could be an appropriate method. A simulation study by Swallow and Searle (1978) indicates that REML, ML and method of moments provide better estimates of the variance components than MIVQUE0.

The values of the MIVQUEs (either evaluated at constants for σ_i^2 , or at a given previous step) are linear combinations of quadratic forms of \mathbf{y} . Thus, the variance can be evaluated since the variance of a quadratic form $\mathbf{y}'B\mathbf{y}$ is $2\text{tr}(B\boldsymbol{\Sigma})^2$. Swallow and Searle (1978) show how to use the expressions to obtain variances for the estimators in an unbalanced one-way model. They computed and compared the variances of the MIVQUE and method of moments estimators for the unbalanced one-way model with varying numbers of populations, sample sizes, and values of the variances. The variances of the MIVQUE estimators were evaluated as if the true values of σ_u^2 and σ_ε^2 were used in the estimation process. The estimate of σ_ε^2 obtained from the method of moments was quite comparable to that from the MIVQUE method where the variance of the MIVQUE was no more than 4% smaller than the variance of the method-of-moments estimator.

For fairly balanced models (n_i not too different), the variance of the MIVQUE of σ_u^2 was no more than 10% smaller than the variance of the method of moments estimator of σ_u^2 . For many unbalanced sample sizes, the variance of the MIVQUE of σ_u^2 was as much as 60% smaller than the corresponding method-of-moments estimator. The variances of the estimators from the two methods would be much more similar if values other than the true σ_u^2 and σ_ε^2 are used in the estimation process. SAS-Mixed has a MIVQUE option, denoted by MIVQUE0, which is a noniterative method.

This section concludes with an example of an unbalanced one-way design using MIVQUE estimators.

19.4.2 Application. Example 19.6: MIVQUE for the Unbalanced One-Way Design

The equations from Swallow and Searle (1978) are presented for a general one-way model and then they are then applied to the data in Example 19.2. The model is

$$y_{ij} = \mu + u_i + \varepsilon_{ij} \quad i = 1, 2, \dots, t \text{ and } j = 1, 2, \dots, n_i$$

where $u \sim N(0, \sigma_u^2 I_t)$, $\varepsilon \sim N(0, \sigma_\varepsilon^2 I_N)$, and where u and ε are independent random variables with $N = \sum_{i=1}^t n_i$.

Define

$$k_i = \frac{n_i}{\sigma_{\varepsilon 0}^2 + n_i \sigma_{u0}^2} \quad \text{and} \quad K = \frac{1}{\sum_{i=1}^t k_i}$$

The elements of the matrix S are

$$\begin{aligned} s_{11} &= \sum_{i=1}^t k_i^2 - 2K \sum_{i=1}^t k_i^3 + K^2 \left(\sum_{i=1}^t k_i^2 \right)^2 \\ s_{12} &= \sum_{i=1}^t \frac{k_i^2}{n_i} - 2K \sum_{i=1}^t \frac{k_i^3}{n_i} + K^2 \left(\sum_{i=1}^t k_i^2 \right) \left(\sum_{i=1}^t \frac{k_i^2}{n_i} \right) \end{aligned}$$

and

$$s_{22} = \frac{N-t}{\sigma_{\varepsilon 0}^4} + \sum_{i=1}^t \frac{k_i^2}{n_i^2} - 2K \sum_{i=1}^t \frac{k_i^3}{n_i^2} + K^2 \left(\sum_{i=1}^t \frac{k_i^2}{n_i} \right)^2$$

The elements of the vector f are

$$f_1 = \sum_{i=1}^t k_i^2 \left(\bar{y}_{i \cdot} - K \sum_{i=1}^t k_i \bar{y}_{i \cdot} \right)^2$$

and

$$f_2 = \frac{\sum_{i=1}^t \sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^t n_i \bar{y}_{i \cdot}^2}{\sigma_{\varepsilon 0}^4} + \sum_{i=1}^t \frac{k_i^2 \left(\bar{y}_{i \cdot} - K \sum_{i=1}^t k_i \bar{y}_{i \cdot} \right)^2}{n_i}$$

For given values of σ_{u0}^2 and $\sigma_{\varepsilon 0}^2$, the MIVQUE estimators of σ_u^2 and σ_ε^2 are $\hat{\sigma}^2 = S^{-1} f$, or

$$\hat{\sigma}_\varepsilon^2 = \frac{s_{11} f_2 - s_{12} f_1}{c}$$

and

$$\hat{\sigma}_u^2 = \frac{s_{22} f_1 - s_{12} f_2}{c}$$

TABLE 19.7

Proc Mixed Code to Compute MIVQUE0 Estimates of the Variance Components and Mean for the Data in Example 19.2

```
proc mixed method=mivque0 data=ex19_2 covtest cl;
class variety;
model damage=/solution;
random variety;
```

Covariance Parameter	Estimate	Standard Error	Lower CL	Upper CL
Variety $\hat{\sigma}_{\text{Var}}^2$	0.05638	0.05739	0.01505	2.4965
Residual $\hat{\sigma}_{\varepsilon}^2$	0.06503	0.03109	0.03050	0.2216
Estimate of Mean	Standard Error	df	t-Value	Pr > t
$\hat{\mu} = 3.9906$	0.1392	3	28.67	>0.0001

TABLE 19.8

Proc Mixed Code to Obtain MIVQUE0 Estimates of the Variance Components for the Two-Way Random Effects Model in Example 19.3

```
Proc Mixed data=ex19_3 covtest cl method=MIVQUE0;
title2 "Using Maximum Likelihood";
class row col;
model y=/solution;
random row col row*col;
```

Covariance Parameter	Estimate	Standard Error	Lower	Upper
Row	0	—	—	—
Col	0	—	—	—
Row \times col	25.4682	9.9894	13.3850	66.1016
Residual	4.0023	1.5698	2.1035	10.3879
Estimate of Mean	Standard Error	df	t-Value	Pr > t
14.7094	2.1309	1	6.90	0.0916

where $c = s_{11} s_{22} - s_{12}^2$. The MIVQUE estimators of the two variance components for the data in Table 19.1 are in Table 19.7 and of the four variance components for the data in Table 19.2 are in Table 19.8. These estimators were obtained using the non iterative solution from SAS-Mixed. The variances of the estimators are

$$\text{Var}(\hat{\sigma}_{\varepsilon}^2) = \frac{2s_{11}}{c}, \quad \text{Var}(\hat{\sigma}_u^2) = \frac{2s_{22}}{c}, \quad \text{and} \quad \text{Cov}(\hat{\sigma}_{\varepsilon}^2, \hat{\sigma}_u^2) = \frac{-2s_{12}}{c}$$

The estimators are not very sensitive to the choice of σ_{u0}^2 and $\sigma_{\varepsilon 0}^2$, but the variances do depend on the choice of the initial values of σ_{u0}^2 and $\sigma_{\varepsilon 0}^2$. Table 19.9 contains the MIVQUE0 estimators and their variances for the data in Table 19.1 using the starting values listed. The variances of the estimators vary greatly and are extremely large when the starting values for σ_{u0}^2 and $\sigma_{\varepsilon 0}^2$ are far from the estimators of $\hat{\sigma}_u^2$ and $\hat{\sigma}_{\varepsilon}^2$. If an iterative procedure is used, the solution converges to $\hat{\sigma}_{\varepsilon}^2 = 0.057003$ and $\hat{\sigma}_u^2 = 0.073155$ with $\text{Var}(\hat{\sigma}_{\varepsilon}^2) = 0.000721$, $\text{Var}(\hat{\sigma}_u^2) = 0.005694$ and $\text{cov}(\hat{\sigma}_{\varepsilon}^2, \hat{\sigma}_u^2) = -0.000235$. The iterative procedure was started at several values ($\sigma_{u0}^2 = 2$ and $\sigma_{\varepsilon 0}^2 = 1$, and $\sigma_{u0}^2 = 50$ and $\sigma_{\varepsilon 0}^2 = 1000$ among others). All choices for starting values converged to the above values in four iterations. After two iterations, the estimators were quite stable but the variances were still changing.

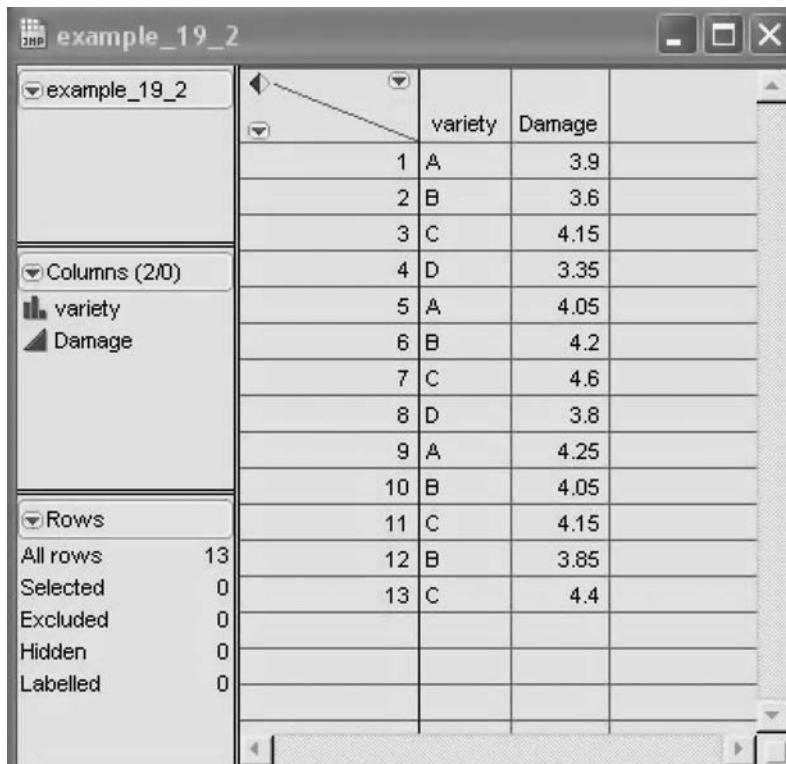
TABLE 19.9

MIVQUE0 Estimates of Variance Components for Example 19.2 for Selected Starting Values

σ_{e0}^2	σ_{u0}^2	σ_e^2	σ_u^2	$\text{Var}(\sigma_e^2)$	$\text{Var}(\sigma_u^2)$	$\text{Cov}(\sigma_e^2, \sigma_u^2)$
1.00000	0.00000	0.05648	0.05893	0.18988	0.07966	-0.05677
1.00000	2.00000	0.05655	0.07495	0.22150	3.63984	-0.07255
10.00000	20.00000	0.05655	0.07495	22.14991	363.98420	-7.25498
1.00000	5.00000	0.05646	0.07718	0.22208	18.97352	-0.07346
0.10000	0.50000	0.05646	0.07718	0.00222	0.18974	-0.00073
0.05644	0.06719	0.05670	0.07294	0.00070	0.00496	-0.00023
0.05700	0.07316	0.05667	0.07326	0.00072	0.00569	-0.00023
0.06503	0.05638	0.05683	0.07145	0.00093	0.00410	-0.00030

19.5 Estimating Variance Components Using JMP

The estimates of the variance components can be obtained using the fit model option of the JMP software (SAS Institute, Inc., 2005). Figure 19.1 gives the data set for Example 19.2 displayed in a JMP data table (which was imported from a SAS data set). On the Analyze menu, select fit model as shown in Figure 19.2. On the fit model screen, select damage to be

**FIGURE 19.1** Data set for Example 19.2 in JMP table.

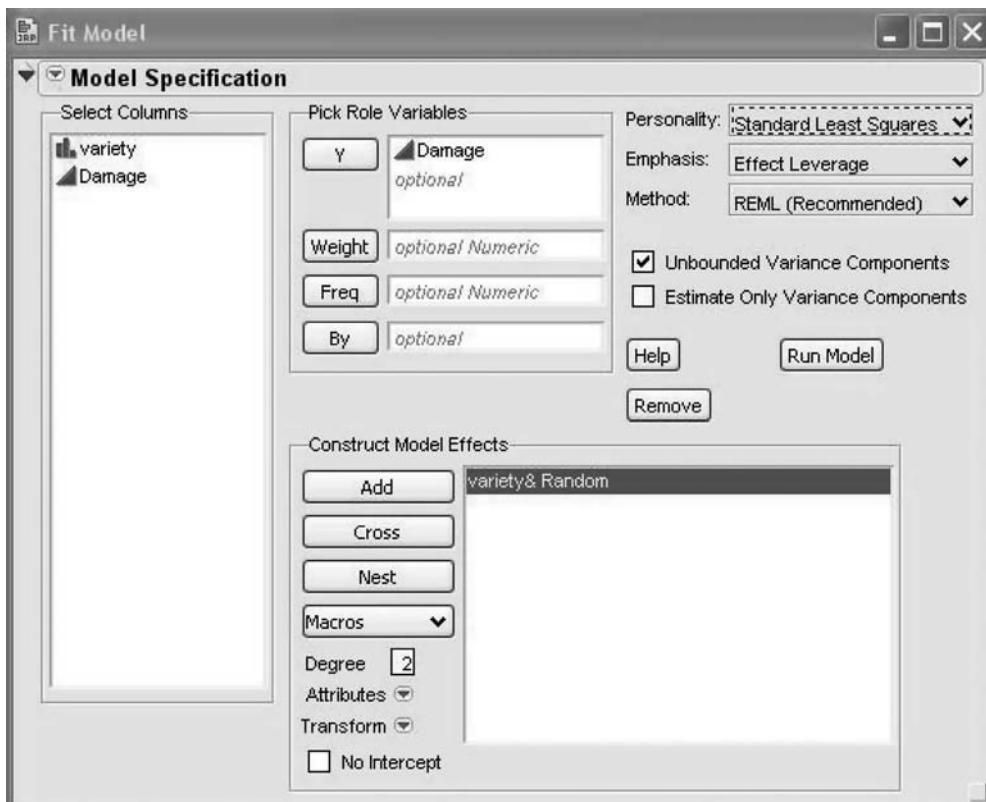


FIGURE 19.2 JMP fit model table for Example 19.2.

the Y variable and select *variety* to be a model effect. Use the attributes menu to specify that *variety* is a random effect. The default estimation method is REML, but EMS can be selected to provide method of moments estimators using type 3 sums of squares. Click the run model button to obtain the results in Figure 19.3. The estimates of the variance components and their estimated standard errors are similar to those from SAS in Table 19.5. The main difference is that the confidence interval about the *variety* variance component is computed using the Wald method instead of the Satterthwaite approximation (see Chapter 20 for details).

Figure 19.4 is the JMP data table for the data of Example 19.3. The fit model screen is shown in Figure 19.5 where row, col and row \times col have been selected as random effects and the REML method is selected for estimation. The REML estimates of the variance components are shown in Figure 19.6, where the results are similar to those from SAS in Table 19.6. The fit model screen has another option that can be used where one can check the unbounded variance components box. This option does not restrict the solution to be in the parameter space (similar to using the unbounded option in SAS-Mixed). Figure 19.7 has the unbounded variance components box checked and the solution is shown in Figure 19.8. The solutions for the row and col variance components are negative and the Wald method is used to compute the confidence intervals (except for Residual). The JMP fit model process provides an appropriate analysis for models with random effects as does SAS-Mixed.

example_19_2- Fit Least Squares

- ▼ Response Damage
- ▼ Whole Model
 - ▶ Actual by Predicted Plot
 - ▶ Summary of Fit
 - ▼ Parameter Estimates

Term	Estimate	Std Error	DFDen	t Ratio	Prob> t
Intercept	3.9862953	0.151998	2.806	26.23	0.0002*
 - ▶ Random Effect Predictions
 - ▼ REML Variance Component Estimates

Random Effect	Var Ratio	Var Component	Std Error	95% Lower	95% Upper	Pct of Total
variety	1.2833611	0.0731554	0.0780195	-0.079763	0.2260736	56.205
Residual		0.057003	0.0271325	0.0268112	0.1928944	43.795
Total		0.1301584				100.000

-2 LogLikelihood = 7.066573344

FIGURE 19.3 Results from REML from JMP for Example 19.2.

example_19_3

	row	col	y
1	1	1	10
2	1	1	12
3	1	1	11
4	1	2	13
5	1	2	15
6	1	3	21
7	1	3	19
8	2	1	16
9	2	1	18
10	2	2	13
11	2	2	19
12	2	2	14
13	2	3	11
14	2	3	13

▼ Columns (3/1)

- row
- col
- y

▼ Rows

All rows	14
Selected	0
Excluded	0

FIGURE 19.4 JMP data table for Example 19.3.

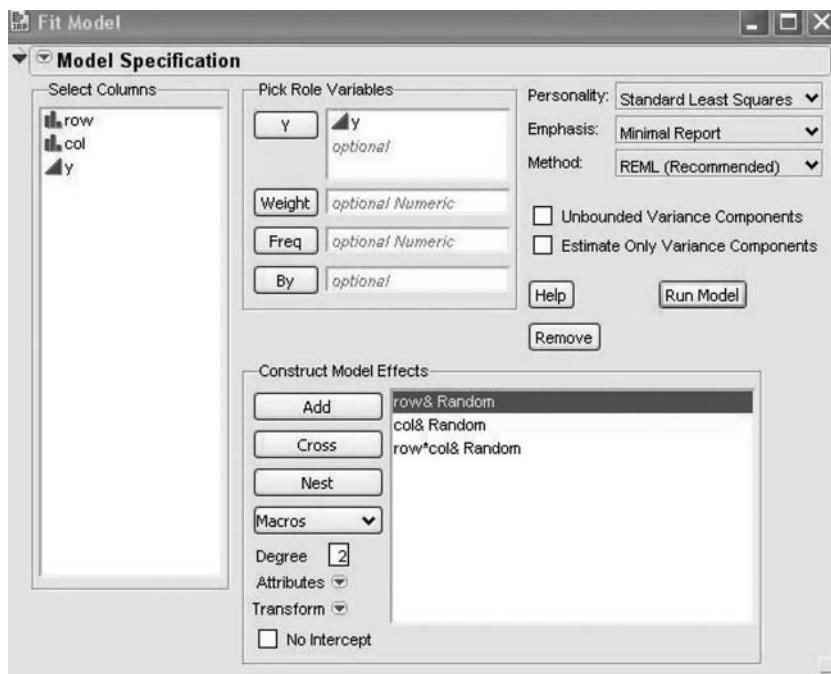


FIGURE 19.5 Fit model screen for JMP for REML estimates for Example 19.3.

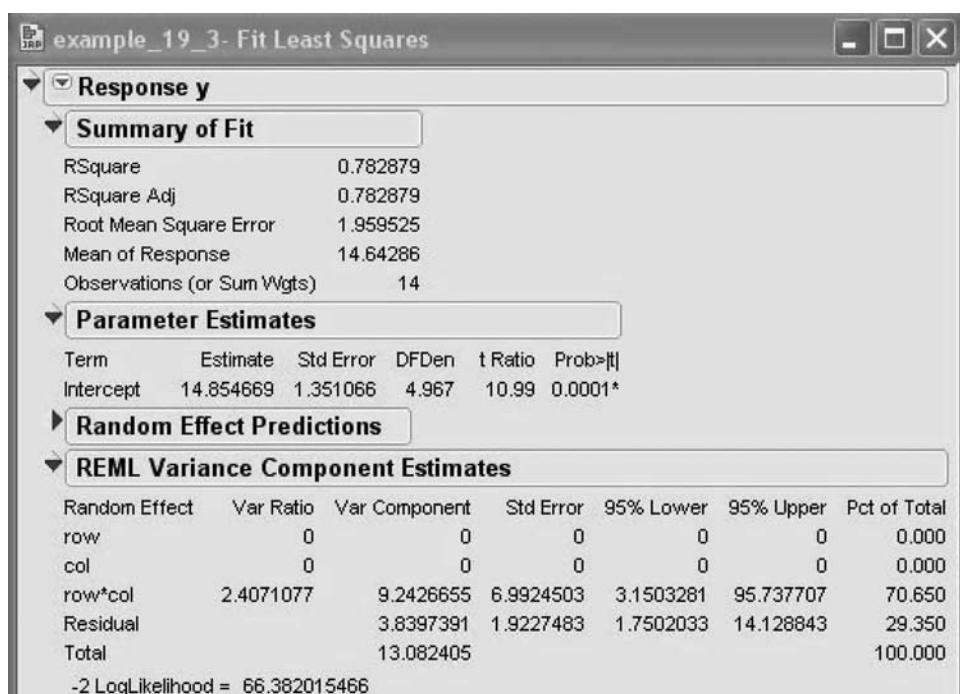


FIGURE 19.6 Results from JMP using REML for Example 19.3.

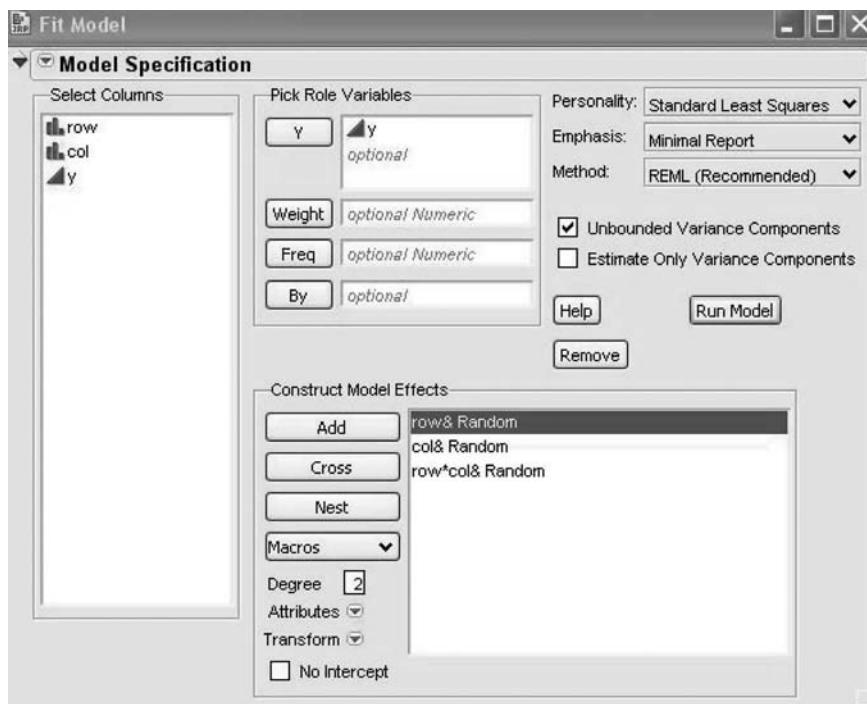


FIGURE 19.7 JMP fit model screen with unbounded variance components selected for Example 19.3.

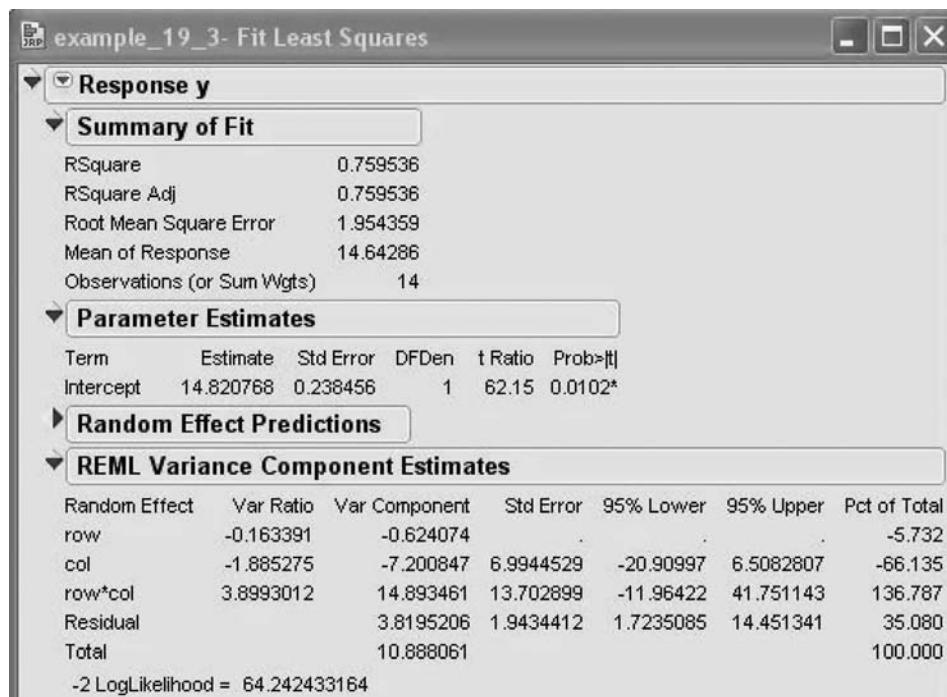


FIGURE 19.8 Unbounded results from REML of JMP for Example 19.3.

19.6 Concluding Remarks

This chapter presented four methods, the method of moments, the maximum likelihood method, the residual maximum likelihood, and the MIVQUE method, for obtaining estimates of the variance components of a random effects model. Two examples, a one-way random effects model and a two-way random effects model, were used to demonstrate each of the methods of estimation. When the data are balanced and the solution provides positive values, the estimates from the REML, MIVQUE0 and method of moments are identical. When the data sets are unbalanced, then each method will produce different estimates. SAS-MIXED and JMP were used to carry out the computations.

19.7 Exercises

- 19.1 The data in the following table are prices of coffee from five randomly selected states from the United States where four, five, or six cities were randomly selected per state and four to six stores per city were randomly selected. The price of brand x coffee was determined at each store. Write out a model to describe the data with states, cities nested within states and stores nested within cities and states. Provide REML, ML, MIVQUE0, and method of moments estimates of the variance components for state, city and store using type I, type II, and type III sums of squares. Also obtain the estimate of the mean price of brand x coffee for the United States.

Coffee Price Data for Exercise 19.1

State	City	Store 1	Store 2	Store 3	Store 4	Store 5	Store 6
1	1	2.78	2.78	2.80	—	—	—
1	2	2.93	2.89	2.91	2.90	—	—
1	3	2.73	2.70	2.74	2.74	2.72	2.74
1	4	2.93	2.92	2.93	2.89	2.93	2.92
1	5	2.94	2.93	2.94	—	—	—
2	1	2.20	2.20	2.16	2.18	—	—
2	2	2.11	2.08	2.11	2.09	—	—
2	3	2.02	2.03	2.06	—	—	—
2	4	1.98	1.96	2.02	1.98	1.99	—
2	5	2.05	2.03	2.04	2.09	2.02	—
2	6	2.10	2.12	2.09	2.09	—	—
3	1	2.18	2.20	2.22	2.22	—	—
3	2	1.97	2.00	1.98	1.99	1.97	—
3	3	2.11	2.12	2.13	2.13	2.10	—
3	4	2.10	2.08	2.07	2.08	—	—
4	3	2.35	2.37	2.40	2.36	2.37	2.40
4	4	2.34	2.41	2.33	2.32	2.32	—

Continued

Coffee Price Data for Exercise 19.1

State	City	Store 1	Store 2	Store 3	Store 4	Store 5	Store 6
4	3	2.35	2.37	2.40	2.36	2.37	2.40
4	4	2.34	2.41	2.33	2.32	2.32	—
4	5	2.34	2.33	2.38	2.34	—	—
5	1	2.24	2.21	2.24	2.19	—	—
5	2	2.23	2.17	2.18	2.18	2.17	—
5	3	2.22	2.21	2.21	2.20	2.21	—
5	4	2.24	2.27	2.26	—	—	—

- 19.2 The data in the following table are from a two-way random effects treatment structure. Write out an appropriate model and provide ML, REML, MIVQUE0, and method of moments estimates of the variance components for rows, columns, row \times column interaction and residual using type I, type II, and type III sums of squares.

Data for Exercise 19.2

Row Treatment	Column Treatment			
	1	2	3	4
1	29	—	30	—
	—	—	29	—
	—	—	31	—
	—	—	29	—
2	22	—	—	17
	—	—	—	16
3	34	28	—	25
	—	26	—	24
	—	26	—	—
	—	26	—	—
4	—	30	—	19
	—	30	—	19
	—	29	—	—
	—	30	—	—
5	22	22	19	24
	22	—	20	21
	19	—	19	21

20

Methods for Making Inferences about Variance Components

When a researcher designs an experiment involving factors that are random effects, she often wishes to make inferences about specific variance components specified in the model. In particular, if σ_u^2 is the variance component corresponding the distribution of the levels of factor A , the experimenter may wish to determine if there is enough evidence to conclude that $\sigma_u^2 > 0$. An appropriate decision can be made by 1) testing the hypothesis $H_0: \sigma_u^2 = 0$ vs $H_a: \sigma_u^2 > 0$; 2) by constructing a confidence interval about σ_u^2 ; or 3) by constructing a lower confidence limit for σ_u^2 . This chapter addresses these kinds of inference procedures for random effects models where methods for hypotheses testing are described in Section 20.1 and the construction of confidence intervals (lower bounds) is described in Section 20.2. The construction of confidence intervals for variance components has been a fertile area of research and many authors have developed specialized confidence intervals for specific functions of the variance component parameters. The methods described in this chapter are available in current software and some are available for specific problems. The discussion is not exhaustive, but rather points to the types of confidence intervals that have been addressed. A more complete discussion is available in Burdick and Graybill (1992) as well as papers in the current statistical journals.

20.1 Testing Hypotheses

There are two basic techniques for testing hypotheses about variance components. The first technique uses sums of squares from the analysis of variance table to construct F -statistics. For most balanced models, the F -statistics are distributed exactly as F -distributions, whereas for unbalanced models the distributions are approximated by F -distributions with the approximations becoming poorer as the designs become more unbalanced. The second technique is based on a likelihood-ratio test which is asymptotically distributed as a chi-square distribution. For balanced designs, the F -statistic approach is probably better than

the likelihood-ratio test, while for unbalanced designs there is no clear-cut choice. The reader may wish to carry out a simulation experiment to study the distributions of test statistics using a data structure similar to the data set of interest before making a decision as to which method to use to test the hypothesis of interest.

20.1.1 Using the Analysis of Variance Table

If the data set is balanced, then sums of squares obtained by the usual analysis of variance are independently distributed as scalar multiples of chi-square random variables. Let Q denote a sum of squares based on v degrees of freedom where its expected mean square is a function of four variance components. That is, suppose

$$E(Q/v) = \sigma_e^2 + k_1\sigma_1^2 + k_2\sigma_2^2 + k_3\sigma_3^2$$

Then, assuming the data follow a normal distribution,

$$W = \frac{Q}{\sigma_e^2 + k_1\sigma_1^2 + k_2\sigma_2^2 + k_3\sigma_3^2}$$

is often distributed as a chi-square random variable with v degrees of freedom. For many hypotheses of the form $H_0: \sigma_1^2 = 0$ vs $H_a: \sigma_1^2 > 0$, there are two independent sums of squares, denoted by Q_1 and Q_2 based on v_1 and v_2 degrees of freedom, respectively, with expectations

$$E(Q_1/v_1) = \sigma_e^2 + k_1\sigma_1^2 + k_2\sigma_2^2 + k_3\sigma_3^2$$

and

$$E(Q_2/v_2) = \sigma_e^2 + k_2\sigma_2^2 + k_3\sigma_3^2$$

The hypothesis $H_0: \sigma_1^2 = 0$ vs $H_a: \sigma_1^2 > 0$ is equivalent to

$$H_0: E(Q_1/v_1) = E(Q_2/v_2) \text{ vs } H_a: E(Q_1/v_1) > E(Q_2/v_2)$$

The statistic used to test this hypothesis is $F = (Q_1/v_1)/(Q_2/v_2)$ which, under the conditions of H_0 , is often distributed as a central F -distribution with v_1 and v_2 degrees of freedom. The hypothesis is rejected for large values of F . This process involves obtaining sums of squares and then using their expected mean squares to determine the appropriate divisor for each hypothesis of interest. The following two examples demonstrate this procedure.

20.1.2 Example 20.1: Two-Way Random Effects TS in a CR DS

A model for the two-way treatment structure with both factors random in a completely randomized design structure is

$$y_{ijk} = \mu + a_i + b_j + c_{ij} + \varepsilon_{ijk} \quad \text{for } i = 1, 2, \dots, a, j = 1, 2, \dots, b, \text{ and } k = 1, 2, \dots, n$$

TABLE 20.1

Analysis of Variance Table for the Two-Way Random Effects Model of Example 20.1

Source of Variation	df	SS	EMS
A	$a - 1$	$nb \sum_{i=1}^a (\bar{y}_{i..} - \bar{y}_{...})^2$	$\sigma_e^2 + n\sigma_c^2 + nb\sigma_a^2$
B	$b - 1$	$na \sum_{j=1}^b (\bar{y}_{.j.} - \bar{y}_{...})^2$	$\sigma_e^2 + n\sigma_c^2 + na\sigma_a^2$
$A \times B$	$(a - 1)(b - 1)$	$n \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2$	$\sigma_e^2 + n\sigma_c^2$
Residual	$ab(n - 1)$	$n \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y}_{ij.})^2$	σ_e^2

Hypothesis	Test Statistic
$H_0: \sigma_a^2 = 0$ vs $H_a: \sigma_a^2 > 0$	$F = MSA/MSAB$
$H_0: \sigma_b^2 = 0$ vs $H_a: \sigma_b^2 > 0$	$F = MSB/MSAB$
$H_0: \sigma_c^2 = 0$ vs $H_a: \sigma_c^2 > 0$	$F = MSAB/MSResidual$

where the $a_i \sim i.i.d. N(0, \sigma_a^2)$, $b_j \sim i.i.d. N(0, \sigma_b^2)$, $c_{ij} \sim i.i.d. N(0, \sigma_c^2)$, $\varepsilon_{ijk} \sim i.i.d. N(0, \sigma_e^2)$, and the random variables a_i , b_j , c_{ij} , and ε_{ijk} , are independently distributed.

The analysis of variance table with sums of squares and expected mean squares for the model is shown in Table 20.1. Test statistics are constructed by examining the expected mean squares to select the proper numerators and denominators. The statistic used to test the hypothesis $H_0: \sigma_a^2 = 0$ vs $H_a: \sigma_a^2 > 0$ is constructed by setting $\sigma_a^2 = 0$ in the expected mean square for MSA . Next find another mean square that has the same expected mean square as does MSA when H_0 is true and use that mean square as the divisor. To test $H_0: \sigma_a^2 = 0$ vs $H_a: \sigma_a^2 > 0$ the appropriate divisor is $MSAB$, to test $H_0: \sigma_b^2 = 0$ vs $H_a: \sigma_b^2 > 0$ the appropriate divisor is $MSAB$, and to test $H_0: \sigma_c^2 = 0$ vs $H_a: \sigma_c^2 > 0$ the appropriate divisor is $MSResidual$. A decision rule is to reject $H_0: \sigma_a^2 = 0$ vs $H_a: \sigma_a^2 > 0$ if $F = MSA/MSAB > F_{\alpha, (a-1), (a-1)(b-1)}$ where α is the selected type I error rate. Test statistics can be determined similarly for σ_b^2 and σ_c^2 . Table 20.1 contains a list of the hypotheses and the corresponding test statistics. Most likely, when the F -statistic does not exceed the specified percentage point, the conclusion is not that the variance component is zero, but rather, the magnitude of the variance component is negligible compared with the other sources of variation in the system.

20.1.3 Example 20.2: Complex Three-Way Random Effects TS

The data in Table 20.2 are from a design where the levels of the three factors in the treatment structure are random effects, the levels of A are crossed with the levels of B , and the levels of C are nested within the levels of B , all in a completely randomized design structure. A general model to describe a larger data set with a structure similar to that in Table 20.2 is

$$y_{ijklm} = \mu + a_i + b_j + (ab)_{ij} + c_{k(j)} + (ac)_{ik(j)} + \varepsilon_{ijklm}$$

for $i = 1, 2, \dots, a$, $j = 1, 2, \dots, b$, $k = 1, 2, \dots, c$, and $m = 1, 2, \dots, n$

TABLE 20.2

Data for Example 20.2

		Factor A					
		A ₁			A ₂		
Factor B	B ₁	C ₁	20	20	25	26	
		C ₂	23	22	26	27	
B ₂	C ₃	36	34	38	36		
		C ₄	39	38	40	39	

The parameter μ denotes an overall mean, a_i denotes the effect of level i of factor A , b_j denotes the effect of level j of factor B , $(ab)_{ij}$ denotes the interaction between the levels of factor A and factor B , $c_{k(j)}$ denotes the effect of level k of factor C nested within the j th level of factor B , $(ac)_{ik(j)}$ denotes the interaction between the levels of factor A and factor C nested within the levels of B , and ε_{ijkm} denotes the experimental unit or sampling error. Under ideal conditions, $a_i \sim i.i.d. N(0, \sigma_a^2)$, $b_j \sim i.i.d. N(0, \sigma_b^2)$, $(ab)_{ij} \sim i.i.d. N(0, \sigma_{ab}^2)$, $(ac)_{ik(j)} \sim i.i.d. N(0, \sigma_{ac(b)}^2)$, and $\varepsilon_{ijkm} \sim i.i.d. N(0, \sigma_e^2)$. Furthermore, $a_i, b_j, (ab)_{ij}, c_{k(j)}, (ac)_{ik(j)}$, and ε_{ijkm} have independent distributions. The analysis of variance table with expected mean squares for the general situation is shown in Table 20.3. F -statistics can be constructed to test hypotheses about each of the variance components by examining the expected mean squares. The statistics used to test the following hypotheses are:

- 1) To test $H_0: \sigma_a^2 = 0$ vs $H_a: \sigma_a^2 > 0$ is $F_a = MSA/MSAB$.
- 2) To test $H_0: \sigma_{ab}^2 = 0$ vs $H_a: \sigma_{ab}^2 > 0$ is $F_{ab} = MSAB/MSAC(B)$.
- 3) To test $H_0: \sigma_{c(b)}^2 = 0$ vs $H_a: \sigma_{c(b)}^2 > 0$ is $F_{c(b)} = MSC(B)/MSAC(B)$.
- 4) To test $H_0: \sigma_{ac(b)}^2 = 0$ vs $H_a: \sigma_{ac(b)}^2 > 0$ is $F_{ac(b)} = MSAC(B)/MSResidual$.

TABLE 20.3

Analysis of Variance Table for Example 20.2

Source of Variation	df	SS	EMS
A	$a - 1$	$nbc \sum_{i=1}^a (\bar{y}_{i...} - \bar{y}_{...})^2$	$\sigma_e^2 + n\sigma_{ac(b)}^2 + nc\sigma_{ab}^2 + nbc\sigma_a^2$
B	$b - 1$	$nac \sum_{j=1}^b (\bar{y}_{.j..} - \bar{y}_{...})^2$	$\sigma_e^2 + n\sigma_{ac(b)}^2 + na\sigma_{c(b)}^2 + nc\sigma_{ab}^2 + nac\sigma_b^2$
AB	$(a - 1)(b - 1)$	$nc \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij..} - \bar{y}_{i...} - \bar{y}_{.j..} + \bar{y}_{...})^2$	$\sigma_e^2 + n\sigma_{ac(b)}^2 + nc\sigma_{ab}^2$
C(B)	$b(c - 1)$	$na \sum_{j=1}^b \sum_{k=1}^c (\bar{y}_{.jk.} - \bar{y}_{.j..})^2$	$\sigma_e^2 + n\sigma_{ac(b)}^2 + na\sigma_{c(b)}^2$
AC(B)	$(a - 1)(c - 1)b$	$n \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^c (\bar{y}_{ijk.} - \bar{y}_{ij..} - \bar{y}_{.jk.} + \bar{y}_{...})^2$	$\sigma_e^2 + n\sigma_{ac(b)}^2$
Residual	$(n - 1)abc$	$\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^c \sum_{m=1}^n (y_{ijkm} - \bar{y}_{ijk.})^2$	σ_e^2

However, there is no F -statistic to test $H_0: \sigma_b^2 = 0$ vs $H_a: \sigma_b^2 > 0$ since none of the mean squares not involving σ_b^2 have the expected value $\sigma_e^2 + n\sigma_{ac(b)}^2 + na\sigma_{c(b)}^2 + nc\sigma_{ab}^2$, which is the expected value of MSB when $\sigma_b^2 = 0$. But there is a linear combination of mean squares (not including MSB) that has the desired expectation, that is, $E[MSC(B) + MSAB - MSAC(B)] = \sigma_e^2 + n\sigma_{ac(b)}^2 + na\sigma_{c(b)}^2 + nc\sigma_{ab}^2$. Let $Q = MSC(B) + MSAB - MSAC(B)$, then the statistic to test $H_0: \sigma_b^2 = 0$ vs $H_a: \sigma_b^2 > 0$ is $F_b = MSB/Q$. The sampling distribution of F_b can be approximated with an F -distribution with $b-1$ and r degrees of freedom. The denominator degrees of freedom, r , are determined by approximating the distribution of $rQ/E(Q)$ by a chi-square distribution using the Satterthwaite (1946) approximation discussed in Chapter 2. The Satterthwaite approximation is used to approximate the sampling distribution of $Q = q_1MS_1 + q_2MS_2 + \dots + q_kMS_k$ where MS_i denotes a mean square based on f_i degrees of freedom, the mean squares are independently distributed, and the q_i are known constants. Then $rQ/E(Q)$ is approximately distributed as a central chi-square random variable based on r degrees of freedom where

$$r = \frac{(Q)^2}{\sum_{i=1}^k \frac{(q_i MS_i)^2}{f_i}}$$

Assume U is a mean square based on f degrees of freedom that is independently distributed of MS_1, MS_2, \dots, MS_k with expectation $E(U) = E(Q) + k_0\sigma_0^2$. The statistic to test $H_0: \sigma_0^2 = 0$ vs $H_a: \sigma_0^2 > 0$ is $F = U/Q$, which is approximately distributed as an F -distribution with f and r degrees of freedom.

The statistic to test $H_0: \sigma_b^2 = 0$ vs $H_a: \sigma_b^2 > 0$ is $F_b = MSB/Q$, which is approximately distributed as an F -distribution with $b-1$ and r degrees of freedom where

$$r = \frac{(Q)^2}{\frac{[MSC(B)]^2}{b(c-1)} + \frac{[MSAB]^2}{(a-1)(b-1)} + \frac{[MSAC(B)]^2}{b(a-1)(c-1)}}$$

Table 20.4 contains the analysis of variance table for the data in Table 20.2, which includes the divisors (denoted by error terms) and the F -statistics for testing the respective hypotheses. To test $H_0: \sigma_b^2 = 0$ vs $H_a: \sigma_b^2 > 0$, let

$$\begin{aligned} Q &= MSC(B) + MSAB - MSAC(B) \\ &= 12.0625 + 10.5625 - 0.8125 = 21.8125 \end{aligned}$$

The degrees of freedom corresponding to Q are

$$\begin{aligned} r &= \frac{(21.8125)^2}{(12.0625)^2/2 + (10.5625)^2/1 + (0.8125)^2/2} \\ &= \frac{475.7852}{184.9785} = 2.57 \end{aligned}$$

The test statistic is $F = 770.0625/21.8125 = 35.304$ is based on 1 and 2.57 degrees of freedom. The significance level of the test is 0.0144 indicating that there is evidence to believe that

TABLE 20.4

Analysis of Variance Table Using Type I Sums of Squares with Proc Mixed Code for the Data from Example 20.2

```
PROC MIXED data=EX_20 METHOD=TYPE1 COVTEST CL; CLASS A B C;
TITLE2 'METHOD=TYPE1';
MODEL Y=;
RANDOM A B A*B C(B) A*C(B);
```

Source	df	SS	MS	Error Term	Error df	F-Value	Pr > F
A	1	39.0625	39.0625	MS(A × B)	1	3.70	0.3053
B	1	770.0625	770.0625	MS(A × B) + MS[C(B)] – MS[A × C(B)]	2.5767	35.30	0.0144
A × B	1	10.5625	10.5625	MS[A × C(B)]	2	13.00	0.0691
C(B)	2	24.1250	12.0625	MS[A × C(B)]	2	14.85	0.0631
A × C(B)	2	1.6250	0.8125	MS(Residual)	8	1.00	0.4096
Residual	8	6.5000	0.8125	—	—	—	—

$\sigma_b^2 > 0$, or that the variation due to the population of levels of factor B is an important part of the total variation in the system.

To test hypotheses about variance components in balanced designs, the F -test constructed from a ratio of two mean squares should be used whenever possible. When the ratio of two mean squares cannot be used, the Satterthwaite approximation is an acceptable alternative.

Some sort of Satterthwaite approximation is almost always necessary to test hypotheses about the variance components when the design is not balanced. Additionally, the sums of squares in the analysis of variance table may not have independent distributions, although sets of sums of squares may be independent for some special cases. The residual or error sum of squares is always independent of the other sums of squares in the analysis of variance table. Thus, for any mean square U with expectation $\sigma_e^2 + k_0\sigma_0^2$, the statistic $F_0 = U/MS_{\text{Residual}}$ provides a test of the hypothesis $H_0: \sigma_0^2 = 0$ vs $H_a: \sigma_0^2 > 0$. Under the conditions of H_0 , F is distributed as a central F -distribution with u and v degrees of freedom where u are the degrees of freedom associated with U and v are the degrees of freedom associated with MS_{Residual} .

Mean squares whose expectations involve more than two variance components generally cannot be used to obtain test statistics about single variance components that have exact F sampling distributions. The exact F -distributions occur for some balanced designs, as demonstrated in the previous two examples. One reason the ratios are not exactly distributed as F is that the respective mean squares are not independently distributed. If the design is not too unbalanced, then using the F -distribution as an approximation should be adequate. Additionally, the sums of squares (other than the residual) are not distributed as scalar multiples of chi-square distributions when the design is unbalanced.

In general, to test $H_0: \sigma_a^2 = 0$ vs $H_a: \sigma_a^2 > 0$, there will be one mean square, denoted by U_1 , with expectation

$$E(U_1) = \sigma_e^2 + k_{1a}\sigma_a^2 + k_{1b}\sigma_b^2 + k_{1c}\sigma_c^2$$

but there will be no other mean square that has expectation $\sigma_e^2 + k_{1b}\sigma_b^2 + k_{1c}\sigma_c^2$; that is, there is no single mean square that is the appropriate divisor. The method is to find a linear combination of other mean squares, say $Q = \sum_{i=1}^k q_i MS_i$ where $E(Q) = \sigma_e^2 + k_{1b}\sigma_b^2 + k_{1c}\sigma_c^2$.

TABLE 20.5Analysis of Variance Table with Expected Mean Squares and *F*-Statistic for the Data in Example 19.2

```
proc mixed method=type3 data=ex19_1 covtest cl;
class variety;
model damage=/solution;
random variety;
```

Source	df	SS	MS	EMS	Error Term	Error df	F-Value	Pr > F
Variety	3	0.810160	0.270053	Var(Residual) + 3.1795 Var(Variety)	MS(Residual)	9	4.79	0.0293
Residual	9	0.507917	0.056435	Var(Residual)	—	—	—	—

The Satterthwaite approximation can be used to approximate the sampling distribution of Q , that is, to find r such that $rQ/E(Q)$ is approximately distributed as a chi-square random variable with r degrees of freedom. The approximation is twofold, since 1) the degrees of freedom are approximated and 2) the mean squares making up Q are not necessarily independently distributed as chi-square random variables as required by the approximation.

The SAS®-Mixed code and the resulting analysis of variance table using type III sums of squares for the wheat insect damage data of Example 19.2 are displayed in Table 20.5. The expected value of the variety mean square is $\sigma_e^2 + 3.1795\sigma_{var}^2$. To test the hypothesis $H_0: \sigma_{var}^2 = 0$ vs $H_a: \sigma_{var}^2 > 0$, the appropriate divisor is the residual mean square which provides an *F*-statistic of 4.79. The computed *F*-statistic is compared to an *F*-distribution with 3 and 9 degrees of freedom; it has a significance level of 0.0293. Since this is a one-way experiment, the type I analysis would have been identical to this type III analysis.

The SAS-Mixed code and analysis of variance table constructed from the type I sums of squares for the two-way random effects data of Example 19.3 are displayed in Table 20.6. To test the hypothesis $H_0: \sigma_{row\times col}^2 = 0$ vs $H_a: \sigma_{row\times col}^2 > 0$, the appropriate divisor is the residual mean square which provides an *F*-statistic of 14.24. The computed *F*-statistic is compared with an *F*-distribution with 2 and 8 degrees of freedom, providing a significance level of 0.0023. There are no exact tests available for testing $H_0: \sigma_{row}^2 = 0$ vs $H_a: \sigma_{row}^2 > 0$ and $H_0: \sigma_{col}^2 = 0$ vs $H_a: \sigma_{col}^2 > 0$, thus approximate tests need be constructed. The appropriate divisor for *MSRow* used to test $H_0: \sigma_{row}^2 = 0$ vs $H_a: \sigma_{row}^2 > 0$ is calculated as

$$\begin{aligned} Q_{row} &= \frac{0.1429}{4.5714} MSCol + \frac{1}{2.2588} \left[2.4286 - \frac{0.1429}{4.5714} \times 2.3126 \right] MSRow \times Col \\ &\quad + \left[1 - \frac{0.1429}{4.5714} - \frac{1}{2.2588} \left(2.4286 - \frac{0.1429}{4.5714} \times 2.3126 \right) \right] MSResidual \\ &= 0.0313 \times MSCol + 1.0432 MSRow \times Col - 0.0744 MSResidual \\ &= 55.7806 \end{aligned}$$

The Satterthwaite approximate degrees of freedom associated with Q_{row} are computed as

$$\begin{aligned} df_{Q_{row}} &= \frac{(Q_{row})^2}{\frac{(0.0313 \times MSCol)^2}{2} + \frac{(1.0432 \times MSRow \times Col)^2}{2} + \frac{(0.0744 \times MSResidual)^2}{8}} \\ &= 1.9961 \end{aligned}$$

TABLE 20.6

Type I Analysis of Variance with Expected Mean Squares and Error Terms Used in Computing the F -Statistics for the Two-Way Random Effects Data of Example 19.3

```
proc mixed data=ex_19_2 method=type1 ic cl covtest asycov;
class row col;
model y=;
random row*col;
```

Source	df	SS	MS	EMS	Error Term	Error df	F-Value	Pr > F
Row	1	0.642857	0.642857	Var(Residual) + 2.4286 Var(row × col) + 0.1429 Var(col) + 7 Var(row)	0.0313 MS(col) + 1.0432 MS(row × col) – 0.0744 MS(Residual)	1.9961	0.01	0.9251
Column	2	14.759664	7.379832	Var(Residual) + 2.3126 Var(row × col) + 4.5714 Var(col)	1.0238 MS(row × col) – 0.0238 MS(Residual)	1.9935	0.13	0.8832
Row × column	2	109.145098	54.572549	Var(Residual) + 2.2588 Var(row × col)	MS(Residual)	8	14.24	0.0023
Residual	8	30.666667	3.833333	Var(Residual)	—	—	—	—

The resulting F -statistic is $F_{\text{row}} = \text{MSRow}/Q_{\text{row}} = 0.0113$ with a significance level 0.9251. The appropriate divisor for MSCol used to test $H_0: \sigma_{\text{col}}^2 = 0$ vs $H_a: \sigma_{\text{col}}^2 > 0$ is calculated as

$$\begin{aligned} Q_{\text{col}} &= \frac{2.3126}{2.2588} \text{MSRow} \times \text{Col} + \left[1 - \frac{2.3126}{2.2588} \right] \text{MSResidual} \\ &= 1.0238 \text{MSRow} \times \text{Col} - 0.0238 \text{MSResidual} \\ &= 56.8729 \end{aligned}$$

The Satterthwaite approximate degrees of freedom associated with Q_{col} are computed as

$$\begin{aligned} df_{Q_{\text{col}}} &= \frac{(Q_{\text{col}})^2}{\frac{(1.0238 \times \text{MSRow} \times \text{Col})^2}{2} + \frac{(0.0238 \times \text{MSResidual})^2}{8}} \\ &= 1.9935 \end{aligned}$$

The resulting F -statistic is $F_{\text{col}} = \text{MSCol}/Q_{\text{col}} = 0.1323$ with significance level 0.8832.

20.1.4 Likelihood Ratio Test

The second method for testing hypotheses about variance components is based on a likelihood ratio procedure which involves evaluating the value of the likelihood function for the complete model and evaluating the value of the likelihood function for the model under the conditions of H_0 .

The general random model of Equation 18.3 is

$$y = \mathbf{j}_n \boldsymbol{\mu} + \mathbf{Z}_1 \mathbf{u}_1 + \mathbf{Z}_2 \mathbf{u}_2 + \cdots + \mathbf{Z}_k \mathbf{u}_k + \boldsymbol{\varepsilon}$$

where $\mathbf{u}_1 \sim N(0, \sigma_1^2 \mathbf{I}_{t_1})$, $\mathbf{u}_2 \sim N(0, \sigma_2^2 \mathbf{I}_{t_2})$, ..., $\mathbf{u}_r \sim N(0, \sigma_r^2 \mathbf{I}_{t_r})$, $\boldsymbol{\epsilon} \sim N(0, \sigma_e^2 \mathbf{I}_N)$ and the random variables are all independently distributed. The distributional assumptions imply that the marginal distribution of \mathbf{y} is $N(\mathbf{j}_n \mu, \Sigma)$ where $\Sigma = \sigma_e^2 \mathbf{I}_n + \sigma_1^2 \mathbf{Z}_1 \mathbf{Z}'_1 + \sigma_2^2 \mathbf{Z}_2 \mathbf{Z}'_2 + \dots + \sigma_k^2 \mathbf{Z}_k \mathbf{Z}'_k$. The likelihood function is

$$L(\mu, \sigma_e^2, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2 | \mathbf{y}) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{y} - \mathbf{j}_n \mu)' \Sigma^{-1} (\mathbf{y} - \mathbf{j}_n \mu)\right]$$

The likelihood function subject to the conditions of $H_0: \sigma_1^2 = 0$ is

$$L_0(\mu, \sigma_e^2, 0, \sigma_2^2, \dots, \sigma_k^2 | \mathbf{y}) = (2\pi)^{-n/2} |\Sigma_0|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{y} - \mathbf{j}_n \mu)' \Sigma_0^{-1} (\mathbf{y} - \mathbf{j}_n \mu)\right]$$

where $\Sigma_0 = \sigma_e^2 \mathbf{I}_n + \sigma_2^2 \mathbf{Z}_2 \mathbf{Z}'_2 + \sigma_3^2 \mathbf{Z}_3 \mathbf{Z}'_3 + \dots + \sigma_k^2 \mathbf{Z}_k \mathbf{Z}'_k$.

The process is to obtain maximum likelihood estimators for the parameters of both likelihood functions and evaluate each likelihood function at the values of its estimators. The likelihood ratio test statistic is

$$LR(\sigma_1^2 = 0) = \frac{L_0(\hat{\mu}_0, \hat{\sigma}_{e0}^2, 0, \hat{\sigma}_{20}^2, \hat{\sigma}_{30}^2, \dots, \hat{\sigma}_{k0}^2 | \mathbf{y})}{L(\hat{\mu}, \hat{\sigma}_e^2, \hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_k^2 | \mathbf{y})}$$

where $\hat{\sigma}_{i0}^2$ denotes the maximum likelihood estimate of σ_i^2 from the likelihood function under the conditions of $H_0: \sigma_i^2 = 0$. When $H_0: \sigma_1^2 = 0$ is true, the asymptotic sampling distribution of

$$\begin{aligned} -2 \log [LR(\sigma_1^2 = 0)] &= -2 \log_e [L_0(\hat{\mu}_0, \hat{\sigma}_{e0}^2, 0, \hat{\sigma}_{20}^2, \hat{\sigma}_{30}^2, \dots, \hat{\sigma}_{k0}^2 | \mathbf{y})] \\ &\quad + 2 \log_e [L(\hat{\mu}, \hat{\sigma}_e^2, \hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_k^2 | \mathbf{y})] \end{aligned}$$

is central chi-square with 1 degree of freedom. The reason there is one degree of freedom is that there is one less parameter in $L_0(\bullet)$ than in $L(\bullet)$. The decision rule is to reject H_0 if $-2 \log_e [LR(\sigma_1^2 = 0)] > \chi_{\alpha,1}^2$. Likelihood ratio test statistics can be computed using SAS-Mixed where METHOD = ML is used as the variance component estimation procedure.

20.1.5 Example 20.3: Wheat Varieties—One-Way Random Effects Model

The SAS-Mixed code with results for obtaining maximum likelihood estimates of the parameters for the full model of Example 19.2 are given in Table 20.7. The maximum likelihood estimates of the parameters for the model describing the data in Example 19.2 are $\hat{\mu} = 3.9909$, $\hat{\sigma}_e^2 = 0.05749$, and $\hat{\sigma}_{\text{var}}^2 = 0.04855$ and the value of $-2 \log_e(\hat{\mu}, \hat{\sigma}_e^2, \hat{\sigma}_{\text{var}}^2 | \mathbf{y})$ is 4.96762. The reduced model is fit to the data using the SAS-Mixed code in Table 20.8 where the “Random Variety;” statement was excluded from the model in Table 20.7. Under the conditions of $H_0: \sigma_{\text{var}}^2 = 0$, the maximum likelihood estimates of the parameters are $\hat{\mu}_0 = 4.0269$, $\hat{\sigma}_{e0}^2 = 0.1014$, and $\hat{\sigma}_{\text{var0}}^2 = 0$. The value of $-2 \log_e(\hat{\mu}_0, \hat{\sigma}_{e0}^2, 0 | \mathbf{y})$ is 7.13832. The value of $-2 \log_e$ of the likelihood ratio test for testing H_0 is $7.13832 - 4.96762 = 2.1707$. The value 2.171 is compared with percentage points of a central chi-square distribution with one degree of

TABLE 20.7

Proc Mixed Code and Results for Fitting the Full Model for Example 19.2 Using Method = ML to Evaluate Likelihood Function

```
Proc mixed method=ml data=ex19_1 covtest cl ic;
class variety;
model damage=/solution;
random variety;
```

Covariance Parameter Estimates

Covariance Parameter	Estimate	Standard Error	Z-Value	Pr Z	α	Lower	Upper
Variety	0.04855	0.05075	0.96	0.1693	0.05	0.01267	2.5700
Residual	0.05749	0.02760	2.08	0.0186	0.05	0.02690	0.1972

Solution for Fixed Effects

Effect	Estimate	Standard Error	df	t-Value	Pr > t
Intercept	3.9909	0.1297	3	30.78	<0.0001
-2loglike	4.967616				

TABLE 20.8

Proc Mixed Code and Results for Fitting the Reduced Model for Example 19.2 Using Method = ML to Evaluate Likelihood Function when $\sigma_{\text{var}}^2 = 0$

```
Proc mixed method=ml data=ex19_1 covtest cl ic;
class variety;
model damage=/solution;
```

Covariance Parameter Estimates

Covariance Parameter	Estimate	Standard Error	Z-Value	Pr Z	α	Lower	Upper
Residual	0.1014	0.03977	2.55	0.0054	0.05	0.05329	0.2632

Solution for Fixed Effects

Effect	Estimate	Standard Error	df	t-Value	Pr > t
Intercept	4.0269	0.08831	12	45.60	<0.0001
-2loglike	7.138320				

freedom. The significance level for this test is 0.1407. The *F*-test constructed using the expected mean squares in Table 20.5 provided a significance level of 0.0293. The sampling distribution of the likelihood ratio test is an asymptotic distribution that is acceptable for large sample sizes whereas, in this example, the sample size is small. The sampling distribution for the *F*-test given previously is exact. *F*-tests for other variance components with models involving more than two variance components are approximate small sample size tests and are generally quite adequate and are likely better for small sample size cases than tests based on asymptotic distributions.

TABLE 20.9

Maximum Likelihood Estimates of the Variance Components for Various Models So That Likelihood Ratio Test Statistics Can Be Computed to Test the Hypothesis That Each Individual Variance Component Is Zero

Proc Mixed data=ex19_3 method=ML; title2 "Using Maximum Likelihood"; class row col; model y=/solution; random row col row*col;		Proc Mixed data=ex19_3 method=ML; title2 "Using Maximum Likelihood without row*col"; class row col; model y=/solution; random row col;		
Covariance Parameter	Full Model Estimate	Model without Row \times Column Estimate	Model without Row Estimate	Model without Column Estimate
Row	0	0	0	0
Column	0	0	0	0
Row \times column	7.4065	0	7.4065	7.4065
Residual	3.8430	0	3.8430	3.8430
Intercept	14.8476	11.0867	14.8476	14.8476
-2loglike	68.7	73.4	68.7	68.7

20.1.6 Example 20.4: Unbalanced Two-Way

The likelihood ratio statistic to test the hypothesis $H_0: \sigma_{\text{row}\times\text{col}}^2 = 0$ vs $H_a: \sigma_{\text{row}\times\text{col}}^2 = 0 > 0$ for the two-way random effects data of Example 19.3 is obtained by fitting a model with all terms in the random statement and then fitting the model without the row \times col term in the random statement, as displayed in Table 20.9. The SAS-Mixed code for fitting the two models is included in Table 20.9. The results from fitting models without row and without column effects are also included without the corresponding SAS-Mixed code. Maximum likelihood estimates of the variance components, estimates of the intercepts, and the $-2\log(\text{likelihood})$ values are given. The likelihood ratio test statistic for the hypothesis $H_0: \sigma_{\text{row}\times\text{col}}^2 = 0$ vs $H_a: \sigma_{\text{row}\times\text{col}}^2 > 0$ is $73.4 - 68.7 = 4.7$, which has a chi-square with one degree of freedom sampling distribution under the conditions of H_0 . The significance level is 0.030, indicating there is sufficient information to believe that $\sigma_{\text{row}\times\text{col}}^2 > 0$. The values of the likelihood ratio statistics are both zero for testing the hypotheses that row and the column variance components are equal to zero. This occurs because the maximum likelihood estimates of both variance components are zero. So for testing the row variance component is zero, the value of the $-2\log(\text{likelihood})$ remains the same whether row is included in the random statement or not.

20.2 Constructing Confidence Intervals

There are a few procedures that provide exact confidence intervals about some of the variance components in some models, but most of the time confidence intervals that are obtained are approximate and rely on some type approximation.

20.2.1 Residual Variance σ_ϵ^2

For the general random model, a $(1 - \alpha)100\%$ confidence interval about σ_ϵ^2 is

$$\frac{v\hat{\sigma}_\epsilon^2}{\chi_{(\alpha/2),v}^2} \leq \sigma_\epsilon^2 \leq \frac{v\hat{\sigma}_\epsilon^2}{\chi_{1-(\alpha/2),v}^2}$$

where $\hat{\sigma}_\epsilon^2 = \text{SSRESIDUAL}/v$, v is the degrees of freedom associated with $\hat{\sigma}_\epsilon^2$ and $\chi_{1-(\alpha/2),v}^2$ and $\chi_{\alpha/2,v}^2$ denote lower and upper $\alpha/2$ percentage points from a chi-square distribution with v degrees of freedom.

20.2.2 General Satterthwaite Approximation

The general Satterthwaite approximation to the sampling distribution associated with a variance component is obtained by equating the first two moments of $r\hat{\sigma}_i^2/E(\hat{\sigma}_i^2)$ to the first two moments of a chi-square distribution based on r degrees of freedom and then solving for r . The first moment of chi-square distribution is equal to its degrees of freedom, r . So equating the first moment of $r\hat{\sigma}_i^2/E(\hat{\sigma}_i^2)$ to the first moment of a chi-square distribution with r degrees of freedom provides no information about r . The variance of a chi-square distribution with r degrees of freedom is equal to $2r$. Equating the variances of $r\hat{\sigma}_i^2/E(\hat{\sigma}_i^2)$ to the variance of a chi-square distribution with r degrees, one obtains

$$\text{Var}\left(\frac{r\hat{\sigma}_i^2}{E(\hat{\sigma}_i^2)}\right) = 2r \quad \text{or} \quad \frac{r^2}{[E(\hat{\sigma}_i^2)]^2} \text{Var}(\hat{\sigma}_i^2) = 2r$$

which implies

$$r = \frac{2[E(\hat{\sigma}_i^2)]^2}{\text{Var}(\hat{\sigma}_i^2)}$$

The value of r is estimated by replacing $E(\hat{\sigma}_i^2)$ by $\hat{\sigma}_i^2$ and $\text{Var}(\hat{\sigma}_i^2)$ by an estimate of its variance. If the REML or ML solutions are obtained, then the inverse of the information matrix can be used to provide the estimates of the variances of the estimated variance components. SAS-Mixed uses the inverse of the information matrix evaluated at the values of the estimates of the variance components to provide asymptotic estimates of the variances and covariances of the estimated variance components.

Table 20.10 contains the SAS-Mixed code using options "covtest," "cl" and "asycov" to yield the estimated standard errors of the estimated variance components, Z-values computed as the ratio of each estimate to its estimated standard error, the Satterthwaite type confidence intervals denoted by Lower and Upper, and the asymptotic covariance matrix of the estimates of the variance components. The data step in Table 20.10 computes the approximate degrees of freedom associated with each estimated variance component and then uses those degrees of freedom to compute the Satterthwaite confidence intervals. The approximate degrees of freedom are 1.758 and 8.828 for $\hat{\sigma}_{\text{var}}^2$ and $\hat{\sigma}_\epsilon^2$, respectively. The recomputed confidence intervals are identical to those provided by SAS-Mixed.

TABLE 20.10

SAS-Mixed Code with Method = REML and a Data Step to Calculate Satterthwaite-Type Confidence Intervals for Variance Components for the Insect Damage in Wheat of Example 19.2

```
Proc mixed method=reml data=ex19_1 covtest cl asycov;
class variety;
model damage=/solution;
random variety;
ods output covparms=cov asycov=asycov;
```

Covariance Parameter Estimates

Covariance Parameter	Estimate	Standard Error	Z-Value	Pr Z	α	Lower	Upper
Variety	0.07316	0.07802	0.94	0.1742	0.05	0.01875	4.4672
Residual	0.05700	0.02713	2.10	0.0178	0.05	0.02681	0.1929

Asymptotic Covariance Matrix of Estimates

Row	Covariance Parameter	CovP1	CovP2
1	Variety	0.006087	-0.00031
2	Residual	-0.00031	0.000736

```
data cov1; set cov;
df=2*zvalue**2;
chi1_alpha=cinv(alpha/2,df);
chialpha=cinv(1-alpha/2,df);
low=df*estimate/chialpha;
up=df*estimate/chi1_alpha;
proc print;
```

Covariance Parameter	df	Chi1_Alpha	Chialpha	Low	Up
Variety	1.75840	0.02880	6.8588	0.018755	4.46721
Residual	8.82768	2.60870	18.7684	0.026811	0.19289

REML estimates of the variance components for the data of Example 20.1 are provided by the SAS-Mixed code in Table 20.11. The confidence intervals about the variance components are computed using $df = 2(Z\text{-value})^2$, which provides 0.50, 0.95, 0.84, 1.73, 0, and 8 degrees of freedom, respectively. One should notice that, when there are very few levels associated with a random effect and consequently very few degrees of freedom, the resulting confidence intervals are going to be extremely wide.

20.2.3 Approximate Confidence Interval for a Function of the Variance Components

Often a researcher wishes to construct a confidence interval about some function of the variance components in the model. As is described in the next section, there are some cases involving balanced models where exact confidence intervals exist. In this section, a first-order Taylor's series (Kendall and Stuart, 1973) is used to provide an estimate of the

TABLE 20.11

SAS-Mixed Code to Provide REML Estimates of and Satterwaite Confidence Intervals about the Variance Components

```
PROC MIXED data=EX_20 METHOD=REML COVTEST CL;
CLASS A B C;
MODEL Y=;
RANDOM A B A*B C(B) A*C(B);
```

Covariance Parameter Estimates

Covariance Parameter	Estimate	Standard Error	Z-Value	Pr Z	α	Lower	Upper
A	3.5625	7.1533	0.50	0.3092	0.05	0.5170	3774318
B	93.5313	136.15	0.69	0.2461	0.05	18.1334	141645
$A \times B$	2.4375	3.7399	0.65	0.2573	0.05	0.4504	8130.42
$C(B)$	2.8125	3.0225	0.93	0.1760	0.05	0.7162	181.58
$A \times C(B)$	0	0.4542	0.00	0.5000	0.05	—	—
Residual	0.8125	0.4063	2.00	0.0228	0.05	0.3707	2.9820

variance of a function of variances which is then used to construct a Satterthwaite type confidence interval. Let σ denote the vector of variance components and let $\text{Var}(\sigma)$ denote the matrix of variances and covariances of the estimates of the variance components. Assume that the function of the variance components of interest is $\varphi(\sigma)$ and that $\varphi(\sigma)$ has continuous first derivatives. The maximum likelihood estimate of $\varphi(\sigma)$ is $\varphi(\hat{\sigma})$, where $\hat{\sigma}$ denotes the maximum likelihood estimate of σ . An estimate of the inverse of the information matrix corresponding to the variance components is $\hat{V}(\hat{\sigma})$ which provides an estimate of $\text{Var}(\sigma)$. Compute the derivatives of $\varphi(\sigma)$ with respect to each of the variance components and evaluate them at $\hat{\sigma}$. Then let

$$\hat{f}' = \left(\frac{\partial \varphi(\sigma)}{\partial \sigma_1^2}, \frac{\partial \varphi(\sigma)}{\partial \sigma_2^2}, \dots, \frac{\partial \varphi(\sigma)}{\partial \sigma_k^2}, \frac{\partial \varphi(\sigma)}{\partial \sigma_\epsilon^2} \right) \Big|_{\sigma = \hat{\sigma}}$$

The approximate variance of $\varphi(\hat{\sigma})$ is $\sigma_{\varphi(\hat{\sigma})}^2 = \hat{f}' \hat{V}(\hat{\sigma}) \hat{f}$. The Satterthwaite approximate degrees of freedom corresponding to this estimate are computed as

$$r = \frac{2[\varphi(\hat{\sigma})]^2}{\hat{\sigma}_{\varphi(\hat{\sigma})}^2}$$

Suppose the researcher is interested in estimating the total variability in the data for the insect damage study of Example 19.2; that is, the researcher wants to estimate $\varphi(\sigma) = \sigma_{\text{var}}^2 + \sigma_\epsilon^2$. The vector of first derivatives of $\varphi(\sigma)$ is

$$f = \begin{bmatrix} \frac{\partial \varphi(\sigma)}{\partial \sigma_{\text{var}}^2} \\ \frac{\partial \varphi(\sigma)}{\partial \sigma_\epsilon^2} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and the covariance matrix of the estimates of the variance components is

$$\hat{V}(\hat{\sigma}) = \begin{bmatrix} \hat{\sigma}_{\hat{\sigma}_{\text{var}}^2}^2 & \hat{\sigma}_{\hat{\sigma}_{\text{var}}^2 \hat{\sigma}_{\varepsilon}^2} \\ \hat{\sigma}_{\hat{\sigma}_{\text{var}}^2 \hat{\sigma}_{\varepsilon}^2} & \hat{\sigma}_{\hat{\sigma}_{\varepsilon}^2}^2 \end{bmatrix}$$

Thus the approximate variance of $\varphi(\hat{\sigma})$ is $\hat{\sigma}_{\varphi(\hat{\sigma})}^2 = \hat{\sigma}_{\hat{\sigma}_{\text{var}}^2}^2 + \hat{\sigma}_{\hat{\sigma}_{\varepsilon}^2}^2 + 2\hat{\sigma}_{\hat{\sigma}_{\text{var}}^2 \hat{\sigma}_{\varepsilon}^2}$. Using the information in Table 20.10 one obtains

$$\varphi(\hat{\sigma}) = 0.07316 + 0.05700 = 0.13010$$

and

$$\hat{\sigma}_{\varphi(\hat{\sigma})}^2 = 0.006087 + 0.000736 - 2(0.00031) = 0.006203$$

The approximate degrees of freedom corresponding to the estimated variance are

$$r = \frac{2(0.13010)^2}{0.006203} = 5.46$$

Finally, a 95% confidence interval about $\varphi(\sigma) = \sigma_{\text{var}}^2 + \sigma_{\varepsilon}^2$ is

$$0.052287 < \sigma_{\text{var}}^2 + \sigma_{\varepsilon}^2 < 0.70253$$

As a second example, the intraclass correlation coefficient is defined to be

$$\rho = \frac{\sigma_{\text{var}}^2}{\sigma_{\varepsilon}^2 + \sigma_{\text{var}}^2}$$

An approximate confidence interval can be constructed using the Taylor's series approach where

$$\varphi(\sigma) = \frac{\sigma_{\text{var}}^2}{\sigma_{\varepsilon}^2 + \sigma_{\text{var}}^2}$$

The transpose of the estimated vector of derivatives of

$$\frac{\sigma_{\text{var}}^2}{\sigma_{\varepsilon}^2 + \sigma_{\text{var}}^2}$$

is

$$\hat{f}' = \left[\frac{(\hat{\sigma}_{\varepsilon}^2 + \hat{\sigma}_{\text{var}}^2) - (\hat{\sigma}_{\text{var}}^2)}{(\hat{\sigma}_{\varepsilon}^2 + \hat{\sigma}_{\text{var}}^2)^2}, \frac{-(\hat{\sigma}_{\text{var}}^2)}{(\hat{\sigma}_{\varepsilon}^2 + \hat{\sigma}_{\text{var}}^2)^2} \right]$$

Using the information in Table 20.10, the estimate of the intraclass correlation is $\hat{\rho} = 0.5621$ and its estimated variance is 0.2156. The corresponding approximated degrees of freedom are $r = 2.93$. The resulting approximate confidence interval for the intraclass correlation coefficient is $(0.1785, 8.214)$ which can be simplified to $(0.1785, 1)$ since the intraclass correlation coefficient cannot be greater than 1. The first-order Taylor's series approach works quite well for linear functions of the variance components and less well for non linear functions such as the intraclass correlation. Better confidence intervals for the intra-class correlation are available and they will be described in Section 20.2.3.

When one uses the sums of squares approach to obtain estimates of the variance components, the usual Satterthwaite approximation can be used to construct a confidence interval. Suppose the estimate of each variance component can be expressed as a linear combination of the mean squares given in an analysis of variance table as, $\hat{\sigma}_s^2 = \sum_{i=1}^k c_{is} MS_i + c_{es} MS_{\text{Residual}}$. Then the approximate number of degrees of freedom is computed as

$$r_s = \frac{(\hat{\sigma}_s^2)^2}{\sum_{i=1}^k \frac{(c_{is} MS_i)^2}{df_{MS_i}} + \frac{(c_{es} MS_{\text{Residual}})^2}{df_{MS_{\text{Residual}}}}}$$

An approximate $(1 - \alpha)100\%$ confidence interval about σ_s^2 is

$$\frac{r_s \hat{\sigma}_s^2}{\chi_{\alpha/2, r_s}^2} \leq \sigma_s^2 \leq \frac{r_s \hat{\sigma}_s^2}{\chi_{1-(\alpha/2), r_s}^2}$$

The coefficients of the mean squares are quite easily obtained from the expected mean squares and the error terms provided by SAS-Mixed when a sum of squares method is used to estimate the variance components. For example, the estimate of σ_a^2 for Example 20.2 is

$$\hat{\sigma}_a^2 = \frac{1}{8} [MSA - MS(A \times B)] = \frac{1}{8} (39.0625 - 10.5625) = 3.5625$$

and the approximate number of degrees of freedom is

$$r = \frac{(3.5625)^2}{\frac{[(1/8)39.0625]^2}{1} + \frac{[(1/8)10.5625]^2}{1}} = 0.305$$

The resulting confidence interval for σ_a^2 is $(0.4226, 27,556,199,203.09)$, a very wide interval caused by the very low number of degrees of freedom (0.305). Also, the estimate of σ_b^2 for Example 20.2 is

$$\begin{aligned}\hat{\sigma}_b^2 &= \frac{1}{8} \{MSB - [MS(A \times B) + MS(C(B)) - MS(A \times C(B))]\} \\ &= \frac{1}{8} (770.0625 - 12.0625 - 10.5625 + 0.8125) \\ &= 93.5313\end{aligned}$$

and the approximate number of degrees of freedom is

$$r = \frac{(93.5313)^2}{\frac{[(1/8)770.0625]^2}{1} + \frac{[(1/8)12.0625]^2}{2} \frac{[(1/8)10.5625]^2}{1} \frac{[(1/8)0.8125]^2}{2}} = 0.944$$

The resulting confidence interval for σ_b^2 is (18.1347, 141, 493.19) which is also a wide interval caused by the small number of degrees of freedom. Basically, when only a small number of degrees of freedom is available to estimate a variance component, you know very little about that variance component.

20.2.4 Wald-Type Confidence Intervals for Variance Components

Wald-type confidence intervals can be computed using the asymptotic normality of maximum likelihood estimates. The $(1 - \alpha)100\%$ Wald confidence interval about a variance component, σ_s^2 , is $\hat{\sigma}_s^2 - Z_{\alpha/2}\sqrt{\hat{\sigma}_{\hat{\sigma}_s^2}^2} \leq \sigma_s^2 \leq \hat{\sigma}_s^2 + Z_{\alpha/2}\sqrt{\hat{\sigma}_{\hat{\sigma}_s^2}^2}$. Wald confidence intervals are symmetric about $\hat{\sigma}_s^2$. Under an assumption of normality of the data, the sampling distribution associated with a variance is a chi-square distribution which is not symmetric. But as the number of degrees of freedom increases, the shape of the chi-square distribution becomes much more symmetric. Thus, when the degrees of freedom associated with a variance component is large, the Wald confidence interval should be adequate, but when the number of degrees of freedom is small, the Satterthwaite type confidence interval will more closely reflect reality. When one uses METHOD = TYPEEx in SAS-Mixed and requests confidence intervals, Wald confidence intervals are provided for all variance components, except for the residual variance component where a Satterthwaite confidence interval is computed.

20.2.5 Some Exact Confidence Intervals

Since all sums of squares (not including *SSRESIDUAL*) in the analysis of variance table for a balanced model are independent of the residual sum of squares, the next result can be used to construct a confidence interval about a variance component, say σ_1^2 , with a confidence coefficient of at least size $1 - \alpha$ when there is a mean square in the model with expectation $\sigma_e^2 + a\sigma_1^2$. Let $Q_1 = MSRESIDUAL$ and suppose it is based on u_1 degrees of freedom, and let Q_2 be a mean square based on u_2 degrees of freedom with expectation $\sigma_e^2 + a\sigma_1^2$. A set of exact simultaneous $(1 - \alpha)100\%$ confidence intervals about σ_e^2 and $\sigma_e^2 + a\sigma_1^2$ is given by

$$\frac{u_1 Q_1}{\chi_{\rho/2, u_1}^2} \leq \sigma_e^2 \leq \frac{u_1 Q_1}{\chi_{1-(\rho/2), u_1}^2}$$

$$\frac{u_2 Q_2}{\chi_{\rho/2, u_2}^2} \leq \sigma_e^2 + a\sigma_1^2 \leq \frac{u_2 Q_2}{\chi_{1-(\rho/2), u_2}^2}$$

where $\rho = 1 - \sqrt{1 - \alpha}$. The intersection of these two regions on the (σ_e^2, σ_1^2) plane provides a $(1 - \alpha)100\%$ simultaneous confidence region for (σ_e^2, σ_1^2) . A graph of the confidence region

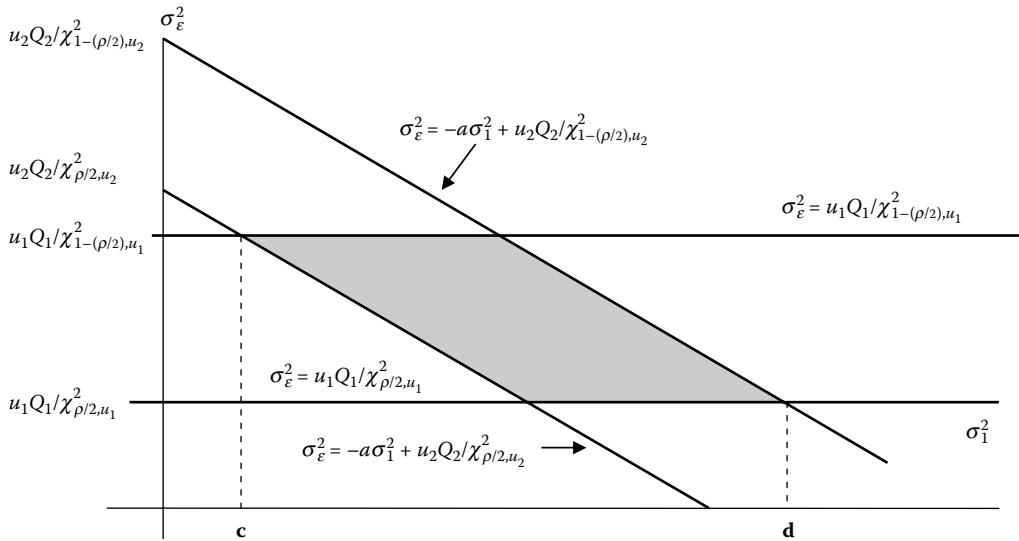


FIGURE 20.1 Graph of confidence region for $(\sigma_\epsilon^2, \sigma_1^2)$.

is shown in Figure 20.1. A $(1 - \alpha)100\%$ confidence interval about σ_1^2 is obtained by determining the maximum and minimum of σ_1^2 over the confidence region. The minimum value of σ_1^2 over the intersection region in Figure 20.1 is given by c and the maximum is given by d . The values of c and d can be determined by solving for the values of σ_1^2 where the respective lines intersect. The two sets of line intersect at

$$c = \frac{u_2 Q_2 / \chi_{\rho/2, u_2}^2 - u_2 Q_2 / \chi_{1-(\rho/2), u_2}^2}{a},$$

and

$$d = \frac{u_2 Q_2 / \chi_{1-(\rho/2), u_2}^2 - u_1 Q_1 / \chi_{\rho/2, u_1}^2}{a}$$

which provides $c \leq \sigma_1^2 \leq d$ as the $(1 - \alpha)100\%$ confidence interval about σ_1^2 . The value of c corresponds to the value of σ_1^2 at which the two lines

$$\sigma_\epsilon^2 = u_1 Q_1 / \chi_{1-(\rho/2), u_1}^2 \quad \text{and} \quad \sigma_\epsilon^2 = -a\sigma_1^2 + u_2 Q_2 / \chi_{\rho/2, u_2}^2$$

intersect, and the value of d corresponds to the value of σ_1^2 at which the two lines

$$\sigma_\epsilon^2 = u_1 Q_1 / \chi_{\rho/2, u_1}^2 \quad \text{and} \quad \sigma_\epsilon^2 = -a\sigma_1^2 + u_2 Q_2 / \chi_{1-(\rho/2), u_2}^2$$

intersect.

The above joint confidence region about $(\sigma_\epsilon^2, \sigma_1^2)$ can be used to construct a confidence interval about any continuous function of σ_ϵ^2 and σ_1^2 . Let $\varphi(\sigma_\epsilon^2, \sigma_1^2)$ denote a continuous function of σ_ϵ^2 and σ_1^2 over the $(\sigma_\epsilon^2, \sigma_1^2)$ space and let $\mathfrak{R}_\alpha(\sigma_\epsilon^2, \sigma_1^2)$ denote the joint confidence region for σ_ϵ^2 and σ_1^2 . The lower confidence limit for $\varphi(\sigma_\epsilon^2, \sigma_1^2)$ is

$$L = \min_{(\sigma_\epsilon^2, \sigma_1^2) \in \mathfrak{R}_\alpha(\sigma_\epsilon^2, \sigma_1^2)} [\varphi(\sigma_\epsilon^2, \sigma_1^2)]$$

and the upper confidence limit for $\varphi(\sigma_\epsilon^2, \sigma_1^2)$ is

$$U = \max_{(\sigma_\epsilon^2, \sigma_1^2) \in \mathfrak{R}_\alpha(\sigma_\epsilon^2, \sigma_1^2)} [\varphi(\sigma_\epsilon^2, \sigma_1^2)]$$

For example, if

$$\varphi(\sigma_\epsilon^2, \sigma_1^2) = \frac{\sigma_1^2}{\sigma_\epsilon^2 + \sigma_1^2}$$

the maximum occurs at $(u_1 Q_1 / \chi^2_{1-(\alpha/2), u_1}, c)$ and the minimum occurs at $(u_1 Q_1 / \chi^2_{\alpha/2, u_1}, d)$. Thus using the data in Example 20.1, the resulting confidence interval about

$$\frac{\sigma_1^2}{\sigma_\epsilon^2 + \sigma_1^2}$$

is

$$-0.048511 < \frac{\sigma_1^2}{\sigma_\epsilon^2 + \sigma_1^2} < 1.89748$$

or

$$0 < \frac{\sigma_1^2}{\sigma_\epsilon^2 + \sigma_1^2} < 1$$

since it is known that

$$\frac{\sigma_1^2}{\sigma_\epsilon^2 + \sigma_1^2}$$

must be greater than 0 and less than 1. When the numbers of degrees of freedom are small the minimization and maximization method can lead to wild results, especially when $\varphi(\sigma_\epsilon^2, \sigma_1^2)$ is a nonlinear function.

Another confidence interval for σ_1^2 , proposed by Williams (1962), has lower and upper confidence limits given by

$$c_1 = \frac{u_2[Q_2 - Q_1 F_{\alpha/2, u_2, u_1}]}{a \chi_{\alpha/2, u_2}^2} \quad \text{and} \quad d_1 = \frac{u_2[Q_2 - Q_1 F_{1-(\alpha/2), u_2, u_1}]}{a \chi_{1-(\alpha/2), u_2}^2}, \text{ respectively}$$

and has a confidence coefficient of at least $1 - 2\alpha$. Simulation studies (Boardman, 1974; and an unpublished study by the authors) have shown that, if ρ is chosen to be α rather than $\rho = 1 - \sqrt{1 - \alpha}$, as described above, the first procedure gives a confidence interval for σ_1^2 that has a confidence coefficients that is no less than $1 - \alpha$. Similar simulations have shown that the interval proposed by Williams has a confidence coefficient that is no less than $1 - \alpha$. The interval of Williams is a little shorter than the one described by $c \leq \sigma_1^2 \leq d$. Thus the (c, d) interval is a little more conservative than Williams's interval.

20.2.6 Example 20.5: Balanced One-Way Random Effects Treatment Structure

The data in Table 20.12 are from a one-way treatment structure where five workers were randomly selected from the population of workers working at a plant. The levels of workers are random effects and the study was conducted as a completely randomized design structure with five workers and three observations per worker where the response is the number of units assembled during a fixed period of time. A model that can be used to describe this data is

$$y_{ij} = \mu + w_i + \varepsilon_{ij}, \quad i = 1, 2, \dots, 5, \quad j = 1, 2, 3, \quad w_i \sim \text{i.i.d. } N(0, \sigma_w^2) \text{ and } \varepsilon_{ij} \sim \text{i.i.d. } N(0, \sigma_e^2)$$

An analysis of variance table including the expected mean squares is provided in Table 20.13. A 95% confidence interval about σ_e^2 is

$$\frac{16}{20.483} \leq \sigma_e^2 \leq \frac{16}{3.247} \quad \text{or} \quad 0.781 \leq \sigma_e^2 \leq 4.928$$

A 95% confidence interval about $\sigma_e^2 + 3\sigma_w^2$ is

$$\frac{128}{11.143} \leq \sigma_e^2 + 3\sigma_w^2 \leq \frac{128}{0.484} \quad \text{or} \quad 11.487 \leq \sigma_e^2 + 3\sigma_w^2 \leq 264.463$$

TABLE 20.12

Data for Example 20.5

Station	Worker 1	Worker 2	Worker 3	Worker 4	Worker 5
1	2	6	7	3	12
2	5	6	9	5	11
3	4	8	8	6	13

TABLE 20.13

SAS-Mixed Code and Resulting Analysis of Variance Table for the Worker Data of Example 20.5

```
PROC MIXED data=EX_20_5 METHOD=TYPE1 COVTEST CL;
CLASS worker;
MODEL num_units=;
RANDOM worker;
```

Type I Analysis of Variance

Source	df	ss	MS	EMS	Error Term	Error df	F-Value	Pr > F
Worker	4	128.0	32.0	Var(Residual) + 3 Var(Worker)	MS(Residual)	10	20.00	<0.0001
Residual	10	16.00	1.6	Var(Residual)	—	—	—	—

By using the joint confidence region resulting from the intersection of these two intervals, an approximate 95% confidence interval about σ_w^2 is $c \leq \sigma_w^2 \leq d$ where

$$c = \frac{11.487 - 4.928}{3} = 2.186 \quad \text{and} \quad d = \frac{264.463 - 0.781}{3} = 87.818$$

Thus one obtains $2.186 \leq \sigma_w^2 \leq 87.818$ as an approximate 95% confidence interval for σ_w^2 .

This type of procedure can be used to construct a confidence interval about any variance component, say σ_1^2 , when there are two independent mean squares Q_1 and Q_2 based on u_1 and u_2 degrees of freedom, respectively, such that $E(Q_2) = E(Q_1) + \sigma_1^2$. Let $\sigma_0^2 = E(Q_1)$ and replace σ_ϵ^2 by σ_0^2 in the previous development; then the $(1 - \alpha)100\%$ approximate confidence interval for σ_1^2 is $c \leq \sigma_1^2 \leq d$. The result of Williams (1962) has also been applied to this case (Graybill, 1976, Theorem 15.3.5).

20.2.7 Example 20.6

To demonstrate the above more general procedure, a 90% confidence interval about σ_a^2 is constructed for the data in Example 20.2 where the data are in Table 20.2, the expected mean squares are in Table 20.3, and the analysis of variance table for the data is in Table 20.4. Let $Q_2 = MSA$ and $Q_1 = MSAB$, both of which are based on 1 degree of freedom. Their expectations are

$$E(Q_1) = \sigma_\epsilon^2 + 2\sigma_{ac(b)}^2 + 4\sigma_{ab}^2 \quad \text{and} \quad E(Q_2) = \sigma_\epsilon^2 + 2\sigma_{ac(b)}^2 + 4\sigma_{ab}^2 + 8\sigma_a^2,$$

respectively. Let $\sigma_0^2 = \sigma_\epsilon^2 + 2\sigma_{ac(b)}^2 + 4\sigma_{ab}^2$, $\sigma_1^2 = \sigma_a^2$, and $a = 8$. Then the limits of the confidence interval for σ_a^2 are

$$c = \frac{(1 \times 39.0625 / 3.8415) - (1 \times 10.5625 / 0.00393)}{8} = -334.69$$

and

$$d = \frac{(1 \times 39.0625 / 0.00393) - (1 \times 10.5625 / 3.8415)}{8} = 1242.10$$

or

$$-334.69 \leq \sigma_a^2 \leq 1242.10$$

In this example, the two lines in Figure 20.1 that determine the value of c intersect outside of the parameter space, thus the lower limit is truncated to zero so the confidence interval only includes values of σ_a^2 in the parameter space. Consequently, the resulting confidence interval for σ_a^2 is $0 \leq \sigma_a^2 \leq 1242.10$. A similar technique can be used to construct confidence intervals about σ_{ab}^2 , $\sigma_{ac(b)}^2$ and $\sigma_{c(b)}^2$, but it cannot be used to obtain a confidence interval about σ_b^2 .

Burdick and Graybill (1992) present several situations where exact confidence intervals can be obtained for the variance components. They also present approximate confidence intervals for several functions of the variance components. Exact confidence intervals can usually be obtained for balanced designs with equal numbers of observations per cell and no missing cells. The intervals are based on having a set of sums of squares in an analysis that are distributed as independent chi-square random variables. Burdick and Graybill (1992) present confidence intervals developed by Graybill and Wang (1980), Lu et al. (1989) and Ting et al. (1990) for various functions of the variance components. The following discussion is about confidence intervals associated with the balanced one-way random effects treatment structure carried out in a completely randomized design structure, which Burdick and Graybill refer to as a one-fold nested design. The model is

$$\begin{aligned} y_{ij} &= \mu + u_i + \varepsilon_{ij}, \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, n \\ u_i &\sim \text{i.i.d. } N(0, \sigma_u^2), \quad \varepsilon_{ij} \sim \text{i.i.d. } N(0, \sigma_\varepsilon^2) \end{aligned}$$

and the u_i and the ε_{ij} are independent random variables.

The analysis of variance table for this model is based on two sums of squares

$$\text{SSBetween} = n \sum_{i=1}^t (\bar{y}_{i\cdot} - \bar{y}_{..})^2 = (t-1)Q_1$$

and

$$\text{SSWithin} = \sum_{i=1}^t \sum_{j=1}^n (y_{ij} - \bar{y}_{i\cdot})^2 = t(n-1)Q_2$$

where Q_1 and Q_2 are the respective mean squares. The expected mean squares corresponding to the sums of squares are

$$E(\text{MSBetween}) = E(Q_1) = \sigma_\varepsilon^2 + n\sigma_u^2 \quad \text{and} \quad E(\text{MSWITHIN}) = E(Q_2) = \sigma_\varepsilon^2, \text{ respectively}$$

The two sums of squares are independent random variables with sampling distributions

$$\frac{(t-1)Q_1}{\sigma_\varepsilon^2 + n\sigma_u^2} \sim \chi_{t-1}^2 \quad \text{and} \quad \frac{t(n-1)Q_2}{\sigma_\varepsilon^2} \sim \chi_{t(n-1)}^2$$

The exact $(1 - \alpha)100\%$ confidence interval about σ_e^2 is

$$\left[\frac{Q_2}{F_{\alpha/2,t(n-1),\infty}} \leq \sigma_e^2 \leq \frac{Q_2}{F_{1-(\alpha/2),t(n-1),\infty}} \right]$$

where $F_{\alpha/2,t(n-1),\infty}$ and $F_{1-(\alpha/2),t(n-1),\infty}$ are the upper and lower critical points from an F -distribution with $t(n-1)$ degrees of freedom for the numerator and with ∞ degrees of freedom for the denominator. It can be noted that $F_{\alpha/2,t(n-1),\infty} = \chi_{\alpha/2,t(n-1)}^2 / [t(n-1)]$ and $F_{1-(\alpha/2),t(n-1),\infty} = \chi_{1-(\alpha/2),t(n-1)}^2 / [t(n-1)]$. Using these representations for the percentage points, the above confidence interval about σ_e^2 is identical to the interval in Section 20.2.1.

An approximate $(1 - \alpha)100\%$ confidence interval about σ_u^2 given in Burdick and Graybill (1992) is

$$\left[\frac{Q_1 - Q_2 - \sqrt{V_L}}{n} \leq \sigma_u^2 \leq \frac{Q_1 - Q_2 + \sqrt{V_u}}{n} \right]$$

where

$$V_L = G_1^2 Q_1^2 + H_2^2 Q_2^2 + G_{12} Q_1 Q_2 \quad V_u = H_1^2 Q_1^2 + G_2^2 Q_2^2 + H_{12} Q_1 Q_2$$

where

$$\begin{aligned} G_1 &= 1 - \frac{1}{F_{\alpha/2,t-1,\infty}}, \quad G_2 = 1 - \frac{1}{F_{\alpha/2,t(n-1),\infty}} \\ H_1 &= \frac{1}{F_{1-(\alpha/2),t-1,\infty}} - 1, \quad H_2 = \frac{1}{F_{1-(\alpha/2),t(n-1),\infty}} - 1 \\ G_{12} &= \frac{(F_{\alpha/2,t-1,t(n-1)} - 1)^2 - G_1^2 F_{\alpha/2,t-1,t(n-1)}^2 - H_2^2}{F_{\alpha/2,t-1,t(n-1)}} \end{aligned}$$

and

$$H_{12} = \frac{(1 - F_{1-(\alpha/2),t-1,t(n-1)})^2 - H_1^2 F_{1-(\alpha/2),t-1,t(n-1)}^2 - G_2^2}{F_{1-(\alpha/2),t-1,t(n-1)}}$$

When the lower bound is negative, the lower bound is set to 0. The lower bound is negative when $Q_1/Q_2 < F_{\alpha/2,t-1,t(n-1)}$. Using $Q_1/Q_2 > F_{\alpha/2,t-1,t(n-1)}$ as a decision rule to reject $H_0: \sigma_u^2 = 0$ vs $H_a: \sigma_u^2 > 0$ provides an exact test that is a uniformly most powerful unbiased test (Lehmann, 1986).

An approximate $(1 - \alpha)100\%$ confidence interval about $\sigma_e^2 + \sigma_u^2$ given in Burdick and Graybill (1992) is

$$\left[\hat{\sigma}_e^2 + \hat{\sigma}_u^2 - \frac{\sqrt{G_1^2 Q_1^2 + G_2^2 (n-1) Q_2^2}}{n} \leq \sigma_e^2 + \sigma_u^2 \leq \hat{\sigma}_e^2 + \hat{\sigma}_u^2 + \frac{\sqrt{H_1^2 Q_1^2 + H_2^2 (n-1) Q_2^2}}{n} \right]$$

where

$$\begin{aligned} G_1 &= 1 - \frac{1}{F_{\alpha/2,t-1,\infty}}, \quad G_2 = 1 - \frac{1}{F_{\alpha/2,t(n-1),\infty}} \\ H_1 &= \frac{1}{F_{1-(\alpha/2),t-1,\infty}} - 1, \quad H_2 = \frac{1}{F_{1-(\alpha/2),t(n-1),\infty}} - 1 \end{aligned}$$

An approximate $(1 - \alpha)100\%$ confidence interval about the intraclass correlation,

$$\rho = \frac{\sigma_u^2}{\sigma_e^2 + \sigma_u^2} \quad \text{is} \quad \left[\frac{L^* - 1}{L^* - 1 + n} \leq \rho \leq \frac{U^* - 1}{U^* - 1 + n} \right]$$

where

$$L^* = \frac{Q_1}{Q_2 F_{\alpha/2,t-1,t(n-1)}} \quad \text{and} \quad U^* = \frac{Q_1}{Q_2 F_{1-(\alpha/2),t-1,t(n-1)}}$$

A $(1 - \alpha)100\%$ confidence interval about the ratio σ_u^2/σ_e^2 is given by

$$\left[\frac{L^* - 1}{n} \leq \frac{\sigma_u^2}{\sigma_e^2} \leq \frac{U^* - 1}{n} \right]$$

Burdick and Graybill (1992) also describe methods for constructing confidence intervals about variance components and functions of variance components for both balanced and unbalanced designs when there are two or more variance components in the model. The discussion in this section has been restricted to models with two variance components. The following example is used to demonstrate the computations of these confidence intervals.

20.2.8 Example 20.6 Continued

The data of Example 20.5 are used to demonstrate the computation of the confidence intervals for σ_e^2 , σ_w^2 , σ_w^2/σ_e^2 , and ρ . The two sums of squares are $Q_1 = 32.0$, $Q_2 = 1.6$, with $n = 3$, $t - 1 = 4$, $t(n - 1) = 10$, $F_{0.025,3,10} = 4.46834$, $F_{0.975,3,10} = 0.11307$, $F_{0.025,3,\infty} = 2.78582$, $F_{0.025,10,\infty} = 2.04832$, $F_{0.975,3,\infty} = 0.12110$, $F_{0.975,10,\infty} = 0.32470$, $G_1 = 0.64104$, $G_2 = 0.51179$, $H_1 = 7.25732$, $H_2 = 2.07979$, $G_{12} = -0.11208$, $H_{12} = -8.15882$, $V_L = 426.129$, $V_U = 53515.71$, $L^* = 4.47593$, and $U^* = 176.878$. The 95% confidence intervals for σ_e^2 , σ_w^2 , σ_w^2/σ_e^2 , and ρ are $[0.78113 \leq \sigma_e^2 \leq 4.92767]$, $[3.25237 \leq \sigma_w^2 \leq 87.2449]$, $[4.78382 \leq \sigma_e^2 + \sigma_w^2 \leq 89.1765]$, $[1.15864 \leq (\sigma_w^2/\sigma_e^2) \leq 58.6259]$, and $[0.53675 \leq \rho \leq 0.98323]$, respectively.

The Satterthwaite type confidence intervals from the REML solution are displayed in Table 20.14. Three methods have been used to construct a confidence interval about σ_w^2 . The three 95% confidence intervals are Burdick and Graybill ($3.25237 \leq \sigma_w^2 \leq 87.2449$), REML ($3.4964 \leq \sigma_w^2 \leq 99.2870$), and min–max over the joint confidence region ($2.186 \leq \sigma_w^2 \leq 87.818$). There is not much difference among the three intervals and the differences

TABLE 20.14

SAS-Mixed Code and REML Estimates of the Variance Components for the Worker Data of Example 20.5

```
PROC MIXED data=EX_20_5 METHOD=REML COVTEST CL;
CLASS worker;
MODEL num_units=;
RANDOM worker;
```

Covariance Parameter Estimates

Covariance Parameter	Estimate	Standard Error	Z-Value	Pr Z	α	Lower	Upper
Worker	10.1333	7.5462	1.34	0.0897	0.05	3.4964	99.2870
Residual	1.6000	0.7155	2.24	0.0127	0.05	0.7811	4.9277

would decrease as the number of levels of the random effects corresponding to the variance components increases.

20.3 Simulation Study

Since there are so many options for computing confidence intervals, one can carry out a simulation study using the structure of one's given data to provide information as to how the confidence interval procedures are performing. For example, the information in Table 20.15 consists of SAS data step code to carry out a simulation for the structure of the data set in Example 20.5, where there are five workers with three observations each. The values of the variance components used in the simulation correspond to the REML estimates from Example 20.5 in Table 20.14 (recall that the method of moments estimates are identical to the REML estimates in this case). The simulations showed that 96.63% of the confidence intervals included the true value of σ_w^2 and 95.26% of the confidence intervals include the true value of σ_e^2 . Both of these empirical confidence rates are very close to the 95% coverage that one expects to have. The coverage of each of these Satterthwaite confidence intervals is quite adequate for this size of design when the data are normally distributed.

The sums of squares from the analysis of variance table for an unbalanced model do not necessarily have independent chi-square distributions. Thus, the techniques used for balanced models cannot be applied directly to unbalanced models without violating the assumptions. If the model is not too unbalanced, the sums of squares will nearly be independent chi-square random variables, and the balanced-model techniques should provide effective confidence intervals. For unbalanced models and large sample sizes, the maximum likelihood estimates and their asymptotic properties can be used to construct confidence intervals about the variance components. The above examples illustrate how different confidence intervals can be for any given problem. When the design is unbalanced and the sample size is not large, one suggestion has been to use the widest confidence interval. Another technique to evaluate the effectiveness of a particular method is to use simulation.

TABLE 20.15

Simulation Code and Results for One-Way Random Effects Check on Coverage of Satterthwaite Confidence Intervals

		Frequency	Percentage
sig2w_check	0	337	3.37
	1	9663	96.63
sig2e_check	0	474	4.74
	1	9526	95.26

```

dm 'log; clear; output; clear;';
options nodate center linesize=75 pagesize=55 PAGENO=1;
ods rtf file="c:\amd_i\chapter 20\Simulation of CI REML.rtf" bodytitle;
TITLE "Simulation of data set like EXAMPLE 20.5-1 WAY AOV - Balanced
Random Effects";
%let sig2e=1.6; %let sig2w=10.133;**set the two variances to values;
ods listing exclude all;**do not print following analyses;
ods rtf exclude all;
ods noresults;
data gen;seed=5494812;*select a seed for the random number generator;
    sige=sqrt(&sig2e);**calculate the standard deviations;
    sigw=sqrt(&sig2w);
do nsim=1 to 10000;**generate 10,000 data sets;
    do work=1 to 5;**five workers in each study;
        w=normal(seed)*sigw;**random effect for workers;
        do j=1 to 3;**3 observations per worker;
            num=7+w+normal(seed)*sige;**generate obs with mean 7;
            output;
        end;
    end;
end;
proc mixed data=gen covtest cl;by nsim;**fit model for each data set;
class work; model num=; random work;
ods output covparms=cov;**output cov parameter estimates;run;
ods listing select all;**print the following;
ods rtf select all;
data work; set cov; if covparm='work';
sig2w_check=(lower<=&sig2w<=upper);**check to see if confidence
interval contains sig2w;
proc freq data=work;table sig2w_check;
data resid; set cov; if covparm='Residual';
sig2e_check=(lower<=&sig2e<=upper);**check to see if confidence
interval contains sig2e;
proc freq data=resid;table sig2e_check;
run;

```

20.4 Concluding Remarks

This chapter describes methods to test hypotheses and construct confidence intervals about variance components and functions of the variance components. Testing hypotheses are accomplished using the sums of squares approach where the test statistics are constructed using the expected mean squares or by using likelihood ratio tests. Tests based on the

sums of squares method seem to have better coverage for small sample size experiments than those based on the likelihood ratio. Confidence intervals provided by SAS-Mixed are described as well as other confidence intervals for special situations. The usefulness of a particular confidence interval for a given data structure can be evaluated by using a simulation study as described in Section 20.3. Approximate methods evolving around the Satterthwaite approximation were described for unbalanced models and for some situations with balanced models. Several examples were included to demonstrate the techniques. The confidence intervals provided by SAS-Mixed seem to be adequate for the individual variance components, but specialized methods are needed for confidence intervals for nonlinear functions of the variance components such as the intraclass correlation coefficient or the ratio of two variance components.

20.5 Exercises

20.1 For the coffee price data in Exercise 19.1,

- 1) Provide tests of the following hypotheses

$$H_0: \sigma_{\text{state}}^2 = 0 \text{ vs } H_a: \sigma_{\text{state}}^2 > 0 \quad \text{and} \quad H_0: \sigma_{\text{city}}^2 = 0 \text{ vs } H_a: \sigma_{\text{city}}^2 > 0$$

- 2) Construct confidence intervals about

$$\sigma_{\text{store}}^2, \sigma_{\text{city}}^2, \sigma_{\text{store}}^2 + \sigma_{\text{city}}^2 + \sigma_{\text{state}}^2, \quad \text{and} \quad \sigma_{\text{store}}^2 / (\sigma_{\text{store}}^2 + \sigma_{\text{city}}^2 + \sigma_{\text{state}}^2)$$

20.2 Use the data in Exercise 19.2 and

- 1) Provide tests of the hypotheses $H_0: \sigma_{\text{row}}^2 = 0$ vs $H_a: \sigma_{\text{row}}^2 > 0$, $H_0: \sigma_{\text{col}}^2 = 0$ vs $H_a: \sigma_{\text{col}}^2 > 0$, and $H_0: \sigma_{\text{row} \times \text{col}}^2 = 0$ vs $H_a: \sigma_{\text{row} \times \text{col}}^2 > 0$ using type I, type II, and type III sums of squares. Discuss the differences observed.
- 2) Provide 95% confidence intervals about each of the variance components in the model.

20.3 For the following balanced one-way random effects data set from a completely randomized design structure, provide confidence intervals about σ_{ε}^2 , σ_u^2 , $\sigma_{\varepsilon}^2 + \sigma_u^2$, $\sigma_u^2 / \sigma_{\varepsilon}^2$, and ρ using the methods of Burdick and Graybill described in Section 20.2. The data in the following table consist of the amount of drug recovered from five random samples of livestock feed from each of six randomly selected batches of feed. It is of interest to determine the sample to sample within batch variability caused by the mixing of the drug into the livestock feed and to determine the batch to batch variability.

Sample Within Batch	Batch_1	Batch_2	Batch_3	Batch_4	Batch_5	Batch_6
1	5.75	5.71	5.93	5.87	6.05	6.43
2	5.78	5.69	5.80	5.75	6.04	6.57
3	5.94	5.87	5.73	5.94	5.72	6.52
4	5.97	5.62	5.87	5.95	5.92	6.45
5	5.76	5.85	5.72	6.06	5.88	6.27

21

Case Study: Analysis of a Random Effects Model

The previous three chapters described methods for analyzing random effects models and provided some examples to demonstrate the various techniques of analysis. This chapter presents the analysis of a more complex experimental situation, which includes estimation, model building, hypothesis testing, and confidence interval estimation.

21.1 Data Set

In this experiment, the efficiency of workers in assembly lines at several plants was studied. Three plants were randomly selected from the population of plants owned by the corporation. Four assembly sites and three workers were randomly selected from each plant. Each worker was expected to work five times at each assembly site in her plant, but because of scheduling problems and other priorities, the number of times that each worker actually worked varied from worker to worker. The order in which a worker was to be at each site was randomized and they adhered to that schedule as much as possible. The response variable was a measure of efficiency in assembling parts as a function of the number of units assembled and the number of errors made. The efficiency scores as well as the plant number, site number, and worker number are listed in Table 21.1 where EFF_1, EFF_2, ..., EFF_5 denote the scores for the five possible days that a worker could have worked.

All three factors in the study are random where sites and workers are nested within a plant. Thus the site and worker effects as well as their interaction effect are nested within plant. The model used to describe the data is

$$y_{ijkl} = \mu + p_i + s_{j(i)} + w_{k(i)} + (sw)_{jk(i)} + \varepsilon_{ijkl} \quad (21.1)$$

for

$$i = 1, 2, 3, \quad j = 1, 2, 3, 4, \quad k = 1, 2, 3, \quad \text{and} \quad l = 1, \dots, n_{ijk}.$$

where p_i is the i th plant effect, $s_{j(i)}$ is the j th site effect within plant i , $w_{k(i)}$ is the k th worker effect within plant i , $(sw)_{jk(i)}$ is the interaction effect between site and worker in plant i , and ε_{ijkl} is the residual error term.

It is assumed that

$$\begin{aligned} p_i &\sim i.i.d. N(0, \sigma_p^2), s_{j(i)} \sim i.i.d. N(0, \sigma_s^2), w_{k(i)} \sim i.i.d. N(0, \sigma_w^2) \\ sw_{jk(i)} &\sim i.i.d. N(0, \sigma_{sw}^2), \varepsilon_{ijkl} \sim i.i.d. N(0, \sigma_e^2), \end{aligned}$$

and all the p_i , $s_{j(i)}$, $sw_{jk(i)}$, and ε_{ijkl} are independent random variables.

TABLE 21.1

The Data for the Case Study Where EFF_ i Denotes the i th Time a Worker Worked at a Site

Plant	Site/Plant	Worker/Plant	EFF_1	EFF_2	EFF_3	EFF_4	EFF_5
1	1	1	100.6	106.8	100.6	—	—
1	1	2	92.3	92.0	97.2	93.9	93.0
1	1	3	96.9	96.1	100.8	—	—
1	2	1	110.0	105.8	—	—	—
1	2	2	103.2	100.5	100.2	97.7	—
1	2	3	92.5	85.9	85.2	89.4	88.7
1	3	1	100.0	102.5	97.6	98.7	98.7
1	3	2	96.4	—	—	—	—
1	3	3	86.8	—	—	—	—
1	4	1	98.2	99.5	—	—	—
1	4	2	108.0	108.9	107.9	—	—
1	4	3	94.4	93.0	91.0	—	—
2	5	4	82.6	—	—	—	—
2	5	5	72.7	—	—	—	—
2	5	6	82.5	82.1	82.0	—	—
2	6	4	96.5	100.1	101.9	97.9	95.9
2	6	5	71.7	72.1	72.4	71.4	—
2	6	6	80.9	84.0	82.2	83.4	81.5
2	7	4	87.9	93.5	88.9	92.8	—
2	7	5	78.4	80.4	83.8	77.7	81.2
2	7	6	96.3	92.4	92.0	95.8	—
2	8	4	83.6	82.7	87.7	88.0	82.5
2	8	5	82.1	79.9	81.9	82.6	78.6
2	8	6	77.7	78.6	77.2	78.8	80.5
3	9	7	107.6	108.8	107.2	104.2	105.4
3	9	8	97.1	94.2	91.5	99.2	—
3	9	9	87.1	—	—	—	—
3	10	7	96.1	98.5	97.3	93.5	—
3	10	8	91.9	—	—	—	—
3	10	9	97.8	95.9	—	—	—
3	11	7	101.1	—	—	—	—
3	11	8	88.0	91.4	90.3	91.5	85.7
3	11	9	95.9	89.7	—	—	—
3	12	7	109.1	—	—	—	—
3	12	8	89.6	86.0	91.2	87.4	—
3	12	9	101.4	100.1	102.1	98.4	—

21.2 Estimation

The method of moments procedure, with type I–III sums of squares, REML, maximum likelihood, and MIVQUE0 with and without the NOBOUND methods, is used to obtain solutions for the variance components and the results are displayed in Table 21.2. The SAS®-Mixed code used to obtain the REML estimates is also included where the other estimation techniques can be selected by incorporating the specific name with the Method = option. The expected mean squares of the type I and III sums of squares are given in Tables 21.3 and 21.4, which can be used to construct the equations required to obtain the method of moments estimators of each of the variance components. The REML, ML, and MIVQUE0 estimates for σ_s^2 are all equal to zero while the MIVQUE0 with the nobound option as well as the type I–III solutions for σ_s^2 are negative. The MIVQUE0 solution is obtained by setting the solution for σ_s^2 equal to zero from the MIVQUE0 solution without the nobound option. The estimates of the variance components are obtained by converting the negative solutions to 0; that is, the estimate of σ_s^2 is $\hat{\sigma}_s^2 = 0$. Since the estimate

TABLE 21.2

Solutions of the Variance Components for Model 21.1 Using Each of the Methods Available in Proc Mixed, Where MVQ Denotes MIVQUE0 and MVQNB Denotes MIVQUE with the No-Bound Option. Proc Mixed Code Is for Method = REML

```
proc mixed data=EX_21_1 method=reml cl covtest;
  class plant worker site;
  model efficiency=;
  random plant worker(plant) site(plant) site*worker(plant);
```

Covariance Parameter	ML	REML	MVQ	MVQNB	Type I	Type II	Type III
Plant σ_p^2	29.6011	50.4898	54.2064	54.2064	48.8063	47.8649	59.6007
Worker(plant) σ_w^2	28.9381	28.9451	24.2302	24.2302	24.2549	27.0455	24.9422
Site(plant) σ_s^2	0.0000	0.0000	0.0000	-4.4155	-4.8780	-4.8780	-4.0644
Worker \times site(plant) σ_{sw}^2	28.7593	28.7707	29.8389	29.8389	35.6167	35.6167	35.6167
Residual σ_e^2	4.9825	4.9818	6.6163	6.6163	4.9831	4.9831	4.9831

TABLE 21.3

Type I Analysis of Variance Table with Expected Mean Squares

Type I Analysis of Variance

Source	df	Mean Square	Expected Mean Square
Plant	2	2313.758960	$\text{Var}(\text{Residual}) + 3.9277 \text{Var}[\text{worker} \times \text{site}(\text{plant})] + 10.307 \text{Var}[\text{site}(\text{plant})] + 13.136 \text{Var}[\text{worker}(\text{plant})] + 38.941 \text{Var}(\text{plant})$
Worker(plant)	6	456.941927	$\text{Var}(\text{Residual}) + 3.9015 \text{Var}[\text{worker} \times \text{site}(\text{plant})] + 0.6563 \text{Var}[\text{site}(\text{plant})] + 13.037 \text{Var}[\text{worker}(\text{plant})]$
Site(plant)	9	84.049170	$\text{Var}(\text{Residual}) + 3.4789 \text{Var}[\text{worker} \times \text{site}(\text{plant})] + 9.1928 \text{Var}[\text{site}(\text{plant})]$
Worker \times site(plant)	18	106.738256	$\text{Var}(\text{Residual}) + 2.8569 \text{Var}[\text{worker} \times \text{site}(\text{plant})]$
Residual	82	4.983134	$\text{Var}(\text{Residual})$

TABLE 21.4

Type III Analysis of Variance Table with Expected Mean Squares

Type III Analysis of Variance

Source	df	Mean Square	Expected Mean Square
Plant	2	1933.164549	$\text{Var}(\text{Residual}) + 2.2998 \text{Var}[\text{worker} \times \text{site}(\text{plant})] + 6.8995 \text{Var}[\text{site}(\text{plant})] + 9.1993 \text{Var}[\text{worker}(\text{plant})] + 27.598 \text{Var}(\text{plant})$
Worker(plant)	6	324.942651	$\text{Var}(\text{Residual}) + 2.3633 \text{Var}[\text{worker} \times \text{site}(\text{plant})] + 9.4533 \text{Var}[\text{worker}(\text{plant})]$
Site(plant)	9	67.811366	$\text{Var}(\text{Residual}) + 2.6823 \text{Var}[\text{worker} \times \text{site}(\text{plant})] + 8.0468 \text{Var}[\text{site}(\text{plant})]$
Worker \times site(plant)	18	106.738256	$\text{Var}(\text{Residual}) + 2.8569 \text{Var}[\text{worker} \times \text{site}(\text{plant})]$
Residual	82	4.983134	$\text{Var}(\text{Residual})$

of σ_s^2 is zero, the next step would be to test hypotheses about the variance components in the model. A stepwise deletion process can be used to remove random components from the model in an attempt to obtain a simpler model that describes the variance in the process. Deleting a term from the model essentially sets the removed variance component equal to zero. Such *model building can occur for any variance components corresponding to factors in the treatment structure*, but there will be *no model building for variance components corresponding to factors in the design structure*.

21.3 Model Building

The analysis of variance table, based on type I sums of squares, is used to build a model by testing hypotheses about the variance components. The process starts by investigating the variance component associated with the last sum of squares and then working up the line to the variance component associated with the first sum of squares. The null hypothesis is that the variance component is equal to zero. The observed and expected mean squares for the type I sums of squares method are listed in Table 21.3.

The first step is to test $H_0: \sigma_{sw}^2 = 0$ vs $H_a: \sigma_{sw}^2 > 0$. On inspecting the expected mean squares in Table 21.3, the Residual is determined to be the appropriate divisor, thus, the test statistic is

$$F_{C_{sw}} = \frac{MS[\text{worker} \times \text{site}(\text{plant})]}{MS(\text{Residual})} = 21.42 \text{ (which is in Table 21.5)}$$

The sampling distribution of $F_{C_{sw}}$ is compared with the critical point from an F -distribution with 18 degrees of freedom for the numerator and 82 degrees of freedom for the denominator. The observed significance level is less than 0.0001, which indicates that σ_{sw}^2 is an important source of variability in the process.

The second step is to test $H_0: \sigma_s^2 = 0$ vs $H_a: \sigma_s^2 > 0$. The $E\{MS[\text{site}(\text{plant})]\}$ equals $\sigma_e^2 + 3.4789\sigma_{sw}^2 + 9.1928\sigma_s^2$, and no other mean square exists with expectation $\sigma_e^2 + 3.4789\sigma_{sw}^2$ that could be used as a divisor for the test statistic. Thus, a mean square Q_s^* needs to be

constructed from $MS(Residual)$ and $MS[worker \times site(plant)]$ such that $E(Q_s^*) = \sigma_e^2 + 3.4789\sigma_{sw}^2$. Such a Q_s^* is

$$\begin{aligned} Q_s^* &= 3.4789 \left[\frac{MS[worker \times site(plant)]}{2.8569} \right] + \left[1 - \frac{3.4989}{2.8569} \right] MS(Residual) \\ &= 1.2177 MS[worker \times site(plant)] - 0.2177 MS(Residual) \\ &= 128.921 \end{aligned}$$

A chi-square distribution can be used to approximate the sampling distribution of $r_s Q_s^*/(\sigma_e^2 + 3.4789\sigma_{sw}^2)$ by using the Satterthwaite approximation to determine the associated degrees of freedom as

$$\begin{aligned} r_s &= \frac{(Q_s^*)^2}{\frac{\{1.2177 MS[worker \times site(plant)]\}^2}{18} + \frac{[0.2177 MS(Residual)]}{82}} \\ &= \frac{(128.921)^2}{\frac{[1.2177 \times 106.738]^2}{18} + \frac{[0.2177 \times 4.983]^2}{82}} = 17.7 \end{aligned}$$

The test statistic is

$$F_{C_s} = \frac{MS[site(plant)]}{Q_s^*} = 0.65 \text{ (see Table 21.5)}$$

which is approximately distributed as an F -distribution with degrees of freedom 9 and 17.7. The observed significance level of the test is 0.7396, which indicates that σ_s^2 is a negligible source of variation in the process; that is, one fails to reject $H_0: \sigma_s^2 = 0$ vs $H_a: \sigma_s^2 > 0$. Tables 21.4 and 21.6 contain the mean squares, expected mean squares, appropriate error terms, approximate denominator degrees of freedom, and test statistics based on the type III sums of squares. The approximate F -statistic to test $H_0: \sigma_s^2 = 0$ vs $H_a: \sigma_s^2 > 0$ has a value of 0.67 with estimated denominator degrees of freedom 18.1 and an observed

TABLE 21.5

Type I Tests of Hypotheses about the Variance Components

Type I Analysis of Variance

Source	df	Error Term	Error df	F-Value	Pr > F
Plant	2	$1.0076 MS[worker(plant)] + 1.0493 MS[site(plant)] - 1.279 MS[worker \times site(plant)] + 0.2221 MS(Residual)$	4.5859	5.60	0.0588
Worker(plant)	6	$0.0714 MS[site(plant)] + 1.2787 MS[worker \times site(plant)] - 0.3501 MS(Residual)$	19.066	3.25	0.0227
Site(plant)	9	$1.2177 MS[worker \times site(plant)] - 0.2177 MS(Residual)$	17.701	0.65	0.7396
Worker \times site(plant)	18	$MS(Residual)$	82	21.42	<0.0001
Residual	82	—	—	—	—

TABLE 21.6

Type III Tests of Hypotheses about the Variance Components

Type III Analysis of Variance

Source	df	Error Term	Error df	F-Value	Pr > F
Plant	2	$0.9731 \text{MS}[worker(plant)] + 0.8574 \text{MS}[site(plant)] - 0.805 \text{MS}[worker \times site(plant)] - 0.0256 \text{MS}(Residual)$	4.7631	6.71	0.0412
Worker(plant)	6	$0.8272 \text{MS}[worker \times site(plant)] + 0.1728 \text{MS}(Residual)$	18.352	3.64	0.0148
Site(plant)	9	$0.9389 \text{MS}[worker \times site(plant)] + 0.0611 \text{MS}(Residual)$	18.11	0.67	0.7217
Worker \times site(plant)	18	$\text{MS}(Residual)$	82	21.42	<0.0001
Residual	82	—	—	—	—

significance level of 0.7217. The conclusion is the same for σ_s^2 from the type III analysis, as was obtained from the type I analysis. Since $site(plant)$ is part of the treatment structure and σ_s^2 is negligible, one strategy is to set σ_s^2 equal to zero, eliminating $s_{j(l)}$ from the model and fitting a reduced model to the data.

21.4 Reduced Model

The reduced model is

$$y_{ijkl} = \mu + p_i + w_{k(i)} + (sw)_{jk(i)} + \varepsilon_{ijkl} \quad (21.2)$$

$$i = 1, 2, 3, j = 1, 2, 3, 4, k = 1, 2, 3, l = 1, \dots, n_{ijk}$$

When $s_{j(i)}$ is eliminated from the model, SAS-Mixed pools the sum of squares due to $site(plant)$ with the sum of squares $worker \times site(plant)$ as indicated by the 27 degrees of freedom in Table 21.7 for $worker \times site(plant)$. This is a reasonable process since the expected mean squares of $MS[site(plant)]$ and $MS[worker \times site(plant)]$ from Tables 21.3 and 21.4 are similar when $\sigma_s^2 = 0$; that is, $E\{MS[site(plant)]\} = \sigma_e^2 + 3.4789\sigma_{sw}^2$ and $E\{MS[worker \times site(plant)]\} = \sigma_e^2 + 2.8569\sigma_{sw}^2$ in Table 21.3, and the coefficients are even closer in the type III analysis.

The degrees of freedom, type I mean squares, and expected mean squares for the reduced model, are listed in Table 21.7. The expected mean square for $worker \times site(plant)$ is $\sigma_e^2 + 3.0643\sigma_{sw}^2$. The coefficient 3.0643 is computed as $3.0643 = [(9 \times 3.4789) + (18 \times 2.8561)]/27$ which is the pooled coefficients of σ_{sw}^2 from the expected mean squares of $site(plant)$ and $worker \times site(plant)$ from the type I analysis. This equivalence occurs since the type I sums of squares are sequential and thus, the pooling process is additive. This type of phenomenon does not occur with other types of sums of squares, as can be observed by comparing the type III analyses in Tables 21.4, 21.6, and 21.8. The estimates of the variance components for the reduced model using REML, ML MIVQUE0, type I-III methods are displayed in Table 21.9. The estimates of the residual variance are approximately equal to 4.98 for all methods except for MIVIQUE0, which yields the largest value of 6.61. The ML estimate of σ_p^2 is the smallest at 29.6 and the type III estimate is the largest at 58.6. The other methods range from 47.6 to 53.1. Since there are only three levels of plant, σ_p^2 is the hardest variance component to estimate; that is, less is known about σ_p^2 than the other variance components

TABLE 21.7

Type I Analysis of Variance Table with Expected Mean Squares for the Reduced Model

Type I Analysis of Variance

Source	df	Mean Square	Expected Mean Square
Plant	2	2313.758960	$\text{Var}(\text{Residual}) + 3.9277 \text{Var}[\text{worker} \times \text{site}(\text{plant})] + 13.136 \text{Var}[\text{worker}(\text{plant})] + 38.941 \text{Var}(\text{plant})$
Worker(plant)	6	456.941927	$\text{Var}(\text{Residual}) + 3.9015 \text{Var}[\text{worker} \times \text{site}(\text{plant})] + 13.037 \text{Var}[\text{worker}(\text{plant})]$
Worker \times site(plant)	27	99.175228	$\text{Var}(\text{Residual}) + 3.0643 \text{Var}[\text{worker} \times \text{site}(\text{plant})]$
Residual	82	4.983134	$\text{Var}(\text{Residual})$

TABLE 21.8

Type III Analysis of Variance Table with Expected Mean Squares for the Reduced Model

Type III Analysis of Variance

Source	df	Mean Square	Expected Mean Square
Plant	2	1933.164549	$\text{Var}(\text{Residual}) + 2.2998 \text{Var}[\text{worker} \times \text{site}(\text{plant})] + 9.1993 \text{Var}[\text{worker}(\text{plant})] + 27.598 \text{Var}(\text{plant})$
Worker(plant)	6	324.942651	$\text{Var}(\text{Residual}) + 2.3633 \text{Var}[\text{worker} \times \text{site}(\text{plant})] + 9.4533 \text{Var}[\text{worker}(\text{plant})]$
Worker \times site(plant)	27	99.175228	$\text{Var}(\text{Residual}) + 3.0643 \text{Var}[\text{worker} \times \text{site}(\text{plant})]$
Residual	82	4.983134	$\text{Var}(\text{Residual})$

TABLE 21.9

Estimates of the Variance Components for the Reduced Model

Covariance Parameters	ML	REML	MQV	Type I	Type II	Type III
Plant σ_p^2	29.6002	50.4898	53.1030	47.5975	47.5975	58.5846
Worker(plant) σ_w^2	28.9415	28.9453	25.3453	25.4692	25.4692	26.1617
Worker \times site(plant) σ_{sw}^2	28.7582	28.7706	25.4195	30.7388	30.7388	30.7388
Residual σ_e^2	4.9825	4.9818	6.6119	4.9831	4.9831	4.9831

in the model. Estimates of σ_w^2 and σ_{sw}^2 are quite consistent across the different methods of estimation with $\hat{\sigma}_w^2$ ranging from 25.5 to 28.9 and $\hat{\sigma}_{sw}^2$ ranging from 25.4 to 30.7.

The analysis can continue by carrying out tests of hypotheses about the remaining variance components to determine whether the model can be simplified further. The statistic to test $H_0: \sigma_{sw}^2 = 0$ vs $H_a: \sigma_{sw}^2 > 0$ is

$$F_{C_{sw}} = \frac{MS[\text{worker} \times \text{site}(\text{plant})]}{MS(\text{Residual})} = 19.90$$

which is distributed as F -distribution with 27 and 82 degrees of freedom. The observed significance level is less than 0.0001. Thus, σ_{sw}^2 is an important source of variance in the process of generating the data and $(sw)_{jk(i)}$ should remain in the model. Again, σ_{sw}^2 is the adaptation variance component where some workers work more effectively at some sites,

while other workers work more effectively at different sites. Tailoring each site to its workers' needs can help reduce the variability in the system.

A Q_w^* needs to be computed to test $H_0: \sigma_w^2 = 0$ vs $H_a: \sigma_w^2 > 0$. The mean square Q_w^* needs to be constructed from $MS[worker \times site(plant)]$ and $MS(Residual)$ such that $E(Q_w^*) = \sigma_e^2 + 3.9015\sigma_{sw}^2$. Then

$$\begin{aligned} Q_w^* &= 3.9015 \left[\frac{MS[site \times worker(plant)]}{3.0643} \right] + \left[1 - \frac{3.9015}{3.0643} \right] MS(Residual) \\ &= 1.2732MS[site \times worker(plant)] - 0.2732MS(Residual) \\ &= 124.9085 \end{aligned}$$

The test statistic is $F_{C_w} = 456.9419/124.9085 = 3.66$, which is distributed as an F -distribution with 6 and r_w degrees of freedom where

$$\begin{aligned} r_w &= \frac{(Q_w^*)}{\{1.2732MS[site \times worker(plant)]\}^2/27 + [0.2732 \times MS(Residual)]^2/82} \\ &= \frac{(124.9085)^2}{[1.2732 \times 99.1752]^2/27 + [0.2732 \times 4.9831]^2/82} \\ &= 26.42 \end{aligned}$$

The observed significance level for this test statistic is 0.0089, indicating that σ_w^2 is an important part of the variation in the system.

Finally, another Q_p^* needs to be constructed from $MS[worker(plant)]$, $MS[worker \times site(plant)]$, and $MS(Residual)$ such that

$$E(Q_p^*) = \sigma_e^2 + 3.9277\sigma_{sw}^2 + 13.136\sigma_w^2$$

so that a statistic can be constructed to test $H_0: \sigma_p^2 = 0$ vs $H_a: \sigma_p^2 > 0$. The required Q_p^* is

$$\begin{aligned} Q_p^* &= \frac{13.136}{13.037} \{MS[worker(plant)]\} \\ &\quad + \left[\frac{3.9277 - \frac{13.136}{13.037}(3.0915)}{3.0643} \right] \{MS[site \times worker(plant)]\} \\ &\quad + \left[1 - \frac{13.136}{13.037} - \frac{3.9277 - \frac{13.136}{13.037}(3.0915)}{3.0643} \right] MS(Residual) \\ &= 1.0076MS[worker(plant)] - 0.0012MS[site \times worker(plant)] - 0.0065MS(Residual) \\ &= 460.2638 \end{aligned}$$

The test statistic is

$$F_{C_p} = \frac{MS(plant)}{Q_p^*} = 5.027$$

which is distributed as an F -distribution with 2 and r_p degrees of freedom where $r_p = (Q_p^*)^2/D_p$ and

$$D_p = \frac{\{1.0076 MS[worker(plant)]\}^2}{6} + \frac{\{0.0012 MS[site \times worker(plant)]\}^2}{27} + \frac{[0.0065 MS(Residual)]^2}{82} = 35,330.281$$

Then $r_p = 5.9961$. The observed significance level associated with this test is 0.0522, which indicates that σ_p^2 is an important contributor to the variation in the system, but it is not as important as either σ_{sw}^2 or σ_w^2 . All of the variance components in model (21.2) are significantly different from zero at $\alpha < 0.10$ and the model cannot be simplified further.

The results in Tables 21.8 and 21.10 are from the type III sums of squares. SAS-Mixed uses the expected mean squares to compute the appropriate divisors and the Satterthwaite approximation to compute the denominator degrees of freedom. The significance levels from the type III sums of squares analysis for testing $H_0: \sigma_p^2 = 0$ vs $H_a: \sigma_p^2 > 0$, $H_0: \sigma_w^2 = 0$ vs $H_a: \sigma_w^2 > 0$, and $H_0: \sigma_{sw}^2 = 0$ vs $H_a: \sigma_{sw}^2 > 0$ are 0.0357, 0.0040, and less than 0.0001, respectively. The significance levels from the type III analysis are a little smaller than those from the type I analysis (for the first two tests), but in other problems the significance levels of the type III analysis can be a little greater than those from the type I analysis. Table 21.11 contains the results from the REML analysis of the reduced model. The significance levels associated with the Z-value test statistics are 0.2118, 0.0869, and 0.0003 for the *plant*, *worker(plant)* and *worker* \times *site(plant)* variance components, respectively. These significance levels are considerably larger than those from the type I and III sums of squares methods. This occurs because the Z-test is asymptotically distributed as a normal random variable, and in these cases, the numbers of degrees of freedom associated with each of the variance component estimates are small. Unless there are a lot of levels associated with an estimate of a variance component, the information associated with the Z-value is not useful for testing hypotheses.

TABLE 21.10

Type III Tests of Hypotheses about the Variance Components for the Reduced Model

Type III Analysis of Variance

Source	df	Error Term	Error df	F-Value	Pr > F
Plant	2	0.9731 $MS[worker(plant)] + 0.0269 MS(Residual)$	6.0051	6.11	0.0357
Worker(plant)	6	0.7713 $MS[worker \times site(plant)] + 0.2287 MS(Residual)$	27.809	4.19	0.0040
Worker \times site(plant)	27	$MS(Residual)$	82	19.90	<0.0001
Residual	82	—	—	—	—

TABLE 21.11

REML Estimates of the Variance Components for the Reduced Model. Satterthwaite-Type Confidence Intervals for the Variance Components

```
proc mixed data=EX_21_1 METHOD=REML cl covtest;
  class plant worker site;
  model efficiency=;
  random plant worker(plant) site*worker(plant);
```

Covariance Parameter Estimates

Covariance Parameter	Estimate	Standard Error	Z-Value	Pr Z	α	Lower	Upper
Plant	50.4898	63.0887	0.80	0.2118	0.05	11.2442	12100
Worker(plant)	28.9453	21.2802	1.36	0.0869	0.05	10.0867	271.39
Worker \times site(plant)	28.7706	8.3753	3.44	0.0003	0.05	17.4775	56.0425
Residual	4.9818	0.7776	6.41	<0.0001	0.05	3.7506	6.9406

The *plant* variance component measures the differences among plants within the population of plants. Since there is a large amount of variability, one of the plants most likely has procedures in place that enable the workers to work more efficiently. The importance of the *worker(plant)* variance component indicates that some workers are more efficient than others, so a training program could help improve the efficiency of workers not performing as well as others. The *worker \times site(plant)* variance component is an adaptation variance component that means that some workers are more adapted to perform at some sites and not as well at other sites where, on the other hand, some other workers do well at the sites the previous workers did not.

21.5 Confidence Intervals

The next step in the analysis is to construct confidence intervals about the variance components of the reduced model in Equation 21.1. The results displayed in Tables 21.11 and 21.12 are the estimates of the variance components for the reduced model using REML and type I sums of squares, respectively. The confidence intervals computed for the REML solution are computed using the method in Section 20.2.2 with $df = 2(Z\text{-value})^2$. The degrees of freedom are 1.28 for $\hat{\sigma}_p^2$, 3.70 for $\hat{\sigma}_w^2$, 23.60 for $\hat{\sigma}_{sw}^2$ and 82.10 for $\hat{\sigma}_e^2$. Table 21.13 contains the computed degrees of freedom for the REML confidence intervals as well as intervals that are expressed in standard deviation units. The confidence intervals computed for the type I analysis are the Wald intervals of Section 20.3.4 (as noted by the fact that the intervals are symmetric about the estimate and some of the lower limits are negative), except for the σ_e^2 , where the interval is computed using the results in Section 20.2.2 based on 81.93 degrees of freedom. Confidence intervals based on the type I sums of squares can be recomputed using the method of Section 20.2.2 with $df = 2(Z\text{-value})^2$. For $\hat{\sigma}_p^2$, the resulting number of degrees of freedom is 1.4677, providing a 95% confidence interval as $[11.2668 \leq \sigma_p^2 \leq 5984.61]$. For $\hat{\sigma}_w^2$, the resulting number of degrees of freedom is 3.9682, providing a 95% confidence interval as $[9.1149 \leq \sigma_w^2 \leq 212.97]$. For $\hat{\sigma}_{sw}^2$, the resulting number of degrees of freedom is

20.8392, providing a 95% confidence interval as $[18.1625 \leq \sigma_{sw}^2 \leq 62.98]$. The results of these recomputed confidence intervals are summarized in Table 21.14.

Next the usual Satterthwaite approximation is used to provide confidence intervals about the variance components using the information from the type I sums of squares. The method in Section 20.2.1 is used construct a confidence interval about σ_e^2 . A 95% confidence interval about σ_e^2 is

$$\frac{82[MS(Residual)]}{\chi_{(\alpha/2),82}^2} \leq \sigma_e^2 \leq \frac{82[MS(Residual)]}{\chi_{1-(\alpha/2),82}^2}$$

TABLE 21.12

Type I Sums of Squares Estimates of Variance Components for the Reduced Model. Wald-Type Confidence Intervals for Variances Other than Residual

Covariance Parameter Estimates							
Covariance Parameter	Estimate	Standard Error	Z-Value	Pr Z	α	Lower	Upper
Plant	47.5975	55.5617	0.86	0.3916	0.05	-61.3015	156.50
Worker(plant)	25.4692	18.0815	1.41	0.1590	0.05	-9.9699	60.9083
Worker \times site(plant)	30.7388	9.5227	3.23	0.0012	0.05	12.0746	49.4030
Residual	4.9831	0.7786	6.40	<0.0001	0.05	3.7509	6.9440

TABLE 21.13

REML Estimates of the Variance Components with the Degrees of Freedom Used in the Computations. The Standard Deviations and Square Roots of the Intervals Are Included

Covariance Parameter	Estimate	Lower	Upper	df	SD	Low SD	High SD
Plant	50.4898	11.2442	12100	1.2810	7.10562	3.35323	109.998
Worker(plant)	28.9453	10.0867	271.39	3.7003	5.38009	3.17596	16.474
Worker \times site(plant)	28.7706	17.4775	56.0425	23.6009	5.36382	4.18061	7.486
Residual	4.9818	3.7506	6.9406	82.1009	2.23200	1.93664	2.635

TABLE 21.14

Recomputed Confidence Intervals Using General Satterthwaite Approximation in Section 20.2.2 Based on Type I Sums of Squares

Covariance Parameter	Estimate	df	New Low	New Up	SD	Low SD	High SD
Plant	47.5975	1.4677	11.2668	5984.61	6.89909	3.35661	77.3602
Worker(plant)	25.4692	3.9682	9.1149	212.97	5.04670	3.01909	14.5936
Worker \times site(plant)	30.7388	20.8392	18.1625	62.98	5.54426	4.26175	7.9360
Residual	4.9831	81.9300	3.7505	6.95	2.23229	1.93662	2.6353

or

$$\frac{82(4.9831)}{108.937} \leq \sigma_{\epsilon}^2 \leq \frac{82(4.9831)}{58.8446}$$

or

$$3.781 \leq \sigma_{\epsilon}^2 \leq 6.944$$

The above 95% confidence interval about σ_{ϵ}^2 is identical to that in Table 21.11 for the *Residual*.

For the Satterthwaite confidence interval about σ_{sw}^2 using the expected mean squares from Table 21.7, the method of moments solution (and estimate) of σ_{sw}^2 can be expressed as

$$\hat{\sigma}_{sw}^2 = \frac{MS[worker \times site(plant)]}{3.064} - \frac{MS(Residual)}{3.064} = 30.7388$$

Then $r_{sw}\hat{\sigma}_{sw}^2/\sigma_{sw}^2$ is approximately distributed as a chi-square random variable with r_{sw} degrees of freedom where

$$r_{sw} = \frac{(\hat{\sigma}_{sw}^2)^2}{\left[\frac{MS[worker \times site(plant)]}{3.064} \right] / 27 + \left[\frac{MS(Residual)}{3.064} \right] / 82} = 24.33$$

An approximate 95% confidence interval about σ_{sw}^2 is

$$\frac{r_{sw}\hat{\sigma}_{sw}^2}{\chi_{025, 24.33}^2} \leq \sigma_{sw}^2 \leq \frac{r_{sw}\hat{\sigma}_{sw}^2}{\chi_{975, 24.33}^2}$$

or

$$18.80 \leq \sigma_{sw}^2 \leq 59.17$$

To construct the confidence interval about σ_w^2 , express the solution for $\hat{\sigma}_w^2$ as (obtain information from the *Error* term of Table 21.15)

$$\begin{aligned} \hat{\sigma}_w^2 &= \{MS(worker) - 1.2732MS[worker \times site(plant)] + 0.2732MS(Residual)\} / 13.037 \\ &= 25.4692 \end{aligned}$$

The approximate degrees of freedom associated with $\hat{\sigma}_w^2$ is $r_w = 3.12$, which leads to an approximate 95% confidence interval as $[8.297 \leq \sigma_w^2 \leq 327.87]$.

To obtain a 95% confidence interval about σ_p^2 one can use the results in Table 21.15 to express the estimate for σ_p^2 as

$$\begin{aligned} \hat{\sigma}_p^2 &= \{MS(plant) - 1.0076MS[worker(plant)] + 0.0012MS[worker \times site(plant)] \\ &\quad + 0.0065MS(Residual)\} / 38.941 \\ &= 47.60 \end{aligned}$$

TABLE 21.15

Type I Tests of Hypotheses about the Variance Components for the Reduced Model

Type I Analysis of Variance

Source	df	Error Term	Error df	F-Value	Pr > F	
Plant	2	$1.0076 MS[worker(plant)] - 0.0012 MS[worker \times site(plant)] - 0.0065 MS(Residual)$		5.9961	5.03	0.0522
Worker(plant)	6	$1.2732 MS[worker \times site(plant)] - 0.2732 MS(Residual)$		26.42	3.66	0.0089
Worker \times site(plant)	27	$MS(Residual)$		82	19.90	<0.0001
Residual	82	—	—	—	—	

The number of degrees of freedom associated with $\hat{\sigma}_p^2$ is $r_p = 1.27$. The resulting approximate 95% confidence interval is $[10.58 \leq \sigma_p^2 \leq 12078.74]$. The set of confidence intervals computed using the Satterthwaite approximation is displayed in Table 21.16.

The final stage in the analysis is to use the confidence region method in Section 20.2 to construct confidence intervals about the variance components, σ_p^2 , σ_w^2 , and σ_{sw}^2 . A 95% confidence interval about σ_{sw}^2 has endpoints c_{sw} and d_{sw} where

$$c_{sw} = \frac{27 MS[worker \times site(plant)]/\chi_{.025,27}^2 - 82 MS(Residual)/\chi_{.975,82}^2}{3.064} = 17.96$$

and

$$d_{sw} = \frac{27 MS[worker \times site(plant)]/\chi_{.975,27}^2 - 82 MS(Residual)/\chi_{.025,82}^2}{3.064} = 58.74$$

Simplifying these two expression one gets the confidence interval $17.96 \leq \sigma_{sw}^2 \leq 58.74$.

The two endpoints of a 95% confidence interval about σ_w^2 are c_w and d_w where

$$c_w = \frac{6 MS[worker(plant)]/\chi_{.025,6}^2 - r_w Q_w^*/\chi_{.975,26.42}^2}{13.037} = -3.335$$

TABLE 21.16

Confidence Intervals Based on the Usual Satterthwaite Approximation Using Type I Sums of Squares

Source	Estimate	RDF	Low s2	High s2	SD	Low SD	High SD
Plant	47.5975	1.2667	10.5468	12078.74	6.89910	3.24759	109.903
Worker(plant)	25.4685	3.1152	8.2972	327.87	5.04664	2.88048	18.107
Worker \times site(plant)	30.7385	24.3347	18.7971	59.17	5.54423	4.33557	7.692
Residual	4.9831	82.0000	3.7509	6.94	2.23229	1.93673	2.635

and

$$d_w = \frac{6MS[\text{worker}(plant)]/\chi^2_{975,6} - r_w Q_p^*/\chi^2_{025,26.42}}{13.037} = 164.00$$

where $r_w = 26.42$ and $Q_p^* = 124.9085$. Simplifying these expressions gives the interval, $0 \leq \sigma_p^2 \leq 164.00$.

The endpoints for a 95% confidence interval about σ_p^2 are c_p and d_p where

$$c_p = \frac{2MS(\text{plant})]/\chi^2_{025,2} - r_p Q_p^*/\chi^2_{975,5.9961}}{38.941} = -41.2471$$

and

$$d_p = \frac{2MS(\text{plant})]/\chi^2_{975,2} - r_p Q_p^*/\chi^2_{025,5.9961}}{38.941} = 2341.94$$

where $r_p = 5.9961$ and $Q_p^* = 460.2638$. Simplifying gives the interval, $0 \leq \sigma_p^2 \leq 2431.94$.

Table 21.17 displays the confidence intervals computed using the confidence region method.

All four of the above methods of computing confidence intervals are summarized in Table 21.18, where the entries in the tables are the lower and upper limits for the respective variance components.

TABLE 21.17

Summary of Confidence Intervals Computed Using the Confidence Region Method of Section 20.2.5

Source	Low v	High v	Low SD	High SD
Plant	0.0000	2341.94	0.00000	48.3936
Worker(plant)	0.0000	164.00	0.00000	12.8061
Worker \times site(plant)	17.9644	58.74	4.23845	7.6641

TABLE 21.18

Summary of the Four Types of 95% Confidence Intervals, with Upper and Lower Limits

Source	REML	General Satterthwaite	Usual Satterthwaite	Confidence Region
Plant	11.24	12,100	11.27	5984.61
Worker	10.09	271.39	9.11	213.00
Worker \times site	17.48	56.04	18.16	62.98
Residual	3.75	6.94	3.75	6.94
			10.56	12079
			8.30	327.87
			18.80	59.17
			3.75	6.94
				17.96
				58.74
				3.75
				6.94

21.6 Computations Using JMP

Figure 21.1 contains a proportion of the JMP® data table for Example 21 where plant, site and worker are specified to be nominal variables. Figure 21.2 contains the fit model screen where *plant*, *site(plant)*, *worker(plant)*, and *site* \times *worker(plant)* are specified as random effects and the REML method is selected. The estimates of the variance components for the full model are shown in Figure 21.3, including the estimated standard errors and 95% confidence intervals, which correspond to the REML estimates in Table 21.2. The reduced model fit model screen is displayed in Figure 21.4, where the *site(plant)* term was removed from the fit model screen in Figure 21.2. The REML estimates from JMP in Figure 21.5 are identical to those from SAS-Mixed that were displayed in Table 21.11. The fit model process in JMP provides excellent results for random effects models.

	plant	worker	site	efficiency
1	1	1	1	100.6
2	1	1	1	106.8
3	1	1	1	100.6
4	1	1	2	110
5	1	1	2	105.8
6	1	1	3	100
7	1	1	3	102.5
8	1	1	3	97.6
9	1	1	3	98.7
10	1	1	3	98.7
11	1	1	4	98.2
12	1	1	4	99.5
13	1	2	1	92.3
14	1	2	1	92
15	1	2	1	97.2
16	1	2	1	93.9
17	1	2	1	93
18	1	2	2	103.2
19	1	2	2	100.5
20	1	2	2	100.2
21	1	2	2	97.7

FIGURE 21.1 Part of the JMP table of data for Example 21.

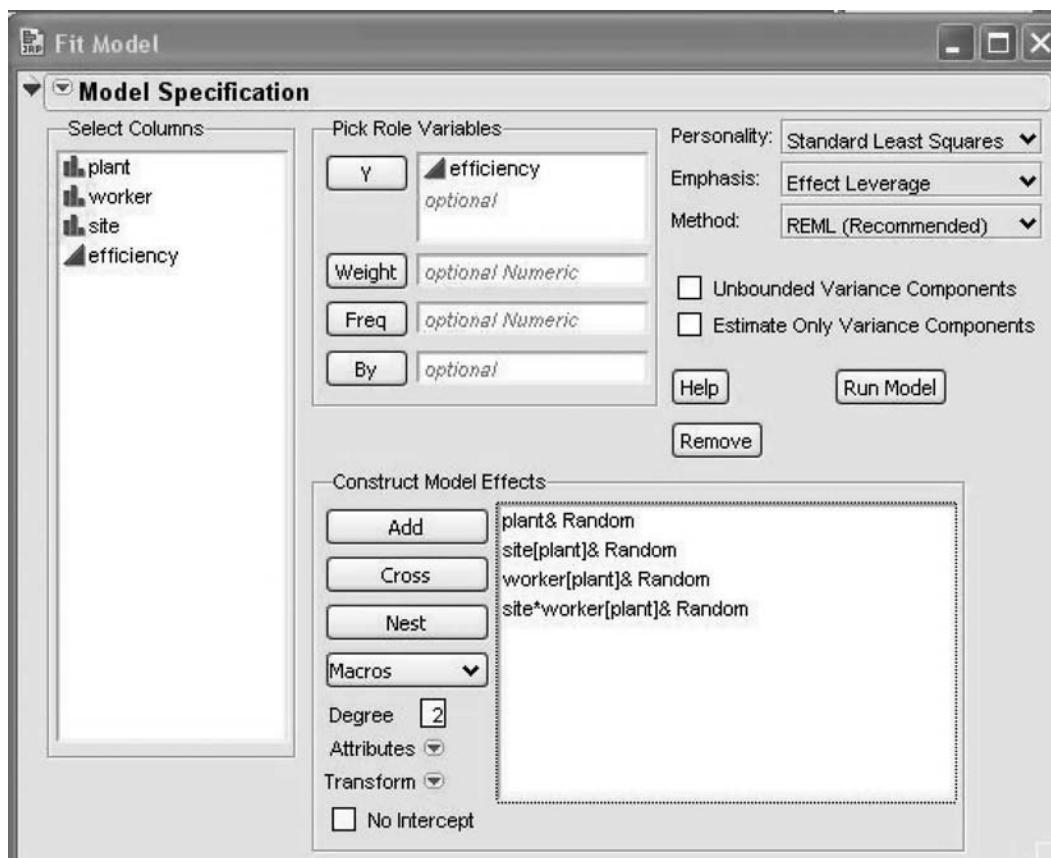


FIGURE 21.2 JMP fit model screen for full model for Example 21.

21.7 Concluding Remarks

The confidence intervals obtained from the four methods are quite different, with the largest discrepancies occurring for variance components having a small number of degrees of freedom. The REML and usual Satterthwaite approximation are quite similar where the general Satterthwaite intervals using the type I sums of squares and confidence regions for the plant and worker variance components are quite variable. In this chapter, a complex unbalanced random effects model was analyzed in detail using all six methods of estimation described in Chapter 19 to estimate the variance components. The methods for testing hypotheses and constructing confidence intervals described in Chapter 20 were demonstrated. One of the variance components in the original model was found to be nonsignificant, thus a model-building approach was used to provide an adequate model with meaningful estimates of the remaining variance components.

At least two methods for estimating variance components and for constructing confidence intervals should be utilized to analyze an unbalanced data set in order to determine how much variability you may have between methods. If both methods yield similar results, then it will not make any difference which method is used and one can have confidence in either method. If the results are very different, then one should investigate what

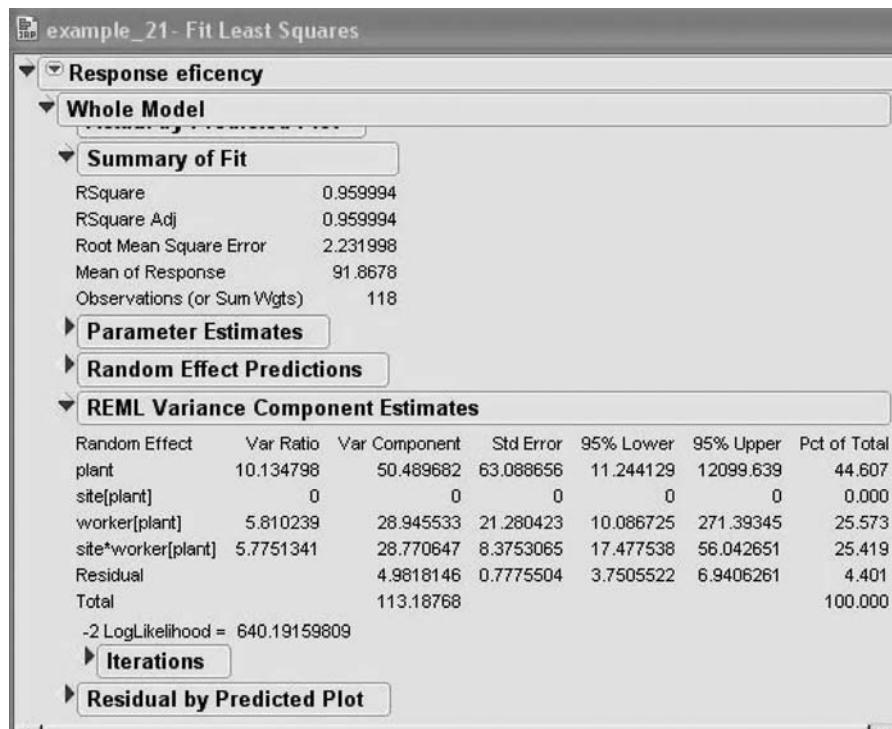


FIGURE 21.3 REML results for Example 21 from JMP.

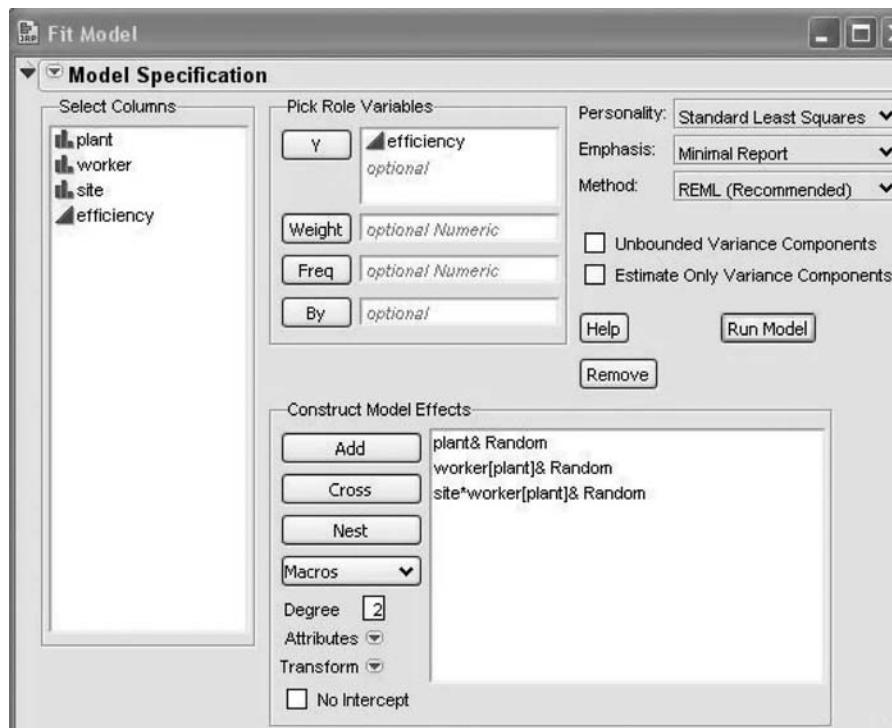


FIGURE 21.4 JMP fit model screen for reduced model for Example 21.

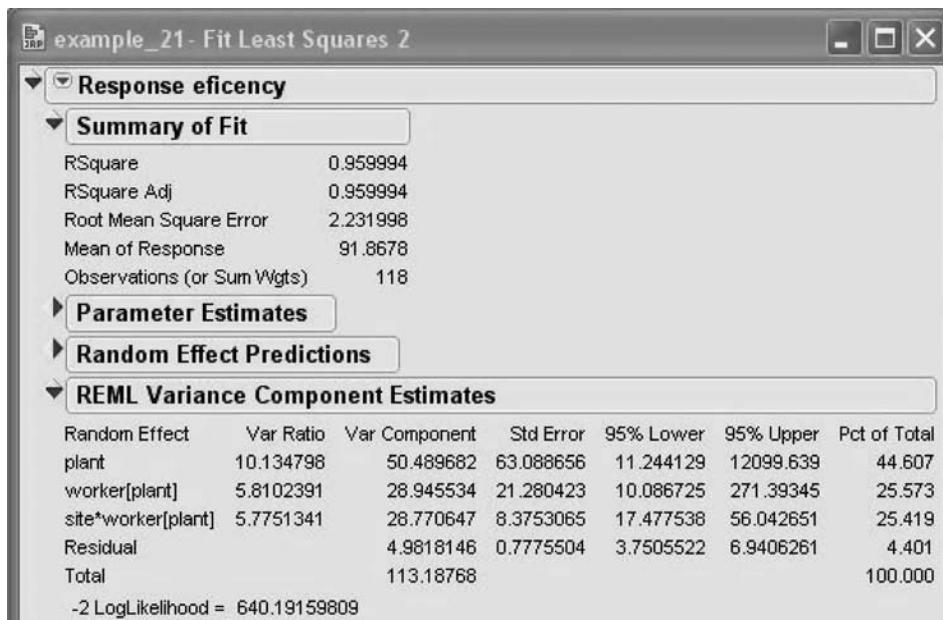


FIGURE 21.5 JMP REML results for reduced model for Example 21.

characteristics of the design are contributing to those differences. The results are described using both SAS-Mixed and the fit model process of JMP.

21.8 Exercises

- 21.1 A group of sheep producers designed a study to evaluate the sources variation in birth weights of lambs. Six producers were selected at random from the group of all producers. Each producer selected three sires (males) at random from their group of sires. Each sire was mated to two to six dams (females). Each dam produced either one, two, or three lambs and the birth weight of each lamb was measured. The data are in the following table where birthwt1–birthwt6 represent data from producers 1 to 6.
- 1) Write out a model that can be used to model the variation in this data set.
 - 2) Provide estimates of the variance components using the six methods discussed in Chapter 19.
 - 3) Provide tests of hypotheses that each variance component is equal to zero (all except the residual).
 - 4) Provide confidence intervals using the REML, the recomputed intervals for type I sums of squares and the usual Satterthwaite approximation.
 - 5) Discuss the similarities and difference provided by the different estimation and confidence interval techniques.

Data Set for Exercise 21.1

Sire	Dam	Lamb	Birthwt1	Birthwt2	Birthwt3	Birthwt4	Birthwt5	Birthwt6
1	1	1	19.10	23.96	17.67	23.10	21.67	16.77
1	2	1	17.67	20.84	16.71	20.14	20.43	19.38
1	2	2	20.15	—	16.75	—	—	—
1	2	3	—	—	15.79	—	—	—
1	3	1	18.67	18.85	16.45	20.35	22.76	18.44
1	3	2	—	—	—	19.83	—	—
1	3	3	—	—	—	22.34	—	—
1	4	1	18.12	20.61	—	—	22.10	23.62
1	4	2	17.98	—	—	—	—	22.33
1	5	1	19.60	—	—	—	20.60	16.65
1	5	2	18.18	—	—	—	—	—
1	5	3	20.64	—	—	—	—	—
1	6	1	—	—	—	—	—	18.48
2	1	1	20.00	22.07	18.37	22.50	21.97	21.47
2	1	2	—	21.07	—	23.54	—	—
2	2	1	19.10	23.32	21.30	19.68	20.46	22.38
2	2	2	—	—	—	22.36	—	—
2	3	1	19.71	23.52	19.08	20.09	18.69	21.99
2	3	2	—	23.41	—	—	18.29	18.57
2	4	1	19.24	21.30	18.14	22.74	—	—
2	4	2	—	21.72	—	19.95	—	—
2	4	3	—	21.91	—	—	—	—
2	5	1	17.18	23.87	20.32	22.87	—	—
3	1	1	19.52	21.65	21.33	19.98	—	18.80
3	1	2	—	—	—	—	—	19.62
3	1	3	—	—	—	—	—	20.68
3	2	1	19.62	19.17	20.45	19.86	—	17.24
3	3	1	17.99	19.84	20.25	—	—	19.88
3	4	1	19.15	19.06	22.01	—	—	—
3	4	2	20.01	19.53	—	—	—	—
3	5	1	19.90	—	22.58	—	—	—
3	6	1	18.46	—	—	—	—	—

21.2 A feed manufacturer makes feed with a supplement mixed in and the amount of supplement should be $3.5 \mu\text{g kg}^{-1}$. The manufacturer designed a study to evaluate the ability of the mixing process to mix the supplement into the feed. Three days were selected at random from the next 30 days of operation; three runs per day were selected from the 24 runs made per day, three batches were randomly selected from each run, and two to four samples were extracted from each batch. The amount of supplement within each sample was determined with the data displayed in the following table.

- 1) Write out a model that can be used to model the variation in this data set.
- 2) Provide estimates of the variance components using the six methods discussed in Chapter 19.
- 3) Provide tests of hypotheses that each variance component is equal to zero (all except the residual).

- 4) Provide confidence intervals using the REML, the recomputed intervals for type I sums of squares and the usual Satterthwaite approximation.
- 5) Discuss the similarities and difference provided by the different estimation and confidence interval techniques.
- 6) Use the model building technique to simplify the model and repeat parts 1–5 for this reduced model.

Drug Concentration Data for Exercise 21.2 Where Amt_{*i,j*} Denotes Amount on Day *i* and Run *j*

Batch	Sample	Amt_1_1	Amt_1_2	Amt_1_3	Amt_2_1	Amt_2_2	Amt_2_3	Amt_3_1	Amt_3_2	Amt_3_3
1	1	3.28	2.78	3.08	3.31	3.35	2.86	3.32	2.61	3.68
1	2	3.14	2.88	3.13	3.38	3.28	2.64	3.61	3.24	3.48
1	3	3.42	2.91	3.14	3.66	3.43	2.24	—	2.92	—
1	4	—	—	—	—	2.43	—	—	—	—
2	1	2.97	2.87	3.04	3.60	3.14	2.74	3.55	2.59	3.68
2	2	3.30	2.78	2.68	3.33	3.09	2.73	3.66	2.65	3.78
2	3	3.21	—	3.10	—	—	2.40	3.50	2.93	3.67
2	4	3.30	—	2.97	—	—	—	3.44	—	—
3	1	3.43	3.01	2.73	3.42	3.06	2.75	3.56	2.91	3.23
3	2	3.17	2.70	2.90	3.25	3.29	2.59	3.35	2.97	3.43
3	3	2.89	2.78	2.80	3.60	3.65	2.51	3.25	3.12	3.47
3	4	—	—	2.93	3.35	3.15	2.45	—	2.62	—

22

Analysis of Mixed Models

Mixed models are used to describe data from experiments or studies that need more than one variance–covariance parameter and involve some fixed effects parameters. The unequal variance models of Chapter 2 are mixed models as they involve more than one variance component. The models described in Chapters 18–21 are called random effects models, but each has an unknown mean parameter, thus the models are essentially mixed models. The definition of a mixed model used in Chapter 18 revolved around having some of the treatment structure with fixed effects and more than one variance component. So the general definition of a mixed model is one with some fixed effect parameters and more than one parameter in the covariance structure. Among the models that are included in this definition are randomized complete blocks models, incomplete blocks models, split-plot-type models, strip-plot-type models, repeated measures type models, random coefficients models, multilevel models, and hierarchical models.

The mixed model involves three parts: 1) the fixed effects part of the model, 2) the random effects part of the model, and 3) the residual part of the model. Consequently, the analysis of a mixed model consists of two types of analyses, an analysis of the random effects and residual parts of the model and an analysis of the fixed effects part of the model. This chapter discusses the construction of mixed models and the necessary steps for analyzing both the random effects part of the model and the fixed effects part of the model. The results of this chapter are provided as a bridge between theoretical results and an understanding of the concepts.

22.1 Introduction to Mixed Models

The model for describing an experiment with both random effects factors and fixed effects factors is called a mixed model. Since there are two types of factors, the resulting model has two parts, a random effects part and a fixed effects part. In order to construct such models, the following rule is used to determine whether a specific interaction is a random or a fixed effect.

Rule: If a main effect is a random effect, then any interaction involving that main effect is also a random effect. The only interactions that are fixed effects are those whose corresponding main effects are all fixed effects.

For example, a model for a three-way treatment structure where the levels of A and B are fixed effects and the levels of C are random effects is

$$y_{ijkm} = \mu + \alpha_i + \beta_j + \gamma_{ij} + c_k + d_{ik} + f_{jk} + g_{ijk} + \varepsilon_{ijkm}$$

where μ denotes the mean response, α_i denotes the effect of the i th level of fixed factor A , β_j denotes the effect of the j th level of fixed factor B , γ_{ij} denotes the interaction effect between the levels of A and the levels of B , c_k denotes the effect of the k th level of random factor C , d_{ik} denotes the interaction between the levels of A and the levels of C , f_{jk} denotes the interaction between the levels of B and the levels of C , g_{ijk} denotes the three-way interaction between the levels of A , B and C , and ε_{ijkm} denotes the residual effect.

The fixed effects part of the model is $\mu + \alpha_i + \beta_j + \gamma_{ij}$, the random effects part of the model is $c_k + d_{ik} + f_{jk} + g_{ijk}$, and the residual part of the model is ε_{ijkm} . Simple assumptions about the terms of the random effects and residual parts of the model are $c_k \sim i.i.d. N(0, \sigma_c^2)$, $d_{ik} \sim i.i.d. N(0, \sigma_d^2)$, $f_{jk} \sim i.i.d. N(0, \sigma_f^2)$, $g_{ijk} \sim i.i.d. N(0, \sigma_g^2)$, $\varepsilon_{ijkm} \sim i.i.d. N(0, \sigma_\varepsilon^2)$ and $c_k, d_{ik}, f_{jk}, g_{ijk}$, and ε_{ijkm} are all independent random variables. The interaction between A and B , denoted by γ_{ij} , is a fixed effect, while all other interactions are random effects since they all involve the random factor C denoted by the subscript k . In general, the mixed model will also involve terms corresponding to the design structure, but such terms are not included in the above model.

The general linear mixed model in matrix notation is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{u}_1 + \mathbf{Z}_2\mathbf{u}_2 + \cdots + \mathbf{Z}_k\mathbf{u}_k + \boldsymbol{\varepsilon}$$

where \mathbf{y} is an observed $N \times 1$ data vector, $\mathbf{X}\boldsymbol{\beta}$ is the fixed effects part of the model, $\mathbf{Z}_1\mathbf{u}_1 + \mathbf{Z}_2\mathbf{u}_2 + \cdots + \mathbf{Z}_k\mathbf{u}_k$ is the random effects part of the model and $\boldsymbol{\varepsilon}$ is the residual part of the model. The ideal conditions are that $\mathbf{u}_i \sim N(\mathbf{0}, \sigma_i^2 \mathbf{I}_{n_i})$, $i = 1, 2, \dots, k$; $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{I}_N)$ and \mathbf{u}_i ($i = 1, 2, \dots, k$) and $\boldsymbol{\varepsilon}$ are independent random variables.

The conditional distribution of \mathbf{y} given $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ is represented by the fixed effects model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{u}_1 + \mathbf{Z}_2\mathbf{u}_2 + \cdots + \mathbf{Z}_k\mathbf{u}_k + \boldsymbol{\varepsilon}$ where $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{I}_N)$ and where $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ are fixed effects (because of the conditioning). The marginal distribution of \mathbf{y} is $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma} = \sigma_1^2 \mathbf{Z}_1 \mathbf{Z}'_1 + \sigma_2^2 \mathbf{Z}_2 \mathbf{Z}'_2 + \cdots + \sigma_k^2 \mathbf{Z}_k \mathbf{Z}'_k + \sigma_\varepsilon^2 \mathbf{I}_N$, or equivalently $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$ where $\mathbf{e} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$.

The population parameters of the mixed model are $\boldsymbol{\beta}, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$ and σ_ε^2 . The analysis of the random effects part of the model consists of estimating, testing hypotheses, and constructing confidence intervals about the variance components $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$ and σ_ε^2 . The analysis of the fixed effects part of the model consists of estimating, testing hypotheses, and constructing confidence intervals about estimable functions of $\boldsymbol{\beta}$. This marginal distribution of \mathbf{y} is based on the assumption that the covariance matrices of the random effects are scalar multiples of identity matrices. A more general form of the mixed effects model is $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{u}_1 + \mathbf{Z}_2\mathbf{u}_2 + \cdots + \mathbf{Z}_k\mathbf{u}_k + \boldsymbol{\varepsilon}$ where the assumptions are that $\mathbf{u}_i \sim N(\mathbf{0}, \mathbf{G}_i)$, $i = 1, 2, \dots, k$, $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{R})$ and \mathbf{u}_i ($i = 1, 2, \dots, k$) and $\boldsymbol{\varepsilon}$ are independent random variables. The marginal distribution of \mathbf{y} can be stated in more general terms as $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma} = \mathbf{Z}_1 \mathbf{G}_1 \mathbf{Z}'_1 + \mathbf{Z}_2 \mathbf{G}_2 \mathbf{Z}'_2 + \cdots + \mathbf{Z}_k \mathbf{G}_k \mathbf{Z}'_k + \mathbf{R}$, or equivalently as $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$ where $\mathbf{e} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$.

Ideally, the matrices $\mathbf{G}_1, \mathbf{G}_2, \dots, \mathbf{G}_k$ and \mathbf{R} are positive definite and consist of the parameters required to model the covariance structure of the random effects and of the residual

parts of the model. The analysis of the mixed model is presented in the next two sections, one section for each part of the analysis. The analysis of the random effects and residual parts of the model is considered first.

22.2 Analysis of the Random Effects Part of the Mixed Model

A random effects model can be constructed from the general linear mixed model by fitting the fixed effects part of the model first, and then computing the residuals of that model. The resulting model does not depend on the fixed effects part of the model. The general linear mixed model expressed as the marginal distribution of \mathbf{y} is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \quad \text{where } \mathbf{e} \sim N(\mathbf{0}, \boldsymbol{\Sigma}) \quad \text{and} \quad \boldsymbol{\Sigma} = \sigma_1^2 \mathbf{Z}_1 \mathbf{Z}'_1 + \sigma_2^2 \mathbf{Z}_2 \mathbf{Z}'_2 + \cdots + \sigma_k^2 \mathbf{Z}_k \mathbf{Z}'_k + \sigma_e^2 \mathbf{I}_N.$$

The (ordinary) least squares estimator of $\boldsymbol{\beta}$ is $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-} \mathbf{X}'\mathbf{y}$ where $(\mathbf{X}'\mathbf{X})^{-}$ denotes a Moore–Penrose generalized inverse of $\mathbf{X}'\mathbf{X}$. The vector of residuals is

$$\mathbf{r} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} = (\mathbf{I} - \mathbf{X}\mathbf{X}^{-})\mathbf{y}$$

A model for the vector of residuals (called the residual model) is

$$\mathbf{r} = (\mathbf{I} - \mathbf{X}\mathbf{X}^{-})\mathbf{Z}_1 \mathbf{u}_1 + (\mathbf{I} - \mathbf{X}\mathbf{X}^{-})\mathbf{Z}_2 \mathbf{u}_2 + \cdots + (\mathbf{I} - \mathbf{X}\mathbf{X}^{-})\mathbf{Z}_k \mathbf{u}_k + (\mathbf{I} - \mathbf{X}\mathbf{X}^{-})\boldsymbol{\epsilon}$$

The residual model does not depend on the fixed effects parameters or on the fixed effects part of the model, $\mathbf{X}\boldsymbol{\beta}$, and thus the residual model is a random effects model. Methods for analyzing random effects models discussed in Chapters 18–20 can be used to analyze the residual random effects model. The four techniques, method of moments, maximum likelihood method, REML method, and the MINQUE method, are the topics of the next four subsections. In Chapter 23, the methods are demonstrated for two examples, one being a balanced data set and one an unbalanced data set.

22.2.1 Method of Moments

As described in Chapter 19 for the analysis of a random effects model, the method of moments technique requires computing sums of squares, determining their expectations, and then estimating the variance components from the system of equations obtained by equating the observed mean squares to their expected values. In order to estimate the variance components, sums of squares must be obtained whose expectations do not depend on the fixed effects part of the model. The method of fitting constants discussed in Section 18.3 provides such sums of squares when the fixed effects part of the model is fitted first, followed by the random effects.

To demonstrate the analysis of the random effects part of the mixed model, consider a two-way treatment structure in a completely randomized design structure with one fixed factor, denoted by B , and the other random factor, denoted by T . The resulting two-way mixed model is

$$y_{ijk} = \mu + \beta_i + t_j + g_{ij} + \varepsilon_{ijk} \quad i = 1, 2, \dots, b \quad j = 1, 2, \dots, t \quad k = 1, 2, \dots, n_{ij}$$

where it is assumed that t_{ij} is distributed *i.i.d.* $N(0, \sigma_t^2)$, g_{ij} is distributed *i.i.d.* $N(0, \sigma_g^2)$, and ε_{ijk} is distributed *i.i.d.* $N(0, \sigma_\varepsilon^2)$. The sums of squares obtained from the method of fitting constants are $R(\beta|\mu)$, $R(t|\mu, \beta)$, $R(g|\mu, \beta, t)$ and SSERROR. The expectations of the last three mean squares have the form

$$E[MSR(t|\mu, \beta)] = \sigma_\varepsilon^2 + k_1\sigma_g^2 + k_2\sigma_t^2$$

$$E[MSR(g|\mu, \beta, t)] = \sigma_\varepsilon^2 + k_3\sigma_g^2$$

and

$$E[MSERROR] = \sigma_\varepsilon^2$$

respectively for some constants k_1 , k_2 , and k_3 . The values of k_1 , k_2 , and k_3 will depend on the sample sizes and the structure of the data. The expectations of these mean squares do not involve the fixed effects parameters and thus the mean squares can be used to estimate the variance components as well to test hypotheses about them.

The system of equations is constructed by equating the observed mean squares to the expected mean squares where $\tilde{\sigma}_\varepsilon^2$, $\tilde{\sigma}_t^2$, and $\tilde{\sigma}_g^2$ denote the solution; that is,

$$MSR(t|\mu, \beta) = \tilde{\sigma}_\varepsilon^2 + k_1\tilde{\sigma}_g^2 + k_2\tilde{\sigma}_t^2$$

$$MSR(g|\mu, \beta, t) = \tilde{\sigma}_\varepsilon^2 + k_3\tilde{\sigma}_g^2$$

and

$$MSERROR = \tilde{\sigma}_\varepsilon^2$$

The solution is

$$\tilde{\sigma}_\varepsilon^2 = MSERROR$$

$$\tilde{\sigma}_g^2 = [MSR(g|\mu, \beta, t) - \tilde{\sigma}_\varepsilon^2]/k_3$$

and

$$\tilde{\sigma}_t^2 = [MSR(t|\mu, \beta) - k_1\tilde{\sigma}_g^2 - \tilde{\sigma}_\varepsilon^2]/k_2$$

The estimates of the variance components are

$$\hat{\sigma}_\varepsilon^2 = \tilde{\sigma}_\varepsilon^2$$

$$\hat{\sigma}_g^2 = \begin{cases} \tilde{\sigma}_g^2 & \text{if } \tilde{\sigma}_g^2 \geq 0 \\ 0 & \text{if } \tilde{\sigma}_g^2 < 0 \end{cases}$$

and

$$\hat{\sigma}_t^2 = \begin{cases} \tilde{\sigma}_t^2 & \text{if } \tilde{\sigma}_t^2 \geq 0 \\ 0 & \text{if } \tilde{\sigma}_t^2 < 0 \end{cases}$$

Approximate $(1-\alpha)100\%$ confidence intervals about the variance components are

$$\frac{df_{\hat{\sigma}_r^2} \hat{\sigma}_r^2}{\chi_{\alpha/2, df_{\hat{\sigma}_r^2}}} \leq \sigma_r^2 \leq \frac{df_{\hat{\sigma}_r^2} \hat{\sigma}_r^2}{\chi_{1-(\alpha/2), df_{\hat{\sigma}_r^2}}}, \quad r = t, g, \varepsilon$$

where

$$df_{\hat{\sigma}_g^2} = df_{MSERROR}, \quad df_{\hat{\sigma}_t^2} = \frac{(\hat{\sigma}_g^2)^2}{\frac{[(1/k_3)MSR(g|\mu, \beta, t)]^2}{df_{MSR(g|\mu, \beta, t)}} + \frac{[(1/k_3)MSERROR]^2}{df_{MSERROR}}}$$

and

$$df_{\hat{\sigma}_t^2} = \frac{(\hat{\sigma}_t^2)^2}{\frac{[(1/k_3)MSR(t|\mu, \beta)]^2}{df_{MSR(t|\mu, \beta)}} + \frac{[(k_1/k_3 k_2)MSR(g|\mu, \beta, t)]^2}{df_{MSR(g|\mu, \beta, t)}} + \frac{[(1/k_2)(1 - (k_1/k_3))MSERROR]^2}{df_{MSERROR}}}$$

The expected means squares can be used to construct tests of hypotheses about the variance components in the model. The statistic to test $H_0: \sigma_t^2 = 0$ vs $H_a: \sigma_t^2 > 0$ is

$$F_{tc} = \frac{MSR(t|\mu, \beta)}{Q} \quad \text{where } Q = \frac{k_1}{k_3} MSR(g|\mu, \beta, t) + \left(1 - \frac{k_1}{k_3}\right) MSERROR$$

which has approximate sampling distribution of F with $df_{MSR(t|\mu, \beta)}$ numerator degrees of freedom and v denominator degrees of freedom where

$$v = \frac{Q^2}{\frac{[(k_1/k_3)MSR(g|\mu, \beta, t)]^2}{df_{MSR(g|\mu, \beta, t)}} + \frac{[(1 - (k_1/k_3))MSERROR]^2}{df_{MSERROR}}}$$

as determined by a Satterthwaite approximation.

The statistic to test $H_0: \sigma_g^2 = 0$ vs $H_a: \sigma_g^2 > 0$ is

$$F_{gc} = \frac{MSR(g|\mu, \beta, t)}{MSERROR}$$

which has sampling distribution of F with $df_{MSR(g|\mu, \beta, t)}$ numerator degrees of freedom and $df_{MSERROR}$ denominator degrees of freedom.

22.2.2 Method of Maximum Likelihood

The method of maximum likelihood can be applied to the complete likelihood function, denoted by $L(\boldsymbol{\beta}, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2, \sigma_e^2) = (2\pi)^{-n/2} |\boldsymbol{\Sigma}|^{-1/2} \exp[-\frac{1}{2}(\mathbf{y} - X\boldsymbol{\beta})' \boldsymbol{\Sigma}^{-1} (\mathbf{y} - X\boldsymbol{\beta})]$. Maximizing this likelihood function or equivalently minimizing $-2\log[L(\boldsymbol{\beta}, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2, \sigma_e^2)]$ with respect to all of the parameters provides equations for simultaneously estimating both the fixed parameters and the random effects parameters (Hartley and Rao, 1967). The required equations are

$$\left. \frac{\partial l}{\partial \boldsymbol{\beta}} \right|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}, \boldsymbol{\sigma}=\hat{\boldsymbol{\sigma}}} = 0, \quad \left. \frac{\partial l}{\partial \sigma_1^2} \right|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}, \boldsymbol{\sigma}=\hat{\boldsymbol{\sigma}}} = 0, \quad \left. \frac{\partial l}{\partial \sigma_2^2} \right|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}, \boldsymbol{\sigma}=\hat{\boldsymbol{\sigma}}} = 0, \dots, \left. \frac{\partial l}{\partial \sigma_k^2} \right|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}, \boldsymbol{\sigma}=\hat{\boldsymbol{\sigma}}} = 0, \quad \left. \frac{\partial l}{\partial \sigma_e^2} \right|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}, \boldsymbol{\sigma}=\hat{\boldsymbol{\sigma}}} = 0$$

where $l = -2\log[L(\boldsymbol{\beta}, \sigma_1^2, \sigma_2^2, \dots, \sigma_k^2, \sigma_e^2)]$, $\boldsymbol{\sigma}' = [\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2, \sigma_e^2]$, and $\hat{\boldsymbol{\sigma}}' = [\hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_k^2, \hat{\sigma}_e^2]$.

The two-way mixed model with equal numbers of observations is used to demonstrate the computation of the maximum likelihood estimators. The model can be expressed as

$$y_{ijk} = \mu_i + a_j + g_{ij} + \varepsilon_{ijk}, \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, a, \quad k = 1, 2, \dots, n$$

where μ_i denotes the mean of the i th level of the fixed effect factor T , a_j denotes the effect of the j th level of the random effect factor A , g_{ij} denotes the random interaction effect, and ε_{ijk} denotes the residual effect. Under the ideal conditions, $a_j \sim i.i.d. N(0, \sigma_a^2)$, $g_{ij} \sim i.i.d. N(0, \sigma_g^2)$, $\varepsilon_{ijk} \sim i.i.d. N(0, \sigma_e^2)$, and the a_j , g_{ij} and ε_{ijk} are all independent random variables. The variance of the data vector is

$$\begin{aligned} \boldsymbol{\Sigma} &= \sigma_a^2 [J_n \otimes I_a \otimes J_t] + \sigma_g^2 [J_n \otimes I_a \otimes I_t] + \sigma_e^2 [I_n \otimes I_a \otimes I_t] \\ &= (\sigma_e^2 + n\sigma_g^2 + nt\sigma_a^2) \left[\frac{1}{n} J_n \otimes I_a \otimes \frac{1}{t} J_t \right] + (\sigma_e^2 + n\sigma_g^2) \left[\frac{1}{n} J_n \otimes I_a \otimes \left(I_t - \frac{1}{t} J_t \right) \right] \\ &\quad + \sigma_e^2 \left[\left(I_n - \frac{1}{n} J_n \right) \otimes I_a \otimes I_t \right] \\ &= \lambda_1 V_1 + \lambda_2 V_2 + \lambda_3 V_3 \text{ (say)} \end{aligned}$$

where $\lambda_1 = \sigma_e^2 + n\sigma_g^2 + nt\sigma_a^2$, $\lambda_2 = \sigma_e^2 + n\sigma_g^2$ and $\lambda_3 = \sigma_e^2$ and V_1 , V_2 , and V_3 are idempotent and pairwise orthogonal. The notation $\mathbf{A} \otimes \mathbf{B}$ denotes the direct product of matrices \mathbf{A} and \mathbf{B} (Graybill 1976). Thus, $|\boldsymbol{\Sigma}| = \lambda_1^a \lambda_2^{a(t-1)} \lambda_3^{at(n-1)}$ and $\boldsymbol{\Sigma}^{-1} = (1/\lambda_1)V_1 + (1/\lambda_2)V_2 + (1/\lambda_3)V_3$. The value of $l = -2\log(L)$ can be expressed as $l = nt\log(2\pi) + a\log(\lambda_1) + a(t-1)\log(\lambda_2) + at(n-1)\log(\lambda_3) + Q$ where $Q = [\mathbf{y} - (J_n \otimes J_a \otimes I_t)\boldsymbol{\mu}]' \boldsymbol{\Sigma}^{-1} [\mathbf{y} - (J_n \otimes J_a \otimes I_t)\boldsymbol{\mu}]$. With some algebra and using the above relationships, the value of Q can be expressed as

$$Q = \frac{1}{\lambda_1} \mathbf{y}' A_1 \mathbf{y} + \frac{1}{\lambda_2} \mathbf{y}' A_2 \mathbf{y} + \frac{1}{\lambda_3} \mathbf{y}' A_3 \mathbf{y} + \frac{1}{\lambda_1} nat(\bar{y}_{...} - \bar{\mu}_{...})^2 + \frac{na}{\lambda_2} \sum_{i=1}^t (\bar{y}_{i..} - \bar{y}_{...} + \bar{\mu}_{..} - \mu_i)^2$$

where

$$\begin{aligned} \mathbf{y}' \mathbf{A}_1 \mathbf{y} &= \mathbf{y}' \left[\frac{1}{n} \mathbf{J}_n \otimes \left(\mathbf{I}_a - \frac{1}{a} \mathbf{J}_a \right) \otimes \frac{1}{t} \mathbf{J}_t \right] \mathbf{y} = SSA \\ \mathbf{y}' \mathbf{A}_2 \mathbf{y} &= \mathbf{y}' \left[\frac{1}{n} \mathbf{J}_n \otimes \left(\mathbf{I}_a - \frac{1}{a} \mathbf{J}_a \right) \otimes \left(\mathbf{I}_t - \frac{1}{t} \mathbf{J}_t \right) \right] \mathbf{y} = SSTA \\ \mathbf{y}' \mathbf{A}_3 \mathbf{y} &= \mathbf{y}' \left[\left(\mathbf{I}_n - \frac{1}{n} \mathbf{J}_n \right) \otimes \mathbf{I}_a \otimes \mathbf{I}_t \right] \mathbf{y} = SSERROR \end{aligned}$$

The value of $-2 \log(\text{likelihood})$ can be expressed as

$$\begin{aligned} l &= nta \log(2\pi) + a \log(\lambda_1) + a(t-1) \log(\lambda_2) + at(n-1) \log(\lambda_3) \\ &\quad + \frac{1}{\lambda_3} SSERROR + \frac{1}{\lambda_1} SSA + \frac{1}{\lambda_2} SSTA + \frac{1}{\lambda_1} nat(\bar{y}_{...} - \bar{\mu}_{.})^2 \\ &\quad + \frac{na}{\lambda_2} \sum_{i=1}^t (\bar{y}_{i..} - \bar{y}_{...} + \bar{\mu}_{.} - \mu_i)^2. \end{aligned}$$

Next differentiate l with respect to λ_1 , λ_2 , λ_3 , $\bar{\mu}_{.}$, and $(\mu_i - \bar{\mu}_{.})$, and set the derivatives equal to zero to provide the system of equations:

$$\begin{aligned} \frac{\partial l}{\partial \lambda_3} \Big|_{\mu=\hat{\mu}, \lambda=\hat{\lambda}} &= -\frac{at(n-1)}{\hat{\lambda}_3} + \frac{SSERROR}{\hat{\lambda}_3^2} = 0 \Rightarrow \hat{\lambda}_3^2 = \frac{SSERROR}{at(n-1)} = \hat{\sigma}_e^2 \\ \frac{\partial l}{\partial \lambda_2} \Big|_{\mu=\hat{\mu}, \lambda=\hat{\lambda}} &= -\frac{a(n-1)}{\hat{\lambda}_2} + \frac{SSTA}{\hat{\lambda}_2^2} = 0 \Rightarrow \hat{\lambda}_2^2 = \frac{SSTA}{a(n-1)} \\ \frac{\partial l}{\partial \lambda_1} \Big|_{\mu=\hat{\mu}, \lambda=\hat{\lambda}} &= -\frac{a}{\hat{\lambda}_1} + \frac{SSA}{\hat{\lambda}_1^2} = 0 \Rightarrow \hat{\lambda}_1^2 = \frac{SSA}{a} \\ \frac{\partial l}{\partial \bar{\mu}_{.}} \Big|_{\mu=\hat{\mu}, \lambda=\hat{\lambda}} &= \frac{-2nat(\bar{y}_{...} - \hat{\mu}_{...})}{\hat{\lambda}_1^2} = 0 \Rightarrow \hat{\mu}_{...} = \bar{y}_{...} \end{aligned}$$

and

$$\frac{\partial l}{\partial \mu_i - \bar{\mu}_{.}} \Big|_{\mu=\hat{\mu}, \lambda=\hat{\lambda}} = \frac{-2na(\bar{y}_{i..} - \bar{y}_{...} + (\hat{\mu}_i - \hat{\mu}_{.}))}{\hat{\lambda}_2} = 0 \Rightarrow (\hat{\mu}_i - \hat{\mu}_{.}) = \bar{y}_{i..} - \bar{y}_{...}$$

or $\hat{\mu}_i = \bar{y}_{i..}$

Using the fact that $\lambda_1 = \sigma_e^2 + n\sigma_g^2 + nt\sigma_a^2$, $\lambda_2 = \sigma_e^2 + n\sigma_g^2$ and $\lambda_3 = \sigma_e^2$, then the maximum likelihood estimators of the variance components are:

$$\hat{\sigma}_e^2 = \hat{\lambda}_3, \quad \hat{\sigma}_g^2 = \begin{cases} \frac{\hat{\lambda}_2 - \hat{\lambda}_3}{n} & \text{if } \hat{\lambda}_2 \geq \hat{\lambda}_3, \\ 0 & \text{if } \hat{\lambda}_2 < \hat{\lambda}_3 \end{cases} \quad \text{and} \quad \hat{\sigma}_a^2 = \begin{cases} \frac{\hat{\lambda}_1 - \hat{\lambda}_2}{nt} & \text{if } \hat{\lambda}_1 \geq \hat{\lambda}_2 \\ 0 & \text{if } \hat{\lambda}_1 < \hat{\lambda}_2 \end{cases}$$

Note that $\hat{\sigma}_e^2$ is an unbiased estimator of σ_e^2 while $\hat{\sigma}_a^2$ and $\hat{\sigma}_g^2$ are biased estimators of σ_a^2 and σ_g^2 , respectively.

22.2.3 Method of Residual Maximum Likelihood

Corbeil and Searle (1976) expressed the likelihood function in two parts, one involving the fixed effects and one free of the fixed effects. They then obtained maximum likelihood estimators of the variance components from that part of the model free of the fixed effects, which they call restricted maximum likelihood estimators, or REML estimators. To illustrate, consider again the general linear mixed model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{u}_1 + \mathbf{Z}_2\mathbf{u}_2 + \dots + \mathbf{Z}_k\mathbf{u}_k + \boldsymbol{\epsilon}$ with the usual assumptions given in Section 22.1. Assume that the rank of \mathbf{X} is equal to p . To construct the restricted likelihood function, let \mathbf{H} be an $N \times (N - p)$ matrix of rank $N - p$ such that $\mathbf{HH}' = (\mathbf{I} - \mathbf{XX}')$. Define the transformation

$$\mathbf{z} = \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}' \\ \mathbf{H}' \end{bmatrix} \mathbf{y}$$

where

$$\begin{bmatrix} \mathbf{X}' \\ \mathbf{H}' \end{bmatrix}$$

is an $N \times N$ nonsingular matrix. Thus the transformation from \mathbf{y} to \mathbf{z} is a one-to-one transformation. The distribution of

$$\mathbf{z} = \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix}$$

is

$$\begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix} \sim N \left[\begin{bmatrix} \mathbf{X}'\mathbf{X}\boldsymbol{\beta} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{X}'\boldsymbol{\Sigma}\mathbf{X} & \mathbf{X}'\boldsymbol{\Sigma}\mathbf{H} \\ \mathbf{H}'\boldsymbol{\Sigma}\mathbf{X} & \mathbf{H}'\boldsymbol{\Sigma}\mathbf{H} \end{bmatrix} \right]$$

The joint likelihood function of

$$\mathbf{z} = \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix}$$

can be partitioned into the likelihood of \mathbf{z}_1 given \mathbf{z}_2 times the marginal likelihood of \mathbf{z}_2 . The marginal likelihood of \mathbf{z}_2 does not depend on the fixed effect parameters of the model, but does depend on the variance components. The marginal likelihood of \mathbf{z}_2 is

$$L_H(\mathbf{z}_2) = L_H(\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2, \sigma_e^2) = (2\pi)^{-(n-p)/2} |\mathbf{H}'\boldsymbol{\Sigma}\mathbf{H}|^{-1/2} \exp[-\frac{1}{2} \mathbf{y}\mathbf{H}(\mathbf{H}'\boldsymbol{\Sigma}\mathbf{H})^{-1} \mathbf{H}'\mathbf{y}]$$

The REML estimators of the variance components are the values of $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$ and σ_e^2 that maximize $L_H(\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2, \sigma_e^2)$. This restricted likelihood function is for the random vector z_2 . But the residuals of the fixed effects model (see beginning of Section 22.2) can be expressed as $r = Hz_2$, a transformation of z_2 . So the restricted likelihood function utilizes the information from the residual model that does not depend on the fixed effects, or the restricted likelihood is a function of the residuals of the model. Hence, the name selected for this method is residual maximum likelihood. The process of maximizing $l_R = \log[L_H(\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2, \sigma_e^2)]$ yields a set of equations that needs to be solved. There is no guarantee that the solution of the set of equations provides a solution set that is in the parameter space. Again, it is important to obtain non-negative estimates of the variance components since these estimates are used in the analysis of the fixed effects part of the model.

On revisiting the balanced two-way mixed model described in Section 22.2.2, the likelihood function can be expressed as

$$\begin{aligned} L(\mu, \lambda_1, \lambda_2, \lambda_3) &= \left[((2\pi)^{1/2} \lambda_1^{1/2})^{-1} \exp\left(-\frac{n\bar{a}}{2\lambda_1} (\bar{y}_{...} - \bar{\mu}_{.})^2\right) \right] \times \left[(2\pi)^{(t-1)/2} \lambda_2^{(t-1)/2} \right]^{-1} \\ &\quad \times \exp\left[-\frac{na}{2\lambda_2} \sum_{i=1}^t (\bar{y}_{i..} - \bar{y}_{...} - \mu_i + \bar{\mu}_{.})^2\right] \times \left[(2\pi)^{(nat-t)/2} \lambda_1^{(a-1)/2} \lambda_2^{(a-1)(t-1)/2} \lambda_3^{at(n-1)/2} \right]^{-1} \\ &\quad \times \exp\left[-\frac{1}{2} \left(\frac{SSERROR}{\lambda_3} + \frac{SSA}{\lambda_1} + \frac{SSTA}{\lambda_2} \right)\right] \\ &= L(\bar{\mu}_{.}) \times L(\mu_i - \bar{\mu}_{.}) \times L(\lambda_1, \lambda_2, \lambda_3) \text{ (say)} \end{aligned}$$

The REML estimates of λ_1, λ_2 , and λ_3 are obtained by minimizing $-2\log[L(\lambda_1, \lambda_2, \lambda_3)] = l_R$ which is accomplished by differentiating l_R with respect to λ_1, λ_2 , and λ_3 and setting the derivatives equal to zero. The derivatives and solutions are:

$$\begin{aligned} \frac{\partial l_R}{\partial \lambda_1} \Big|_{\lambda=\hat{\lambda}} &= \frac{(a-1)}{\hat{\lambda}_1} - \frac{SSA}{\hat{\lambda}_1^2} = 0 \Rightarrow \hat{\lambda}_1 = \frac{SSA}{a-1} \\ \frac{\partial l_R}{\partial \lambda_2} \Big|_{\lambda=\hat{\lambda}} &= \frac{(a-1)(t-1)}{\hat{\lambda}_2} - \frac{SSTA}{\hat{\lambda}_2^2} = 0 \Rightarrow \hat{\lambda}_2 = \frac{SSTA}{(a-1)(t-1)} \\ \frac{\partial l_R}{\partial \lambda_3} \Big|_{\lambda=\hat{\lambda}} &= \frac{at(n-1)}{\hat{\lambda}_3} - \frac{SSERROR}{\hat{\lambda}_3^2} = 0 \Rightarrow \hat{\lambda}_3 = \frac{SSERROR}{at(n-1)} \end{aligned}$$

The estimates of the variance components are computed as

$$\hat{\sigma}_e^2 = \hat{\lambda}_3 \quad \hat{\sigma}_g^2 = \begin{cases} \frac{\hat{\lambda}_2 - \hat{\lambda}_3}{n} & \text{if } \hat{\lambda}_2 \geq \hat{\lambda}_3, \\ 0 & \text{if } \hat{\lambda}_2 < \hat{\lambda}_3 \end{cases} \quad \text{and} \quad \hat{\sigma}_a^2 = \begin{cases} \frac{\hat{\lambda}_1 - \hat{\lambda}_2}{n} & \text{if } \hat{\lambda}_1 \geq \hat{\lambda}_2 \\ 0 & \text{if } \hat{\lambda}_1 < \hat{\lambda}_2 \end{cases}$$

The REML estimates are unbiased estimates of the variance components and they are identical to those one would obtain using the method of moments.

22.2.4 MINQUE Method

The MINQUE method for a random model is described in Section 19.3. In that application, the mean of the model is $j_n\mu$ while the mean of this mixed model is $X\beta$. The estimator can be generalized to the general linear mixed model (Swallow and Searle, 1978) where the MINQUE estimator of

$$\boldsymbol{\sigma} = (\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2, \sigma_e^2)' \text{ is } \hat{\boldsymbol{\sigma}} = \mathbf{S}^{-1}\mathbf{q} \text{ where the matrix } \mathbf{S} \text{ has elements}$$

$$s_{ii'} = \text{tr}[\mathbf{Z}_i \mathbf{Z}'_i \mathbf{R} \mathbf{Z}_{i'} \mathbf{Z}'_{i'}], i, i' = 1, 2, \dots, k+1 \text{ where } i = k+1 \text{ corresponds to } \boldsymbol{\varepsilon} \text{ and } \mathbf{Z}_{k+1} = \mathbf{I}_N,$$

$$\mathbf{R} = \boldsymbol{\Sigma}^{-1}[\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}']\boldsymbol{\Sigma}^{-1}$$

and \mathbf{q} has elements $q_i = \mathbf{y}'\mathbf{R}\mathbf{Z}_i \mathbf{Z}'_i \mathbf{R} \mathbf{y}, i = 1, 2, \dots, k+1$.

The solution depends on the elements of $\boldsymbol{\Sigma}$, which for a MINQUE solution are generally selected to be 1 for variances and 0 for covariances. If the model is balanced, the solution generally does not depend on the values selected for $\boldsymbol{\Sigma}$ and the process converges in one iteration. There is no guarantee that the solution to the system provides values in the parameter space. Be sure to use a solution that provides non-negative estimators, since the estimators need to be used in analyzing the fixed part of the model. One-iteration MINQUE estimators are implemented in some software packages (called zero iteration) and are computed for the two examples to be discussed in Chapter 23.

22.3 Analysis of the Fixed Effects Part of the Model

The analysis of the fixed effects part of a mixed model consists of all aspects of the analysis of a fixed effects part of the model as described below.

22.3.1 Estimation

There are several methods for estimating estimable functions of β in the mixed model. The general linear mixed model can be expressed as

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{e} \quad \text{where } \text{Var}(\mathbf{e}) = \boldsymbol{\Sigma}$$

and

$$\boldsymbol{\Sigma} = \sigma_1^2 \mathbf{Z}_1 \mathbf{Z}'_1 + \sigma_2^2 \mathbf{Z}_2 \mathbf{Z}'_2 + \dots + \sigma_k^2 \mathbf{Z}_k \mathbf{Z}'_k + \sigma_e^2 \mathbf{I}_N$$

A linear combination, $a'\beta$ is estimable for this mixed model if and only if there exists a vector c such that $E(c'y) = a'\beta$. This definition is the same as that for the general linear model (see Chapter 6). The estimate of $a'\beta$, an estimable function of β , is $a'\hat{\beta}$ where $\hat{\beta}$ is any estimator of β .

The ordinary least squares estimator of $a'\beta$ is $a'\hat{\beta}_{LS}$ where $\hat{\beta}_{LS} = (X'X)^{-1}X'y$ or some other solution for $\hat{\beta}$ to the normal equations $X'X\hat{\beta} = X'y$. The least squares estimator of β does not depend on the covariance matrix of y ; that is, it does not depend on Σ .

If the elements of Σ are known (that is, if $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$ and σ_e^2 are known), the best linear unbiased estimator (BLUE) of $a'\beta$ is $a'\hat{\beta}_{BLUE}$ where $\hat{\beta}_{BLUE} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}y$ or any other solution for $\hat{\beta}_{BLUE}$ in $X'\Sigma^{-1}X\hat{\beta}_{BLUE} = X'\Sigma^{-1}y$.

For most balanced designs and for some simple unbalanced designs $\hat{\beta}_{BLUE} = \hat{\beta}_{LS}$. Thus for these designs, the BLUE of $a'\beta$ is $a'\hat{\beta}_{LS}$ where $\hat{\beta}_{LS} = (X'X)^{-1}X'y$, which does not depend on the variance components.

When the designs are unbalanced and the variance components are unknown, life is not so easy. Because the BLUE does not exist (since it depends on the unknown variances), a weighted least squares estimator must be obtained where $\hat{\Sigma}$ is used as the weighting matrix. The estimated covariance matrix is

$$\hat{\Sigma} = \hat{\sigma}_1^2 Z_1 Z_1' + \hat{\sigma}_2^2 Z_2 Z_2' + \dots + \hat{\sigma}_k^2 Z_k Z_k' + \hat{\sigma}_e^2 I_N$$

where $\hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_k^2$, and $\hat{\sigma}_e^2$ are the estimators of the variance components obtained using one of the methods discussed in Section 22.2. The weighted least squares estimator of $a'\beta$ or estimated BLUE (EBLUE) of $a'\beta$ is $a'\hat{\beta}_W$ where $\hat{\beta}_W = (X'\hat{\Sigma}^{-1}X)^{-1}X'\hat{\Sigma}^{-1}y$ or some other solution for $\hat{\beta}_W$ in $X'\hat{\Sigma}^{-1}X\hat{\beta}_W = X'\hat{\Sigma}^{-1}y$. For most designs, $a'\hat{\beta}_W$ converges to $a'\beta$ as the sample size increases. For convergence, some care must be taken so that as the sample size increases, the number of parameters does not go to infinity. The large sample variance of $a'\hat{\beta}_W$ is equal to $\text{Var}(a'\hat{\beta}_W) = a'(X'\hat{\Sigma}^{-1}X)^{-1}a$. This approximation to the variance of $a'\hat{\beta}_W$ does not take into account the variability in the estimates of the variance components. Kackar and Harville (1984) showed that $a'(X'\hat{\Sigma}^{-1}X)^{-1}a$ is too small and needs to be adjusted for the fact that the true values of the variance components are replaced by their respective estimates. Kackar and Harville (1984) and Kenward and Roger (1997) use a Taylor series expansion about the unknown variance components to provide an adjustment to the estimated standard errors of the fixed effects (that approximation is beyond the scope of this text). Using the DDFM = KR option in SAS-Mixed provides the adjusted estimated standard error for the estimates of the fixed effects parameters. The option also estimates the degrees of freedom to associate with a standard error by applying Satterthwaite's method to the KR adjusted standard error estimate.

The method of maximum likelihood can also be used to obtain an estimator of $a'\beta$. The maximum likelihood estimate of $a'\beta$ is $a'\hat{\beta}_{ML}$ where $\hat{\beta}_{ML}$ is a solution to the unrestricted likelihood equations. The variance of $a'\hat{\beta}_{ML}$ is $a'Wa$ where W is the partition of the generalized inverse of the information matrix corresponding to $\hat{\beta}_{ML}$. The examples in Chapter 23 demonstrate the above estimators for both balanced and an unbalanced designs.

22.3.2 Construction of Confidence Intervals

A $(1 - \alpha)100\%$ confidence about an estimable function of $a'\beta$ where $a'\beta$ is an estimable function of β is obtained by using the asymptotic sampling distribution of $a'\hat{\beta}_W$ which is $a'\hat{\beta}_W \sim N[a'\beta, a'(X'\hat{\Sigma}^{-1}X)^{-1}a]$. The estimate of the standard error of $a'\hat{\beta}_W$ is $\widehat{s.e}(a'\hat{\beta}_W) = \sqrt{[a'(X'\hat{\Sigma}^{-1}X)^{-1}a]}$. A Kackar-Harville type of adjustment to the estimated standard error

should be used at this point. Approximate degrees of freedom are obtained by using the generalization of the Satterthwaite approximation (Geisbrecht and Burns, 1985) as

$$\hat{v} = \frac{2[a'(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}a]^2}{\widehat{\text{Var}}[a'(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}a]}$$

Thus, an approximate $(1 - \alpha)100\%$ confidence interval about $a'\beta$ is

$$a'\hat{\beta}_W \pm (t_{\alpha/2, \hat{v}}) \sqrt{[a'(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}a]}$$

A simultaneous set of confidence intervals can be constructed about a set of estimable linear combinations $a'_1\beta, a'_2\beta, \dots, a'_m\beta$, but the degrees of freedom need to be approximated for each linear combination. Adjustments for multiple comparisons can then be made using the techniques of Chapter 3.

22.3.3 Testing Hypotheses

Often a test of hypothesis provides an appropriate inference procedure for linear combinations of the fixed effects. To test $H_0: H\beta = b$ vs $H_a: H\beta \neq b$ compute the statistic

$$Q = (\mathbf{H}\hat{\beta}_W - b)' [H(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}H']^{-1} (\mathbf{H}\hat{\beta}_W - b)$$

Under the conditions of the null hypothesis, the asymptotic sampling distribution of Q is χ_q^2 where $q = \text{Rank}(\mathbf{H})$; that is, q is the number of linear independent linear combinations of β in H . A small sample tests statistic is $F_c = Q/v$ where v is the approximate degrees of freedom associated with $H(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}H'$. The approximate degrees of freedom (SAS Institute, Inc., 1999) are computed by performing a spectral decomposition on $H(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}H'$ as $H(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}H' = P'\Delta P$ where Δ is the $q \times q$ diagonal matrix with the characteristic roots of $H(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}H'$ on the diagonal and P is the $q \times q$ matrix of corresponding characteristic vectors. Let h_s be the s th row of PH , then $v_s = 2d_s^2/(\xi_s'\Omega\xi_s)$ where

$$\xi_s = \frac{\partial[h_s'(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}h_s]}{\partial\sigma}$$

(the vector of derivatives with respect to each of the variance components) and Ω is the asymptotic covariance matrix of the estimates of the variance components, $\hat{\sigma}$. Then

$$\varsigma = \sum_{s=1}^q \frac{v_s}{v_s - 2} \text{In}(v_s > 2)$$

where the indicator function $\text{In}(v_s > 2)$ deletes terms where $v_s \leq 2$. Then the approximate number of denominator degrees of freedom associated with F_c is

$$v = \begin{cases} \frac{2\varsigma}{\varsigma - q} & \text{if } \varsigma > q \\ 0 & \text{if } \varsigma \leq q \end{cases}$$

Examples are discussed in Chapter 23.

22.4 Best Linear Unbiased Prediction

There are some situations where it is of interest to predict the value of a random variable or to predict the values of the random effects used in a study. Suppose you have the linear model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$ where $\text{Var}(\mathbf{e}) = \boldsymbol{\Sigma}$. Suppose there is a random variable ω whose value is not known, but it is assumed that $\omega \sim N(k\boldsymbol{\beta}, \sigma_\omega^2)$ and $\text{Cov}(\mathbf{y}, \omega) = \mathbf{c}$. The object is to predict the value of ω . The predicted value of ω is called a best linear unbiased predictor (BLUP) when

- 1) $\tilde{\omega} = \mathbf{a}'\mathbf{y} + b$;
- 2) $E(\tilde{\omega}) = k'\boldsymbol{\beta}$; and
- 3) $E(\tilde{\omega} - \omega)^2$ is minimized.

The resulting BLUP of ω is $\tilde{\omega} = \mathbf{c}'\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{\text{BLUE}}) + k'\hat{\boldsymbol{\beta}}_{\text{BLUE}}$. When the elements of $\boldsymbol{\Sigma}$ are not known and need to be estimated, then the estimated BLUP (EBLUP) of ω is

$$\hat{\omega} = \mathbf{c}'\hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_W) + k'\hat{\boldsymbol{\beta}}_W. \text{ If } k'\boldsymbol{\beta} = 0, \text{ then } \hat{\omega} = \mathbf{c}'\hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_W)$$

For the general linear mixed model, $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}$ where $\mathbf{u} \sim N(\mathbf{0}, \mathbf{G})$ and $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{R})$, the covariance between \mathbf{y} and \mathbf{u} is $\text{Cov}(\mathbf{y}, \mathbf{u}) = \mathbf{G}\mathbf{Z}'$ and the BLUP of \mathbf{u} is $\tilde{\mathbf{u}} = \mathbf{G}\mathbf{Z}'\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$.

22.5 Mixed Model Equations

Henderson (1984) developed the mixed model equations where the solution simultaneously yields the BLUE of estimable functions of $\boldsymbol{\beta}$, $\mathbf{a}'\boldsymbol{\beta}$, and best linear unbiased predictors of the random effects $\mathbf{u} = (\mathbf{u}_1', \mathbf{u}_2', \dots, \mathbf{u}_k')'$. The mixed model is expressed as $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}$ where

$$\begin{aligned} \mathbf{Z}\mathbf{u} &= \mathbf{Z}_1\mathbf{u}_1 + \mathbf{Z}_2\mathbf{u}_2 + \cdots + \mathbf{Z}_k\mathbf{u}_k \\ \mathbf{u}_i &\sim N(\mathbf{0}, \sigma_i^2 \mathbf{I}_{n_i}), \quad i = 1, 2, \dots, k \\ \boldsymbol{\varepsilon} &\sim N(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{I}_N) \end{aligned}$$

where $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k, \boldsymbol{\varepsilon}$ are independent random variables.

Note that

$$\mathbf{u} = \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_k \end{bmatrix} \sim N(\mathbf{0}, \mathbf{G})$$

where

$$\mathbf{G} = \begin{bmatrix} \sigma_1^2 \mathbf{I}_{n_1} & 0 & \cdots & 0 \\ 0 & \sigma_2^2 \mathbf{I}_{n_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_k^2 \mathbf{I}_{n_k} \end{bmatrix}$$

Using these assumptions, the marginal distribution of \mathbf{y} can be expressed as

$$\mathbf{y} \sim N(X\boldsymbol{\beta}, \boldsymbol{\Sigma})$$

where

$$\text{Var}(\mathbf{y}) = \boldsymbol{\Sigma} = \sum_{i=1}^k \sigma_i^2 \mathbf{Z}_i \mathbf{Z}'_i + \sigma_\epsilon^2 \mathbf{I} = \mathbf{Z} \mathbf{G} \mathbf{Z}' = \sigma_\epsilon^2 \mathbf{I}_N$$

This model with its assumptions implies that $\text{Cov}(\mathbf{y}, \mathbf{u}) = \mathbf{G}\mathbf{Z}'$.

Using the above information, the joint distribution of \mathbf{y} and \mathbf{u} is

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{y} \end{bmatrix} \sim N \left(\begin{bmatrix} \mathbf{0} \\ X\boldsymbol{\beta} \end{bmatrix}', \begin{bmatrix} \mathbf{G} & \mathbf{G}\mathbf{Z}' \\ \mathbf{Z}\mathbf{G}' & \boldsymbol{\Sigma} \end{bmatrix} \right)$$

The conditional distribution of \mathbf{y} given \mathbf{u} is $\mathbf{y}|\mathbf{u} \sim N(X\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}, \sigma_\epsilon^2 \mathbf{I})$ and the marginal distribution of \mathbf{u} is $\mathbf{u} \sim N(\mathbf{0}, \mathbf{G})$. Thus the joint distribution of \mathbf{y} and \mathbf{u} can be expressed as the product of the conditional distribution of \mathbf{y} given \mathbf{u} and the marginal distribution of \mathbf{u} , as

$$\begin{aligned} h(\mathbf{y}, \mathbf{u}) &= f(\mathbf{y}|\mathbf{u})g(\mathbf{u}) \\ &= [2\pi\sigma_\epsilon^2]^{-(N/2)} \exp \left[-\frac{1}{2\sigma_\epsilon^2} (\mathbf{y} - X\boldsymbol{\beta} - \mathbf{Z}\mathbf{u})' (\mathbf{y} - X\boldsymbol{\beta} - \mathbf{Z}\mathbf{u}) \right] [2\pi]^{-(q/2)} |\mathbf{G}|^{-1/2} \exp(-\frac{1}{2}\mathbf{u}'\mathbf{G}^{-1}\mathbf{u}) \end{aligned}$$

where $q = n_1 + n_2 + \dots + n_k$.

Henderson (1984) differentiated $-2 \log[f(\mathbf{y}|\mathbf{u})g(\mathbf{u})]$ with respect to $\boldsymbol{\beta}$ and \mathbf{u} to derive the mixed model equations whose solution provides the BLUE of estimable functions of $\boldsymbol{\beta}$ and BLUPs of \mathbf{u} as:

$$\begin{aligned} -2 \log[f(\mathbf{y}|\mathbf{u})g(\mathbf{u})] &= (N+q)\log(2\pi) + N\log(\sigma_\epsilon^2) + \log|\mathbf{G}| \\ &\quad + \frac{1}{\sigma_\epsilon^2} (\mathbf{y} - X\boldsymbol{\beta} - \mathbf{Z}\mathbf{u})' (\mathbf{y} - X\boldsymbol{\beta} - \mathbf{Z}\mathbf{u}) + \mathbf{u}'\mathbf{G}^{-1}\mathbf{u} \\ &= h \text{ (say)} \end{aligned}$$

$$\frac{\partial h}{\partial \boldsymbol{\beta}} = \frac{2}{\sigma_\epsilon^2} \mathbf{X}'(\mathbf{y} - X\boldsymbol{\beta} - \mathbf{Z}\mathbf{u}), \quad \frac{\partial h}{\partial \mathbf{u}} = \frac{2}{\sigma_\epsilon^2} \mathbf{Z}'(\mathbf{y} - X\boldsymbol{\beta} - \mathbf{Z}\mathbf{u}) + 2\mathbf{G}^{-1}\mathbf{u}$$

Setting the derivatives equal to zero provides the mixed model equations,

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & (\mathbf{Z}'\mathbf{Z} + \sigma_\epsilon^2 \mathbf{G}^{-1}) \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}}_{\text{BLUE}} \\ \tilde{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{u} \end{bmatrix}$$

If the variance of $\boldsymbol{\epsilon}$ (the residual) is \mathbf{R} then the mixed model equations become

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1}) \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}}_{\text{BLUE}} \\ \tilde{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{u} \end{bmatrix}$$

The solutions to the mixed models equations are

$$\hat{\beta} = (X'\hat{\Sigma}^{-1}X)^{-1}X'\hat{\Sigma}^{-1}y \quad \text{and} \quad \tilde{u} = GZ'\hat{\Sigma}^{-1}(y - X\hat{\beta})$$

In general the elements of Σ are not known and need to be estimated, thus the estimated BLUE (EBLUE) and the estimated BLUP (EBLUP) are

$$\hat{\beta}_W = (X'\hat{\Sigma}^{-1}X)^{-1}X'\hat{\Sigma}^{-1}y \quad \text{and} \quad \hat{u} = \hat{G}Z'\hat{\Sigma}^{-1}(y - X\hat{\beta}), \text{ respectively.}$$

22.6 Concluding Remarks

This chapter presents a theoretical development of a mixed model to describe experiments where the factors in the treatment and design structures involve both fixed and random effects. The analysis of the mixed model involves estimating and making inferences about the variance components and functions of the fixed effect parameters. As for the random effects part of the model, tests of hypotheses can be carried out using a sums of squares method rather than using the asymptotic sampling distribution of the estimates of the variance components. The analysis of the fixed effects part of the model was examined in order to make inferences about the fixed effects. Confidence intervals and tests of hypotheses can be carried out using the asymptotic sampling distributions of the estimates of estimable functions or approximate F -statistics where the denominator degrees of freedom are determined via a Satterthwaite type approximation. Best linear unbiased predictions of random effects are discussed as well as the mixed model equations.

22.7 Exercises

- 22.1 For the model used to describe data from a one-way treatment structure in a randomized complete block design structure,

$$y_{ij} = \mu_i + b_j + \varepsilon_{ij}, \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, b, \quad b_j \sim i.i.d. N(0, \sigma_{blk}^2), \quad \varepsilon_{ij} \sim i.i.d. N(0, \sigma_\varepsilon^2)$$

- 1) Determine the ML estimates of the model's parameters.
 - 2) Determine the REML estimates of the model's parameters.
 - 3) Determine the BLUE of μ , the vector of treatment means.
 - 4) Determine the BLUP of b , the vector of block effects.
- 22.2 Five school districts were carrying out a study to evaluate the effectiveness of elementary school math teachers. Within each district, three elementary schools were randomly selected. Within each of the selected elementary schools, four teachers were randomly selected. Each teacher was given a questionnaire and the response was the total number of positive answers. Assume the districts are

fixed effects, the schools within a district are random effects and the teachers within a school are random effects.

- 1) Write out the model with the assumptions.
- 2) Obtain the BLUEs of the district means.
- 3) Obtain the REML estimates of the variance components.
- 4) Obtain the BLUP of the school effects and of the teacher effects.

23

Case Studies of a Mixed Model

Chapter 22 discussed methods for analyzing data for balanced and unbalanced mixed models. This chapter presents detailed analyses using examples for each situation. The data for the unbalanced case is obtained by randomly deleting some observations from the data for the balanced situation. The study involved a company wanting to replace machines used to make a certain component in one of its factories. Three different brands of machines were available, so the management designed an experiment to evaluate the productivity of the three machines when operated by the company's own personnel. Six employees (persons) were randomly selected from the population of employees that are trained to operate such machines. Each selected employee was required to operate each machine during three different shifts. The data recorded were overall productivity scores that took into account the number and quality of components produced. The data are given in Table 23.1.

23.1 Two-Way Mixed Model

The treatment structure for this experiment is a two-way with machines being a fixed effect and persons being a random effect. The design structure is a completely randomized design. The two-way mixed model used to describe data from b persons each operating the t machines during n different shifts is

$$y_{ijk} = \mu + \tau_i + p_j + (\tau p)_{ij} + \varepsilon_{ijk}, \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, b, \quad k = 1, 2, \dots, n$$

where μ denotes the average productivity score of the set of brands of machines as operated by the population of workers, τ_i denotes the effect of the i th machine on productivity scores, p_j denotes the random effect of the j th person on productivity scores, and $(\tau p)_{ij}$ denotes the random interaction effect on productivity scores that is specific to the j th person operating the i th machine, and ε_{ijk} denotes the random error term associated with the k th time the j th person operates the i th machine.

TABLE 23.1

Productivity Scores for Machine–Person Example

Machine	Person	Data for Balanced Case (Section 23.1)			Data for Unbalanced Case (Section 23.2)		
		Rate_1	Rate_2	Rate_3	Mrate_1	Mrate_2	Mrate_3
1	1	52.0	52.8	53.1	52.0	—	—
1	2	51.8	52.8	53.1	51.8	52.8	—
1	3	60.0	60.2	58.4	60.0	—	—
1	4	51.1	52.3	50.3	51.1	52.3	—
1	5	50.9	51.8	51.4	50.9	51.8	51.4
1	6	46.4	44.8	49.2	46.4	44.8	49.2
2	1	62.1	62.6	64.0	—	—	64.0
2	2	59.7	60.0	59.0	59.7	60.0	59.0
2	3	68.6	65.8	69.7	68.6	65.8	—
2	4	63.2	62.8	62.2	63.2	62.8	62.2
2	5	64.8	65.0	65.4	64.8	65.0	—
2	6	43.7	44.2	43.0	43.7	44.2	43.0
3	1	67.5	67.2	66.9	67.5	67.2	66.9
3	2	61.5	61.7	62.3	61.5	61.7	62.3
3	3	70.8	70.6	71.0	70.8	70.6	71.0
3	4	64.1	66.2	64.0	64.1	66.2	64.0
3	5	72.1	72.0	71.1	72.1	72.0	71.1
3	6	62.0	61.4	60.5	62.0	61.4	60.5

The additional assumptions about the random variables in this model are

$$\begin{aligned} p_j &\sim i.i.d. N(0, \sigma_{person}^2) \\ (\tau p)_{ij} &\sim i.i.d. N(0, \sigma_{m \times p}^2) \\ \varepsilon_{ijk} &\sim i.i.d. N(0, \sigma_\varepsilon^2) \end{aligned}$$

and the p_j , $(\tau p)_{ij}$ and ε_{ijk} are all independent random variables.

Assuming $\varepsilon_{ijk} \sim i.i.d. N(0, \sigma_\varepsilon^2)$ implies the time intervals between the different shifts when measurements are made on a *person* \times *machine* combination are long enough so that the error terms are uncorrelated.

The first step is to analyze the random effects part of the model. SAS®-Mixed code was used to obtain the method-of-moments, REML, maximum likelihood, and MINQUE0 estimates of the three variance components. The type III sums of squares, mean squares, and their corresponding expected mean squares are shown in Table 23.2. The maximum likelihood, REML, method of moments and MINQUE0 estimators of the variance components are given in Table 23.3. The SAS-Mixed code in Table 23.4 was used to fit the two-way mixed model and obtain the REML estimates of the variance components. The estimates from the other three methods were obtained by specifying Method = ML, Method = MIVQUE0 and Method = type3.

TABLE 23.2

Analysis of Variance Table for the Balanced Data Using Type III Sums of Squares

Source	df	Sum of Squares	Mean Square	Expected Mean Square	F-Value	Pr > F
Machine	2	1755.263333	877.631667	Var(Residual) + 3 Var(person × machine) + Q(machine)	20.58	0.0003
Person	5	1241.895000	248.379000	Var(Residual) + 3 Var(person × machine) + 9 Var(person)	5.82	0.0089
Person × machine	10	426.530000	42.653000	Var(Residual) + 3 Var(person × machine)	46.13	<0.0001
Residual	36	33.286667	0.924630	Var(Residual)	—	—

TABLE 23.3

Estimates of the Variance Components Using REML, ML, MIVQUE0, and Method of Moments (type III)

Covariance Parameter	REML	ML	MIVQUE0	Type III
Person	22.8584	19.0487	22.8584	22.8584
Person × machine	13.9095	11.5398	13.9095	13.9095
Residual	0.9246	0.9246	0.9246	0.9246

TABLE 23.4

SAS-Mixed Code to Obtain REML Estimates of the Variance Components for the Balanced Data Set

```
proc mixed method=reml cl covtest data=ex23bal;
  class person machine;
  model rating=machine/DDFM=KR;
  random person person*machine;
  LSMEANS MACHINE/ diff;
```

The first hypothesis to be tested is $H_0: \sigma_{m \times p}^2 = 0$ vs $H_a: \sigma_{m \times p}^2 > 0$. The statistic to test this hypothesis, constructed by using the expected mean squares in Table 23.2, is

$$F_{m \times p} = \frac{MS_{Person \times Machine}}{MS_{Residual}} = 46.13$$

The significance level associated with this F test is less than 0.0001, indicating there is strong evidence that $\sigma_{m \times p}^2$ is nonzero and is an important contributor to the variability in the data. If the $machine \times person$ interaction variance component was equal to zero, it would indicate that productivity differences among all employees (not just those in the sample) are similar for the three machines. It would also indicate that the productivity differences among the three machines are similar for all employees (not just those in the sample). In other words, the inferences are to the populations of all possible employees and not just to those employees that were randomly selected from all of the company's employees.

The interpretation of a significant $machine \times person$ interaction variance component is that productivity differences among machines vary depending upon the person using the machine. One interpretation of a large machine by person interaction variance component is that some persons are better adapted to some machines than to other machines. In the data, one can see that person 6 does not perform as well with machines 1 and 2 as the other persons in the sample, but person 6 does about as well as the other persons in the sample on machine 3. The second hypothesis to be tested is $H_0: \sigma_p^2 = 0$ vs $H_a: \sigma_p^2 > 0$. Again, the statistic to test this hypothesis is constructed by using the expected mean squares in Table 23.2 and is given by

$$F_p = \frac{MSPerson}{MSPerson \times Machine}$$

The significance level associated with $H_0: \sigma_p^2 = 0$ vs $H_a: \sigma_p^2 > 0$ is 0.0089, indicating that there is considerable variation in productivity scores among employees at the plant. A training program may decrease the variability among the employees.

The estimates of the variance components from the four methods are shown in Table 23.3. the estimates of the variances components are identical for REML, MIVQUE0, and type III, but those from ML are a little smaller for the *person* and $machine \times person$ variance components.

The Satterthwaite approximation can be used to determine the degrees of freedom to use when constructing confidence intervals about the variance components. The first step is to express the estimate of a variance component as a linear combination of the mean squares in Table 23.2. The method of moments estimate of σ_p^2 is $\hat{\sigma}_p^2 = \frac{1}{9} MSPerson - \frac{1}{9} MSPerson \times Machine = 22.8584$. The number of degrees of freedom of the approximating chi-square distribution obtained through the Satterthwaite approximation is

$$v = \frac{(\hat{\sigma}_p^2)^2}{\frac{[\frac{1}{9} MSPerson]^2}{5} + \frac{[\frac{1}{9} MSPerson \times Machine]^2}{10}} = 3.38035$$

The percentage points for a 95% confidence interval are

$$\chi_{.025, 3.38}^2 = 10.0467 \quad \text{and} \quad \chi_{.975, 3.38}^2 = 0.30725$$

Thus, an approximate 95% confidence interval about σ_p^2 is

$$\frac{3.38035(22.8584)}{10.0467} \leq \sigma_p^2 \leq \frac{3.38035(22.8584)}{0.30725}$$

or

$$7.69102 \leq \sigma_p^2 \leq 251.486$$

A 95% confidence interval about σ_p is $2.773 \leq \sigma_p \leq 15.858$.

A confidence interval for $\sigma_{m \times p}^2$ can be obtained in a similar manner where the approximate number of degrees of freedom is 9.56989 and the 95% confidence interval is obtained as

$$6.70314 \leq \sigma_{m \times p}^2 \leq 44.2384$$

The estimates of the covariance parameters from the type III analysis are given in Table 23.5. The estimates of the variance components and estimates of the standard errors are in the first two columns. The Z-value column gives the ratio of each estimate to its corresponding estimated standard error and the Pr Z column is the two-sided significance level corresponding to the computed Z-value. The Lower and Upper columns are confidence intervals computed using the Wald interval; that is, $\hat{\sigma}^2 \pm Z_{0.025}[\widehat{s.e.}(\hat{\sigma}^2)]$. The Wald confidence intervals are only appropriate when the number of degrees of freedom associated with the estimated variance component is large. The confidence interval about σ_e^2 (residual) is based on the chi-square distribution with 36 degrees of freedom. Table 23.6 contains the estimates of the covariance parameters from the REML analysis.

The estimates of the variance components and their estimated standard errors in Tables 23.5 and 23.6 are identical, but the Pr Z columns are different. The significance levels in Table 23.6 are one-sided (as they should be). The computed number of degrees of freedom, $df = 2(Z\text{-value})^2$, are the same as those computed above for the method of moments estimators. The confidence interval about the residual variance component is based on the chi-square distribution using 36 degrees of freedom. Therefore, for balanced designs, the information about the variance components from REML, MIVQUE0 and method of moments is identical up to the approximate degrees of freedom associated with the individual estimates, but the Wald method is used to construct confidence intervals for the person and machine by person interaction variance components from method of moments and the more appropriate Satterthwaite approximation is used with REML and MIVQUE0.

TABLE 23.5

Estimates of the Covariance Parameters from the Type III Analysis

Covariance Parameter Estimates

Covariance Parameter	Estimate	Standard Error	Z-Value	Pr Z	α	Lower	Upper
Person	22.8584	17.5825	1.30	0.1936	0.05	-11.6026	57.3195
Person \times machine	13.9095	6.3587	2.19	0.0287	0.05	1.4465	26.3724
Residual	0.9246	0.2179	4.24	<0.0001	0.05	0.6115	1.5601

TABLE 23.6

Estimates of the Covariance Parameters from the REML Analysis Including the Computed Degrees of Freedom Used in Computing the Confidence Intervals

Covariance Parameter	Estimate	Standard Error	Z-Value	Pr Z	α	Lower	Upper	df
Person	22.8584	17.5825	1.30	0.0968	0.05	7.6910	251.49	3.3804
Person \times machine	13.9095	6.3587	2.19	0.0144	0.05	6.7031	44.2384	9.5699
Residual	0.9246	0.2179	4.24	<0.0001	0.05	0.6115	1.5601	36.0000

The methods necessary to estimate and test hypotheses about the fixed effects in a mixed model, depend on whether the data are balanced or unbalanced. The methods for the unbalanced design are discussed in Section 22.3.

The general mixed model can be expressed as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{u}_1 + \mathbf{Z}_2\mathbf{u}_2 + \cdots + \mathbf{Z}_k\mathbf{u}_k + \boldsymbol{\varepsilon}$$

where $\mathbf{X} = [j, \mathbf{X}_1]$ and $\boldsymbol{\beta} = (\mu, \tau_1, \tau_2, \dots, \tau_t)'$.

If the elements of the covariance matrix,

$$\boldsymbol{\Sigma} = \sigma_1^2 \mathbf{Z}_1 \mathbf{Z}_1' + \sigma_2^2 \mathbf{Z}_2 \mathbf{Z}_2' + \cdots + \sigma_k^2 \mathbf{Z}_k \mathbf{Z}_k' + \sigma_e^2 \mathbf{I}$$

are known; that is, the variance components are known, the BLUE (best linear unbiased estimate) of an estimable function $\mathbf{a}'\boldsymbol{\beta}$ is

$$\mathbf{a}'\hat{\boldsymbol{\beta}}_{\text{BLUE}} = \mathbf{a}'(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{y}$$

For most balanced mixed models, the estimator of $\mathbf{a}'\boldsymbol{\beta}$ simplifies to

$$\mathbf{a}'\hat{\boldsymbol{\beta}}_{\text{BLUE}} = \mathbf{a}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

The next example shows this simplification for the balanced two-way mixed model. The model can be reparameterized as

$$y_{ijk} = \mu_i + p_j + (\tau p)_{ij} + \varepsilon_{ijk}$$

where $\mu_i = \mu + \tau_i$ and \mathbf{X}_1 is the $nbt \times t$ design matrix corresponding to the fixed effects part of the model $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_t)'$

Note that

$$\mathbf{X}_1 = \begin{bmatrix} j_{nb} & 0 & 0 & \cdots & 0 \\ 0 & j_{nb} & 0 & \cdots & 0 \\ 0 & 0 & j_{nb} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & j_{nb} \end{bmatrix} = \mathbf{j}_n \otimes \mathbf{j}_b \otimes \mathbf{I}_t$$

The covariance matrix for the balanced model is

$$\boldsymbol{\Sigma} = \sigma_p^2 (\mathbf{J}_n \otimes \mathbf{I}_b \otimes \mathbf{J}_t) + \sigma_{m \times p}^2 (\mathbf{J}_n \otimes \mathbf{I}_b \otimes \mathbf{I}_t) + \sigma_e^2 (\mathbf{I}_n \otimes \mathbf{I}_b \otimes \mathbf{I}_t)$$

which can be expressed as

$$\boldsymbol{\Sigma} = \lambda_1 \left(\frac{1}{n} \mathbf{J}_n \otimes \mathbf{I}_b \otimes \frac{1}{t} \mathbf{J}_t \right) + \lambda_2 \left[\frac{1}{n} \mathbf{J}_n \otimes \mathbf{I}_b \otimes \left(\mathbf{I}_t - \frac{1}{t} \mathbf{J}_t \right) \right] + \lambda_3 \left[\left(\mathbf{I}_n - \frac{1}{n} \mathbf{J}_n \right) \otimes \mathbf{I}_b \otimes \mathbf{I}_t \right]$$

where $\lambda_1 = nt\sigma_p^2 + n\sigma_{m \times p}^2 + \sigma_e^2$, $\lambda_2 = n\sigma_{m \times p}^2 + \sigma_e^2$, and $\lambda_3 = \sigma_e^2$. It can be shown that

$$\boldsymbol{\Sigma}^{-1} = \frac{1}{\lambda_1} \left(\frac{1}{n} \mathbf{J}_n \otimes \mathbf{I}_b \otimes \frac{1}{t} \mathbf{J}_t \right) + \frac{1}{\lambda_2} \left[\frac{1}{n} \mathbf{J}_n \otimes \mathbf{I}_b \otimes \left(\mathbf{I}_t - \frac{1}{t} \mathbf{J}_t \right) \right] + \frac{1}{\lambda_3} \left[\left(\mathbf{I}_n - \frac{1}{n} \mathbf{J}_n \right) \otimes \mathbf{I}_b \otimes \mathbf{I}_t \right]$$

Then the matrices $\mathbf{X}_1' \boldsymbol{\Sigma}^{-1} \mathbf{X}_1$ and $\mathbf{X}_1' \boldsymbol{\Sigma}^{-1} \mathbf{y}$ simplify to

$$\mathbf{X}_1' \boldsymbol{\Sigma}^{-1} \mathbf{X}_1 = \frac{nb}{\lambda_1} \left(\frac{1}{t} \mathbf{J}_t \right) + \frac{nb}{\lambda_2} \left(\mathbf{I}_t - \frac{1}{t} \mathbf{J}_t \right)$$

and

$$\mathbf{X}_1' \boldsymbol{\Sigma}^{-1} = \frac{1}{\lambda_1} \mathbf{j}'_n \otimes \mathbf{j}'_b \otimes \frac{1}{t} \mathbf{J}_t + \frac{1}{\lambda_2} \mathbf{j}'_n \otimes \mathbf{j}'_b \otimes \left(\mathbf{I}_t - \frac{1}{t} \mathbf{J}_t \right)$$

Next one can show that

$$(\mathbf{X}_1' \boldsymbol{\Sigma}^{-1} \mathbf{X}_1)^{-1} = \frac{\lambda_1}{nb} \left(\frac{1}{t} \mathbf{J}_t \right) + \frac{\lambda_2}{nb} \left(\mathbf{I}_t - \frac{1}{t} \mathbf{J}_t \right)$$

Thus the estimator of $\boldsymbol{\mu}$ is

$$\hat{\boldsymbol{\mu}} = (\mathbf{X}_1' \boldsymbol{\Sigma}^{-1} \mathbf{X}_1)^{-1} \mathbf{X}_1' \boldsymbol{\Sigma}^{-1} \mathbf{y} = \left(\frac{1}{n} \mathbf{j}'_n \otimes \frac{1}{b} \mathbf{j}'_b \otimes \mathbf{I}_t \right) \mathbf{y} = (\mathbf{X}_1' \mathbf{X}_1)^{-1} \mathbf{X}_1' \mathbf{y}$$

Therefore the estimate of $\boldsymbol{\mu}_i$ is $\hat{\boldsymbol{\mu}}_i = \bar{y}_{1..}, i = 1, 2, \dots, t$.

The variance of $\hat{\boldsymbol{\mu}}_i$ is

$$\text{Var}(\hat{\boldsymbol{\mu}}_i) = (\mathbf{X}_1' \boldsymbol{\Sigma}^{-1} \mathbf{X}_1)^{-1} = \frac{\sigma_e^2 + n\sigma_{m \times p}^2 + nt\sigma_p^2}{nb} \left(\frac{1}{t} \mathbf{J}_t \right) + \frac{\sigma_e^2 + n\sigma_{m \times p}^2}{nb} \left(\mathbf{I}_t - \frac{1}{t} \mathbf{J}_t \right)$$

Thus the variance of $\hat{\mu}_i$ is

$$\text{Var}(\hat{\mu}_i) = \frac{\sigma_e^2 + n\sigma_{m \times p}^2 + n\sigma_p^2}{nb}$$

For the machine–person example, the variance of each machine mean is

$$\text{Var}(\hat{\mu}_i) = \frac{\sigma_e^2 + 3\sigma_{m \times p}^2 + 3\sigma_p^2}{18}$$

The estimate of a contrast $\mathbf{a}' \boldsymbol{\mu}$ (where $\mathbf{a}' \mathbf{j}_t = 0$) is $\mathbf{a}' \hat{\boldsymbol{\mu}}$ with variance

$$\text{Var}(\mathbf{a}' \hat{\boldsymbol{\mu}}) = \frac{\sigma_e^2 + n\sigma_{m \times p}^2}{nb} \mathbf{a}' \mathbf{a}$$

The estimate the difference $\mu_i - \mu_{i'} (i \neq i')$ is $\hat{\mu}_i - \hat{\mu}_{i'}$ with variance

$$\text{Var}(\hat{\mu}_i - \hat{\mu}_{i'}) = 2 \left[\frac{\sigma_e^2 + n\sigma_{m \times p}^2}{nb} \right]$$

The estimate of the standard error of the difference between the two machine means is

$$\widehat{s.e.}_{\mu_i - \mu_j} = \sqrt{\frac{2(MSPerson \times Machine)}{18}} = 2.177$$

Table 23.7 contains the estimated machine means, estimated standard errors, *t*-values, and significance levels for testing the individual means are equal to zero. Pairwise differences among the machine means are in Table 23.8, which includes the estimated differences, estimated standard errors, *t*-values, and significance levels. An *LSD*_{0.05} value for comparing the two means is computed as

$$LSD_{0.05} = (t_{0.025,10})(\widehat{s.e.}_{\hat{\mu}_i - \hat{\mu}_j}) = 2.228(2.177) = 4.85$$

All of the differences are greater than 4.85, so the *LSD* shows the machine means are all significantly different from one another.

To test the hypothesis that the means are equal, $H_0: \mu_1 = \mu_2 = \mu_3$ vs $H_a: (\text{not } H_0)$, use the expected mean squares from Table 23.2 to construct the test statistic

$$F_{mc} = \frac{MSMachine}{MSPerson \times Machine} = 20.58$$

The significance level corresponding to the equal means hypothesis is 0.0004. This example shows that the analysis is quite straightforward for the balanced case. However, the analysis of the unbalanced case is not quite as easy.

TABLE 23.7

Machine Means and Estimated Standard Errors for the Balanced Data Set

Least Squares Means

Effect	Machine	Estimate	Standard Error	df	t-Value	Pr > t
Machine	1	52.3556	2.4858	8.52	21.06	<0.0001
Machine	2	60.3222	2.4858	8.52	24.27	<0.0001
Machine	3	66.2722	2.4858	8.52	26.66	<0.0001

TABLE 23.8

Pairwise Difference among the Machine Means for the Balanced Data Set

Differences of Least Squares Means

Effect	Machine	Machine	Estimate	Standard Error	df	t-Value	Pr > t
Machine	1	2	-7.9667	2.1770	10	-3.66	0.0044
Machine	1	3	-13.9167	2.1770	10	-6.39	<0.0001
Machine	2	3	-5.9500	2.1770	10	-2.73	0.0211

23.2 Unbalanced Two-Way Mixed Model

The data in Table 23.1 for this example are the same as those for the balanced example except that some observations have been randomly deleted. This has been done to demonstrate the problems that occur when analyzing unbalanced data sets and to compare the estimation procedures for the balanced and unbalanced cases. The model to describe this data is identical to that for the balanced data set in Section 23.1, or

$$y_{ijk} = \mu + \tau_i + p_j + (\tau p)_{ij} + \varepsilon_{ijk}, \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, b, \quad k = 1, 2, \dots, n_{ij}$$

The analysis of variance table for the unbalanced data is shown in Table 23.9. The sums of squares are those obtained by the method of fitting constants, or type I sum of squares; that is,

$$\begin{aligned} SSMachine &= R(\tau | \mu) \\ SSPerson &= R(p | \mu, \tau) \end{aligned}$$

and

$$SSPerson \times Machine = R((\tau p) | \mu, \tau, p)$$

Their corresponding expected mean squares are included in Table 23.9.

First, consider estimating σ_p^2 . In the unbalanced case the coefficients on $\sigma_{m \times p}^2$ in the expected mean squares of the *MSPerson* and *MSPerson* \times *Machine* are not the same. Thus in

TABLE 23.9

Analysis of Variance Table Based on Type I Sums of Squares for Unbalanced Data Set

Type I Analysis of Variance

Source	df	Sum of Squares	Mean Square	Expected Mean Square	Error Term	Error df	F-Value
Machine	2	1648.664722	824.332361	Var(Residual) + 2.6115 Var(person \times machine) + 0.1569 Var(person) + Q(machine)	0.0217 MS(person) + 1.1032 MS(person \times machine) - 0.1249 MS(Residual)	11.782	16.86
Person	5	1008.763583	201.752717	Var(Residual) + 2.5866 Var(person \times machine) + 7.219 Var(person)	1.1167 MS(person \times machine) - 0.1167 MS(Residual)	9.9549	4.48
Person \times machine	10	404.315028	40.431503	Var(Residual) + 2.3162 Var(person \times machine)	MS(Residual)	26	46.34
Residual	26	22.686667	0.872564	Var(Residual)	—	—	—

order to estimate σ_p^2 , one will need to find a function of all three expected mean squares that is equal to σ_p^2 . It can be shown that

$$\begin{aligned}\tilde{\sigma}_p^2 &= \frac{1}{7.219} \left[(\tilde{\sigma}_\epsilon^2 + 2.5866\tilde{\sigma}_{m \times p}^2 + 7.2190\tilde{\sigma}_p^2) - \frac{2.5866}{2.3162}(\tilde{\sigma}_\epsilon^2 + 2.3162\tilde{\sigma}_{m \times p}^2) + \left(\frac{2.5866}{2.3162} - 1 \right) \tilde{\sigma}_\epsilon^2 \right] \\ &= \frac{1}{7.219} \left[(\tilde{\sigma}_\epsilon^2 + 2.5866\tilde{\sigma}_{m \times p}^2 + 7.2190\tilde{\sigma}_p^2) - 1.1167(\tilde{\sigma}_\epsilon^2 + 2.3162\tilde{\sigma}_{m \times p}^2) + (0.1167)\tilde{\sigma}_\epsilon^2 \right] \\ &= \frac{1}{7.219} \left[E(MSPerson) - 1.1167E(MSPerson \times MSMachine) + (0.1167)E(MSResidual) \right]\end{aligned}$$

The method of moments equations for the type I sums of squares are

$$\begin{aligned}MSPerson &= \tilde{\sigma}_\epsilon^2 + 2.5866\tilde{\sigma}_{m \times p}^2 + 7.2190\tilde{\sigma}_p^2 \\ MSPerson \times Machine &= \tilde{\sigma}_\epsilon^2 + 2.3162\tilde{\sigma}_{m \times p}^2\end{aligned}$$

and

$$MSResidual = \tilde{\sigma}_\epsilon^2$$

Solving the above equations simultaneously give the type I method of moments estimators of each of the variance components as

$$\begin{aligned}\tilde{\sigma}_p^2 &= \frac{1}{7.219} [MSPerson - 1.1167 MSPerson \times Machine + 0.1167 MSResidual] \\ &= \frac{1}{7.219} [201.7527 - 1.1167(40.4315) + 0.1167(0.8726)] = 21.7073 \\ \tilde{\sigma}_{m \times p}^2 &= \frac{1}{2.3162} [MSPerson \times Machine - MSResidual] \\ &= \frac{1}{2.3162} [40.4315 - 0.8726] = 17.0792\end{aligned}$$

and

$$\tilde{\sigma}_\epsilon^2 = MSResidual = 0.8726$$

The estimates of the variance components obtained from the REML, ML, MIVQUE0, type I, type II, and type III methods are given in Tables 23.10–23.13. The solution for the residual variance component using the MIVQUE0 method is 0. (If one uses the no-bound option in SAS-Mixed, the solution is negative. If there is one variance component we should always be able to estimate, it is the residual.) The summary in Table 23.13 indicates that the REML method seems to be in the middle of the other methods' estimates.

The SAS-Mixed code used to obtain these results is given in Table 23.14, where the Method = REML can be replaced by ML, MIVQUE0, type1, type2 or type3 to produce the other results. The tables contain the estimates of the variance components, estimated standard errors, recomputed degrees of freedom, confidence intervals, and recomputed

TABLE 23.10

Estimates of the Residual Variance Component for the Unbalanced Data Set

Method	Estimate	Standard Error	Z-Value	df	Lower	Upper	Newlower	Newupper
REML	0.87	0.24	3.61	26.10	0.54	1.63	0.54	1.63
MIVQUE0	0.00	0.00	—	—	—	—	—	—
ML	0.87	0.24	3.62	26.14	0.54	1.63	0.54	1.63
Type III	0.87	0.24	3.58	25.66	0.54	1.64	0.54	1.65
Type II	0.87	0.24	3.59	25.71	0.54	1.64	0.54	1.65
Type I	0.87	0.24	3.59	25.71	0.54	1.64	0.54	1.65

TABLE 23.11

Estimates of the Person Variance Component for the Unbalanced Data Set

Method	Estimate	Standard Error	Z-Value	df	Lower	Upper	Newlower	Newupper
REML	22.46	17.41	1.29	3.33	7.51	254.64	7.51	254.64
MIVQUE0	24.34	10.79	2.26	10.18	11.95	74.00	11.95	74.00
ML	18.70	13.24	1.41	3.99	6.71	155.14	6.71	155.14
Type III	24.26	21.27	1.14	2.60	-17.44	65.95	7.34	464.28
Type II	21.71	17.81	1.22	2.97	-13.20	56.62	6.94	308.07
Type I	21.71	17.81	1.22	2.97	-13.20	56.62	6.94	308.07

TABLE 23.12Estimates of the Person \times Machine Variance Component for the Unbalanced Data Set

Method	Estimate	Standard Error	Z-Value	df	Lower	Upper	Newlower	Newupper
REML	14.23	6.52	2.18	9.55	6.85	45.35	6.85	45.35
MIVQUE0	16.15	8.53	1.89	7.17	7.12	65.44	7.12	65.44
ML	11.82	4.96	2.38	11.35	5.98	33.37	5.98	33.37
Type III	17.08	9.57	1.78	6.37	-1.68	35.84	7.24	77.75
Type II	17.08	9.57	1.78	6.37	-1.68	35.84	7.24	77.76
Type I	17.08	9.57	1.78	6.37	-1.68	35.84	7.24	77.76

TABLE 23.13

Summary of the Estimates of the Variance Components for the Unbalanced Data Set

Covariance Parameter	REML	ML	MIVQUE0	Type I	Type II	Type III
Person	22.4551	18.7009	24.3389	21.7069	21.7069	24.2571
Person \times machine	14.2340	11.8176	16.1548	17.0791	17.0791	17.0791
Residual	0.8709	0.8701	0.0000	0.8726	0.8726	0.8726

TABLE 23.14

SAS-Mixed Code to Produce the REML Estimates and Analysis for the Unbalanced Data Set

```
proc mixed method=reml cl covtest data=ex23unb;
class person machine;
model rating=machine/DDFM=KR;
random person person*machine;
LSMEANS MACHINE/ diff;
```

Satterthwaite confidence intervals. Recall that the confidence intervals for the variance components for REML, ML and MIVQUE0 are based on the chi-square distribution using $df = 2(Z\text{-value})^2$. The lower and newlower and upper and newupper are identical for these methods. Intervals are provided for variance components (other than residual) using the Wald interval when Method is type I or type II or type III. The newlower and newupper are the recomputed confidence intervals, and would be the intervals of choice.

The F -statistics in Table 23.9 can be used to test hypotheses about the variance components. The statistic to test $H_0: \sigma_p^2 = 0$ vs $H_a: \sigma_p^2 > 0$ is $F = 4.48$ with 5 numerator and 9.9549 denominator degrees. The significance level is 0.0003. The statistic to test $H_0: \sigma_{m \times p}^2 = 0$ vs $H_a: \sigma_{m \times p}^2 > 0$ is $F = 46.34$ with 10 numerator and 26 denominator degrees of freedom. The significance level is <0.0001. The likelihood ratio test can also be used to provide tests of the hypotheses about the variance components. The process involves obtaining ML estimates of the variance components for the full model and computing $-2\log(\text{full likelihood function})$, as is shown in Table 23.15. To test $H_0: \sigma_{m \times p}^2 = 0$ vs $H_a: \sigma_{m \times p}^2 > 0$, fit a model without *person* \times *machine* and compute the $-2\log(\text{reduced likelihood function})$. The likelihood ratio test statistic is computed as $\text{LR test} = -2\log(\text{reduced likelihood function}) - [-2\log(\text{full likelihood function})]$, which equals 55.5665 and is asymptotically distributed as a chi-square distribution with 1 degree of freedom. The significance level for this test is <0.0001. To test $H_0: \sigma_p^2 = 0$ vs $H_a: \sigma_p^2 > 0$, fit a model without person and compute $-2\log(\text{reduced likelihood function})$. The likelihood ratio test statistic is computed as $\text{LR test} = -2\log(\text{reduced likelihood function}) - [-2\log(\text{full likelihood function})]$ which equals 6.4226 and is asymptotically distributed as a chi-square distribution with 1 degree of freedom. The significance level for this test is 0.0113.

The estimates of the fixed effects parameters are computed by

$$\hat{\beta}_W = (X' \hat{\Sigma}^{-1} X)^{-1} X' \hat{\Sigma}^{-1} y$$

where $\hat{\Sigma}$ is the estimated covariance matrix evaluated at the estimates of the variance components from the specified method of estimating the variance components, and the variance-covariance matrix of $\hat{\beta}$ is given by $\text{Var}(\hat{\beta}) = (X' \Sigma^{-1} X)^{-1}$. The test for the equality of the machine means was obtained for each method of estimating the variance components. The type III tests for fixed effects are summarized in Table 23.16. The MIVQUE0 results are not useful. The number of denominator degrees of freedom from REML is 10.1, which is close to the 10 one would use for the balanced data set. The ML method has too many denominator degrees of freedom and the type I–III methods have too few.

Tables 23.17–23.19 contain the estimates of the machines' means. There is not much difference between the REML and the type I, II, or III results. Tables 23.20–23.22 contain the

TABLE 23.15

ML Estimates of the Variance Components for Three Models for Constructing Likelihood Ratio Tests for Hypotheses about σ_p^2 and σ_g^2

Type	Covariance Parameter	Estimate	$-2\log(\text{LH})$	LR test	Pr chi
Full	Person	18.7009	191.506		
Full	Person \times machine	11.8176			
Full	Residual	0.8701			
No machine \times person	Person	20.1456	247.073	55.5665	<0.0001
No machine \times person	Residual	11.2266	Test for $H_0: \sigma_{m \times p}^2 = 0$ vs $H_a: \sigma_{m \times p}^2 > 0$		
No person	Person \times machine	30.5079	197.929	6.4226	0.0113
No person	Residual	0.8719	Test for $H_0: \sigma_p^2 = 0$ vs $H_a: \sigma_p^2 > 0$		

TABLE 23.16

Summary of the Type III Tests for Fixed Effects for the Unbalanced Data Set

Method	Effect	Num <i>df</i>	Den <i>df</i>	F-Value	Pr <i>F</i>
REML	Machine	2	10.1	19.96	0.0003
MIVQUE0	Machine	2	6.58	2.4×10^{14}	<0.0001
ML	Machine	2	12.2	23.91	<0.0001
Type III	Machine	2	6.7	16.72	0.0025
Type II	Machine	2	6.7	16.72	0.0025
Type I	Machine	2	6.7	16.72	0.0025

TABLE 23.17

Estimates of the Mean for Machine 1

Method	Estimate	Standard Error	<i>df</i>	t-Value	Pr <i>t</i>
REML	52.35	2.49	8.72	21.02	<0.0001
MIVQUE0	46.80	0.00	6.58	I	<0.0001
ML	52.35	2.27	10.5	23.01	<0.0001
Type III	52.35	2.64	7.22	19.82	<0.0001
Type II	52.35	2.56	8.88	20.45	<0.0001
Type I	52.35	2.56	8.88	20.45	<0.0001

TABLE 23.18

Estimates of the Mean for Machine 2

Method	Estimate	Standard Error	<i>df</i>	t-Value	Pr <i>t</i>
REML	60.32	2.49	8.68	24.25	<0.0001
MIVQUE0	43.63	0.00	6.58	I	<0.0001
ML	60.31	2.27	10.4	26.55	<0.0001
Type III	60.32	2.64	7.18	22.86	<0.0001
Type II	60.32	2.56	8.83	23.59	<0.0001
Type I	60.32	2.56	8.83	23.59	<0.0001

TABLE 23.19

Estimates of the Mean for Machine 3

Method	Estimate	Standard Error	<i>df</i>	t-Value	Pr <i>t</i>
REML	66.27	2.48	8.61	26.69	<0.0001
MIVQUE0	61.30	0.00	6.58	I	<0.0001
ML	66.27	2.27	10.3	29.25	<0.0001
Type III	66.27	2.63	7.13	25.16	<0.0001
Type II	66.27	2.55	8.77	25.97	<0.0001
Type I	66.27	2.55	8.77	25.97	<0.0001

TABLE 23.20

Estimates of the Difference between the Means for Machines 1 and 2

Method	Estimate	Standard Error	df	t-Value	Pr t
REML	-7.96	2.21	10.2	-3.60	0.0047
MIVQUE0	3.17	0.00	6.58	I	<0.0001
ML	-7.96	2.02	12.3	-3.93	0.0019
Type III	-7.97	2.42	6.75	-3.29	0.0140
Type II	-7.97	2.42	6.75	-3.29	0.0140
Type I	-7.97	2.42	6.75	-3.29	0.0140

TABLE 23.21

Estimates of the Difference between the Means for Machines 1 and 3

Method	Estimate	Standard Error	df	t-Value	Pr t
REML	-13.92	2.21	10.1	-6.30	<0.0001
MIVQUE0	-14.50	0.00	6.58	M	<0.0001
ML	-13.92	2.02	12.1	-6.89	<0.0001
Type III	-13.92	2.41	6.7	-5.76	0.0008
Type II	-13.92	2.41	6.7	-5.76	0.0008
Type I	-13.92	2.41	6.7	-5.76	0.0008

TABLE 23.22

Estimates of the Difference between the Means for Machines 2 and 3

Method	Estimate	Standard Error	df	t-Value	Pr t
REML	-5.96	2.21	10	-2.70	0.0222
MIVQUE0	-17.67	0.00	6.58	M	<0.0001
ML	-5.96	2.01	12.1	-2.96	0.0119
Type III	-5.95	2.41	6.66	-2.47	0.0446
Type II	-5.95	2.41	6.66	-2.47	0.0446
Type I	-5.95	2.41	6.66	-2.47	0.0446

TABLE 23.23

Type III Tests of Equal Machine Means for the Unbalanced Data Set

Type III Tests of Fixed Effects

Effect	Num df	Den df	F-Value	Pr > F
Machine	2	10.1	19.96	0.0003

estimates of the differences between each pair of machine means for each of the methods of estimating the variance components. There is a difference between the results for the methods as the denominator degrees of freedom for REML are approximately equal to 10, as expected if the data set was balanced. The REML method seems to provide the best all-around results.

The test of the hypothesis that the means are equal, $H_0: \mu_1 = \mu_2 = \mu_3$ vs $H_a: (\text{not } H_0)$ based on REML estimates is displayed in Table 23.23, where the computed F -value is 19.96 based on 2 numerator degrees of freedom and 10.1 denominator degrees of freedom. The denominator degrees of freedom from the balanced data set analysis is 10 (which corresponds to the degrees of freedom for the machine by person interaction), so 10.1 is a close approximation.

The results in Tables 23.20–23.22 are the pairwise differences between each pair of machine means. Again, the results based on the REML method of estimating the variance components seem to be the best as the degrees of freedom for the pairwise comparisons are 10.2, 10.1 and 10, values that are close to 10, as was the case for the balanced analysis.

23.3 JMP Analysis of the Unbalanced Two-Way Data Set

The analysis of the unbalanced data set using JMP® involves constructing a data table as shown in Figure 23.1. The variables machine, person and rep are declared to be nominal and rating is continuous. The fit model screen is displayed in Figure 23.2, where the rating has been selected as the “Y” variable, machine is a fixed effect, and person and machine by person are random effects. The REML method was selected as the method to estimate the variance components. The REML estimates of the variance components are shown in Figure 23.3 along with the estimated standard errors and Wald confidence intervals. The test of the hypothesis that the means are equal, $H_0: \mu_1 = \mu_2 = \mu_3$ vs $H_a: (\text{not } H_0)$ is displayed in Figure 23.4, where the computed F -value is 19.9639 based on 2 numerator degrees of freedom and 10.11 denominator degrees of freedom.

The least squares means and their estimated standard errors are shown in Figure 23.5. An LSD multiple comparison method was used to do pairwise comparisons among the means and the results which include the differences, their estimated standard errors and 95% confidence intervals, are shown in Figure 23.6. The results of the estimation process using JMP provide the same results as those from SAS-Mixed using the REML method, except that the confidence intervals are constructed using the Wald method instead of the using a Satterthwaite approximation.

	machine	person	rating	rep
1	1	1	52	1
2	1	1	▪	2
3	1	1	▪	3
4	1	2	51.8	1
5	1	2	52.8	2
6	1	2	▪	3
7	1	3	60	1
8	1	3	▪	2
9	1	3	▪	3
10	1	4	51.1	1
11	1	4	52.3	2
12	1	4	▪	3

FIGURE 23.1 JMP data table screen.

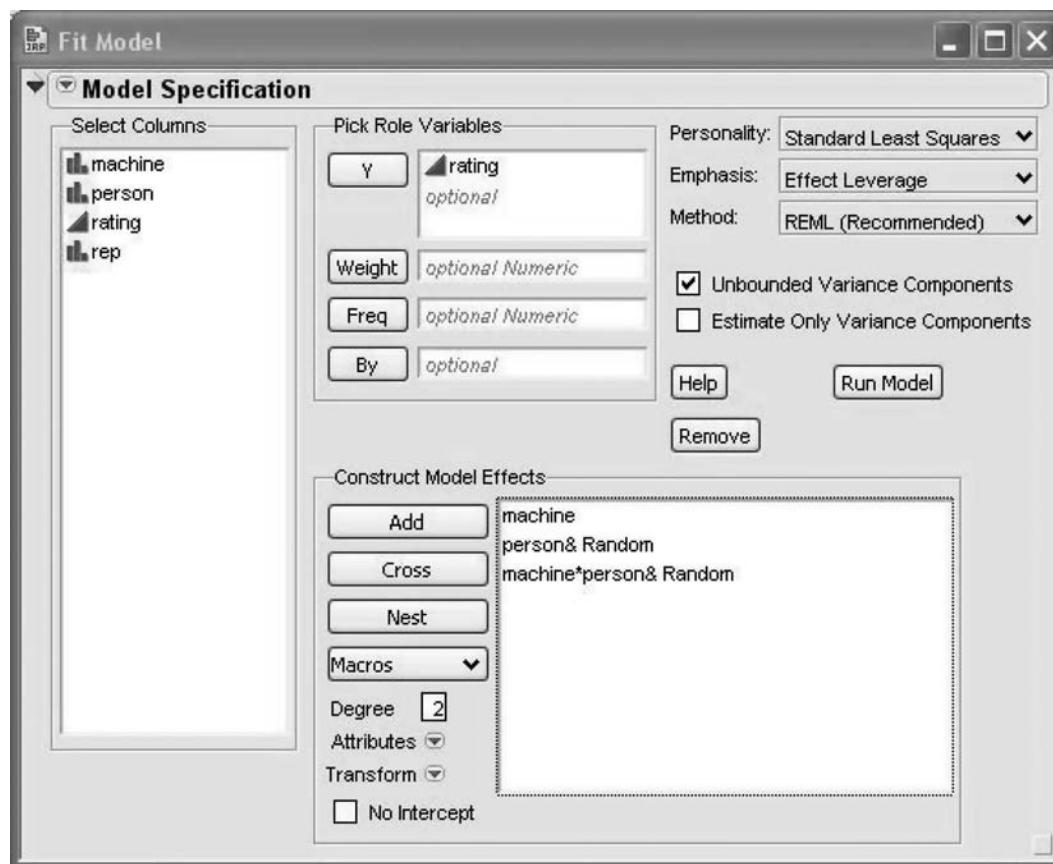


FIGURE 23.2 JMP fit model screen with response variable and model effects.

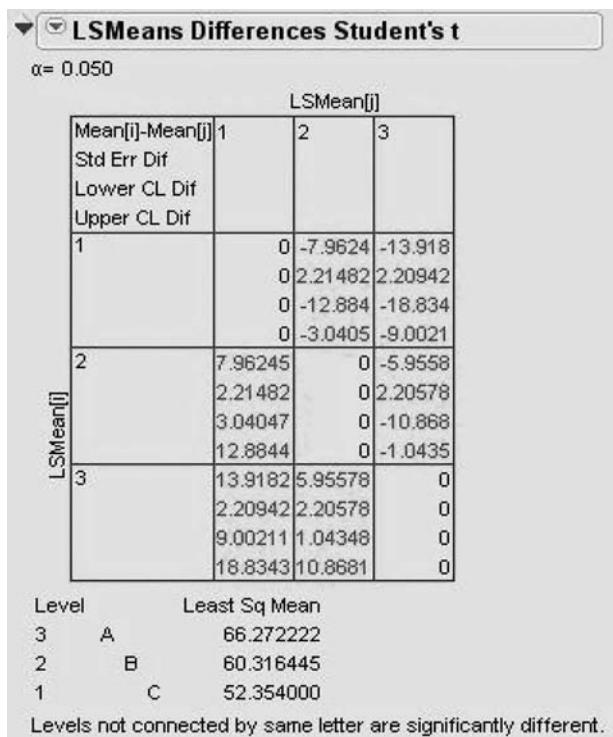
Random Effect	Var Ratio	Var Component	Std Error	95% Lower	95% Upper	Pct of Total
person	25.785476	22.455765	17.414105	-11.67588	56.58741	59.785
machine*person	16.344588	14.233991	6.5151618	1.4642735	27.003708	37.896
Residual		0.8708687	0.2410664	0.5405379	1.6332462	2.319
Total		37.560625				100.000

FIGURE 23.3 REML estimates of the variance components from JMP.

Fixed Effect Tests					
Source	Nparm	DF	DFDen	F Ratio	Prob > F
machine	2	2	10.11	19.9639	0.0003*

FIGURE 23.4 Tests of the machine effects from JMP.

Least Squares Means Table		
Level	Least Sq Mean	Std Error
1	52.354000	2.4906713
2	60.316445	2.4874448
3	66.272222	2.4826077

FIGURE 23.5 Machine least squares means and estimated standard errors.**FIGURE 23.6** LSD multiple comparisons of machine means.

23.4 Concluding Remarks

The REML, MIVQUE0, and method of moments methods of estimating the variance components provide identical results for balanced data sets as long as the solutions for the variance components are positive. The methods also provide identical results for the analysis of the fixed effects. The REML method seems to provide the best results for the analysis of the fixed effects when the data set is unbalanced. The main indicator is that the error degrees of freedom for comparing the machine means using the REML estimates of the variance components were similar to those for the balanced data set. The results from JMP using the REML option provided the same estimates of the variance components and of the fixed effects as obtained from SAS-Mixed using the REML method. Unfortunately, the confidence intervals constructed about the variance components use the Wald method, which is only appropriate when there are a large number of degrees of freedom associated

with each of the estimates of the variance components. For small-size data sets, the chi-square confidence intervals using the Satterthwaite approximation to the degrees of freedom are superior to those provided by the Wald method.

23.5 Exercises

- 23.1 Compute the likelihood ratio tests to test $H_0: \sigma_p^2 = 0$ vs $H_a: \sigma_p^2 > 0$ and $H_0: \sigma_p^2 = 0$ vs $H_a: \sigma_p^2 > 0$ for the balanced data set in Table 23.1.
- 23.2 The data in the following table are from a study where five states were selected at random and three to five cities were randomly selected from the cities within a selected state. Within each city, six stores were selected, two of the stores were randomly selected from all of the locally owned stores in the city, two of the stores were randomly selected from all of the convenience stores in the city, and two of the stores were randomly selected from all of the chain stores in the city. The price of one pound of coffee was determined for each store.
- 1) What type of model should be used to describe this data? What are the assumptions?
 - 2) Carry out a random effects analysis by estimating the variance components, testing the respective variance components are equal to zero, and construct 95% confidence intervals about the variance components.
 - 3) Carry out the fixed effects analysis by testing the appropriate hypothesis, estimating the means, constructing a 95% confidence intervals about the differences between the store means, and carry out the appropriate multiple comparisons of the fixed effects means using the Tukey method.

State	City	Loc1	Loc2	Conv1	Conv2	Chain1	Chain2
1	1	2.87	2.60	2.55	2.79	2.51	2.72
1	2	2.40	2.55	2.53	2.32	2.23	2.46
1	3	2.65	2.60	2.60	3.10	2.65	2.54
1	4	2.75	2.82	3.03	2.95	2.75	3.01
1	5	2.47	2.42	2.54	2.71	2.52	2.52
2	1	2.60	2.79	2.63	2.87	2.70	2.60
2	2	2.30	2.18	2.18	2.28	2.39	2.51
2	3	2.14	2.25	2.26	2.35	2.33	2.17
3	1	2.37	2.37	2.44	2.37	2.31	2.31
3	2	2.34	2.38	2.33	2.46	2.38	2.30
3	3	2.31	2.19	2.42	2.21	2.10	2.13
3	4	2.49	2.31	2.45	2.34	2.39	2.33
4	1	2.85	2.41	2.53	2.49	2.79	2.57
4	2	2.65	2.48	2.72	2.83	2.75	2.59
4	3	2.62	2.61	2.56	2.84	2.85	2.49
4	4	2.52	2.62	2.79	2.61	2.40	2.54
4	5	2.45	2.33	2.25	2.39	2.36	2.35
5	1	2.69	2.48	2.44	2.81	2.35	2.41
5	2	2.52	2.55	2.33	2.53	2.65	2.36
5	3	2.50	2.13	2.49	2.37	2.25	2.14

- 23.3 The data in the following table are from the data set for Exercise 23.2 where some of the observations have been set to missing. Carry out the analysis of this price of one pound of coffee data by answering the following questions.

- 1) What type of model should be used to describe this data? What are the assumptions?
- 2) Carry out a random effects analysis by estimating the variance components, testing the respective variance components are equal to zero, and construct 95% confidence intervals about the variance components.
- 3) Carry out the fixed effects analysis by testing the appropriate hypothesis, estimating the means, constructing a 95% confidence intervals about the differences between the store means, and carry out the appropriate multiple comparisons of the fixed effects means using the Tukey method.

State	City	Loc1	Loc2	Conv1	Conv2	Chain1	Chain2
1	1	2.87	2.60	—	2.79	—	—
1	2	2.40	2.55	2.53	2.32	2.23	2.46
1	3	—	2.60	2.60	—	2.65	2.54
1	4	2.75	2.82	3.03	2.95	—	3.01
1	5	2.47	—	2.54	—	2.52	2.52
2	1	—	2.79	2.63	2.87	2.70	—
2	2	2.30	—	2.18	—	—	2.51
2	3	2.14	—	2.26	—	—	—
3	1	2.37	2.37	2.44	2.37	2.31	2.31
3	2	2.34	2.38	2.33	2.46	—	2.30
3	3	—	—	2.42	—	2.10	2.13
3	4	2.49	2.31	—	—	—	2.33
4	1	2.85	2.41	2.53	—	—	2.57
4	2	2.65	—	—	—	2.75	2.59
4	3	—	—	2.56	2.84	2.85	—
4	4	2.52	2.62	—	2.61	2.40	2.54
4	5	—	2.33	2.25	2.39	.	—
5	1	—	2.48	—	—	2.35	2.41
5	2	2.52	2.55	2.33	2.53	2.65	2.36
5	3	2.50	—	2.49	2.37	2.25	—

- 23.4 Carry out an analysis for the data in Table 12.1 assuming that the flour blocking factor is a random effect.
- 23.5 Carry out an analysis for the data in Table 15.1 assuming that the flour blocking factor is a random effect.
- 23.6 Carry out an analysis of the data in Exercise 15.1 assuming that gyms are a random effect.

24

Methods for Analyzing Split-Plot Type Designs

24.1 Introduction

The split-plot type design involves a design structure with more than one size of experimental unit where the smaller-size experimental units are nested within the larger-size experimental units. Some examples of split-plot design structures were presented in Chapter 5, including the class of hierachal design structures. Two main problems occur in the design and analysis of split-plot type design structures. The first problem consists of the selection and/or identification of the different sizes of experimental units used in the design structure followed by the assignment of treatments from the treatment structure to the different experimental unit sizes in the design structure. The successful identification of the different sizes of experimental units is paramount in the specification of an appropriate model that can describe the resulting data. The second problem is constructing the appropriate model that describes the pertinent features of the treatment and design structures. It is important to be able to identify the sources of variation that measure the variability associated with each size of experimental unit. These sources of variability are used to compute the respective error terms which are used to compute estimates of the standard errors of estimated means and for pairwise comparisons among means. Since these design structures involve more than one size of experimental units, the estimates of the standard errors of the fixed effect parameters and their comparisons among them involve one or more sources of variation. A very important characteristic of the model for the split-plot type designs is that they are the basic model for starting the construction of the repeated measures models discussed in Chapter 26. Examples of several of the concepts were presented in Chapter 5.

The design and analysis of a split-plot or hierachal design structures with two sizes of experimental units are explained in Section 24.1 and the determination and estimation of standard errors associated with the fixed effects are described in Section 24.2. A general method for determining the appropriate standard errors and their estimates for fixed effects parameter estimates in a general split-plot design structure is discussed in Section 24.3. The computations of standard errors of contrasts of means are discussed in Section 24.4. Four examples of split-plot design structures are presented in Section 24.6, where each

example demonstrates some salient features of analyzing such designs. A discussion of sample size determination and power computations for split-plot design structures is given in Section 24.7. Analyses using both SAS®-Mixed and JMP® are given in this chapter with the JMP analyses shown in Section 24.8.

The key concepts in constructing models for split-plot designs are recognizing the different sizes of experimental units and then identifying the corresponding design structures and treatment structures. The overall model is constructed by incorporating models developed for each size of experimental unit. Several examples of model construction were presented in Chapter 5, but the assumptions underlying the models were not stated. The assumptions are that the components denoting the error terms for the various experimental units are all distributed independently with zero means and an associated variance (see Chapter 26 for assumptions that are more general). Under ideal conditions, the error terms are normally distributed. The objective of an analysis is to use the model assumptions to obtain estimates of the population parameters and to make inferences about them. Both method of moments and REML are used in the following examples to demonstrate the computations needed to estimate the standard errors of the fixed effects. In practice, REML is the method that can be recommended in most cases.

24.1.1 Example 24.1: Bread Recipes and Baking Temperatures

The process of baking bread involves mixing a batch of bread dough according to the specifications of a recipe, putting the dough into a pan (container), letting the bread rise, and then putting the pan of bread dough into an oven to be baked at a specific temperature and time combination. Each oven is large enough so that four pans of bread dough can be put into one oven at the same time. An experiment was designed to evaluate how four different recipes of bread respond to baking at three different temperatures where the response measured is the volume of the resulting loaf of bread. The process is to make dough from each of the four recipes and place one loaf from each recipe into a single oven that is set to a specific temperature. The batches are left in the oven for a specified length of time and then the loaves are cooled to room temperature before measuring their volumes. The data in Table 24.1 are the volumes of loaves of bread made from the four recipes and three temperatures where the process was repeated on three different days. Days are considered as a blocking factor in the design structure.

TABLE 24.1

Loaf Volumes (cm³) for Various Recipes and Temperatures for Example 24.1

Day	Temperature	Recipe A	Recipe B	Recipe C	Recipe D
1	325	1143	1148	1181	1165
1	340	1420	1425	1340	1404
1	355	1222	1166	1231	1274
2	325	1225	1208	1177	1193
2	340	1447	1402	1353	1414
2	355	1209	1293	1322	1285
3	325	1133	1115	1122	1142
3	340	1298	1261	1190	1321
3	355	1179	1175	1236	1257

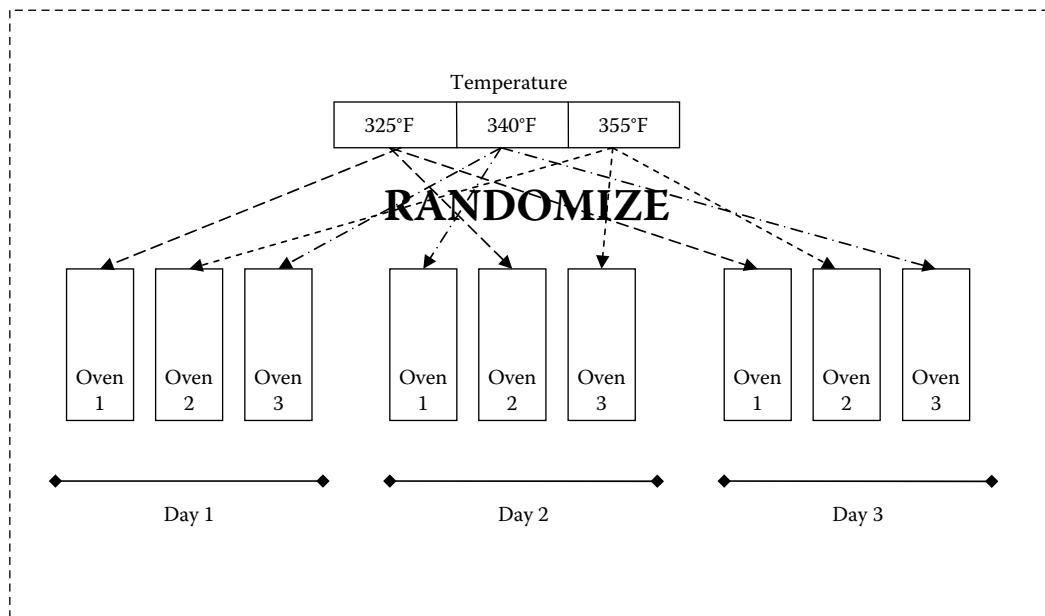


FIGURE 24.1 Schematic showing the randomization of temperatures to ovens within each day.

The diagram in Figure 24.1 demonstrates the process of assigning temperatures to the ovens, thus the ovens are the experimental units for levels of temperatures (notice that the recipes are not included in this step of the process). The design associated with the oven size experimental unit is a one-way treatment structure (three levels of temperature) in a randomized complete block design structure (three days). A model that could be used to describe the mean loaf volume of the four loaves within each oven (i.e. one observation per oven) is

$$y_{ik}^o = \mu_i^o + d_k^o + o_{ik}^o$$

where y_{ik}^o denotes the observed mean loaf volume, μ_i^o denotes the mean loaf volume from the i th level of temperature, d_k^o denotes the random effect of the k th day, and o_{ik}^o denotes the random oven effect from the i th temperature and the k th day. It is assumed that $d_k^o \sim i.i.d. N(0, \sigma_d^2)$, $o_{ik}^o \sim i.i.d. N(0, \sigma_o^2)$, and all of the d_k^o and o_{ik}^o are independent. The analysis of variance associated with the oven model is displayed in Table 24.2, where the oven error term is computed as the temperature by day interaction mean square.

TABLE 24.2

Analysis of Variance Table for the Oven Level Analysis

Source	df
Day	2
Temperature	2
Error(oven) = Day × Temperature	4

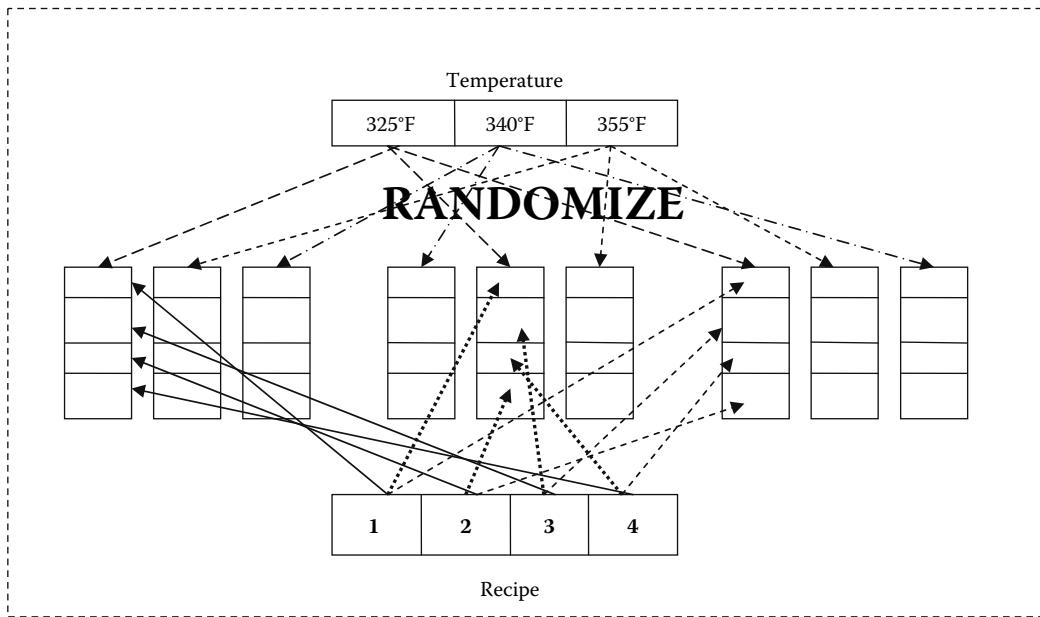


FIGURE 24.2 Schematic showing the randomization of recipes to positions within one oven within each day.

The diagram in Figure 24.2 shows the assignment of a recipe to a position within each oven. The oven within a day is a block for the four recipes. Each recipe provides one loaf of bread. The loaf design is a one-way treatment structure (four recipes) in a randomized complete block design structure with nine blocks (three blocks or ovens on each of three days).

If all ovens were at the same temperature, then the data structure would be a one-way treatment structure in a randomized complete block design structure. But not all ovens are treated alike, three are set at 325°F, three are set at 340°F, and three are set at 355°F. Next simplify the design by considering all data from 325°F, as shown in Figure 24.3. The resulting data are from a one-way treatment structure in a randomized complete block design structure with three blocks. The error term for this data is the recipe by day interaction. A model that can be used to describe the data from the 325°F ovens is

$$y_{jk}^+ = \mu_j^+ + o_k^+ + \varepsilon_{jk}^+$$

where y_{jk}^+ denotes the observed volume of the loaf in the 325°F oven of the j th recipe on the k th day, μ_j^+ denotes the mean loaf volume from the j th recipe, o_k^+ denotes the random effect of the oven used on the k th day, and ε_{jk}^+ denotes the random loaf effect from the j th recipe and the k th day. It is assumed that $o_k^+ \sim i.i.d. N(0, \sigma_o^2)$, $\varepsilon_{jk}^+ \sim i.i.d. N(0, \sigma_\varepsilon^2)$, and all o_k^+ and ε_{jk}^+ are independently distributed. The analysis of variance associated with the loaf model is displayed in Table 24.3, where the loaf error term is estimated by the recipe by day interaction mean square.

The error term that measures the loaf to loaf variability is computed by pooling the recipe by day interaction across the three temperatures, obtaining $Error(\text{loaf}) = \text{Recipe} \times \text{Day}(\text{Temperature})$. Putting the two models together provides the model

$$y_{ijk} = \mu_{ij} + d_k + o_{ik} + \varepsilon_{ijk}, \quad i = 1, 2, 3, \quad j = 1, 2, 3, 4, \quad k = 1, 2, 3$$

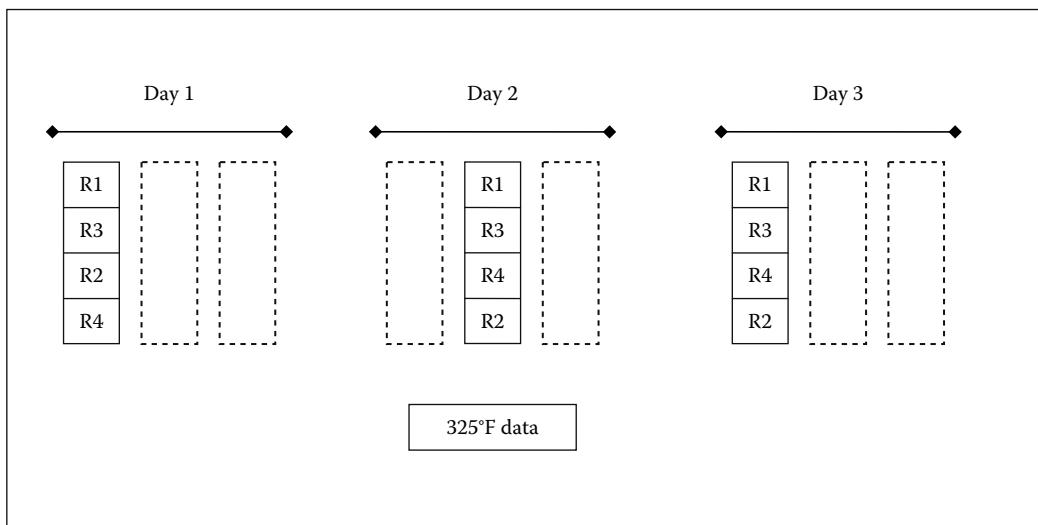


FIGURE 24.3 Part of the design with ovens at 325°F.

where $\mu_{ij} = \mu + T_i + R_j + (TR)_{ij}$, T_i denotes the effect of the i th temperature, R_j denotes the j th recipe effect, $(TR)_{ij}$ denotes the temperature by recipe interaction, d_k denotes the k th day, o_{ik} denotes the i th oven effect on the k th day, and ε_{ijk} represents the error term. Under the ideal conditions $d_k \sim i.i.d. N(0, \sigma_{\text{day}}^2)$, $o_{ik} \sim i.i.d. N(0, \sigma_{\text{oven}}^2)$, $\varepsilon_{ijk} \sim i.i.d. N(0, \sigma_{\text{loaf}}^2)$, and all d_k , o_{ik} , and ε_{ijk} are independently distributed. The oven is the whole plot or larger-sized experimental unit and the loaf is the subplot or split-plot or smaller sized experimental unit.

The model can be expressed by size of experimental unit as

$$\begin{aligned} y_{ijk} = & \mu_{ij} + d_k + T_i + o_{ik} \} \quad \text{whole-plot or oven part of the model} \\ & + R_j + (TR)_{ij} + \varepsilon_{ijk} \} \quad \text{subplot or loaf part of the model} \end{aligned}$$

Combining the analyses of variance tables in Tables 24.2 and 24.3 gives the analysis of variance table for this model in Table 24.4. The expected mean squares dictate the appropriate denominator for computing test statistics for the fixed effects. The *Error(oven)* is used as the error to test for temperature main effects and the *Error(loaf)* is used as the error to test for the recipe main effect and the temperature by recipe interaction effects. The SAS-Mixed code and resulting analysis of variance table for the loaf volume data are

TABLE 24.3
Analysis of Variance Table for the Loaf
Volume Data at 325°F

Source	df
Day	2
Recipe	3
<i>Error(loaf) = Day × Recipe</i>	6

TABLE 24.4

Analysis of Variance Table for the Loaf Volume Data of Example 24.1

Source	df	Expected Mean Square
Day	2	$\sigma_{\text{loaf}}^2 + 4\sigma_{\text{oven}}^2 + 12\sigma_{\text{day}}^2$
Temperature (T)	2	$\sigma_{\text{loaf}}^2 + 4\sigma_{\text{oven}}^2 + \phi^2(T)$
Error(oven)	4	$\sigma_{\text{loaf}}^2 + 4\sigma_{\text{oven}}^2$
Recipe (R)	3	$\sigma_{\text{loaf}}^2 + \phi^2(R)$
T × R	6	$\sigma_{\text{loaf}}^2 + \phi^2(T \times R)$
Error(loaf)	18	σ_{loaf}^2

TABLE 24.5

Analysis of Variance Table and SAS-Mixed Code for the Loaf Volume Data of Example 24.1

Source	df	SS	MS	EMS	F-Value	Pr F
Temperature	2	228756.6	114378.3	$\text{Var}(\text{Residual}) + 4 \text{Var}(\text{day} \times \text{temperature}) + Q(\text{temp}, \text{temp} \times \text{recipe})$	27.92	0.0045
Recipe	3	6041.4	2013.8	$\text{Var}(\text{Residual}) + Q(\text{recipe}, \text{temp} \times \text{recipe})$	3.06	0.0547
Temperature × recipe	6	21790.3	3631.7	$\text{Var}(\text{Residual}) + Q(\text{temp} \times \text{recipe})$	5.52	0.0021
Day	2	51777.7	25888.8	$\text{Var}(\text{Residual}) + 4 \text{Var}(\text{day} \times \text{temp}) + 12 \text{Var}(\text{day})$	6.32	0.0578
Day × temperature	4	16385.7	4096.4	$\text{Var}(\text{Residual}) + 4 \text{Var}(\text{day} \times \text{temp})$	6.23	0.0025
Residual	18	11841.6	657.9	$\text{Var}(\text{Residual})$	—	—

in Table 24.5. There is indication of a significant temperature by recipe interaction ($p = 0.0021$), thus further comparisons among the $\text{temperature} \times \text{recipe}$ two-way means should follow (see Section 24.2).

24.1.2 Example 24.2: Wheat Varieties Grown in Different Fertility Regimes

The data in Figure 24.4 are the yields in pounds of two varieties of wheat (B_1 and B_2) grown in four different fertility regimes (A_1, A_2, A_3 , and A_4). The field was divided into two blocks, each with four whole plots. Each of the four fertilizer levels was randomly assigned to one whole plot within each block. Thus, the whole plot design consists of a one-way treatment structure (four levels of fertilizer) in a randomized complete block design structure with two blocks. Each block contains four whole plot experimental units which were split into two parts (called subplots). Each variety of wheat was randomly assigned to one subplot within each whole plot. The subplot design consists of a one-way treatment structure (two varieties) in a randomized complete block design with eight blocks where each block contains two subplot experimental units. A model that can be used to describe this data is

$$y_{ijk} = \mu_{ij} + b_k + w_{ik} + \varepsilon_{ijk}, \quad i = 1, 2, 3, 4, \quad j = 1, 2, \quad k = 1, 2$$

		Block 1		Block 2		
		Variety		Variety		
Fertility regime	V ₁	V ₂	V ₁	V ₂		
	A ₁	35.4	37.9	A ₁	41.6	40.3
	A ₂	36.7	38.2	A ₂	42.7	41.6
	A ₃	34.8	36.4	A ₃	43.6	42.8
	A ₄	39.5	40.0	A ₄	44.5	47.6

Subplot Whole-plot

FIGURE 24.4 Data for the variety by fertility regime split-plot example.

where μ_{ij} denotes the expected response (yield) for the i th fertility level and j th variety, and y_{ijk} denotes the observed yield (response) from the k th block with the i th fertility level and j th variety, b_k denotes the block effect which is assumed to be distributed iid $N(0, \sigma_{\text{block}}^2)$, w_{ik} denotes the whole plot error which is assumed to be distributed *i.i.d.* $N(0, \sigma_{\text{wp}}^2)$ and ε_{ijk} denotes the subplot error which is assumed to be distributed *i.i.d.* $N(0, \sigma_{\varepsilon}^2)$. It is also assumed that all of the b_k , w_{ik} , and ε_{ijk} are distributed independently. The mean response can be expressed using an effects model representation as $\mu_{ij} = \mu + F_i + V_j + (FV)_{ij}$. The analysis of variance table for this effects model is in Table 24.6. The denominators for the F -tests for the three fixed effect comparisons are determined by the expected mean squares; that is, the *Error(whole plot)* is used to test for fertility main effects and the *Error(subplot)* is used to test for variety main effects and fertility by variety interaction effects. The SAS-Mixed code and the numerical results using type III sums of squares are given in Table 24.7, and results using the REML option are given in Table 24.8. The F -tests for the fixed effects are identical for these two analyses since the data set is balanced and both estimates of the variance components are greater than zero.

TABLE 24.6
Analysis of Variance Table for the Wheat Yield Data
of Example 24.2

Source	df	Expected Mean Square
Block	1	$\sigma_{\varepsilon}^2 + 2\sigma_{\text{wp}}^2 + 8\sigma_{\text{block}}^2$
Fertility (F)	3	$\sigma_{\varepsilon}^2 + 2\sigma_{\text{wp}}^2 + \phi^2(F)$
Error(whole plot)	3	$\sigma_{\varepsilon}^2 + 2\sigma_{\text{wp}}^2$
Variety (V)	1	$\sigma_{\varepsilon}^2 + \phi^2(V)$
$F \times V$	3	$\sigma_{\varepsilon}^2 + \phi^2(F \times V)$
Error(subplot)	4	σ_{ε}^2

TABLE 24.7

Analysis of Variance Table Using Type III Sums of Squares with the Wheat Yield Data of Example 24.2

```
proc mixed data=ex24_1 method=type3;
class block a b;
model y = a|b/ddfm=kr;
random block a*block;
lsmeans a|b/diff;
```

Source	df	SS	MS	EMS	F-Value	Pr F
a	3	40.2	13.4	Var(Residual) + 2 Var(block \times a) + Q(a,a \times b)	5.80	0.0914
b	1	2.3	2.3	Var(Residual) + Q(b,a \times b)	1.07	0.3599
a \times b	3	1.6	0.5	Var(Residual) + Q(a \times b)	0.25	0.8612
Block	1	131.1	131.1	Var(Residual) + 2 Var(block \times a) + 8 Var(block)	56.77	0.0048
Block \times a	3	6.9	2.3	Var(Residual) + 2 Var(block \times a)	1.10	0.4476
Residual	4	8.4	2.1	Var(Residual)	—	—

TABLE 24.8

Analysis of Variance Table Using REML with the Wheat Yield Data of Example 24.2

```
proc mixed data=ex24_1;
class block a b;
model y=a|b/ddfm=kr;
random block a*block;
lsmeans a|b/ diff;
```

Covariance Parameter Estimates

Covariance Parameter	Estimate
Block	16.0992
Block \times a	0.1008
Residual	2.1075

Type III Tests of Fixed Effects

Effect	Num df	Den df	F-Value	Pr > F
a	3	3	5.80	0.0914
b	1	4	1.07	0.3599
a \times b	3	4	0.25	0.8612

24.2 Model Definition and Parameter Estimation

The general model for the split-plot design with a randomized complete block whole plot design structure with r blocks, the whole-plot factor (A) with a levels and the subplot factor (C) with c levels is

$$y_{ijk} = \mu + \alpha_i + b_k + w_{ik} + \gamma_j + (\alpha\gamma)_{ij} + \varepsilon_{ijk}, \quad i = 1, 2, \dots, a, \quad j = 1, 2, \dots, c, \quad k = 1, 2, \dots, r$$

where y_{ijk} is the observed response, b_k denotes the k th block effect which is assumed to be distributed $N(0, \sigma_B^2)$, w_{ik} denotes the whole plot error which is assumed to be distributed $N(0, \sigma_w^2)$, and ε_{ijk} denotes the subplot error and is assumed to be distributed $N(0, \sigma_\varepsilon^2)$. It is also assumed that all of the b_k , w_{ik} and ε_{ijk} are distributed independently. It should be noted that the most important assumption is that the all of the b_k , w_{ik} and ε_{ijk} are distributed independently. Fortunately, this assumption can be guaranteed by the randomization process as the fixed effect factors are randomly assigned to their appropriate sized experimental units. The fixed effects in this model are the overall mean, μ , the effect of the whole plot factor (A), α_i , the effect of the subplot factor (C), γ_j , and the effect of the interaction between the levels of the whole plot factor and the levels of the subplot factor, $(\alpha\gamma)_{ij}$. The means model can be represented in terms of the effects model as $\mu_{ij} = \mu + \alpha_i + \gamma_j + (\alpha\gamma)_{ij}$. The analysis of variance table for this general model with sources of variability, degrees of freedom and expected mean squares is in Table 24.9.

The whole plot error is computed as the $Block \times A$ interaction and the subplot error is computed as the $Block \times C$ interaction pooled across the levels of A , denoted as $Block \times C(A)$. The equations from which to obtain the method of moments estimates of the variance components are

$$\begin{aligned} MSBlock &= \tilde{\sigma}_\varepsilon^2 + c\tilde{\sigma}_w^2 + ac\tilde{\sigma}_B^2 \\ MSError(whole\ plot) &= \tilde{\sigma}_\varepsilon^2 + c\tilde{\sigma}_w^2 \end{aligned}$$

and

$$MSError(subplot) = \tilde{\sigma}_\varepsilon^2$$

The method of moments solution to these equations is

$$\begin{aligned} \tilde{\sigma}_\varepsilon^2 &= MSError(subplot) \\ \tilde{\sigma}_w^2 &= \frac{MSError(whole\ plot) - MSError(subplot)}{c} \end{aligned}$$

and

$$\tilde{\sigma}_B^2 = \frac{MSBlock - MSError(whole\ plot)}{ac}$$

TABLE 24.9

Analysis of Variance Table for the General Split-Plot Model in Section 24.1

Source	df	Expected Mean Square
Block	$r - 1$	$\sigma_\varepsilon^2 + c\sigma_w^2 + ac\sigma_B^2$
A	$a - 1$	$\sigma_\varepsilon^2 + c\sigma_w^2 + \phi^2(\alpha)$
Error(whole plot)	$(r - 1)(a - 1)$	$\sigma_\varepsilon^2 + c\sigma_w^2$
C	$c - 1$	$\sigma_\varepsilon^2 + \phi^2(\gamma)$
$A \times C$	$(a - 1)(c - 1)$	$\sigma_\varepsilon^2 + \phi^2(\alpha\gamma)$
Error(subplot)	$a(r - 1)(c - 1)$	σ_ε^2

The method of moments estimates of the variance components are

$$\hat{\sigma}_\epsilon^2 = \tilde{\sigma}_\epsilon^2$$

$$\hat{\sigma}_w^2 = \begin{cases} \tilde{\sigma}_w^2 & \text{if } \tilde{\sigma}_w^2 > 0 \\ 0 & \text{if } \tilde{\sigma}_w^2 \leq 0 \end{cases}$$

and

$$\hat{\sigma}_B^2 = \begin{cases} \tilde{\sigma}_B^2 & \text{if } \tilde{\sigma}_B^2 > 0 \\ 0 & \text{if } \tilde{\sigma}_B^2 \leq 0 \end{cases}$$

The estimators of μ_{ij} , $\bar{\mu}_{i..}$, and $\bar{\mu}_{.j}$ are $\bar{y}_{ij..}$, $\bar{y}_{i..}$, and $\bar{y}_{.j..}$ respectively. The comparisons among the levels of A are between whole plot comparisons and the proper F -statistic for testing the equality of levels of A means is $F_A = MSA/MSError(\text{whole plot})$. The comparisons among the levels of C and for the $A \times C$ interaction are within whole plot comparisons or between subplot comparisons within a whole plot and the proper F -statistics are $F_C = MSC/MSError(\text{subplot})$ and $F_{A \times C} = MSA \times C/MSError(\text{subplot})$. These F -statistics were constructed by looking at the expected mean squares in Table 24.9.

Once the F -tests have been computed to determine if there are significant differences between means, the next step is to carry out multiple comparisons to determine where the differences occur. The following section presents methods to compute standard errors of various differences of means for split-plot designs.

24.3 Standard Errors for Comparisons among Means

Contrasts of treatment means or treatment combination means are used to study the treatment effects, particularly when the analysis of variance indicates that one or more of the fixed effects are significantly different from zero. The standard error of a contrast of sample means is necessary to determine if a contrast in the means is equal to zero or to construct a confidence interval about the contrast in the means. More often than not, contrasts involve a comparison of two means. Consequently, comparisons of two means are discussed in this section and general contrasts are discussed in Section 24.4.

To demonstrate methods for determining appropriate standard errors, the general split-plot design model is used where the whole plot design is a one-way treatment structure in a randomized complete block design and the subplot treatment structure involves a one-way treatment structure. The corresponding effects model for this situation can be expressed as

$$y_{ijk} = \mu + \alpha_i + b_k + w_{ik} + \gamma_j + (\alpha\gamma)_{ij} + \varepsilon_{ijk}, \quad i = 1, 2, \dots, a, \quad j = 1, 2, \dots, c, \quad k = 1, 2, \dots, r$$

A means model can be expressed as

$$y_{ijk} = \mu_{ij} + b_k + w_{ik} + \varepsilon_{ijk}, \quad i = 1, 2, \dots, a, \quad j = 1, 2, \dots, c, \quad k = 1, 2, \dots, r$$

where the terms are as described in Section 24.1.

Four types of comparisons may be of interest, depending on whether any interaction exists between the levels of A and the levels of C . If there is no interaction, then it is of interest to compare the levels of A to one another and to compare the levels of C to one another.

To compare the levels of C , one needs to compare two of the $\bar{\mu}_{j\cdot}$, which are estimated by the $\bar{y}_{j\cdot}$. The process of determining the appropriate standard error involves expressing $\bar{y}_{j\cdot}$ in terms of the quantities in the model obtained by summing over i and k . The model for the j th main effect mean of C is $\bar{y}_{j\cdot} = \bar{\mu}_{j\cdot} + \bar{b}_{\cdot} + \bar{w}_{\cdot\cdot} + \bar{\epsilon}_{j\cdot}$. Consider the difference $\bar{\mu}_{1\cdot} - \bar{\mu}_{2\cdot}$. The estimate of $\bar{\mu}_{1\cdot} - \bar{\mu}_{2\cdot}$ is $\bar{y}_{1\cdot} - \bar{y}_{2\cdot}$, which can be expressed in terms of the C mean model as $\bar{y}_{1\cdot} - \bar{y}_{2\cdot} = \bar{\mu}_{1\cdot} - \bar{\mu}_{2\cdot} + \bar{\epsilon}_{1\cdot} - \bar{\epsilon}_{2\cdot}$, as the terms involving \bar{b}_{\cdot} and $\bar{w}_{\cdot\cdot}$ cancel out; that is, the comparison $\bar{y}_{1\cdot} - \bar{y}_{2\cdot}$ does not depend on the whole plot error, nor does it depend on the block error. The variance of $\bar{y}_{1\cdot} - \bar{y}_{2\cdot}$ can be shown to be equal to

$$\text{Var}(\bar{y}_{1\cdot} - \bar{y}_{2\cdot}) = \text{Var}(\bar{\epsilon}_{1\cdot} - \bar{\epsilon}_{2\cdot}) = \frac{2\sigma_{\epsilon}^2}{ar}$$

where the variance of the mean $\bar{\epsilon}_{1\cdot}$ is $\text{Var}(\bar{\epsilon}_{1\cdot}) = \sigma_{\epsilon}^2/ar$ and where ar is the number of observations in the mean. Similarly, one obtains $\text{Var}(\bar{y}_{j\cdot} - \bar{y}_{j'\cdot}) = 2\sigma_{\epsilon}^2/ar$ for any $j \neq j'$. The estimate of the standard error of $\bar{y}_{j\cdot} - \bar{y}_{j'\cdot}$ is

$$\widehat{s.e.}(\bar{y}_{j\cdot} - \bar{y}_{j'\cdot}) = \sqrt{\frac{2\hat{\sigma}_{\epsilon}^2}{ar}} = \sqrt{\frac{2\text{MSError}(\text{subplot})}{ar}} \quad \text{for all } j \neq j'$$

which is based on $a(c-1)(r-1)$ degrees of freedom. If one wants to carry out multiple comparisons (see Chapter 3), $a(c-1)(r-1)$ is the number of degrees of freedom needed to be used when determining the percentage point of the desired multiple comparison procedure. For simplicity, the LSD values are computed, but the LSD values may not be the appropriate method for a given situation. The LSD value for comparing two subplot treatment means is

$$\text{LSD}_{\alpha} = [t_{\alpha/2, a(c-1)(r-1)}] \widehat{s.e.}(\bar{y}_{j\cdot} - \bar{y}_{j'\cdot})$$

To compare the levels of A , one needs to compare the $\bar{\mu}_{i\cdot}$, which are estimated by the $\bar{y}_{i\cdot}$. The quantity $\bar{y}_{i\cdot}$ can be expressed in terms of the general model by summing over j and k , obtaining $\bar{y}_{i\cdot} = \bar{\mu}_{i\cdot} + \bar{b}_{\cdot} + \bar{w}_{\cdot\cdot} + \bar{\epsilon}_{ij\cdot}$.

The estimate of the contrast $\bar{\mu}_{1\cdot} - \bar{\mu}_{2\cdot}$ is $\bar{y}_{1\cdot} - \bar{y}_{2\cdot}$, which can be expressed in terms of the $\bar{y}_{i\cdot}$ model as $\bar{y}_{1\cdot} - \bar{y}_{2\cdot} = \bar{\mu}_{1\cdot} - \bar{\mu}_{2\cdot} + \bar{w}_{1\cdot} - \bar{w}_{2\cdot} + \bar{\epsilon}_{1\cdot} - \bar{\epsilon}_{2\cdot}$. This comparison depends on both the whole plot and subplot variance components. The variance of $\bar{y}_{1\cdot} - \bar{y}_{2\cdot}$ is

$$\begin{aligned} \text{Var}(\bar{y}_{1\cdot} - \bar{y}_{2\cdot}) &= \text{Var}(\bar{w}_{1\cdot} - \bar{w}_{2\cdot} + \bar{\epsilon}_{1\cdot} - \bar{\epsilon}_{2\cdot}) \\ &= \frac{2\sigma_w^2}{r} + \frac{2\sigma_{\epsilon}^2}{rc} \\ &= \frac{2(\sigma_{\epsilon}^2 + c\sigma_w^2)}{rc} \end{aligned}$$

The estimate of the standard error of $\bar{y}_{1..} - \bar{y}_{2..}$ is

$$\widehat{s.e.}(\bar{y}_{1..} - \bar{y}_{2..}) = \sqrt{\frac{2(\hat{\sigma}_e^2 + c\hat{\sigma}_w^2)}{rc}} = \sqrt{\frac{2MSError(wholeplot)}{rc}}$$

which is based on $(r - 1)(a - 1)$ degrees of freedom. The LSD value for comparing two whole plot treatment means is $LSD_\alpha = [t_{\alpha/2, (a-1)(r-1)}] \widehat{s.e.}(\bar{y}_{1..} - \bar{y}_{2..})$.

When there is a significant $A \times C$ interaction, comparisons usually must be based on the set of two-way cell means. There are two different types of comparisons one must consider when studying these cell means. The first type arises when two subplot treatment (C) means are compared at the same level of a whole plot treatment (A), such as $\mu_{11} - \mu_{12}$. The best estimator of $\mu_{11} - \mu_{12}$ is $\bar{y}_{11..} - \bar{y}_{12..}$. The term $\bar{y}_{ij..}$ can be expressed in terms of the general model by summing over k as $\bar{y}_{ij..} = \mu_{ij..} + \bar{b}_{..} + \bar{w}_{i..} + \bar{\varepsilon}_{ij..}$, and the estimate $\bar{y}_{11..} - \bar{y}_{12..}$ can be expressed as $\bar{y}_{11..} - \bar{y}_{12..} = \mu_{11} - \mu_{12} + \bar{\varepsilon}_{11..} - \bar{\varepsilon}_{12..}$. The variance of $\bar{y}_{11..} - \bar{y}_{12..}$ is $\text{Var}(\bar{y}_{11..} - \bar{y}_{12..}) = \text{Var}(\bar{w}_{1..} + \bar{\varepsilon}_{11..} - \bar{w}_{1..} - \bar{\varepsilon}_{12..}) = 2\sigma_e^2/r$. Thus the variance of comparisons between subplot treatments at the same level of the whole plot treatment depends only on the subplot error. The estimate of the standard error of $\bar{y}_{11..} - \bar{y}_{12..}$ is

$$\widehat{s.e.}(\bar{y}_{11..} - \bar{y}_{12..}) = \sqrt{\frac{2\hat{\sigma}_e^2}{r}} = \sqrt{\frac{2MSError(subplot)}{r}}$$

and the corresponding LSD value is $LSD_\alpha = [t_{\alpha/2, a(r-1)(c-1)}] \widehat{s.e.}(\bar{y}_{1..} - \bar{y}_{12..})$. This LSD value can be used to compare any pair of subplot treatments at the same level of a whole plot treatment.

The second type of comparison occurs when two whole plot treatments are compared at the same level or different levels of the subplot treatments, such as $\mu_{11} - \mu_{21}$ or $\mu_{11} - \mu_{22}$. These two types of comparisons have the same standard errors. The best estimate of $\mu_{11} - \mu_{21}$ is $\bar{y}_{11..} - \bar{y}_{21..}$, which can be expressed in terms of the general model as $\bar{y}_{11..} - \bar{y}_{21..} = \mu_{11} - \mu_{21} + \bar{w}_{1..} - \bar{w}_{2..} + \bar{\varepsilon}_{11..} - \bar{\varepsilon}_{21..}$. Then

$$\begin{aligned} \text{Var}(\bar{y}_{11..} - \bar{y}_{21..}) &= \text{Var}(\bar{w}_{1..} - \bar{w}_{2..} + \bar{\varepsilon}_{11..} - \bar{\varepsilon}_{21..}) \\ &= \frac{2\sigma_w^2}{r} + \frac{2\sigma_e^2}{r} \\ &= \frac{2(\sigma_e^2 + \sigma_w^2)}{r} \end{aligned}$$

This comparison depends on both the whole plot and the subplot variance components. An unbiased estimate of $\sigma_e^2 + \sigma_w^2$ is

$$\widehat{\sigma_e^2 + \sigma_w^2} = \frac{MSError(wholeplot) + (c - 1)MSError(subplot)}{c}$$

The sampling distribution associated with $\widehat{\sigma_e^2 + \sigma_w^2}$ is not a chi-square distribution, but is a linear combination of chi-square distributions. The degrees of freedom associated with $\widehat{\sigma_e^2 + \sigma_w^2}$ can be estimated using the Satterthwaite approximation as

$$\hat{v} = \frac{\widehat{(\sigma_e^2 + \sigma_w^2)^2}}{\frac{[MSE(wholeplot)/c]^2}{(r-1)(a-1)} + \frac{\left[\frac{c-1}{c} MSE(subplot)\right]^2}{a(r-1)(c-1)}}$$

An approximate LSD value for comparing two whole plot treatments at the same or different subplot treatment is

$$LSD_{\alpha} = (t_{\alpha/2, \hat{v}}) \sqrt{\frac{2(\widehat{\sigma_e^2 + \sigma_w^2})}{r}}$$

For the data in Example 24.1 (Table 24.1), the temperature means, recipe means and temperature by recipe means are given in Table 24.10. The estimated standard errors for the four types of comparisons are computed as follows:

1) For comparing recipe main effect means: the estimated standard error is

$$\widehat{s.e.}(\bar{y}_{1..} - \bar{y}_{2..}) = \sqrt{\frac{2(657.87)}{3(3)}} = 12.09$$

and this estimated standard error is based on 18 degrees of freedom. Note that $t_{0.025, 18} = 2.101$, so the 5% LSD value for comparing recipe main effect means is $LSD_{0.05} = 2.101(12.09) = 25.40$.

2) For comparing temperature main effect means: The estimated standard error is

$$\widehat{s.e.}(\bar{y}_{1..} - \bar{y}_{2..}) = \sqrt{\frac{2(4096.42)}{3(4)}} = 26.13$$

TABLE 24.10

Means for Recipes, Temperatures and for Combinations of Recipe and Temperature for the Data in Table 24.1

Temperature	Recipe A	Recipe B	Recipe C	Recipe D	Temperature Mean
325	1167.09	1157.00	1160.02	1166.68	1162.70
340	1388.47	1362.76	1294.17	1379.76	1356.29
355	1203.39	1211.07	1263.17	1272.06	1237.42
Recipe mean	1252.98	1243.61	1239.12	1272.83	

and this standard error is based on 4 degrees of freedom. Then $t_{0.025,4} = 2.776$ and the 5% LSD value for comparing temperature main effect means is $LSD_{0.05} = 2.776(26.13) = 72.55$.

3) *For comparing recipe means at same temperature level:* The estimated standard error is

$$\widehat{s.e.}(\bar{y}_{11.} - \bar{y}_{12.}) = \sqrt{\frac{2(657.87)}{3}} = 20.94$$

and this estimated standard error is based on 18 degrees of freedom. Hence, the 5% LSD value for comparing two recipe means at the same temperature level is $LSD_{0.05} = 2.101(20.94) = 43.99$.

4) *For comparing temperature means at the same or different recipe levels:* The estimated standard error is

$$\widehat{s.e.}(\bar{y}_{11.} - \bar{y}_{21.}) = \sqrt{\frac{2\left[\frac{1}{4}4096.42 + \frac{(4-1)}{4}657.87\right]}{3}} = \sqrt{\frac{2[1517.51]}{3}} = 31.80$$

and it is based on

$$\hat{v} = \frac{\frac{(1517.51)^2}{(4-1)4096.42^2}}{4} + \frac{\frac{(3)}{4}657.87^2}{18} = 8.35 \text{ degrees of freedom}$$

Note that $t_{0.025,8.35} = 2.289$, and thus, the 5% LSD for comparing two means that are at different temperature levels is $LSD_{0.05} = 2.289(31.80) = 72.79$.

Next a general process for computing standard errors split-plot type designs is described.

24.4 A General Method for Computing Standard Errors of Differences of Means

Applying the techniques presented above for computing standard errors for comparisons involving more than one size of experimental unit is not always straightforward. In this section, a general method for computing standard errors and approximate degrees of freedom values is described that can be applied to more complex situations. This section may be skipped by the casual reader. The method is described by using an example with the model

$$y_{ijk} = \mu + \alpha_i + b_k + w_{ik} + \gamma_j + (\alpha\gamma)_{ij} + \varepsilon_{ijk}, \quad i = 1, 2, \dots, a, \quad j = 1, 2, \dots, c, \quad k = 1, 2, \dots, r$$

Basically, the technique consists of expressing a given comparison of means as the sum of components where each component is a comparison of means involving only one size of experimental unit. Then the components of the comparison are independently distributed,

and the variance of the comparison is obtained by summing the variances of the components. For example, the comparison $\mu_{11} - \mu_{21}$ can be expressed as

$$\mu_{11} - \mu_{21} = (\bar{\mu}_{1..} - \bar{\mu}_{2..}) + [(\mu_{11} - \bar{\mu}_{1..}) - (\mu_{21} - \bar{\mu}_{2..})]$$

The component $(\bar{\mu}_{1..} - \bar{\mu}_{2..})$ is a whole plot comparison, and the component $[(\mu_{11} - \bar{\mu}_{1..}) - (\mu_{21} - \bar{\mu}_{2..})]$ is a subplot comparison. The estimates of these components are $\hat{\mu}_{1..} - \hat{\mu}_{2..} = \bar{y}_{1..} - \bar{y}_{2..}$ and $(\hat{\mu}_{11} - \hat{\mu}_{1..}) - (\hat{\mu}_{21} - \hat{\mu}_{2..}) = (\bar{y}_{11..} - \bar{y}_{1..}) - (\bar{y}_{21..} - \bar{y}_{2..})$, respectively. The estimate of $\mu_{11} - \mu_{21}$ is

$$\begin{aligned}\hat{\mu}_{11} - \hat{\mu}_{21} &= \hat{\mu}_{1..} - \hat{\mu}_{2..} + (\hat{\mu}_{11} - \hat{\mu}_{1..}) - (\hat{\mu}_{21} - \hat{\mu}_{2..}) \\ &= \bar{y}_{11..} - \bar{y}_{21..}\end{aligned}$$

Since comparisons computed from the whole plot part of the model are independently distributed of comparisons computed from the subplot part, the variance of $\hat{\mu}_{11} - \hat{\mu}_{21}$ is

$$\text{Var}(\hat{\mu}_{11} - \hat{\mu}_{21}) = \text{Var}(\bar{y}_{1..} - \bar{y}_{2..}) + \text{Var}[(\bar{y}_{11..} - \bar{y}_{1..}) - (\bar{y}_{21..} - \bar{y}_{2..})]$$

The quantity $\bar{y}_{1..} - \bar{y}_{2..}$ expressed in terms of the general model is $\bar{y}_{1..} - \bar{y}_{2..} = \bar{\mu}_{1..} - \bar{\mu}_{2..} + \bar{w}_{1..} - \bar{w}_{2..} + \bar{\varepsilon}_{1..} - \bar{\varepsilon}_{2..}$ and its variance is

$$\begin{aligned}\text{Var}(\bar{y}_{1..} - \bar{y}_{2..}) &= \text{Var}(\bar{w}_{1..} - \bar{w}_{2..} + \bar{\varepsilon}_{1..} - \bar{\varepsilon}_{2..}) \\ &= \frac{2}{rc}(\sigma_e^2 + c\sigma_w^2)\end{aligned}$$

An estimate of $\text{Var}(\bar{y}_{1..} - \bar{y}_{2..})$ is $(2/rc) \text{MSError(WholePlot)}$.

The quantity $(\bar{y}_{11..} - \bar{y}_{1..}) - (\bar{y}_{21..} - \bar{y}_{2..})$ expressed in terms of the general model is

$$\begin{aligned}(\bar{y}_{11..} - \bar{y}_{1..}) - (\bar{y}_{21..} - \bar{y}_{2..}) &= (\mu_{11} + \bar{b}_{1..} + \bar{w}_{1..} + \bar{\varepsilon}_{11..}) - (\bar{\mu}_{1..} + \bar{b}_{1..} + \bar{w}_{1..} + \bar{\varepsilon}_{1..}) \\ &\quad - (\mu_{21} + \bar{b}_{2..} + \bar{w}_{2..} + \bar{\varepsilon}_{21..}) + (\bar{\mu}_{2..} + \bar{b}_{2..} + \bar{w}_{2..} + \bar{\varepsilon}_{2..}) \\ &= [(\mu_{11} - \bar{\mu}_{1..}) - (\mu_{21} - \bar{\mu}_{2..})] + [(\bar{\varepsilon}_{11..} - \bar{\varepsilon}_{1..}) - (\bar{\varepsilon}_{21..} - \bar{\varepsilon}_{2..})]\end{aligned}$$

and its variance is

$$\begin{aligned}\text{Var}[(\bar{y}_{11..} - \bar{y}_{1..}) - (\bar{y}_{21..} - \bar{y}_{2..})] &= \text{Var}[(\bar{\varepsilon}_{11..} - \bar{\varepsilon}_{1..}) - (\bar{\varepsilon}_{21..} - \bar{\varepsilon}_{2..})] \\ &= \frac{2(c-1)}{cr}\sigma_e^2\end{aligned}$$

An estimate of $\text{Var}[(\bar{y}_{11..} - \bar{y}_{1..}) - (\bar{y}_{21..} - \bar{y}_{2..})]$ is

$$\widehat{\text{Var}}[(\bar{y}_{11..} - \bar{y}_{1..}) - (\bar{y}_{21..} - \bar{y}_{2..})] = \frac{2(c-1)}{rc} \text{MSError(subplot)}$$

Combining the whole plot component variance and the subplot component variance yields

$$\text{Var}[\bar{y}_{11\cdot} - \bar{y}_{21\cdot}] = \left(\frac{2}{rc}\right)(\sigma_e^2 + c\sigma_w^2) + \frac{2(c-1)}{rc}\sigma_e^2$$

and the estimate of the standard error of $\bar{y}_{11\cdot} - \bar{y}_{21\cdot}$ is

$$s.e.[\bar{y}_{11\cdot} - \bar{y}_{21\cdot}] = \sqrt{\left(\frac{2}{rc}\right)\text{MSE}(wholeplot) + \frac{2(c-1)}{rc}\text{MSE}(subplot)}$$

The approximate number of degrees of freedom associated with the $\widehat{s.e.}[\bar{y}_{11\cdot} - \bar{y}_{21\cdot}]$ can be obtained using the Satterthwaite approximation as:

$$\hat{v} = \frac{\left\{\left(\frac{2}{rc}\right)[\text{MSE}(wholeplot)] + \frac{2(c-1)}{rc}[\text{MSE}(subplot)]\right\}^2}{\left\{\left(\frac{2}{rc}\right)[\text{MSE}(wholeplot)]\right\}^2 + \frac{\left\{\frac{2(c-1)}{rc}[\text{MSE}(subplot)]\right\}^2}{a(r-1)(c-1)}}$$

With some simple algebra one can show the above $\widehat{s.e.}[\bar{y}_{11\cdot} - \bar{y}_{12\cdot}]$ is identical to that obtained in Section 24.2.

24.5 Comparison via General Contrasts

General contrasts of means can be constructed for each of the comparisons discussed in Section 24.2. A contrast between the levels of C or the $\bar{\mu}_{\cdot j}$ is

$$\theta = d_1\bar{\mu}_{\cdot 1} + d_2\bar{\mu}_{\cdot 2} + \cdots + d_c\bar{\mu}_{\cdot c} \text{ where } \sum_{j=1}^c d_j = 0$$

An estimate of the contrast is $\hat{\theta} = d_1\bar{y}_{\cdot 1} + d_2\bar{y}_{\cdot 2} + \cdots + d_c\bar{y}_{\cdot c}$ with variance

$$\text{Var}(\hat{\theta}) = \frac{\sigma_e^2}{ar} \sum_{j=1}^c d_j^2$$

The estimate of the standard error of $\hat{\theta}$ is

$$\widehat{s.e.}(\hat{\theta}) = \sqrt{\frac{\text{MSError}(subplot)}{ar} \sum_{j=1}^c d_j^2}$$

A contrast between the levels of A or the $\bar{\mu}_i$, is $\tau = h_1\bar{\mu}_{1.} + h_2\bar{\mu}_{2.} + \cdots + h_a\bar{\mu}_{a.}$ where $\sum_{i=1}^a h_i = 0$; its estimate is $\hat{\tau} = h_1\bar{y}_{1..} + h_2\bar{y}_{2..} + \cdots + h_a\bar{y}_{a..}$ with variance

$$\text{Var}(\hat{\tau}) = \left(\frac{\sigma_e^2 + c\sigma_w^2}{rc} \right) \sum_{i=1}^a h_i^2$$

The estimate of the standard error of $\hat{\tau}$ is

$$\widehat{s.e.}(\hat{\tau}) = \sqrt{\left(\frac{\text{MSError(wholeplot)}}{rc} \right) \sum_{i=1}^a h_i^2}$$

A contrast of subplot treatments at the same level of a whole plot treatment; that is, a contrast of the $\mu_{i1}, \mu_{i2}, \dots, \mu_{ic}$ is $\delta_i = s_1\mu_{i1} + s_2\mu_{i2} + \cdots + s_c\mu_{ic}$ where $\sum_{j=1}^c s_j = 0$. Its estimate is $\hat{\delta}_i = s_1\bar{y}_{i1.} + s_2\bar{y}_{i2.} + \cdots + s_c\bar{y}_{ic.}$ with variance

$$\text{Var}(\hat{\delta}_i) = \frac{\sigma_e^2}{r} \sum_{j=1}^c s_j^2$$

The estimate of the standard error of $\hat{\delta}_i$ is

$$\widehat{s.e.}(\hat{\delta}_i) = \sqrt{\frac{\text{MSError(subplot)}}{r} \sum_{j=1}^c s_j^2}$$

A contrast of whole plot treatments at the same subplot treatment, that is, a contrast of $\mu_{1j}, \mu_{2j}, \dots, \mu_{aj}$ is $\lambda_j = u_1\mu_{1j} + u_2\mu_{2j} + \cdots + u_a\mu_{aj}$ where $\sum_{i=1}^a u_i = 0$. It is estimated by $\hat{\lambda}_j = u_1\bar{y}_{1j.} + u_2\bar{y}_{2j.} + \cdots + u_a\bar{y}_{aj.}$, and the variance of $\hat{\lambda}_j$ is

$$\text{Var}(\hat{\lambda}_j) = \frac{(\sigma_e^2 + \sigma_w^2)}{r} \sum_{i=1}^a u_i^2$$

The estimated standard error of $\hat{\lambda}_k$ is

$$\widehat{s.e.}(\hat{\lambda}_j) = \sqrt{\frac{(\hat{\sigma}_e^2 + \hat{\sigma}_w^2)}{r} \sum_{i=1}^a u_i^2}$$

where

$$\hat{\sigma}_e^2 + \hat{\sigma}_w^2 = \frac{\text{MSError(wholeplot)} + (c-1)\text{MSError(subplot)}}{c}$$

The number of degrees of freedom associated with $\hat{\sigma}_e^2 + \hat{\sigma}_w^2$ is obtained using the Satterthwaite approximation (see Section 24.2) as

$$\hat{v} = \frac{\frac{(\hat{\sigma}_e^2 + \hat{\sigma}_w^2)^2}{[MSE(wholeplot)/c]^2} + \frac{[\frac{c-1}{c} MSE(subplot)]^2}{a(r-1)(c-1)}}{a(r-1)(c-1)}$$

Next consider any linear combination of the μ_{ij} 's such as $\sum_{i=1}^a \sum_{j=1}^c g_{ij} \mu_{ij}$. From Chapter 8, we know that such a contrast is an interaction contrast if $\sum_{i=1}^a g_{ij} = 0$ for every j and if $\sum_{j=1}^c g_{ij} = 0$ for every i . One can show that

$$\begin{aligned} \text{Var}\left(\sum_{i=1}^a \sum_{j=1}^c g_{ij} \bar{y}_{ij\cdot}\right) &= \text{Var}\left(\sum_{i=1}^a \sum_{j=1}^c g_{ij} (\bar{b}_{\cdot i} + \bar{w}_{i\cdot} + \bar{\epsilon}_{ij\cdot})\right) = \text{Var}\left[\sum_{i=1}^a \sum_{j=1}^c (g_{ij} \bar{b}_{\cdot i} + g_{ij} \bar{w}_{i\cdot} + g_{ij} \bar{\epsilon}_{ij\cdot})\right] \\ &= \text{Var}\left[\sum_{i=1}^a \sum_{j=1}^c (g_{ij} \bar{b}_{\cdot i})\right] + \text{Var}\left[\sum_{i=1}^a \sum_{j=1}^c (g_{ij} \bar{w}_{i\cdot})\right] + \text{Var}\left[\sum_{i=1}^a \sum_{j=1}^c (g_{ij} \bar{\epsilon}_{ij\cdot})\right] \\ &= (g_{..})^2 \frac{\sigma_B^2}{r} + \left[\sum_{i=1}^a (g_{i\cdot})^2\right] \frac{\sigma_W^2}{r} + \left[\sum_{i=1}^a \sum_{j=1}^c (g_{ij})^2\right] \frac{\sigma_e^2}{r} \end{aligned}$$

If $\sum_{i=1}^a \sum_{j=1}^c g_{ij} \mu_{ij}$ is an interaction contrast, then

$$\text{Var}\left(\sum_{i=1}^a \sum_{j=1}^c g_{ij} \bar{y}_{ij\cdot}\right) = \left[\sum_{i=1}^a \sum_{j=1}^c (g_{ij})^2\right] \frac{\sigma_e^2}{r}$$

and this variance can be estimated by

$$\text{Var}\left(\sum_{i=1}^a \sum_{j=1}^c g_{ij} \bar{y}_{ij\cdot}\right) = \left[\sum_{i=1}^a \sum_{j=1}^c (g_{ij})^2\right] \frac{\hat{\sigma}_e^2}{r},$$

and its corresponding degrees of freedom are $a(r-1)(c-1)$.

Next consider the $A_{i\cdot}$ main effect mean defined by $\bar{y}_{i\cdot}$. One can obtain this main effect mean as a special case of $\sum_{i=1}^a \sum_{j=1}^c g_{ij} \bar{y}_{ij\cdot}$ by taking

$$\begin{aligned} g_{ij} &= \frac{1}{c} \quad \text{if } i = i' \quad \text{for } j = 1, 2, \dots, c \\ &= 0 \quad \text{otherwise} \end{aligned}$$

Then $(g_{..})^2 = 1$, and $\sum_{i=1}^a (g_{i\cdot})^2 = (g_{i'\cdot})^2 = 1$, and

$$\sum_{i=1}^a \sum_{j=1}^c (g_{ij})^2 = \sum_{j=1}^c (g_{i'j})^2 = \sum_{j=1}^c \left(\frac{1}{c}\right)^2 = \frac{1}{c}$$

Thus, the

$$\text{Var}(\bar{y}_{i..}) = \frac{\sigma_B^2}{r} + \frac{\sigma_W^2}{r} + \frac{\sigma_\epsilon^2}{rc} = \frac{1}{rc}(\sigma_\epsilon^2 + c\sigma_W^2 + c\sigma_B^2)$$

Note that the variance of the estimate of an A main effect mean depends on the block variance component as well as the whole plot and subplot variance components. Similarly, one can show that the variance of the estimate of C main effect mean, $\bar{y}_{.j}$, is

$$\text{Var}(\bar{y}_{.j}) = \frac{\sigma_B^2}{r} + \frac{\sigma_W^2}{r} + \frac{\sigma_\epsilon^2}{ra} = \frac{1}{ra}(\sigma_\epsilon^2 + a\sigma_W^2 + a\sigma_B^2)$$

Finally, the estimate of μ_{ij} is $\bar{y}_{ij..}$, and its variance is

$$\text{Var}(\bar{y}_{ij..}) = \frac{\sigma_B^2}{r} + \frac{\sigma_W^2}{r} + \frac{\sigma_\epsilon^2}{r} = \frac{1}{r}(\sigma_\epsilon^2 + \sigma_W^2 + \sigma_B^2)$$

Note that the degrees of freedom associated with each of the estimates of these functions of the three variance components will need to be estimated by the Satterthwaite method.

Many researchers use effects models rather than means models to describe their data. It is important to be able to express contrasts of means as their corresponding contrasts in effects model parameters. For example, the expression for δ_i is

$$\begin{aligned}\delta_i &= \sum_{j=1}^c s_j \mu_{ij} \\ &= \sum_{j=1}^c s_j [\mu + \alpha_i + \gamma_j + (\alpha\gamma)_{ij}] \\ &= \sum_{j=1}^c s_j \mu + \sum_{j=1}^c s_j \alpha_i + \sum_{j=1}^c s_j \gamma_j + \sum_{j=1}^c s_j (\alpha\gamma)_{ij} \\ &= \mu \sum_{j=1}^c s_j + \alpha_i \sum_{j=1}^c s_j + \sum_{j=1}^c s_j \gamma_j + \sum_{j=1}^c s_j (\alpha\gamma)_{ij} \\ &= \sum_{j=1}^c s_j \gamma_j + \sum_{j=1}^c s_j (\alpha\gamma)_{ij} \quad \text{since } \sum_{j=1}^c s_j = 0\end{aligned}$$

Thus, this contrast of the means involves both the γ_j and the $(\alpha\gamma)_{ij}$. Examples of such contrasts using SAS-Mixed will be illustrated later in this chapter.

The estimates and estimated standard errors given above can be used to test hypotheses via t -tests and/or construct confidence intervals about estimable functions of the fixed effect parameters. There are many possible choices for contrasts of interest. For example, if the levels of a factor are quantitative, one can use the linear, quadratic, and so on, contrasts to investigate possible trends in the mean parameters. Other contrasts could involve comparing a set of controls to each of the treatments.

24.6 Additional Examples

24.6.1 Example 24.3: Moisture and Fertilizer

The data in Table 24.11 are from an experiment where the amount of dry matter was measured on wheat plants grown in different levels of moisture and with different amounts of fertilizer. There were 48 different peat pots and 12 plastic trays; four pots could be put into each tray. The moisture treatments consisted of adding 10, 20, 30, or 40 ml of water per pot per day to the tray where the water was absorbed by the peat pots. The levels of moisture were randomly assigned to the trays. The trays are the large size of experimental unit or whole plot, and the whole plot design is a one-way treatment structure (the four levels of moisture) in a completely randomized design structure. The levels of fertilizer were 2, 4, 6, or 8 mg per pot. The four levels of fertilizer were randomly assigned to the four pots in each tray so that each fertilizer occurred once in each tray. The pot is the smallest size of experimental unit or split-plot or subplot, and the subplot design is a one-way treatment structure (the four levels of fertilizer) in a randomized complete block design structure where the 12 trays are the blocks. The wheat seeds were planted in each pot and after 30 days the dry matter of the wheat plants growing in each pot was measured. A model that can be used to describe the dry matter from a pot in the k th tray assigned to the i th level of moisture and j th level of fertilizer is

$$y_{ijk} = \mu_{ij} + t_{ik} + p_{ijk}, \quad i = 1, 2, 3, 4; j = 1, 2, 3, 4; k = 1, 2, 3$$

where μ_{ij} is the mean dry matter of level i of moisture with level j of fertilizer, t_{ik} is the tray error term distributed *i.i.d.* $N(0, \sigma_{\text{tray}}^2)$, and p_{ijk} is the pot error term distributed *i.i.d.* $N(0, \sigma_{\text{pot}}^2)$. Note that p_{ijk} is equivalent to the residual error in this model. The analysis of variance table is in Table 24.12, which shows a significant *Moisture* \times *Fertilizer* interaction. Because there is an interaction between moisture and fertilizer, the treatment combination means, which are given in Table 24.13, are used when making inferences. Since the levels of moisture and the levels of fertilizer are equally spaced quantitative levels, orthogonal

TABLE 24.11

Dry Matter Measurements per Pot for Example 24.3

MST	Tray	Fertilizer 2	Fertilizer 4	Fertilizer 6	Fertilizer 8
10	1	3.3458	4.3170	4.5572	5.8794
10	2	4.0444	4.1413	6.5173	7.3776
10	3	1.9758	3.8397	4.4730	5.1180
20	1	5.0490	7.9419	10.7697	13.5168
20	2	5.9131	8.5129	10.3934	13.9157
20	3	6.9511	7.0265	10.9334	15.2750
30	1	6.5693	10.7348	12.2626	15.7133
30	2	8.2974	8.9081	13.4373	14.9575
30	3	5.2785	8.6654	11.1372	15.6332
40	1	6.8393	9.0842	10.3654	12.5144
40	2	6.4997	6.0702	10.7486	12.5034
40	3	4.0482	3.8376	9.4367	10.2811

TABLE 24.12

Analysis of Variance Table for Dry Matter for the Moisture and Fertilizer Example

```
proc mixed cl covtest method=type3 data=ex_243;
  class mst tray fr;
  model dry_matter=mst|fr / ddfm=KR;
  random tray(mst);
```

Source	df	Sum of Squares	Mean Square	F-Value	Pr > F	Expected Mean Square
MST	3	269.189429	89.729810	26.34	0.0002	Var(Residual) + 4 Var[TRAY(MST)] + Q(MST, MST × FR)
FR	3	297.054856	99.018285	131.65	<0.0001	Var(Residual) + Q(FR, MST × FR)
MST × FR	9	38.056379	4.228487	5.62	0.0003	Var(Residual) + Q(MST × FR)
TRAY(MST)	8	27.251576	3.406447	4.53	0.0019	Var(Residual) + 4 Var[TRAY(MST)]
Residual	24	18.051379	0.752141			Var(Residual)

TABLE 24.13

Fertilizer by Moisture Cell Means for Dry Matter

Fertilizer	Moisture			
	10	20	30	40
2	3.1220	5.9711	6.7151	5.7957
4	4.0993	7.8271	9.4361	6.3307
6	5.1825	10.6988	12.2790	10.1836
8	6.1250	14.2358	15.4347	11.76630

polynomials can easily be used to investigate the trends over the levels of fertilizer for each level of moisture and the trends over the levels of moisture for each level of fertilizer. The contrasts that measure the linear and quadratic trends of fertilizer at the i th level of moisture (a contrast of the subplot treatments at the same whole plot treatment) are

$$\delta_{LinF|M_i} = -3\mu_{i1} - \mu_{i2} + \mu_{i3} + 3\mu_{i4}, \quad i = 1, 2, 3, 4$$

and

$$\delta_{QuadF|M_i} = \mu_{i1} - \mu_{i2} - \mu_{i3} + \mu_{i4}, \quad i = 1, 2, 3, 4$$

Their estimates are

$$\hat{\delta}_{LinF|M_i} = -3\bar{y}_{i,1} - \bar{y}_{i,2} + \bar{y}_{i,3} + 3\bar{y}_{i,4}, \quad i = 1, 2, 3, 4,$$

and

$$\hat{\delta}_{QuadF|M_i} = \bar{y}_{i,1} - \bar{y}_{i,2} - \bar{y}_{i,3} + \bar{y}_{i,4}, \quad i = 1, 2, 3, 4, \text{ respectively}$$

The variances of these contrasts are

$$\begin{aligned}\text{Var}(\hat{\delta}_{\text{LinF}|M_i}) &= \frac{\sigma_p^2}{3}[(-3)^2 + (-1)^2 + 1^2 + 3^2] \\ &= \frac{20\sigma_p^2}{3}\end{aligned}$$

and

$$\begin{aligned}\text{Var}(\hat{\delta}_{\text{QuadF}|M_i}) &= \frac{\sigma_p^2}{3}[(-1)^2 + (-1)^2 + 1^2 + 1^2] \\ &= \frac{4\sigma_p^2}{3}, \text{ respectively}\end{aligned}$$

The estimated standard errors are obtained by replacing σ_p^2 with $MSError(pot)$ in the square root of the respective variances. The SAS-Mixed code to evaluate these contrasts is given in Table 24.14. The estimates of the linear and quadratic trends of fertilizer for each level of moisture and the corresponding estimated standard errors and *t*-statistics (testing that the trends are zero) are given in Table 24.15.

The contrasts that measure the linear and quadratic trends of moisture at each fertilizer level (comparisons of whole plot treatments at the same subplot treatment) are

$$\lambda_{\text{LinM}|F_j} = -3\mu_{1j} - \mu_{2j} + \mu_{3j} + 3\mu_{4j}, \quad j = 1, 2, 3, 4$$

TABLE 24.14

SAS-Mixed Code with Estimate Statements Used to Evaluate the Linear and Quadratic Trends

```
proc mixed cl covtest method=reml data=ex_243;
  class mst tray fr;
  model dry_matter=mst|fr / ddfm=KR;
  random tray / sub=mst;
  lsmeans mst|fr/diffs;
  estimate 'LF M10' fr -3 -1 1 3 mst*fr -3 -1 1 3;
  estimate 'LF M20' fr -3 -1 1 3 mst*fr 0 0 0 0 -3 -1 1 3;
  estimate 'LF M30' fr -3 -1 1 3 mst*fr 0 0 0 0 0 0 0 -3 -1 1 3;
  estimate 'LF M40' fr -3 -1 1 3 mst*fr 0 0 0 0 0 0 0 0 0 0 -3 -1 1 3;
  estimate 'QF M10' fr 1 -1 -1 1 mst*fr 1 -1 -1 1;
  estimate 'QF M20' fr 1 -1 -1 1 mst*fr 0 0 0 0 1 -1 -1 1;
  estimate 'QF M30' fr 1 -1 -1 1 mst*fr 0 0 0 0 0 0 0 1 -1 -1 1;
  estimate 'QF M40' fr 1 -1 -1 1 mst*fr 0 0 0 0 0 0 0 0 0 0 1 -1 -1 1;
  estimate 'LM F2' mst -3 -1 1 3 mst*fr -3 0 0 0 -1 0 0 0 1 0 0 0 3;
  estimate 'LM F4' mst -3 -1 1 3 mst*fr 0 -3 0 0 0 -1 0 0 0 1 0 0 0 3;
  estimate 'LM F6' mst -3 -1 1 3 mst*fr 0 0 -3 0 0 0 -1 0 0 0 1 0 0 0 3;
  estimate 'LM F8' mst -3 -1 1 3 mst*fr 0 0 0 -3 0 0 0 -1 0 0 0 1 0 0 0 3;
  estimate 'QM F2' mst 1 -1 -1 1 mst*fr 1 0 0 0 -1 0 0 0 -1 0 0 0 1;
  estimate 'QM F4' mst 1 -1 -1 1 mst*fr 0 1 0 0 0 -1 0 0 0 -1 0 0 0 1;
  estimate 'QM F6' mst 1 -1 -1 1 mst*fr 0 0 1 0 0 0 -1 0 0 0 -1 0 0 0 1;
  estimate 'QM F8' mst 1 -1 -1 1 mst*fr 0 0 0 1 0 0 0 -1 0 0 0 -1 0 0 0 1;
```

TABLE 24.15

Estimates of the Linear and Quadratic Trends of Fertilizer for Each Level of Moisture

Moisture	Linear		Quadratic	
	Estimate	t-Value	Estimate	t-Value
10	10.0922	4.51	-0.03483	-0.03
20	27.6660	12.36	1.6810	1.68
30	29.0017	12.95	0.4346	0.43
40	21.7646	9.72	1.0478	1.05

Note: The estimated standard errors for the linear and quadratic trends are 2.239 and 1.001, respectively. Compare t-values with $t_{\alpha/2, 24}$.

and

$$\lambda_{QuadM|F_j} = \mu_{1j} - \mu_{2j} - \mu_{3j} + \mu_{4j}, \quad j = 1, 2, 3, 4, \text{ respectively}$$

The estimates and corresponding variances of these comparisons of the levels of fertilizer within each level of moisture are

$$\begin{aligned}\hat{\lambda}_{LinM|F_j} &= -3\bar{y}_{1j} - \bar{y}_{2j} + \bar{y}_{3j} + 3\bar{y}_{4j}, \quad j = 1, 2, 3, 4 \\ \hat{\lambda}_{QuadM|F_k} &= \bar{y}_{1j} - \bar{y}_{2j} - \bar{y}_{3j} + \bar{y}_{4j}, \quad j = 1, 2, 3, 4\end{aligned}$$

and

$$\begin{aligned}\text{Var}(\hat{\lambda}_{LinM|F_k}) &= \frac{\sigma_{pot}^2 + \sigma_{tray}^2}{3}(3^2 + 1^2 + 1^2 + 3^2) \\ &= \frac{20(\sigma_{pot}^2 + \sigma_{tray}^2)}{3}\end{aligned}$$

and

$$\begin{aligned}\text{Var}(\hat{\lambda}_{QuadM|F_k}) &= \frac{\sigma_{pot}^2 + \sigma_{tray}^2}{3}(1^2 + 1^2 + 1^2 + 1^2) \\ &= \frac{4(\sigma_{pot}^2 + \sigma_{tray}^2)}{3}\end{aligned}$$

The estimated standard errors are obtained by replacing $\sigma_{pot}^2 + \sigma_{tray}^2$ with

$$\hat{\sigma}_{pot}^2 + \hat{\sigma}_{tray}^2 = \frac{MSError(tray) + (4-1)MSError(pot)}{4}$$

in the square root of the respective variances. The Satterthwaite approximation can be used to obtain the approximate degrees of freedom to be associated with $\hat{\sigma}_{\text{pot}}^2 + \hat{\sigma}_{\text{tray}}^2$. For this example

$$\hat{\sigma}_{\text{pot}}^2 + \hat{\sigma}_{\text{tray}}^2 = \frac{3.406 + (4 - 1)0.752}{4} = 1.416$$

which is based on \hat{v} degrees of freedom where

$$\hat{v} = \frac{\frac{[3.406 + (4 - 1)0.752]^2}{8}}{\frac{[3.406]^2}{24} + \frac{[(4 - 1)0.752]^2}{24}} = 19.3$$

The estimates of the linear and quadratic trends of fertilizer for each level of moisture and the corresponding estimated standard errors and t -statistics are given in Table 24.15. Table 24.15 shows that there is a linear response to fertilizer for each level of moisture and no significant quadratic trend. Table 24.16 shows that there are both linear and quadratic responses to moisture for each level of fertilizer. The graphs in Figures 24.5 and 24.6 show the response to fertilizer for each moisture level and the response to moisture for each fertilizer level. A Satterthwaite approximation was used to determine the approximate degrees of freedom given in Table 24.16.

24.6.2 Example 24.4: Regression with Split-Plot Errors

The levels of moisture and fertilizer in Example 24.3 are quantitative levels, so it may be of interest to investigate whether a regression model can be constructed that will describe the data. This is a different model from the usual regression situation since the data are correlated. Mixed models software is required to carry out the computations. To obtain an appropriate analysis, one needs to use moisture as a continuous variable in the regression model and as a class variable in the random statement. *Mst* is used to denote moisture as a continuous variable and *mstblk* (which is equal to *mst*) is used to denote the class variable. Table 24.17 contains the SAS-Mixed code where the model statement is a general cubic regression model for moisture (*mst*) and fertilizer (*fr*). The random statement inserts the

TABLE 24.16

Estimates of the Linear and Quadratic Trends of Moisture for Each Level of Fertilizer

Fertilizer	Linear		Quadratic	
	Estimate	<i>t</i> -Value	Estimate	<i>t</i> -Value
2	-2.74	2.85	-3.7684	-2.74
4	-4.97	2.70	-6.8332	-4.97
6	-5.54	5.40	-7.6118	-5.54
8	-8.57	5.90	-11.7792	-8.57

Note: The estimated standard errors for the linear and quadratic trends are 3.072 and 1.374, respectively. Compare *t*-values with $t_{\alpha/2, 19.3}$.

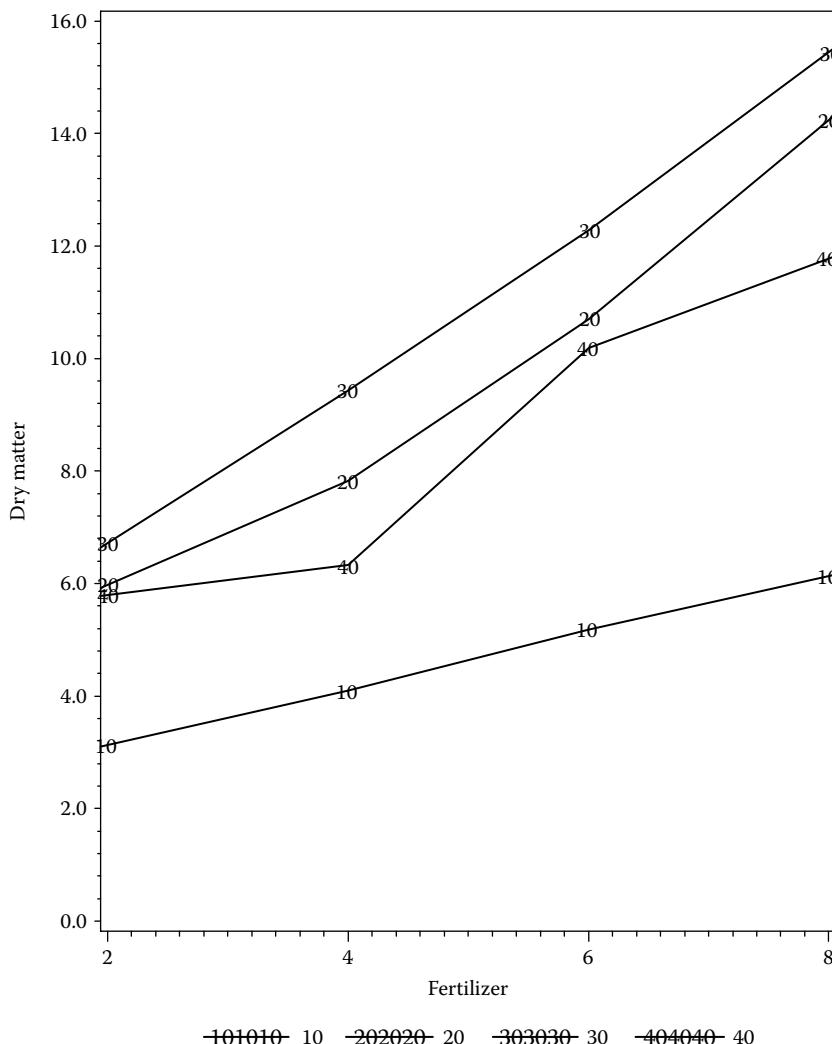


FIGURE 24.5 Graphic of dry matter means against the level of fertilizer for each level of moisture.

tray error term and imposes the correlation structure on the data. The estimates of variance components are given in Table 24.17 and the solutions for the regression coefficients are given in Table 24.18. Most of the significance levels are quite large, so several steps of a deletion process (there are no automatic processes in SAS-Mixed) were carried out until all of the remaining variables had coefficients that were significantly different from zero. The reduced model is

$$DM_{ijk} = \beta_0 + \beta_1 mst_i + \beta_2 fr_j + \beta_3 (mst_i)(fr_j) + \beta_4 (mst_i)^2 (fr_j) + \beta_5 (mst_i)(fr_j)^2 + t_{ik} + p_{ijk}$$

It is of interest to determine if the reduced model adequately describes the data, so a lack of fit test is constructed. Let $frx = fr$ and include frx in the class statement. Include $mstblk \times frx$ in the reduced model, as shown in Table 24.19. The F -test corresponding to $mstblk \times frx$ in the

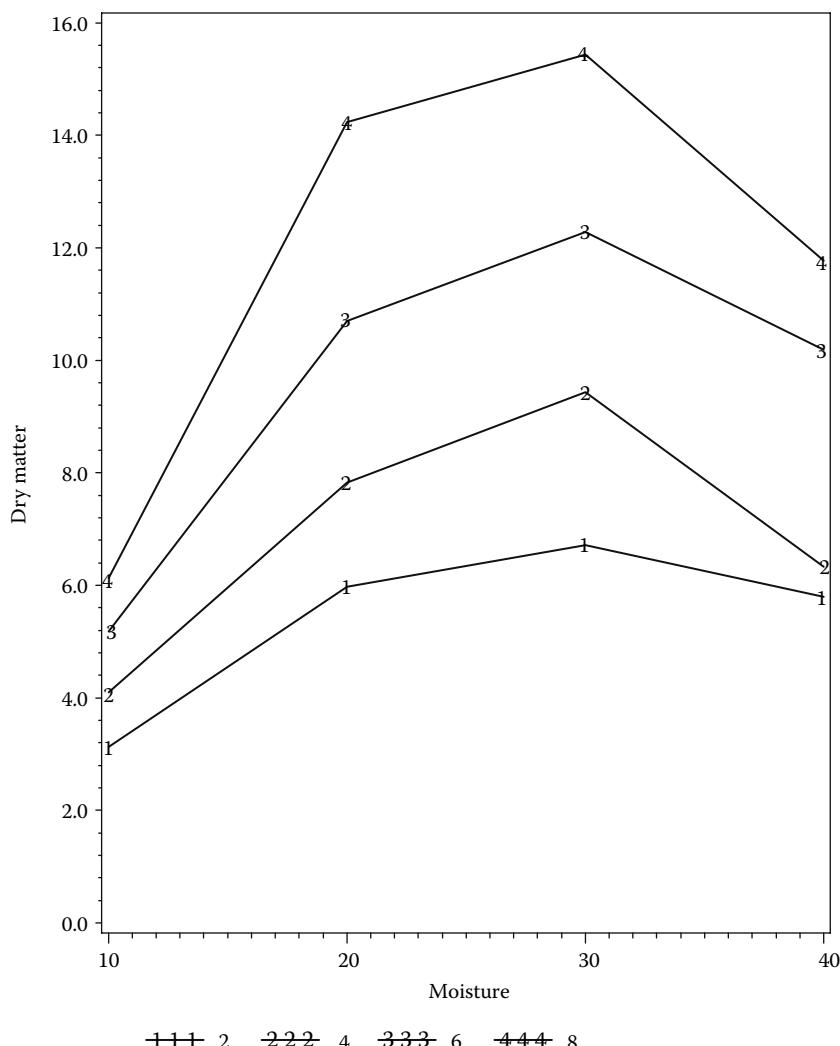


FIGURE 24.6 Graphic of dry matter means against the level of moisture for each level of fertilizer.

TABLE 24.17

SAS-Mixed Code and Covariance Parameter Estimates for Full Regression Model

```
proc mixed data=ex_243 cl covtest;
class mstblk tray;
model dry_matter=mst mst*mst mst*mst*mst fr fr*fr fr*fr*fr mst*fr fr*mst*mst
fr*mst*mst*mst mst*mst*fr*fr mst*mst*mst*fr*fr mst*mst*mst*mst*fr*fr
mst*fr*fr*fr mst*mst*fr*fr*fr mst*mst*mst*fr*fr*fr
/solution ddfm=KR;
random tray(mstblk);
```

Covariance Parameter Estimates

Covariance Parameter	Estimate	Standard Error	Z-Value	Pr Z	α	Lower	Upper
Tray(MSTBLK)	0.6636	0.4293	1.55	0.0611	0.05	0.2544	4.2338
Residual	0.7521	0.2171	3.46	0.0003	0.05	0.4586	1.4556

TABLE 24.18

Solution for the Parameters for the Full Response Surface Model

Solution for Fixed Effects

Effect	Estimate	Standard Error	df	t-Value	Pr > t
Intercept	-21.3654	34.7678	24.8	-0.61	0.5445
MST	3.9747	5.3173	24.8	0.75	0.4618
MST × MST	-0.1861	0.2349	24.8	-0.79	0.4358
MST × MST × MST	0.002723	0.003120	24.8	0.87	0.3911
FR	14.7827	26.4179	24	0.56	0.5810
FR × FR	-2.9296	5.8356	24	-0.50	0.6202
FR × FR × FR	0.1556	0.3875	24	0.40	0.6916
MST × FR	-2.5838	4.0403	24	-0.64	0.5286
MST × MST × FR	0.1329	0.1785	24	0.74	0.4637
MST × MST × MST × FR	-0.00205	0.002370	24	-0.87	0.3948
MST × FR × FR	0.5263	0.8925	24	0.59	0.5609
MST × MST × FR × FR	-0.02671	0.03943	24	-0.68	0.5046
MST × MST × MST × FR × FR	0.000413	0.000524	24	0.79	0.4380
MST × FR × FR × FR	-0.02884	0.05926	24	-0.49	0.6310
MST × MST × FR × FR × FR	0.001518	0.002618	24	0.58	0.5674
MST × MST × MST × FR × FR × FR	-0.00002	0.000035	24	-0.70	0.4934

TABLE 24.19

SAS-Mixed Code to Test for the Lack of Fit of the Reduced Regression Model

```
proc mixed data=ex_243 cl covtest;
class mstblk FRX tray;
model dry_matter=mst fr mst*fr fr*mst*mst mst*fr*fr FRX*MSTBLK
/ddfm=KR solution;
random TRAY(mstblk);
```

Covariance Parameter Estimates

Covariance Parameter	Estimate	Standard Error	Z-Value	Pr Z	α	Lower	Upper
Tray(MSTBLK)	0.6636	0.4293	1.55	0.0611	0.05	0.2544	4.2338
Residual	0.7521	0.2171	3.46	0.0003	0.05	0.4586	1.4556

analysis of variance table in Table 24.20 provides the test of lack of fit of the reduced model. In this case, the significance level corresponding to the test of lack of fit is 0.5561, indicating the reduced model adequately describes the data. The code in Table 24.21 fits the reduced model and displays the estimators of the covariance parameters and the estimates of the regression coefficients are in Table 24.22. The tray error term from the reduced model is somewhat less than that of the full model while the residual or pot variances are similar. Predicted values from the reduced model are displayed in Figure 24.7 and the moisture by fertilizer cell means are graphed in Figure 24.8. The regression model does an adequate job of describing the cell means, as would be expected from the results of the test for lack of fit.

TABLE 24.20Test for Lack of Fit in MSTBLK \times FXR

Type III Tests of Fixed Effects				
Effect	Num df	Den df	F-Value	Pr > F
MST	0	—	—	—
FR	0	—	—	—
MST \times FR	0	—	—	—
MST \times MST \times FR	0	—	—	—
MST \times FR \times FR	0	—	—	—
MSTBLK \times FRX	10	24.7	0.89	0.5561

TABLE 24.21

SAS-Mixed Code for Final Regression Model and Covariance Parameter Estimates

```
proc mixed data=ex_243 cl covtest;
  title3 'reduced regression model with split-plot errors';
  title4 'Method=REML';
  class mstblk tray;
  model dry_matter=mst fr mst*fr fr*mst*mst mst*fr*fr
    / solution ddfm=KR;
  RANDOM TRAY(mstblk);
```

Covariance Parameter Estimates

Covariance Parameter	Estimate	Standard Error	Z-Value	Pr Z	α	Lower	Upper
Tray(MSTBLK)	0.5582	0.3498	1.60	0.0553	0.05	0.2189	3.2798
Residual	0.7575	0.1881	4.03	<0.0001	0.05	0.4912	1.3195

24.6.3 Example 24.5: Mixed-Up Split-Plot Design

Two researchers designed a study to evaluate how three varieties of soybeans respond to four different types of herbicides. At site 1, the first researcher used a split-plot design in two replications where the varieties of soybeans were the levels of the whole-plot and the herbicides were the levels of the subplots. At site 2, the second researcher used a split-plot design in two replications, but assigned the levels of herbicides to the whole-plots and the levels of the varieties to the subplots. A combined analysis was desired, but since the two designs are very different, a combined analysis did not seem possible. However, remember from Chapter 5, the split-plot design structure is nothing other than an incomplete block design where the whole-plots are the blocks of the incomplete block design. The treatment combinations in each of the 14 whole-plots or incomplete blocks are shown in Table 24.23, where $Vx \times y$ denotes variety x with herbicide y . The first eight whole-plots (from site 1) are incomplete blocks of size three and the last six whole-plots (from site 2) are incomplete blocks of size four. A combined analysis can be accomplished using the model

$$y_{ijkl} = \mu_{kl} + b_i + w_{ij} + \varepsilon_{ijkl}$$

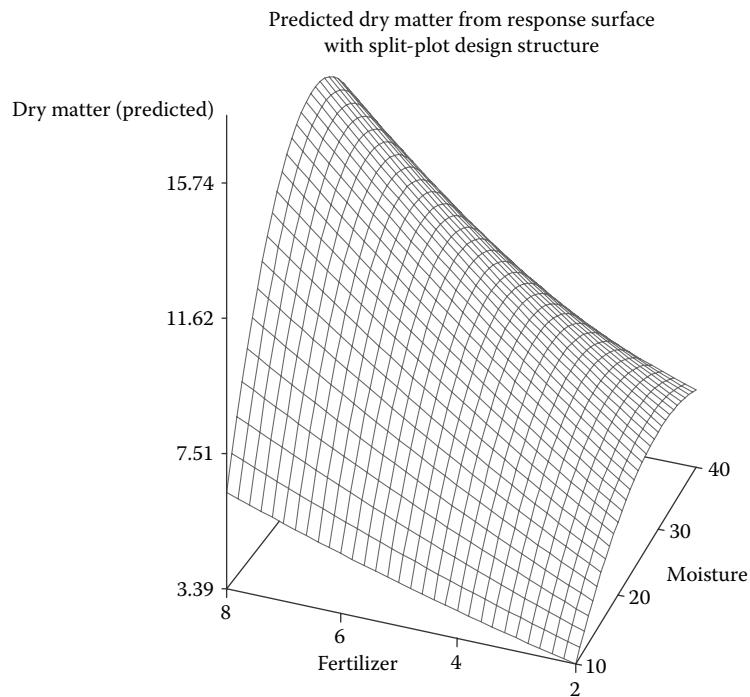


FIGURE 24.7 Prediction surface for dry matter as a function of moisture and fertilizer.

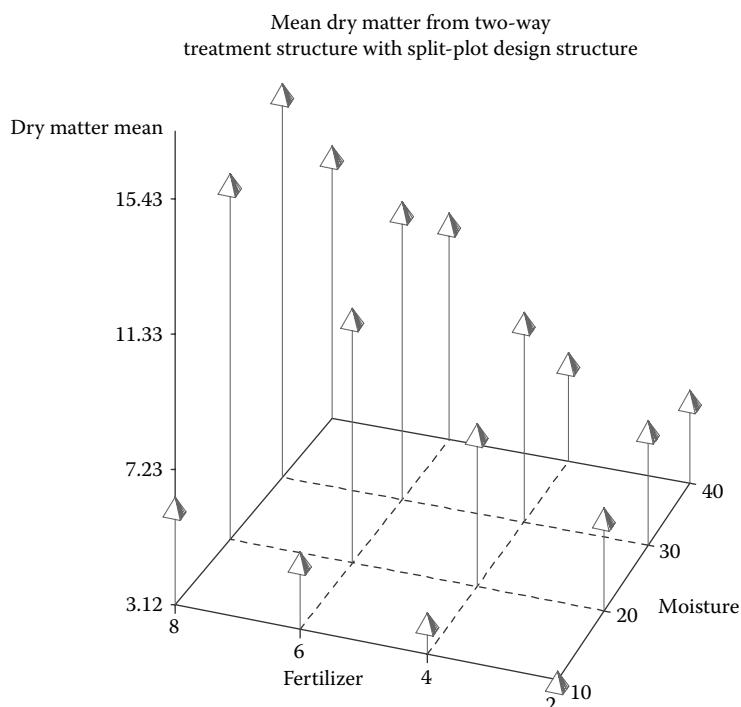


FIGURE 24.8 Means for the combinations of fertilizer and moisture.

TABLE 24.22

Solution for the Parameters of the Reduced Regression Model

Solution for Fixed Effects

Effect	Estimate	Standard Error	df	t-Value	Pr > t
Intercept	1.9536	0.9204	33	2.12	0.0414
MST	0.07531	0.04069	40.6	1.85	0.0715
FR	-1.0795	0.2419	37.8	-4.46	<0.0001
MST × FR	0.1730	0.02351	36.5	7.36	<0.0001
MST × MST × FR	-0.00346	0.000398	25.4	-8.70	<0.0001
MST × FR × FR	0.001838	0.001147	32.4	1.60	0.1187

where μ_{kl} is the mean response of the k th variety and the l th herbicide, $b_i \sim i.i.d. N(0, \sigma_{blk}^2)$ are the large block effects within each site, $w_{ij} \sim i.i.d. N(0, \sigma_{wp}^2)$ denote the whole-plot or incomplete block effects and $\varepsilon_{ijkl} \sim i.i.d. N(0, \sigma_e^2)$ are the sub-plot effects. The data are in Tables 24.24 and 24.25 and the SAS-Mixed code and covariance parameter estimates are given in Table 24.26. The model statement contains the treatment structure with varieties (V), herbicides (H) and variety by herbicide interaction. The random statement contains the blocks (replications within a site) and whole-plots within a block (the incomplete blocks). The tests of the fixed effects are given in Table 24.27 and the means are given in Table 24.28. There is a significant variety by herbicide interaction, thus comparisons need to be made using the cell means. For comparisons of the levels of herbicide within each of the varieties, the estimated standard error of the difference is 3.33 with 26.1 df. For comparisons of the varieties within each of the herbicides, the estimated standard error of the difference is 3.26 with df = 24.7.

TABLE 24.23

Combinations of Varieties and Herbicides Assigned to Each of the Whole Plots within Each Block

Block	Whole Plot	
1	1	V1H1, V2H1, V3H1
	2	V1H2, V2H2, V3H2
	3	V1H3, V2H3, V3H3
	4	V1H4, V2H4, V3H4
2	5	V1H1, V2H1, V3H1
	6	V1H2, V2H2, V3H2
	7	V1H3, V2H3, V3H3
	8	V1H4, V2H4, V3H4
3	9	V1H1, V1H2, V1H3, V1H4
	10	V2H1, V2H2, V2H3, V2H4
	11	V3H1, V3H2, V3H3, V3H4
4	12	V1H1, V1H2, V1H3, V1H4
	14	V2H1, V2H2, V2H3, V2H4
	14	V3H1, V3H2, V3H3, V3H4

TABLE 24.24

Data from First Location for Example 24.4 Where H and V Denote Herbicides and Varieties, Respectively, and the Entries in the Table Are Weights of Soybeans

Block	WP	H	V1	V2	V3
1	1	1	22.4	39.7	30.4
1	2	2	36.6	38.3	33.7
1	3	3	38.2	33.0	34.6
1	4	4	19.9	28.0	19.6
2	5	1	25.7	32.5	21.7
2	6	2	26.7	25.4	13.2
2	7	3	26.6	28.2	34.6
2	8	4	22.5	18.5	18.0

TABLE 24.25

Data from Second Location for Example 24.4 Where H and V Denote Herbicides and Varieties, Respectively, and the Entries in the Table Are Weights of Soybeans

Block	WP	V	H1	H2	H3	H4
3	9	1	21.5	24.6	30.0	28.8
3	10	2	29.3	33.8	35.1	35.1
3	11	3	28.9	29.0	41.2	38.4
4	12	1	17.3	25.5	23.7	18.6
4	13	2	29.2	29.9	26.8	20.4
4	14	3	14.3	27.7	43.3	30.2

TABLE 24.26

SAS-Mixed Code and Estimates of the Covariance Parameters for Example 24.5

```
proc mixed data=one covtest cl;
class block wp V H;
model yield=V|H/ddfm=kr;
random block wp(block);
lsmeans V H V*H/diff;
```

Covariance Parameter Estimates

Covariance

Parameter	Estimate	Standard Error	Z-Value	Pr Z	α	Lower	Upper
Block	6.7190	11.3440	0.59	0.2768	0.05	1.1382	120756
WP(block)	18.2370	12.8963	1.41	0.0787	0.05	6.5461	150.62
Residual	17.4321	5.3451	3.26	0.0006	0.05	10.3483	35.4086

24.6.4 Example 24.6: Split-Split-Plot Design

A split-split-plot design has three sizes of experimental units, whole-plot, subplot, and sub-subplot, where the whole-plot, subplot, and sub-subplot are the largest, medium, and smallest

TABLE 24.27

Tests of the Fixed Effects for Example 24.5

Type II Tests of Fixed Effects

Effect	Num df	Den df	F-Value	Pr > F
V	2	30.4	2.17	0.1320
H	3	31.3	5.07	0.0056
V × H	6	21.3	3.64	0.0122

TABLE 24.28

Estimates of the Variety by Herbicide Cell Means, the Variety Means and the Herbicide Means for Example 24.5

V	H	Standard		V	Standard		H	Standard	
		Estimate	Error		Estimate	Error		Estimate	Error
1	1	21.31	2.99	1	26.94	2.18	1	24.25	2.32
1	2	29.39	2.99	2	30.56	2.18	2	28.32	2.32
1	3	31.09	2.99	3	26.87	2.18	3	32.99	2.32
1	4	25.95	2.99				4	26.92	2.32
2	1	31.22	2.99						
2	2	31.85	2.99						
2	3	31.20	2.99						
2	4	27.96	2.99						
3	1	20.21	2.99						
3	2	23.73	2.99						
3	3	36.68	2.99						
3	4	26.85	2.99						

size experimental units, respectively. The present study involves evaluating the effects of combinations of two rations (regular corn and high oil corn), two temperature (3 and 6°C), and three types of packaging (vacuum, CO₂, and low O₂) on the tenderness of meat. Twenty steers were randomly assigned to the two rations (10 per ration). At slaughter, the animal was split in halves and a loin from each side was extracted. The loin was randomly assigned to one of the two storage temperatures. After 10 days of storage, three steaks were cut from each loin and the levels of packaging were assigned to one of those steaks. The steak was put into a display case for seven more days. After the seven days in the display case, five cores were obtained from each steak and the force required to shear each core was measured. The data in Table 24.29 are the means of the five cores. Some packaging problems caused some of the steaks not to be representative of the true state of nature, and they were deleted from the data set.

This study involves three sizes of experimental units. The animal is the experimental unit to which the levels of ration are assigned. The side of an animal is the experimental unit for the levels of temperature, and the steak is the experimental unit for the levels of packaging. A model that can be used to describe the shear forces (f_{ijkl}) is

$$f_{ijkl} = \mu_{ikl} + a_{ij} + s_{ijk} + \varepsilon_{ijkl} \quad i = 1, 2, \quad j = 1, 2, \dots, 10, \quad k = 1, 2, \quad l = 1, 2, 3$$

TABLE 24.29

Shear Force Data for Example 24.6

Ration	Animal	Temperature 3°C			Temperature 6°C		
		Vacuum	CO ₂	Low O ₂	Vacuum	CO ₂	Low O ₂
1	1	9.287	8.697	9.708	6.592	—	7.771
1	2	8.192	8.782	9.961	—	8.782	—
1	3	4.487	4.234	—	4.739	2.887	3.055
1	4	—	—	8.613	—	5.329	2.718
1	5	2.382	4.824	4.992	2.971	6.339	4.487
1	6	6.508	5.582	8.445	3.982	7.182	7.855
1	7	6.171	4.908	5.497	3.561	6.171	2.297
1	8	7.266	—	6.171	8.950	5.413	—
1	9	4.487	5.413	6.424	5.329	4.571	—
1	10	—	7.771	6.845	—	4.403	—
2	1	2.718	5.918	6.171	5.918	7.266	8.697
2	2	—	7.350	10.80	4.655	—	—
2	3	5.750	5.076	7.687	3.813	4.066	7.687
2	4	5.666	4.739	5.918	6.339	—	7.434
2	5	6.761	7.687	—	4.234	5.413	8.866
2	6	6.003	6.508	—	4.487	9.455	—
2	7	5.918	8.782	—	7.182	8.950	—
2	8	4.571	5.582	11.06	4.824	9.287	9.455
2	9	—	5.918	8.024	—	7.182	7.434
2	10	2.803	—	—	—	7.434	7.097

where μ_{ikl} denotes the mean response of the i th ration, k th temperature and l th packaging, $a_{ij} \sim i.i.d. N(0, \sigma_{\text{animal}}^2)$ denote the animal errors, $s_{ijk} \sim i.i.d. N(0, \sigma_{\text{side}}^2)$ denote the side errors and $\varepsilon_{ijkl} \sim i.i.d. N(0, \sigma_{\text{steak}}^2)$ denote the steak errors.

The animal error term can be obtained by ignoring the temperature and packaging parts of the treatment structure. The animal design is a one-way treatment structure in a completely randomized design structure where there are 10 animals assigned to each ration. Thus the animal error term is computed as the variation among animals treated alike within a ration pooled across rations, denoted by animal(ration) . The side error term is obtained by considering only those animals assigned to ration 1 and ignoring the packaging effects. Each animal within ration 1 forms a block of size 2 (two sides) for the two temperatures, thus the side design is a one-way treatment structure (two temperatures) in a randomized complete block design structure (10 animals within ration 1). So the side error term is computed as the temperature by animal interaction within a ration pooled across rations, denoted by $\text{temp} \times \text{animal(ration)}$. The residual is the steak error term, but its form can be obtained by considering that data from ration 1 and temperature 3°C. The side of an animal from ration 1 assigned to 3°C is the blocking factor for the three types of packaging. The steak design is a one-way treatment structure (three levels of packaging) in a randomized complete design structure (10 sides of animals assigned to ration 1 stored at 3°C temperature). The package by animal interaction is the error term for this part of the data. The computational form of the residual is $\text{package} \times \text{animal(temp ration)}$. If the data set were balanced, there would be 18, 18, 72 degrees of freedom for the animal error, side error, and steak error terms, respectively.

The SAS-Mixed code for fitting the split-split-plot model to the shear force data is in Table 24.30 where the random statement contains *animal(ration)* and *temp × animal(ration)*, the whole-plot error and the sub-plot error, respectively. The estimates of the covariance parameters are also included in Table 24.30. The tests for the fixed effects are given in Table 24.31 where there are significant *ration × temperature* and *ration × packaging* interactions. The *ration × temperature* means are given in Table 24.32 and pairwise comparisons among them are given in Table 24.33. The *adjust = Tukey* option was used on the lsmeans statements to provide adjusted *p*-values from the Tukey multiple comparison method. The estimated standard errors for comparing temperatures within a ration are close to 0.45, while the estimated standard errors for comparing rations within a temperature are around 0.679. Comparing temperatures within a ration are within-animal comparisons while comparing rations within a temperature are between animal comparisons and thus have the larger estimated standard errors. Tables 24.34 and 24.35 contain the means and pairwise comparisons for the *ration × packaging* means. The estimated standard errors for comparing packaging within a ration are about 0.49 while the estimated standard errors for comparing rations with a package type are about 0.71. The packaging comparisons are

TABLE 24.30

SAS-Mixed Code and Estimates of the Covariance Parameters for Example 24.6

```
proc mixed data=one covtest cl;
class ration animal temp pack;
model wbsf=ration|temp|pack/ddfm=kr;
random animal(ration) temp*animal(ration);
lsmeans ration*temp/diff adjust=tukey;
lsmeans ration*pack/diff adjust=tukey;
```

Covariance Parameter Estimates

Covariance Parameter	Estimate	Standard Error	Z-Value	Pr Z	α	Lower	Upper
Animal (ration)	1.2434	0.6102	2.04	0.0208	0.05	0.5740	4.4215
Animal × temperature (ration)	0.1992	0.3612	0.55	0.2906	0.05	0.03165	15986
Residual	1.7529	0.3600	4.87	<0.0001	0.05	1.2166	2.7440

TABLE 24.31

Tests of the Fixed Effects for Example 24.6

Type III Tests of Fixed Effects

Effect	Num df	Den df	F-Value	Pr > F
Ration	1	18	1.89	0.1864
Temperature	1	16.7	1.63	0.2188
Ration × temperature	1	16.7	5.47	0.0321
Pack	2	53.7	13.80	<0.0001
Ration × pack	2	53.7	8.34	0.0007
Temperature × pack	2	54.6	2.24	0.1157
Ration × temperature × pack	2	54.6	0.53	0.5899

TABLE 24.32

Ration and Temperature Force Means for Example 24.6

Ration	Temperature	Estimate	Standard Error
1	3	6.5512	0.4682
1	6	5.3947	0.4852
2	3	6.6181	0.4777
2	6	6.9574	0.4748

TABLE 24.33

Pairwise Comparisons of the Ration by Temperature Means with the Tukey Adjustment for Multiple Comparisons

Ration	Temperature	Ration	Temperature	Estimate	Standard Error	df	t-Value	Adj p
1	3	1	6	1.1565	0.4529	16.5	2.55	0.0874
1	3	2	3	-0.06693	0.6689	27	-0.10	0.9996
1	3	2	6	-0.4062	0.6669	26.7	-0.61	0.9278
1	6	2	3	-1.2234	0.6809	28.5	-1.80	0.3091
1	6	2	6	-1.5627	0.6789	28.2	-2.30	0.1374
2	3	2	6	-0.3392	0.4517	16.9	-0.75	0.8750

TABLE 24.34

Ration and Packaging Force Means for Example 24.6

Ration	Pack	Estimate	Standard Error
1	1	5.7044	0.5158
1	2	6.0996	0.4927
1	3	6.1149	0.5166
2	1	5.0572	0.5014
2	2	6.8470	0.4928
2	3	8.4591	0.5357

made within a side of animal while the ration comparisons are made between animals, explaining the differences in magnitude of the two estimated standard errors.

24.7 Sample Size and Power Considerations

The sample size needed for a particular study depends on the comparisons that are most interesting. The basic sample size equation for detecting a difference of δ between two means where the variance of the difference is $2\sigma^2/n$, $\hat{\sigma}^2$ is the estimate of σ^2 based on v degrees of freedom, the type I error rate is α , and the type II error rate is β is

$$n = \frac{2\hat{\sigma}^2}{\delta^2} [t_{\alpha/2,v} + t_{\beta,v}]^2$$

TABLE 24.35

Pairwise Comparisons of the Ration by Packaging Force Means with the Tukey Adjustment for Multiple Comparisons

Ration	Pack	Ration	Pack	Estimate	Standard Error	df	t-Value	Adj p
1	1	1	2	-0.3952	0.4881	52.2	-0.81	0.9645
1	1	1	3	-0.4104	0.5089	52.1	-0.81	0.9651
1	1	2	1	0.6472	0.7193	34.7	0.90	0.9449
1	1	2	2	-1.1425	0.7134	34	-1.60	0.6011
1	1	2	3	-2.7547	0.7437	38	-3.70	0.0063
1	2	1	3	-0.01522	0.4952	56.1	-0.03	1.0000
1	2	2	1	1.0424	0.7030	33	1.48	0.6764
1	2	2	2	-0.7473	0.6969	32.2	-1.07	0.8901
1	2	2	3	-2.3595	0.7279	36.4	-3.24	0.0237
1	3	2	1	1.0577	0.7199	35.3	1.47	0.6849
1	3	2	2	-0.7321	0.7139	34.6	-1.03	0.9073
1	3	2	3	-2.3443	0.7442	38.7	-3.15	0.0303
2	1	2	2	-1.7898	0.4751	54	-3.77	0.0053
2	1	2	3	-3.4019	0.5269	56.8	-6.46	<0.0001
2	2	2	3	-1.6122	0.5083	50.8	-3.17	0.0286

The sample size equation can be used to evaluate the power of a test by solving for $t_{\beta,v}$ and then determining the power as the value of $1 - \beta$. The equation for $t_{\beta,v}$ is

$$t_{\beta,v} = \sqrt{\frac{nb^2}{2\hat{\sigma}^2}} - t_{\alpha/2,v}$$

Suppose for Example 24.6 that it is of interest to detect a difference of 1.5 force units between ration means within a level of temperature. The parameter of interest representing the difference of the two rations at the first temperature is $\bar{\mu}_{11} - \bar{\mu}_{21}$. If the data set were balanced, the best estimate of $\bar{\mu}_{11} - \bar{\mu}_{21}$ is $\bar{f}_{1\cdot 1} - \bar{f}_{2\cdot 1}$. The model for $\bar{f}_{i\cdot k}$ is $\bar{f}_{i\cdot k} = \bar{\mu}_{ik} + \bar{\alpha}_i + \bar{s}_{i\cdot k} + \bar{\epsilon}_{i\cdot k}$, and the model for $\bar{f}_{1\cdot 1} - \bar{f}_{2\cdot 1}$ is $\bar{f}_{1\cdot 1} - \bar{f}_{2\cdot 1} = \bar{\mu}_{11} - \bar{\mu}_{21} + \bar{\alpha}_1 - \bar{\alpha}_2 + \bar{s}_{1\cdot 1} - \bar{s}_{2\cdot 1} + \bar{\epsilon}_{1\cdot 1} - \bar{\epsilon}_{2\cdot 1}$.

The sample size needed is the number of animals per ration since the number of sides and the number of steaks per side are fixed to 2 and 3, respectively. The variance of $\bar{f}_{1\cdot 1} - \bar{f}_{2\cdot 1}$ is

$$\begin{aligned}\text{Var}(\bar{f}_{1\cdot 1} - \bar{f}_{2\cdot 1}) &= \frac{2\sigma_{\text{animal}}^2}{n} + \frac{2\sigma_{\text{side}}^2}{n} + \frac{2\sigma_{\text{steak}}^2}{3n} \\ &= \frac{2}{3n}(\sigma_{\text{steak}}^2 + 3\sigma_{\text{side}}^2 + 3\sigma_{\text{animal}}^2)\end{aligned}$$

The number of animals per ration needed to detect a difference of δ between the two means with type I and type II error rates of α and β , respectively is approximately

$$n = \frac{2}{3}(\hat{\sigma}_{\text{steak}}^2 + 3\hat{\sigma}_{\text{side}}^2 + 3\hat{\sigma}_{\text{animal}}^2)[t_{\alpha/2,v} + t_{\beta,v}]^2 / \delta^2$$

where

$$\hat{\sigma}_{\text{steak}}^2 + 3\hat{\sigma}_{\text{side}}^2 + 3\hat{\sigma}_{\text{animal}}^2$$

is an estimate of $\sigma_{\text{steak}}^2 + 3\sigma_{\text{side}}^2 + 3\sigma_{\text{animal}}^2$ based on \hat{v} degrees of freedom. The expected mean squares of the three error terms using a balanced data set are

$$\begin{aligned} E[MSAnimal(Ration)] &= \sigma_{\text{steak}}^2 + 3\sigma_{\text{side}}^2 + 6\sigma_{\text{animal}}^2 \\ E[MSTemp \times Animal(Ration)] &= \sigma_{\text{steak}}^2 + 3\sigma_{\text{side}}^2 \end{aligned}$$

and

$$E[MSPack \times Animal(Temp * Ration)] = \sigma_{\text{steak}}^2$$

An unbiased estimate of $\sigma_{\text{steak}}^2 + 3\sigma_{\text{side}}^2 + 3\sigma_{\text{animal}}^2$ is

$$\hat{\sigma}_{\text{steak}}^2 + 3\hat{\sigma}_{\text{side}}^2 + 3\hat{\sigma}_{\text{animal}}^2 = \frac{1}{2}MSAnimal(Ration) + \frac{1}{2}MSTemp \times Animal(Ration)$$

with approximate degrees of freedom obtained by the Satterthwaite as

$$\hat{v} = \frac{(\hat{\sigma}_{\text{steak}}^2 + 3\hat{\sigma}_{\text{side}}^2 + 3\hat{\sigma}_{\text{animal}}^2)^2}{\frac{[\frac{1}{2}MSAnimal(Ration)]^2}{18} + \frac{[\frac{1}{2}MSTemp \times Animal(Ration)]^2}{18}}$$

When the data set is not balanced, the approximate degrees of freedom can be evaluated using the estimates of the variance components and their corresponding asymptotic covariance matrix from REML. Let $f(\sigma_{\epsilon}^2, \sigma_{\text{side}}^2, \sigma_{\text{animal}}^2)$ denote the function of the variance components that is to be estimated. The estimate of $f(\sigma_{\epsilon}^2, \sigma_{\text{side}}^2, \sigma_{\text{animal}}^2)$ is $f(\hat{\sigma}_{\epsilon}^2, \hat{\sigma}_{\text{side}}^2, \hat{\sigma}_{\text{animal}}^2)$ and the approximate variance is

$$\text{Var}[f(\hat{\sigma}_{\epsilon}^2, \hat{\sigma}_{\text{side}}^2, \hat{\sigma}_{\text{animal}}^2)] = \begin{bmatrix} \frac{\partial f}{\partial \sigma_{\epsilon}^2} & \frac{\partial f}{\partial \sigma_{\text{side}}^2} & \frac{\partial f}{\partial \sigma_{\text{animal}}^2} \end{bmatrix} \text{Var}[\hat{\sigma}_{\epsilon}^2, \hat{\sigma}_{\text{side}}^2, \hat{\sigma}_{\text{animal}}^2] \begin{bmatrix} \frac{\partial f}{\partial \sigma_{\epsilon}^2} \\ \frac{\partial f}{\partial \sigma_{\text{side}}^2} \\ \frac{\partial f}{\partial \sigma_{\text{animal}}^2} \end{bmatrix}$$

Evaluate $\text{Var}[f(\hat{\sigma}_{\epsilon}^2, \sigma_{\text{side}}^2, \hat{\sigma}_{\text{animal}}^2)]$ at the REML estimates of the variance components and compute

$$Z = \frac{f(\hat{\sigma}_{\epsilon}^2, \hat{\sigma}_{\text{side}}^2, \hat{\sigma}_{\text{animal}}^2)}{\sqrt{\text{Var}[f(\hat{\sigma}_{\epsilon}^2, \hat{\sigma}_{\text{side}}^2, \hat{\sigma}_{\text{animal}}^2)]}}$$

then the Satterthwaite approximation to the number of degrees of freedom associated with $f(\hat{\sigma}_e^2, \hat{\sigma}_{\text{side}}^2, \hat{\sigma}_{\text{animal}}^2)$ is $\hat{v} = 2Z^2$.

For Example 24.6, use the results of the analysis and determine the number of animals that are needed to detect a difference of 1.5 shear units with type I and type II error rates of 0.05 and 0.05, respectively. The estimates of the variance components from Table 24.30 are $\hat{\sigma}_{\text{animal}}^2 = 1.2434$, $\hat{\sigma}_{\text{side}}^2 = 0.1992$ and $\hat{\sigma}_{\text{steak}}^2 = 1.7529$. The estimate of $\sigma_{\text{steak}}^2 + 3\sigma_{\text{side}}^2 + 3\sigma_{\text{animal}}^2$ is 6.08086. The estimate of the asymptotic covariance matrix of the variance components is in Table 24.36 and is

$$\widetilde{\text{Var}} \begin{bmatrix} \hat{\sigma}_{\text{animal}}^2 \\ \hat{\sigma}_{\text{side}}^2 \\ \hat{\sigma}_{\text{steak}}^2 \end{bmatrix} = \begin{bmatrix} 0.3723 & -0.06154 & 0.002209 \\ -0.06154 & 0.1305 & -0.05370 \\ 0.002209 & -0.05370 & 0.1296 \end{bmatrix} = \hat{V}$$

The approximate variance of $\hat{\sigma}_{\text{steak}}^2 + 3\hat{\sigma}_{\text{side}}^2 + 3\hat{\sigma}_{\text{animal}}^2$ is then

$$\widehat{\text{Var}}(\hat{\sigma}_{\text{steak}}^2 + 3\hat{\sigma}_{\text{side}}^2 + 3\hat{\sigma}_{\text{animal}}^2) = [3 \quad 3 \quad 1] \hat{V} \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix} = 3.23805$$

and the Z-score is

$$Z = \frac{\hat{\sigma}_{\text{steak}}^2 + 3\hat{\sigma}_{\text{side}}^2 + 3\hat{\sigma}_{\text{animal}}^2}{\sqrt{\widehat{\text{Var}}(\hat{\sigma}_{\text{steak}}^2 + 3\hat{\sigma}_{\text{side}}^2 + 3\hat{\sigma}_{\text{animal}}^2)}} = 3.379$$

thus there are $2(3.379)^2 = 22.8$ degrees of freedom associated with $\hat{\sigma}_{\text{steak}}^2 + 3\hat{\sigma}_{\text{side}}^2 + 3\hat{\sigma}_{\text{animal}}^2$. The estimated sample size is $n = \frac{2}{3}(6.08086)[t_{0.025, 22.8} + t_{0.05, 22.8}]^2 / 1.5^2 = 25.8$ or $n = 26$ animals per ration.

The power of the study with 10 animals per ration to detect a shear force of 1.5 units between rations within a temperature is determined by computing

$$t_{\beta, \hat{v}} = \sqrt{\frac{10\delta^2}{\frac{2}{3}(\hat{\sigma}_{\text{steak}}^2 + 3\hat{\sigma}_{\text{side}}^2 + 3\hat{\sigma}_{\text{animal}}^2)}} - t_{\alpha/2, \hat{v}} = \sqrt{\frac{10(1.5)^2}{\frac{2}{3}(6.08086)}} - 2.069 = 0.2869$$

TABLE 24.36

Asymptotic Covariance Matrix for the Estimates of the Variance Components of Example 24.6

Asymptotic Covariance Matrix of Estimates

Row	Covariance Parameter	CovP1	CovP2	CovP3
1	Animal (ration)	0.3723	-0.06154	0.002209
2	Animal \times temperature (ration)	-0.06154	0.1305	-0.05370
3	Residual	0.002209	-0.05370	0.1296

For $\alpha = 0.05$, $\beta = 0.388$, providing power = $1 - \beta = 0.612$; that is, with 10 animals per ration there is 61% chance of determining that two ration means are significantly different when they actually differ by 1.5 units.

The second sample size problem is to determine the number of animals required to detect a difference of 1.5 shear force units between two temperature means at the same level of packaging averaged over the two levels of ration. The comparison of interest is $\bar{\mu}_{..11} - \bar{\mu}_{..21}$. The model needed to evaluate the variance of the difference is $\bar{f}_{..kl} = \bar{\mu}_{..kl} + \bar{\sigma}_{..} + \bar{s}_{..k} + \bar{\epsilon}_{..jk}$. The variance of the difference is

$$\begin{aligned}\text{Var}(\bar{f}_{..11} - \bar{f}_{..21}) &= \text{Var}(\bar{s}_{..1} - \bar{s}_{..2} + \bar{\epsilon}_{..11} - \bar{\epsilon}_{..21}) \\ &= \frac{2\sigma_{\text{side}}^2}{2n} + \frac{2\sigma_{\text{steak}}^2}{2n} \\ &= \frac{2}{2n}(\sigma_{\text{side}}^2 + \sigma_{\text{steak}}^2)\end{aligned}$$

The number of animals per ration is $n = \frac{2}{2}(\hat{\sigma}_{\text{steak}}^2 + \sigma_{\text{side}}^2)[t_{0.025,49.9} + t_{0.05,49.9}]^2 / 1.5^2 = 11.8$ where the number of degrees of freedom was determined to be 49.9 using the procedure outlined above.

There are several comparisons that could be important, so sample sizes need to be determined for each comparison and the number of animals or number of whole plots needed can be based on the maximum number of animals required for each of the comparisons deemed to be important.

24.8 Computations Using JMP—Example 24.7

The data set in Figures 24.9 and 24.10 are from a JMP table with nominal variables, *block*, *variety*, *fert*, and *rate*. The design of the experiment consists of four blocks of four whole plots where the combinations of two varieties of wheat and two levels of fertilizer are randomly assigned to the whole plots. Each whole plot is split into three subplots and the three levels of seeding rate were randomly assigned to the three subplots. The whole plot design is a two-way treatment structure (varieties by levels of fertilizer) in a randomized complete block design structure (four blocks). There are some missing data points due to uncontrollable insect damage, but the following evaluation of the structures are carried out assuming there are no missing data points. The whole plot design analysis of variance table is given in Table 24.37. The whole-plot error is computed from the treatment structure by design structure interaction. In this case, there is a two-way treatment structure with four treatment combinations, thus the whole plot error is made up of $(4 - 1)(4 - 1) = 9$ degrees of freedom. Next, compare the levels of seeding rates at variety one and fertilizer 1. The resulting design is a one-way treatment structure in a randomized complete block design structure and the analysis of variance table is given in Table 24.38. The subplot error degrees of freedom are computed from the block by rate interaction, providing 6 degrees of freedom. There are four combinations of variety and fertilizer where each provides 6 degrees of freedom for the subplot error. Pooling these four sets of 6 degrees of

example_24_7

	block	Variety	fert	rate	yield	
1	1	1	1	4	23.3	
2	1	1	1	6	27.7	
3	1	1	1	10	▪	
4	1	1	2	4	24.2	
5	1	1	2	6	31	
6	1	1	2	10	31.4	
7	1	2	1	4	25.9	
8	1	2	1	6	26.7	
9	1	2	1	10	▪	
10	1	2	2	4	31.3	
11	1	2	2	6	32.4	
12	1	2	2	10	37.2	
13	2	1	1	4	▪	
14	2	1	1	6	19.3	
15	2	1	1	10	17.2	
16	2	1	2	4	13.7	
17	2	1	2	6	▪	
18	2	1	2	10	19.8	
19	2	2	1	4	7.7	
20	2	2	1	6	13.4	
21	2	2	1	10	10.2	
22	2	2	2	4	20.3	
23	2	2	2	6	16.2	
24	2	2	2	10	26.3	
25	3	1	1	4	24.9	
26	3	1	1	6	30.3	
27	3	1	1	10	30.6	
28	3	1	2	4	23.5	
29	3	1	2	6	28.1	
30	3	1	2	10	28.9	
31	3	2	1	4	24.2	
32	3	2	1	6	32.9	

Columns (5/0)

- block
- Variety
- fert
- rate
- yield

Rows

All rows	48
Selected	0
Excluded	0
Hidden	0
Labelled	0

FIGURE 24.9 Data set with first 32 observations for Example 24.7.

freedom provides $(4)(6) = 24$ degrees of freedom for the subplot error. A model that can be used to describe this data is

$$y_{ijkl} = \mu_{ijl} + b_k + w_{ijk} + \varepsilon_{ijkl} \quad i = 1, 2, j = 1, 2, k = 1, 2, 3, 4, l = 1, 2, 3$$

where $b_k \sim i.i.d. N(0, \sigma_{blk}^2)$, $w_{ijk} \sim i.i.d. N(0, \sigma_{wp}^2)$, $\varepsilon_{ijkl} \sim i.i.d. N(0, \sigma_{sp}^2)$, and μ_{ijl} denotes the mean of variety i , fertilizer j and seeding rate l .

	33	3	2	1	10	33.3	
	34	3	2	2	4	▪	
	35	3	2	2	6	23.8	
	36	3	2	2	10	▪	
	37	4	1	1	4	32.2	
	38	4	1	1	6	37.7	
	39	4	1	1	10	▪	
	40	4	1	2	4	35.5	
	41	4	1	2	6	▪	
All rows	48						
Selected	0						
Excluded	0						
Hidden	0						
Labelled	0						

FIGURE 24.10 Data set with last 16 observations for Example 24.7.**TABLE 24.37**

Whole-Plot Analysis for Example 24.7

Source	df
Block	3
Variety	1
Fertilizer	1
Variety \times fertilizer	1
Error(whole-plot)	9

TABLE 24.38

Analysis of Variance for Comparing Rates at Variety 1 and Fertilizer 1

Source	df
Block	3
Rate	2
Error(subplot)	6

The analysis of variance table with the two error terms and the expected mean squares is displayed in Table 24.39. The JMP fit model screen is presented in Figure 24.11. The *block* and *block \times variety \times fert* terms have been declared to be random effects by using the attributes button. The default method of estimation is REML, but the type III sums of squares can be used by changing the method button. Click the run model button to obtain the analysis. Figure 24.12 contains the estimates of the variance components and the analysis

TABLE 24.39

Analysis of Variance Table for Example 24.7 Assuming No Missing Data

Source	df	Expected Mean Squares
Block	3	$\sigma_{sp}^2 + 3\sigma_{wp}^2 + 12\sigma_{blk}^2$
Variety	1	$\sigma_{sp}^2 + 3\sigma_{wp}^2 + \phi^2(V)$
Fertilizer	1	$\sigma_{sp}^2 + 3\sigma_{wp}^2 + \phi^2(F)$
Variety \times fertilizer	1	$\sigma_{sp}^2 + 3\sigma_{wp}^2 + \phi^2(V \times F)$
Error(whole-plot) = variety \times fertilizer \times block	9	$\sigma_{sp}^2 + 3\sigma_{wp}^2$
Rate	2	$\sigma_{sp}^2 + \phi^2(R)$
Rate \times variety	2	$\sigma_{sp}^2 + \phi^2(V \times R)$
Rate \times fertilizer	2	$\sigma_{sp}^2 + \phi^2(F \times R)$
Rate \times variety \times fertilizer	2	$\sigma_{sp}^2 + \phi^2(V \times F \times R)$
Error(subplot) = rate \times block(variety \times fertilizer)	24	σ_{sp}^2

FIGURE 24.11 Fit model table for Example 24.7.

of the fixed effects. The confidence intervals about the variance components are computed using the Wald interval instead of a Satterthwaite approximation interval, as done by SAS-Mixed. All of the effects involving rate are significant, including the three-way interaction, thus one needs to address the *variety \times fert \times rate* three-way means. Clicking on the

REML Variance Component Estimates							
Random Effect	Var Ratio	Var Component	Std Error	95% Lower	95% Upper	Pct of Total	
block	23.119242	64.838059	56.44131	-45.78691	175.46303	77.805	
block*Variety*fert	5.5949103	15.690961	7.8980547	0.2107738	31.171148	18.829	
Residual		2.8045063	0.9820348	1.5631132	6.4330899	3.365	
Total		83.333527				100.000	
-2 LogLikelihood = 187.42860736							

► Iterations

▼ Fixed Effect Tests

Source	Nparm	DF	DDFden	F Ratio	Prob > F
Variety	1	1	9.313	0.0030	0.9577
fert	1	1	9.308	0.3425	0.5723
Variety*fert	1	1	9.304	0.0054	0.9431
rate	2	2	16.6	27.9065	<.0001*
fert*rate	2	2	16.57	4.9021	0.0213*
Variety*rate	2	2	16.66	4.6008	0.0257*
Variety*fert*rate	2	2	16.66	4.1246	0.0351*

FIGURE 24.12 Estimates of the variance components (REML) and tests of the fixed effects for Example 24.7.

▼ (▼ Variety
▼ Least Squares Means Table
Level Least Sq Mean Std Error
1 27.434210 4.2835179
2 27.546696 4.2819455
▼ (▼ fert
▼ Least Squares Means Table
Level Least Sq Mean Std Error
1 26.886153 4.2812613
2 28.094752 4.2841854
▼ (▼ Variety*fert
▼ Least Squares Means Table
Level Least Sq Mean Std Error
1,1 26.754198 4.5269151
1,2 28.114222 4.5257066
2,1 27.018109 4.5171506
2,2 28.075282 4.5295541

FIGURE 24.13 Least squares means for varieties, fertilizers and their interaction for Example 24.7.

effects button provides the tables of least squares means for each effect, as shown in Figures 24.13—24.15. The least square means options are displayed in Figure 24.16. The tables option is the default. The plots option provides a line plot of the means. The contrast option provides the ability to construct contrasts of the least squares means that are of interest. The Tukey HSD option provides the Tukey adjustment for multiple comparisons, and the

▼ **rate**

▼ **Least Squares Means Table**

Level	Least Sq Mean	Std Error
4	24.600850	4.1734394
6	28.091917	4.1750166
10	29.778591	4.1816862

▼ **fert*rate**

▼ **Least Squares Means Table**

Level	Least Sq Mean	Std Error
1,4	23.424406	4.3146250
1,6	28.762500	4.3036552
1,10	28.471554	4.3465763
2,4	25.777295	4.3168902
2,6	27.421334	4.3338869
2,10	31.085628	4.3168902

▼ **Variety*rate**

▼ **Least Squares Means Table**

Level	Least Sq Mean	Std Error
1,4	23.811906	4.3146250
1,6	29.246334	4.3338869
1,10	29.244391	4.3364808
2,4	25.389795	4.3168902
2,6	26.937500	4.3036552
2,10	30.312792	4.3269197

FIGURE 24.14 Least squares means for rate and its interactions with variety and fertilizer.

▼ **Variety*fert*rate**

▼ **Least Squares Means Table**

Level	Least Sq Mean	Std Error
1,1,4	23.398811	4.6055992
1,1,6	28.750000	4.5643599
1,1,10	28.113782	4.6869873
1,2,4	24.225000	4.5643599
1,2,6	29.742667	4.6773808
1,2,10	30.375000	4.5643599
2,1,4	23.450000	4.5643599
2,1,6	28.775000	4.5643599
2,1,10	28.829326	4.6023424
2,2,4	27.329590	4.6140820
2,2,6	25.100000	4.5643599
2,2,10	31.796257	4.6140820

FIGURE 24.15 Least squares means for variety by fertilizer by rate interaction for Example 24.7.

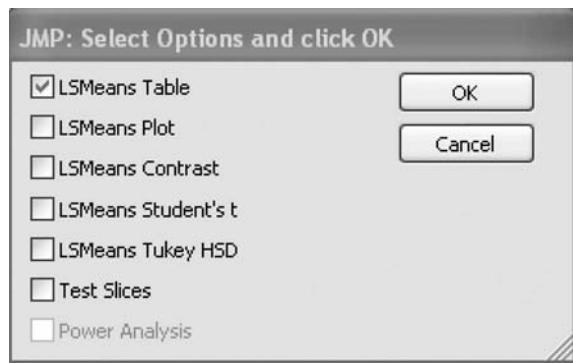


FIGURE 24.16 Least squares means options of an effect.

Level	Least Sq Mean
2,2,10 A C D E G H	31.796257
1,2,10 A B D E H	30.375000
1,2,6 A B C D E F G H	29.742667
2,1,10 A B C D F	28.829326
2,1,6 A B C D F	28.775000
1,1,6 A B C E F G	28.750000
1,1,10 A B C D E F G H	28.113782
2,2,4 A B C D E F G H	27.329590
2,2,6 B F	25.100000
1,2,4 C F G	24.225000
2,1,4 E G H	23.450000
1,1,4 D H	23.398811

Levels not connected by same letter are significantly different.

FIGURE 24.17 Least squares means for the three way interaction letters denoting significant groupings via Tukey's method.

slice option provides a test of the equality of means for each level of each effect in an interaction. The results in Figure 24.17 are part of the Tukey HSD option results with lines or letters provided. Since there are missing data in the study, the letters may result in some contradictions. For example, mean 2, 2, 10 is significantly different from mean 2, 2, 6 (see A and B) but 2, 2, 10 is not significantly different from 1, 2, 4 (see C) even though the mean for 2, 2, 6 is larger than the mean for 1, 2, 4. Missing data cause the variances of comparisons to be different, which in turn can provide multiple comparisons that are not completely ordered as demonstrated above.

Figure 24.18 contains the complete set of results of applying the Tukey HSD option to the *Variety* \times *Fert* means. The table contains the estimated differences, the estimated standard errors of each difference and upper and lower 95% confidence intervals about each difference. When a confidence interval does not include zero, the results are displayed in red. All of these confidence intervals include zero, so they are displayed in black. Contrasts of the least squares means can be evaluated by selecting coefficients for each mean on the contrast table displayed in Figure 24.19.

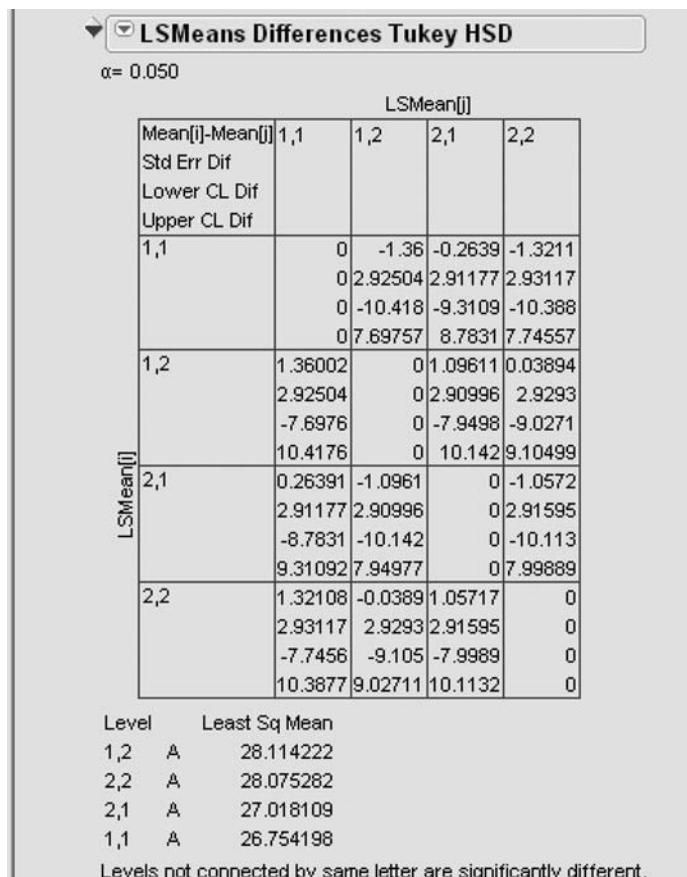


FIGURE 24.18 Computations for Tukey's method of multiple comparisons for the variety by fertilizer means for Example 24.7.

This example shows some of the options that are available for analyzing split-plot type designs using the JMP software. All of the data sets and models described in this chapter can be analyzed using the fit model screen of JMP.

24.9 Concluding Remarks

In this chapter, split-plot designs were discussed in detail with the analyses being carried out using SAS-Mixed and JMP. The analysis of variance tables for the designs discussed involve more than one error term and methods are described that enable one to construct the appropriate representation of each error term. Estimated standard errors of various estimates can involve more than one error term. Methods were presented for computing estimated standard errors for the various types of comparisons that may be interesting to an experimenter. Estimated standard errors can also be used for carrying out multiple comparisons. Methods are described for estimating linear, quadratic, and similar trends when the levels of a factor are quantitative. Sample size and power methods are demonstrated. Seven examples were discussed in detail.

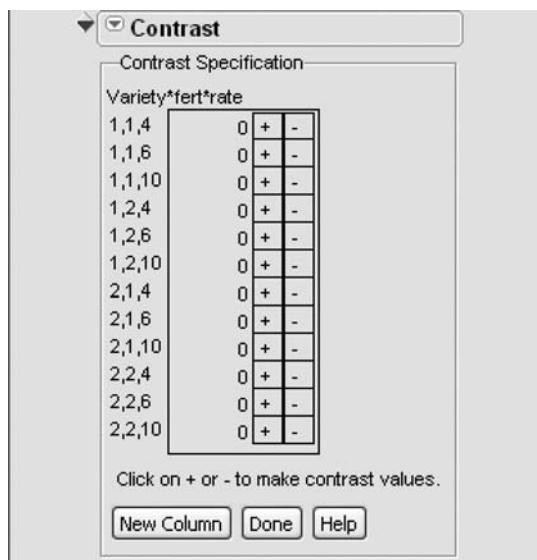


FIGURE 24.19 Contrast option window for least squares means.

24.10 Exercises

- 24.1 A baker wanted to determine the effect that the amount of fat in a recipe of cookie dough would have on the texture of the cookie. She also wanted to determine if the temperature (°F) at which the cookies were baked would have an influence on the texture of the surface. The texture of the cookie is measured by determining the amount of force (g) required to penetrate the cookie surface. The process she used was to make a batch of cookie dough for each of the four recipes and bake one cookie from each batch in the oven at one time. She carried this process out each of four days when she baked cookies at three different temperatures. The following are the forces required to penetrate the cookie surface

Data for Exercise 24.1. Data Are Force (g) to Penetrate a Cookie Surface

Day	Temperature	2% Fat	4% Fat	6% Fat	8% Fat
1	350	7.4	7.1	7.2	6.7
1	375	11.2	11.1	10.6	10.8
1	400	11.8	11.2	11.3	11.4
2	350	8.7	8.1	7.4	7.2
2	375	10.8	10.1	10.0	10.2
2	400	12.9	14.0	13.6	12.7
3	350	7.0	6.4	6.2	6.1
3	375	8.3	7.7	7.4	8.4
3	400	12.1	12.1	12.2	12.0
4	350	6.6	6.7	6.9	6.0
4	375	9.1	8.1	8.3	8.0
4	400	11.7	12.1	11.9	11.4

- 1) Key out the analysis of variance table of this study.
 - 2) Carry out an analysis of variance and determine the important fixed effects in the model.
 - 3) Complete the analysis by carrying out multiple comparisons on the important effects in the model.
 - 4) Fit a response surface model and determine a good model to predict force as a function of percentage fat and temperature.
 - 5) Determine the number of times (days) she would have to repeat this experiment so that she could detect a difference of one unit in force in the temperature means within a given recipe using the estimate of the variance components from this data and type I and type II errors of 0.05 and 0.20, respectively.
- 24.2 An agriculture engineer studied the amount of water that would flow through a particular type of nozzle at a fixed pressure in the line to which the nozzle was attached. Three different nozzle types were used where one had a round hole, one an oval hole and the other a narrow slit. The areas of the openings were equal in size. The study involved setting the pressure in the system to a fixed value and then attaching each nozzle to the line (in a random order). The amount of water that flowed through the nozzle in 20 s was measured (ounces). The pressure was selected at random (10, 20, 40, and 80 psi) and all four pressures were evaluated. The whole process was repeated three times (blocks). The data are given in the following table.

Data for Exercise 24.2 with Three Nozzle Types
(R, O, and S) and Ounces of Water per Trial

Block	Pressure	R	O	S
1	10	4.4	4.5	3.2
1	20	5.7	5.0	3.9
1	40	7.5	6.2	5.2
1	80	10.7	11.0	9.9
2	10	6.1	6.7	6.5
2	20	9.4	7.4	8.2
2	40	9.6	7.3	6.3
2	80	14.6	13.2	12.1
3	10	2.4	3.0	1.6
3	20	3.4	2.5	2.3
3	40	5.0	2.3	1.9
3	80	11.5	11.5	10.2

- 1) Key out the analysis of variance table for this study.
 - 2) Carry out an analysis of variance and evaluate the presence of interaction among the levels of pressure and the three nozzle types.
 - 3) Determine if there is a linear or quadratic effect of pressure for each of the three nozzle types.
- 24.3 The agricultural engineer who evaluated the nozzles in Exercise 24.2 decided that more data were needed and ran some additional trials. In this study he used three pressures, 40, 60, and 80 psi, as well as the three shapes of holes in the

nozzles, R, O, and S. The process for this study was to place a nozzle type on the line and then determine the amount of water through the nozzle at each of the three pressures where the order of the pressures was selected at random for each nozzle type. He repeated the study for two blocks. The amount of water in the 20 s time interval was recorded as in the following table.

Data for Exercise 24.3 with Three Nozzle Types (R, O, and S) and Three Line Pressures (P-40, P-60, and P-80). The Data Are in Ounces of Water per Trial

Block	Nozzle	P-40	P-60	P-80
1	O	6.9	8.7	12.2
1	R	9.7	11.1	12.6
1	S	6.7	8.6	12.5
2	O	8.0	9.7	12.8
2	R	9.2	11.9	12.3
2	S	7.1	9.1	13.4
3	O	6.7	8.1	11.6
3	R	8.1	10.0	11.8
3	S	4.0	5.6	9.5

- 1) Key out the analysis of variance table for this study.
 - 2) Carry out an analysis of variance and evaluate the presence of interaction among the levels of pressure and the three nozzle types.
 - 3) Determine if there is a linear or quadratic effect of pressure for each of the three nozzle types.
- 24.4 Carry out a combined analysis of the data in Exercises 24.2 and 24.3. Discuss the method of analysis and determine if there is a linear or quadratic effect of pressure on the nozzle type.
- 24.5 An enhanced after-school exercise program was used to see if the participating children would do more moderate or better exercise than those students in a regular after-school program. The students of interest are those classified as coming from low-income families. Ten schools were selected, six large and four small, and three of the large and two of the small schools were randomly selected to receive the enhanced exercise after-school program. The other five schools carried out their regular after-school program. Both female and male students were from either the second, third, or fourth grades. Schools numbered 1–3 and

Data for Exercise 24.5 for Female Students Using the Regular Exercise Program

School 1			School 2			School 3			School 4			School 5		
cl4	cl2	cl3												
19	15	18	18	18	17	21	17	17	13	14	11	13	20	14
18	18	17	18	17	17	16	15	18	11	15	13	13	17	16
18	18	17	15	20	14	18	—	13	13	—	12	13	—	18
18	15	—	18	14	—	—	—	12	10	—	12	14	—	16
14	15	—	—	—	—	—	—	12	13	—	12	—	—	—
19	—	—	—	—	—	—	—	—	—	—	13	—	—	—

Data for Exercise 24.5 for Male Students Using the Regular Exercise Program

School 1			School 2			School 3			School 4			School 5		
cl4	cl4	cl4												
12	20	16	10	16	13	14	16	13	9	14	11	12	15	17
13	18	14	14	16	11	11	16	11	10	14	10	10	17	14
16	13	12	13	15	16	—	13	10	8	11	12	12	18	16
14	20	—	12	17	—	—	14	12	9	16	9	—	—	—
—	—	—	13	—	—	—	12	10	7	13	10	—	—	—
—	—	—	11	—	—	—	—	—	6	14	—	—	—	—

Data for Exercise 24.5 for Female Students Using the Enhanced Exercise Program

School 6			School 7			School 8			School 9			School 10		
cl4	cl4	cl4	cl4	cl4	cl4									
22	17	18	25	17	22	20	20	24	22	17	24	17	18	17
21	21	16	21	22	22	21	24	19	22	21	20	22	19	20
18	23	19	21	20	22	17	—	22	18	18	—	21	22	22
—	23	19	26	20	18	—	—	22	25	20	—	—	—	—
—	—	21	20	—	20	—	—	24	21	—	—	—	—	—
—	—	—	—	—	23	—	—	—	19	—	—	—	—	—

Data for Exercise 24.5 for Male Students Using the Enhanced Exercise Program

School 6			School 7			School 8			School 9			School 10		
cl4	cl4	cl4	cl4	cl4	cl4									
16	19	18	13	12	15	13	14	22	19	13	15	12	17	16
12	17	13	16	14	13	13	19	19	14	16	16	8	15	15
13	20	10	12	18	15	—	18	16	—	15	19	—	20	—
16	—	—	16	15	18	—	20	—	—	—	14	—	16	—
14	—	—	—	—	18	—	16	—	—	—	17	—	19	—
—	—	—	—	—	—	—	—	—	—	—	15	—	19	—

6–8 denote the large schools and schools numbered 4–5 and 9–10 denote the small schools. The data are given above.

- 1) Assuming the data set is such that there are six males and six females per class, key out the analysis of variance table and include the numbers of degrees of freedom.
- 2) For the analysis of variance table in part (1) determine the computational representation of the three error terms.
- 3) Carry out an analysis of variance using the data set and include all needed comparisons where the Tukey adjustment is used to account for multiple comparisons.

25

Methods for Analyzing Strip-Plot Type Designs

Strip-plot design structures are used in situations where the experimental units are arranged in rectangles, such as in a field, on a piece of fabric, or with a stack of cages. The strip-plot design was described in Chapter 5 as one of the basic design structures and the details of the analysis are discussed in this chapter. A strip-plot design structure occurs when the levels of one factor are applied to the rows of each rectangle and the levels of another factor are applied to the columns of each rectangle. Thus, this design is useful when the method of application of treatments does not allow the experimental units to be treated individually but rather in groups such as a whole row of units or a whole column of units. The strip-plot design structure is also useful for experiments where the experimental units are processed in a series of steps where groups of experimental units are treated together within each step. Six examples are discussed to demonstrate the analyses of complex design structures. The first example involving levels of both irrigation and nitrogen is used to demonstrate the basic analysis of the simplest strip-plot design. Most of the remaining examples involve various combinations of strip-plot and split-plot design structures. The last example involves a strip-strip-plot design structure.

25.1 Description of the Strip-Plot Design and Model

A strip-plot design structure is similar to a split-plot design structure, but the experimental units are constructed differently. The strip-plot involves at least a two-way treatment structure where the basic experimental units are arranged in sets of rectangles. Each rectangle has a rows where a is the number of levels of the first factor (A) and c columns where c is the number of levels of the second factor (C). The levels of factor A are randomly assigned to rows of the rectangle where all of the experimental units within a row receive the same level of A and are treated together. The levels of factor C are randomly assigned to the columns of the rectangle so that all of the experimental units within a column receive the same level of C and are treated together. A schematic showing the assignment of the levels of A to the rows and the levels of C to the columns is shown in Figure 25.1, where

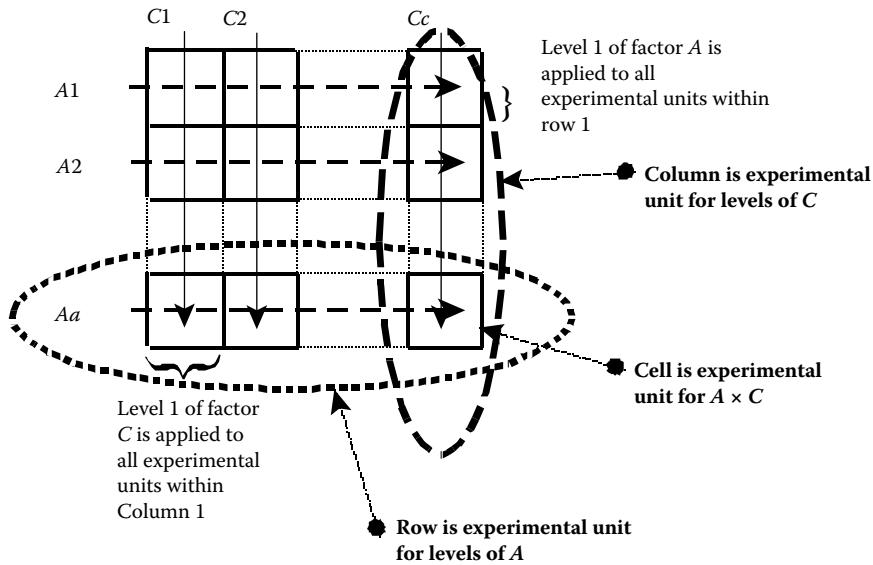


FIGURE 25.1 The levels of two factors are assigned to strips (row and columns) of experimental units arranged in a rectangle.

there are r such rectangles representing blocked replicates. This assignment process generates a design with three sizes of experimental units. The rows are the experimental units for factor A , the columns are the experimental units for factor C , and cells are the experimental units for the $A \times C$ interaction.

An easy way to visualize the appropriate analysis is to first ignore the levels of factor C and just look at the rectangles with the levels of A assigned to the rows, as displayed in Figure 25.2. The row design consists of a one-way treatment structure in a randomized

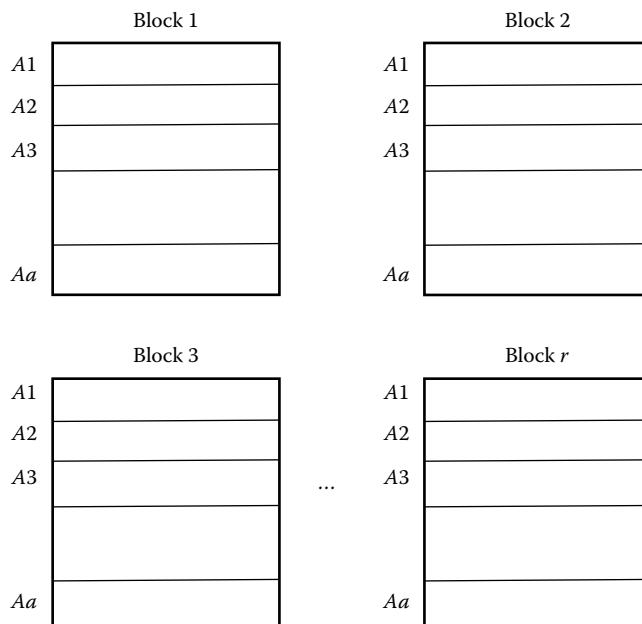


FIGURE 25.2 Assignment of the levels of A to the rows of each rectangle.

complete block design structure. A model that can be used to describe data from the row design is

$$y_{ik} = \mu_i^r + b_k^r + e_{ik}^r \quad i = 1, 2, \dots, a, \quad k = 1, 2, \dots, r$$

where the block effects are $b_k^r \sim i.i.d. N(0, \sigma_{\text{block}}^2)$ and the row effects are $e_{ik}^r \sim i.i.d. N(0, \sigma_{\text{row}}^2)$.

The analysis of variance table for the row design is given in Table 25.1 where the row error term is computed as the $\text{rectangle} \times A$ interaction.

Next, ignore the levels of factor A and look at the rectangles with the levels of C assigned to the columns, as displayed in Figure 25.3. The column design consists of a one-way treatment structure in a randomized complete block design structure. A model that can be used to describe data from the column design is

$$y_{jk} = \mu_j^c + b_k^c + e_{jk}^c \quad j = 1, 2, \dots, c, \quad k = 1, 2, \dots, r$$

TABLE 25.1

Analysis of Variance for the Row Part of the Strip-Plot Design Structure

Source	df
Block	$r - 1$
A	$a - 1$
Error(row) = $A \times \text{block}$	$(a - 1)(r - 1)$

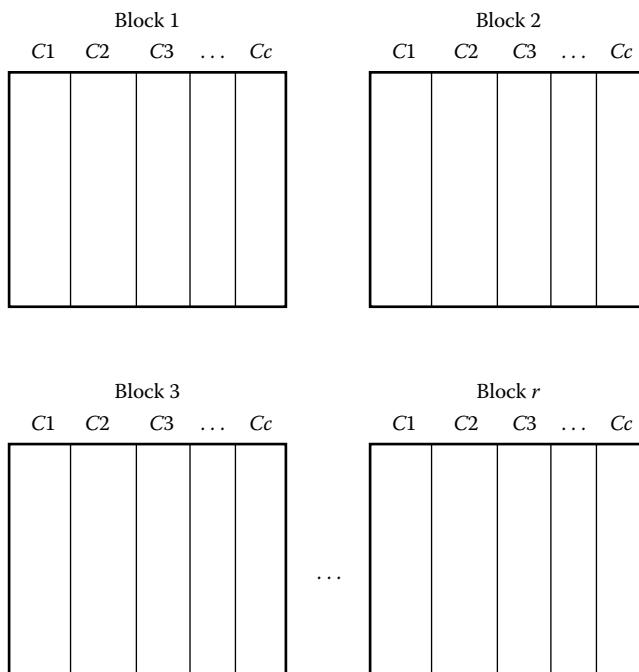


FIGURE 25.3 Assignment of the levels of C to the rows of each rectangle.

where the block effects are $b_k^c \sim i.i.d. N(0, \sigma_{\text{block}}^2)$ and the column effects are $e_{jk}^c \sim i.i.d. N(0, \sigma_{\text{column}}^2)$. The analysis of variance table for the column design is given in Table 25.2, where the column error term is computed as the $\text{rectangle} \times C$ interaction. Comparisons of the levels of factor A are between-row comparisons, those of factor C are between-column comparisons, but measures of $A \times C$ interaction are within-row comparisons by within-column comparisons. Thus the size of the experimental unit on which the $A \times C$ interaction is measured is the intersection of a row and a column which corresponds to a cell within the rectangle.

A model that describes the strip-plot design structure with two factors in the treatment structure is

$$y_{ijk} = \mu + b_k + \alpha_i + r_{ik} + \gamma_j + c_{jk} + (\alpha\gamma)_{ij} + \varepsilon_{ijk} \quad i = 1, 2, \dots, a, j = 1, 2, \dots, c, k = 1, 2, \dots, r$$

where μ denotes the overall mean effect, α_i denotes the effect of the i th level of factor A , γ_j denotes the j th level of factor C , $(\alpha\gamma)_{ij}$ denotes the interaction effect between the levels of factors A and C , b_k denotes the effect of the k th block assumed to be distributed $b_k \sim i.i.d. N(0, \sigma_{\text{block}}^2)$, r_{ik} denotes the effect of the i th row within the k th block assumed to be distributed $r_{ik} \sim i.i.d. N(0, \sigma_{\text{row}}^2)$, c_{jk} denotes the effect of the j th column within the k th block assumed to be distributed $c_{jk} \sim i.i.d. N(0, \sigma_{\text{column}}^2)$, and ε_{ijk} denotes the random effect associated with the ijk th cell within the k th block assumed to be distributed as $\varepsilon_{ijk} \sim i.i.d. N(0, \sigma_{\text{cell}}^2)$. In addition, all random terms are assumed to be independently distributed.

The row error is computed as the A by *block* interaction, the column error term is computed as the C by *block* interaction and the cell error term is computed as the A by C by *block* three-way interaction. The analysis of variance table is shown in Table 25.3, which is

TABLE 25.2

Analysis of Variance for the Column Part
of the Strip-Plot Design Structure

Source	df
Block	$r - 1$
C	$c - 1$
Error(column) = $C \times Block$	$(c - 1)(r - 1)$

TABLE 25.3

Analysis of Variance Table for the Strip-Plot Design Structure with a
Two-Way Treatment Structure

Source	df	EMS
Block	$r - 1$	$\sigma_{\text{cell}}^2 + c\sigma_{\text{row}}^2 + a\sigma_{\text{column}}^2 + ac\sigma_{\text{block}}^2$
A	$a - 1$	$\sigma_{\text{cell}}^2 + c\sigma_{\text{row}}^2 + \varphi^2(\alpha)$
Error(row)	$(a - 1)(r - 1)$	$\sigma_{\text{cell}}^2 + c\sigma_{\text{row}}^2$
C	$c - 1$	$\sigma_{\text{cell}}^2 + a\sigma_{\text{column}}^2 + \varphi^2(\gamma)$
Error(column)	$(c - 1)(r - 1)$	$\sigma_{\text{cell}}^2 + a\sigma_{\text{column}}^2$
AC	$(a - 1)(c - 1)$	$\sigma_{\text{cell}}^2 + \varphi^2(\alpha\gamma)$
Error(cell)	$(a - 1)(c - 1)(r - 1)$	σ_{cell}^2

constructed by taking a combination of the entries from Tables 25.1 and 25.2 with the A by C interaction and the cell error. The strip-plot model can be expressed as a means model as

$$y_{ijk} = \mu_{ij} + b_k + r_{ik} + c_{jk} + \varepsilon_{ijk}, \quad i = 1, 2, \dots, a, \quad j = 1, 2, \dots, c, \quad k = 1, 2, \dots, r$$

where $\mu_{ij} = \mu + \alpha_i + \gamma_j + (\alpha\gamma)_{ij}$.

For a balanced data set, the estimates of the cell means are $\hat{\mu}_{ij} = \bar{y}_{ij}$, and various contrasts, sums or means involving the μ_{ij} are estimated by taking the corresponding contrasts, sums or means of the $\hat{\mu}_{ij}$. The method of moments solution for the variance components is obtained from the equations generated by equating the mean squares to the expected mean squares in Table 25.3 giving

$$\begin{aligned}\tilde{\sigma}_{\text{cell}}^2 &= \text{MSError}(cell) \\ \tilde{\sigma}_{\text{row}}^2 &= \frac{\text{MSError}(row) - \text{MSError}(cell)}{c} \\ \tilde{\sigma}_{\text{column}}^2 &= \frac{\text{MSError}(column) - \text{MSError}(cell)}{a} \\ \tilde{\sigma}_{\text{block}}^2 &= \frac{\text{MSBlock} - \text{MSError}(column) - \text{MSError}(row) + \text{MSError}(cell)}{ac}\end{aligned}$$

The estimates of the variance components are

$$\begin{aligned}\hat{\sigma}_{\text{cell}}^2 &= \tilde{\sigma}_{\text{cell}}^2 \\ \hat{\sigma}_{\text{row}}^2 &= \begin{cases} \tilde{\sigma}_{\text{row}}^2 & \text{if } \tilde{\sigma}_{\text{row}}^2 > 0 \\ 0 & \text{if } \tilde{\sigma}_{\text{row}}^2 \leq 0 \end{cases} \\ \hat{\sigma}_{\text{column}}^2 &= \begin{cases} \tilde{\sigma}_{\text{column}}^2 & \text{if } \tilde{\sigma}_{\text{column}}^2 > 0 \\ 0 & \text{if } \tilde{\sigma}_{\text{column}}^2 \leq 0 \end{cases}\end{aligned}$$

and

$$\hat{\sigma}_{\text{block}}^2 = \begin{cases} \tilde{\sigma}_{\text{block}}^2 & \text{if } \tilde{\sigma}_{\text{block}}^2 > 0 \\ 0 & \text{if } \tilde{\sigma}_{\text{block}}^2 \leq 0 \end{cases}$$

Care must be taken when making inferences about the parameters for this model since the model does involve three error terms, one for each size of experimental unit. The necessary tests of hypotheses and appropriate standard errors for making comparisons among the means are discussed in the next section.

25.2 Techniques for Making Inferences

The column of expected mean squares in Table 25.3 provides the information necessary to construct test statistics to make inferences about the factorial effects. To test the no-interaction hypothesis $H_{0AC}: \mu_{ij} - \mu_{ij'} - \mu_{i'j} + \mu_{i'j'} = 0$ for all $i \neq i'$ and $j \neq j'$ vs H_{aAC} : (not H_{0AC}), divide the

mean square $A \times C$ by the $MSError(cell)$. To test the equal main effect means hypothesis for the levels of A , H_{0A} : $\bar{\mu}_1 = \bar{\mu}_2 = \dots = \bar{\mu}_a$ vs H_{aA} : (not H_{0A}), divide the mean square for factor A by $MSError(row)$. To test the equal means hypothesis for the levels of C , H_{0C} : $\bar{\mu}_{.1} = \bar{\mu}_{.2} = \dots = \bar{\mu}_{.c}$ vs H_{aC} : (not H_{0C}), divide the mean square for factor C by $MSError(column)$.

As for the split-plot design discussed in Chapter 24, where there were four comparisons of the means with different standard errors, there are six comparisons of the means with different standard errors for the strip-plot design. The comparisons, estimators of the comparisons based on a balanced data set and the variances of the comparisons are given in Table 25.4. The method described in Chapter 24 was used to obtain the specific variances for each contrast of the cell means. To demonstrate, the variance of the difference of two means where the levels of A and C are different for both means, such as $\mu_{12} - \mu_{34}$, is derived by first obtaining the model representation for $\hat{\mu}_{ij} = \bar{y}_{ij}.$ as $\hat{\mu}_{ij} = \mu_{ij} + \bar{b}_{.i} + \bar{r}_{.j} + \bar{c}_{.i} + \bar{\epsilon}_{ij}.$ Then the model for $\hat{\mu}_{12} - \hat{\mu}_{34}$ is

$$\hat{\mu}_{12} - \hat{\mu}_{34} = \bar{y}_{12.} - \bar{y}_{34.} = \mu_{12} - \mu_{34} + \bar{r}_{1.} - \bar{r}_{3.} + \bar{c}_{2.} - \bar{c}_{4.} + \bar{\epsilon}_{12.} - \bar{\epsilon}_{34.}$$

The estimate of $\mu_{12} - \mu_{34}$ is $\bar{y}_{12.} - \bar{y}_{34.}$ and the variance of $\hat{\mu}_{12} - \hat{\mu}_{34}$ is

$$\begin{aligned} \text{Var}(\hat{\mu}_{12} - \hat{\mu}_{34}) &= \text{Var}(\bar{r}_{1.} - \bar{r}_{3.} + \bar{c}_{2.} - \bar{c}_{4.} + \bar{\epsilon}_{12.} - \bar{\epsilon}_{34.}) \\ &= \frac{2\sigma_{\text{row}}^2}{r} + \frac{2\sigma_{\text{column}}^2}{r} + \frac{2\sigma_{\text{cell}}^2}{r} \\ &= \frac{2(\sigma_{\text{cell}}^2 + \sigma_{\text{column}}^2 + \sigma_{\text{row}}^2)}{r} \end{aligned}$$

The estimate of the variance of $\hat{\mu}_{12} - \hat{\mu}_{34}$ is

$$\widehat{\text{Var}}(\hat{\mu}_{12} - \hat{\mu}_{34}) = \frac{2(\hat{\sigma}_{\text{cell}}^2 + \hat{\sigma}_{\text{column}}^2 + \hat{\sigma}_{\text{row}}^2)}{r}$$

When $\hat{\sigma}_{\text{row}}^2 > 0$ and $\hat{\sigma}_{\text{column}}^2 > 0$ the estimate of the variance is

$$\widehat{\text{Var}}(\hat{\mu}_{12} - \hat{\mu}_{34}) = \frac{2}{abc} [aMSError(row) + cMSError(column) + (ac - a - c)MSError(cell)]$$

The six types of comparisons, their estimators and their corresponding variances are given in Table 25.4. The estimates of the standard errors for the comparisons are given in Table 25.5

TABLE 25.4

Comparisons, Estimators, and the Variance of the Estimators for the Strip-Plot Design for $i \neq m$ and $j \neq n$

Comparison	Estimator	Variance
$\bar{\mu}_{i.} - \bar{\mu}_{m.}$	$\bar{y}_{i..} - \bar{y}_{m..}$	$2(\sigma_{\text{cell}}^2 + c\sigma_{\text{row}}^2)/cr$
$\bar{\mu}_{.j} - \bar{\mu}_{.n}$	$\bar{y}_{.ji} - \bar{y}_{.ni}$	$2(\sigma_{\text{cell}}^2 + a\sigma_{\text{column}}^2)/ar$
$\mu_{ij} - \mu_{in} - \mu_{mj} + \mu_{mn}$	$\bar{y}_{ij.} - \bar{y}_{in.} - \bar{y}_{mj.} + \bar{y}_{mn.}$	$2\sigma_{\text{cell}}^2/r$
$\mu_{ij} - \mu_{in}$	$\bar{y}_{ij.} - \bar{y}_{in.}$	$2(\sigma_{\text{cell}}^2 + \sigma_{\text{column}}^2)/r$
$\mu_{ij} - \mu_{mj}$	$\bar{y}_{ij.} - \bar{y}_{mj.}$	$2(\sigma_{\text{cell}}^2 + \sigma_{\text{row}}^2)/r$
$\mu_{ij} - \mu_{mn}$	$\bar{y}_{ij.} - \bar{y}_{mn.}$	$2(\sigma_{\text{cell}}^2 + \sigma_{\text{column}}^2 + \sigma_{\text{row}}^2)/r$

TABLE 25.5

Estimated Standard Errors of the Comparisons for the Strip-Plot Design for $i \neq m$ and $j \neq n$ with Associated Degrees of Freedom

Comparison	Estimator	df
$\bar{\mu}_i - \bar{\mu}_m$	$\sqrt{\frac{2}{cr} [MSError(row)]}$	$(a-1)(r-1)$
$\bar{\mu}_{.j} - \bar{\mu}_{.n}$	$\sqrt{\frac{2}{ar} [MSError(column)]}$	$(c-1)(r-1)$
$\mu_{ij} - \mu_{in} - \mu_{mj} + \mu_{mn}$	$\sqrt{\frac{2}{r} [MSError(cell)]}$	$(a-1)(c-1)(r-1)$
$\mu_{ij} - \mu_{in}$	$\sqrt{\frac{2}{cr} [MSError(column) + (c-1)MSError(cell)]}$	$\hat{\omega}_1$
$\bar{\mu}_{ij} - \bar{\mu}_{mj}$	$\sqrt{\frac{2}{ar} [MSError(row) + (a-1)MSError(cell)]}$	$\hat{\omega}_2$
$\bar{\mu}_{ij} - \bar{\mu}_{mn}$	$\sqrt{\frac{2}{acr} [aMSError(row) + cMSError(column) + (ac-a-c)MSError(cell)]}$	$\hat{\omega}_3$

where the degrees of freedom correspond to those either from a single mean square or for a combination of mean squares computed by using the Satterthwaite approximation as

$$\hat{\omega}_1 = \frac{[MSError(column) + (c-1)MSError(cell)]^2}{\frac{[MSError(column)]^2}{(c-1)(r-1)} + \frac{[(c-1)MSError(cell)]^2}{(a-1)(c-1)(r-1)}}$$

$$\hat{\omega}_2 = \frac{[MSError(row) + (a-1)MSError(cell)]^2}{\frac{[MSError(row)]^2}{(a-1)(r-1)} + \frac{[(a-1)MSError(cell)]^2}{(a-1)(c-1)(r-1)}}$$

and

$$\hat{\omega}_3 = \frac{[aMSError(row) + cMSError(column) + (ac-a-c)MSError(cell)]^2}{\frac{[aMSError(row)]^2}{(a-1)(r-1)} + \frac{[cMSError(column)]^2}{(c-1)(r-1)} + \frac{[(ac-a-c)MSError(cell)]^2}{(a-1)(c-1)(r-1)}}$$

When there are missing data, there will be many more different types of standard errors depending on the pattern of the missing data.

25.3 Example: Nitrogen by Irrigation

An experiment was conducted to study the effects and relationships between two irrigation methods and three levels of nitrogen on the yield of wheat. The field layout utilized four blocks or rectangles and the randomization plan for treatment assignment, and the yields are shown in Figure 25.4. The SAS®-Mixed code and type III sums of squares analysis of variance table for these data are in Table 25.6. The F -values for comparing the

		Irrigation method		Irrigation method	
		I ₁	I ₂	I ₁	I ₂
Nitrogen level	N ₁	55	71	N ₂	70
	N ₃	69	78	N ₃	79
	N ₂	62	77	N ₁	63
Block 1					
		Irrigation method		Irrigation method	
		I ₂	I ₁	I ₁	I ₂
Nitrogen level	N ₃	81	77	N ₃	76
	N ₁	77	63	N ₂	66
	N ₂	79	66	N ₁	65
Block 2					
		Irrigation method		Irrigation method	
		I ₂	I ₁	I ₁	I ₂
Nitrogen level	N ₃	81	77	N ₃	79
	N ₁	77	63	N ₂	76
	N ₂	79	66	N ₁	75
Block 3					
		Irrigation method		Irrigation method	
		I ₂	I ₁	I ₁	I ₂
Nitrogen level	N ₃	81	77	N ₃	76
	N ₁	77	63	N ₂	66
	N ₂	79	66	N ₁	75
Block 4					

FIGURE 25.4 Field layout, treatment assignment and data for the nitrogen and irrigation strip-plot.

levels of nitrogen, comparing the levels of irrigation, and for evaluating the nitrogen by irrigation interaction are 60.13, 52.18, and 33.12, respectively. The estimates of the standard errors and associated degrees of freedom or approximated degrees of freedom for the six types of comparisons are given in Table 25.7. The last three comparisons involve combinations of error terms and thus the degrees of freedom are approximated. The SAS-Mixed code to provide the REML estimates of the variance components is in Table 25.8 along with the estimates of the variance components. The tests for the fixed effects are shown in Table 25.9. The results in Table 25.9 are identical to those from the type III analysis in Table 25.6. These tests are identical since the data set is balanced and the estimates of the variance

TABLE 25.6

SAS-Mixed Code to Provide Type III Sums of Squares Analysis for the Nitrogen and Irrigation Example

```
PROC MIXED CL COVTEST METHOD=TYPE3 ;
  CLASS BLK N IRR ;
  MODEL Yield=N|IRR/DDFM=KR;
  RANDOM BLK BLK*N IRR*BLK;
```

Source	df	SS	Error df	F-Value	Pr F
N	2	339.083333	6	60.13	0.0001
IRR	1	570.375000	3	52.18	0.0055
N × IRR	2	94.750000	6	33.12	0.0006
BLK	3	123.458333	3.6578	3.34	0.1479
BLK × N	6	16.916667	6	1.97	0.2147
BLK × IRR	3	32.791667	6	7.64	0.0179
Residual	6	8.583333	—	—	—

TABLE 25.7

Six Comparisons, Estimated Standard Errors and the Degrees of Freedom for the Nitrogen and Irrigation Data for $i \neq m$ and $j \neq n$

Comparison	Estimated Standard Error	<i>df</i> or Approximate <i>df</i>
$\bar{\mu}_i - \bar{\mu}_m$	0.8396	6
$\bar{\mu}_j - \bar{\mu}_n$	1.3497	3
$\mu_{ij} - \mu_{in} - \mu_{mj} - \mu_{mn}$	1.1961	6
$\mu_{ij} - \mu_{in}$	1.5161	4.62
$\mu_{ij} - \mu_{mj}$	1.0308	10.8
$\mu_{ij} - \mu_{mn}$	1.6266	5.88

TABLE 25.8

SAS-Mixed Code to Provide the REML Analysis of the Nitrogen and Irrigation Example

```
PROC MIXED CL COVTEST METHOD=REML;
  CLASS BLK N IRR;
  MODEL Yield= N|IRR/DDFM=KR;
  RANDOM BLK BLK*N IRR*BLK;
```

Covariance Parameter Estimates

Covariance Parameter	Estimate	Standard Error	Z-Value	Pr Z	α	Lower	Upper
BLK	4.8056	5.8023	0.83	0.2038	0.05	1.1037	822.35
BLK \times N	0.6944	0.9127	0.76	0.2234	0.05	0.1478	286.67
BLK \times IRR	3.1667	2.9876	1.06	0.1446	0.05	0.9021	88.5030
Residual	1.4306	0.8259	1.73	0.0416	0.05	0.5940	6.9369

TABLE 25.9

Fixed Effects Tests for the REML Analysis for the Nitrogen and Irrigation Example

Type III Tests of Fixed Effects

Effect	Num df	Den df	F-Value	Pr > F
N	2	6	60.13	0.0001
IRR	1	3	52.18	0.0055
N \times IRR	2	6	33.12	0.0006

components are all greater than zero. The SAS-Mixed code with estimate statements for providing estimates to compare the levels of nitrogen, the levels of irrigation, and the other four types of comparisons in Table 25.7 are given in Table 25.10. Again, there are six different types of comparisons with six different estimated standard errors.

One of the inference problems with designs that involve more than one size of experimental unit is that various comparisons of the cell means involve different estimated standard errors. It is not possible to summarize these type of data by providing a single LSD value, as was the case for completely randomized and randomized complete block design structures.

TABLE 25.10

Estimate Statements Used with SAS-Mixed REML Analysis to Evaluate the Six Types of Comparisons

```
ESTIMATE 'N1 - N2' N 1 -1 0;
ESTIMATE 'N1 - N3' N 1 0 -1;
ESTIMATE 'N2 - N3' N 0 1 -1;
ESTIMATE 'IRR1 - IRR2' IRR 1 -1;
ESTIMATE 'M11 - M12 - M21 + M22' N*IRR 1 -1 -1 1 0 0;
ESTIMATE 'M11 - M12' IRR 1 -1 N*IRR 1 -1 0 0 0 0;
ESTIMATE 'M11 - M21' N 1 -1 0 N*IRR 1 0 -1 0 0 0;
ESTIMATE 'M11 - M22' N 1 -1 0 IRR 1 -1 N*IRR 1 0 0 -1 0 0;
```

Estimates

Label	Estimate	Standard Error	df	t-Value	Pr > t
N1 - N2	-3.5000	0.8396	6	-4.17	0.0059
N1 - N3	-9.1250	0.8396	6	-10.87	<0.0001
N2 - N3	-5.6250	0.8396	6	-6.70	0.0005
IRR1 - IRR2	-9.7500	1.3497	3	-7.22	0.0055
M11 - M12 - M21 + M22	-2.0000	1.1961	6	-1.67	0.1455
M11 - M12	-13.5000	1.5161	4.62	-8.90	0.0004
M11 - M21	-4.5000	1.0308	10.8	-4.37	0.0012
M11 - M22	-16.0000	1.6266	5.88	-9.84	<0.0001

25.4 Example: Strip-Plot with Split-Plot 1

The design in Section 25.3 can be extended to the case where within each of the cells there are three varieties of wheat, as shown in Figure 25.5. The randomization process is not displayed in Figure 25.5, but the levels of irrigation were randomly assigned to the columns of a rectangle, the levels of nitrogen were randomly assigned to the rows of a rectangle, and the varieties were randomly assigned to the three plots in a cell created by the intersection of a row and column in each rectangle. If one averages over the levels of variety, then the nitrogen by fertilizer analysis is identical to that in Table 25.3 or Table 25.6. The levels of variety are subplots within each cell and any comparisons of varieties are within-cell or between-subplot comparisons. A model that can be used to describe this data is

$$y_{ijkl} = \mu_{ijk} + b_l + r_{il} + c_{jl} + d_{ijl} + \varepsilon_{ijkl}, \quad i = 1, 2, \quad j = 1, 2, 3, \quad k = 1, 2, 3, \quad l = 1, 2, 3, 4$$

where μ_{ijk} denotes the mean of the i th level of irrigation with the j th level of nitrogen and k th variety, b_l denotes the effect of the l th block (rectangle) assumed to be distributed $b_l \sim i.i.d. N(0, \sigma^2_{\text{block}})$, r_{il} denotes the effect of the i th row within the l th block assumed to be distributed $r_{il} \sim i.i.d. N(0, \sigma^2_{\text{row}})$, c_{jl} denotes the effect of the j th column within the l th block assumed to be distributed $c_{jl} \sim i.i.d. N(0, \sigma^2_{\text{column}})$, d_{ijl} denotes the random effect associated with the ijl th cell within the l th block assumed to be distributed as $d_{ijl} \sim i.i.d. N(0, \sigma^2_{\text{cell}})$, and ε_{ijkl} denotes the subplot random effect associated with the k th variety assigned to the ijl th cell within the

Irrigation levels			
		I ₁	I ₂
Nitrogen levels	N ₁	V ₁	V ₁
		V ₂	V ₂
		V ₃	V ₃
	N ₂	V ₁	V ₁
		V ₂	V ₂
		V ₃	V ₃
	N ₃	V ₁	V ₁
		V ₂	V ₂
		V ₃	V ₃

FIGURE 25.5 One of the blocks of the field layout with nitrogen and irrigation as strip-plot and varieties as split-plot.

*l*th block assumed to be distributed as $\varepsilon_{ijkl} \sim i.i.d. N(0, \sigma_{\text{subplot}}^2)$. The analysis of variance for this model is given in Table 25.11 and the SAS-Mixed code that can be used to fit the model is given in Table 25.12. The row error term is *block* \times *nit*, the column error term is *block* \times *irr* and the cell error term is *block* \times *nit* \times *irr*. The residual is the subplot error term and would be computed as the *variety* by *block* interaction pooled across the levels of *nit* and *irr*.

TABLE 25.11

Analysis of Variance Table for the Strip-Plot with Split-Plot Design Structure

Source	df	EMS
Block	3	$\sigma_{\text{subplot}}^2 + 3\sigma_{\text{cell}}^2 + 6\sigma_{\text{row}}^2 + 9\sigma_{\text{column}}^2 + 18\sigma_{\text{block}}^2$
Irrigation	1	$\sigma_{\text{subplot}}^2 + 3\sigma_{\text{cell}}^2 + 6\sigma_{\text{row}}^2 + \varphi^2(I)$
Error(row)	3	$\sigma_{\text{subplot}}^2 + 3\sigma_{\text{cell}}^2 + 6\sigma_{\text{row}}^2$
Nitrogen	2	$\sigma_{\text{subplot}}^2 + 3\sigma_{\text{cell}}^2 + 9\sigma_{\text{column}}^2 + \varphi^2(N)$
Error(column)	6	$\sigma_{\text{subplot}}^2 + 3\sigma_{\text{cell}}^2 + 9\sigma_{\text{column}}^2$
Nitrogen \times irrigation	2	$\sigma_{\text{subplot}}^2 + 3\sigma_{\text{cell}}^2 + \varphi^2(N \times I)$
Error(cell)	6	$\sigma_{\text{subplot}}^2 + 3\sigma_{\text{cell}}^2$
Variety	2	$\sigma_{\text{subplot}}^2 + \varphi^2(V)$
Irrigation \times variety	2	$\sigma_{\text{subplot}}^2 + \varphi^2(I \times V)$
Nitrogen \times variety	4	$\sigma_{\text{subplot}}^2 + \varphi^2(N \times V)$
Nit \times irr \times variety	4	$\sigma_{\text{subplot}}^2 + \varphi^2(N \times I \times V)$
Error(subplot)	36	$\sigma_{\text{subplot}}^2$

TABLE 25.12

SAS-Mixed Code to Fit the Model to the Strip-Split-Plot Data Set in Section 25.4

```
Proc Mixed CL Covtest;
  Class block nit irr var;
  Model Yield=nit|irr|var/ddfm=KR;
  Random block block*nit irr*block irr*nit*block;
  Lsmeans nit|irr|var/diff;
```

25.5 Example: Strip-Plot with Split-Plot 2

Another way (different from that in Section 25.4) that the levels of variety can be included in the study are displayed in Figure 25.6. The levels of variety are striped across both levels of irrigation within each level of nitrogen. The levels of variety are subplots for the levels of nitrogen, but the levels of variety and the levels of nitrogen are strip-plots with the levels of irrigation. The irrigation by nitrogen analysis (obtained by summing over the varieties within each cell) is identical to that in Section 25.3. The nitrogen by variety part of the design can be obtained by summing over the levels of irrigation within each variety by nitrogen combination. The resulting design is a split-plot where the levels of nitrogen form the whole-plot factor and the levels of variety form the subplot factor. The split-plot analysis of variance is displayed in Table 25.13. The whole plot error is identical to the row error and is computed as the *block* \times *nitrogen* interaction. The subplot or subrow or 1/3 row error is computed as the *block* \times *variety* interaction pooled across the levels of nitrogen. Next consider the data from nitrogen level 1 only. In this case, the resulting design is a strip-plot involving the levels of irrigation and varieties. The corresponding strip-plot analysis of variance table is in Table 25.14. The error term of interest is the subcell error for

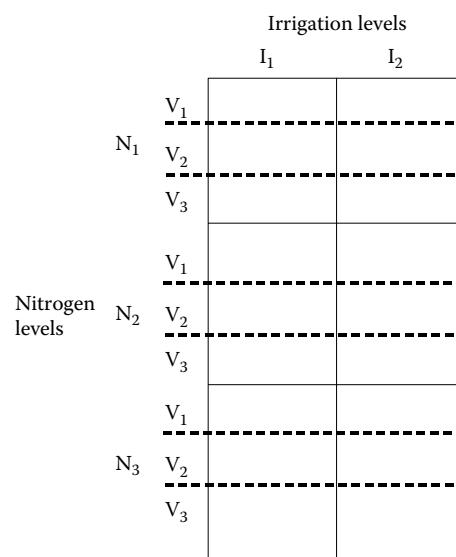


FIGURE 25.6 One of the blocks of the field layout with nitrogen and irrigation strip-plot, varieties as strip-plots for levels of nitrogen and varieties as strip-plot with levels of irrigation.

TABLE 25.13

Split-Plot Part of the Analysis for the Examples
in Section 25.5

Source	<i>df</i>
<i>Block</i>	3
<i>Nitrogen</i>	2
<i>Error(row) = block × nitrogen</i>	6
<i>Variety</i>	2
<i>Nitrogen × variety</i>	4
<i>Error(subrow) = block × variety(nitrogen)</i>	18

TABLE 25.14

Irrigation and Variety Data for the First Level of
Nitrogen in Section 25.5

Source	<i>df</i>
<i>Block</i>	3
<i>Irrigation</i>	1
<i>Error(column) = block × irrigation</i>	3
<i>Variety</i>	2
<i>Error(subrow) = block × variety</i>	6
<i>Irrigation × variety</i>	2
<i>Error(subcell) = block × variety × irrigation</i>	6

the intersection of variety and irrigation, which is computed as $block \times irrigation \times variety$ interaction with six degrees of freedom. These errors are pooled across the levels of nitrogen, yielding 18 degrees of freedom. The $nitrogen \times irrigation \times variety$ interaction is also a subcell comparison. Thus the $block \times nitrogen \times irrigation \times variety$ interaction provides an additional 12 degrees of freedom for subcell error yielding a subcell error term with 30 degrees of freedom. The analysis of variance table in Table 25.15 displays the sources of variation and the respective degrees of freedom.

25.6 Strip-Plot with Split-Plot 3

The graphic in Figure 25.7 shows that three varieties are assigned to the columns of the rectangle, the combinations of herbicide and nitrogen are assigned to the rows of the rectangle and the levels of seeding rate are assigned to the subplots within the cell. Assume the design consists of four such blocks. The varieties form a strip plot with the combination of levels of nitrogen and herbicide and the seeding rates are the subplots within the cells. The two-way treatment structure for the rows is the unique feature of this design. This structure points out that any type of treatment structure can be associated with any of the features of the design structure. The analysis of variance table is in Table 25.16 where the row error is computed by pooling the $block \times herbicide$ interaction, the $block \times nitrogen$ interaction, and the $block \times herbicide \times nitrogen$ interaction, providing nine degrees of freedom. The column error term is computed by the $block \times variety$ interaction. The cell

TABLE 25.15

Analysis of Variance Table for the Strip-Split-Plot Design in Section 25.5

Source	df
Block	3
Irrigation	1
Error(column)	3
Nitrogen	2
Error(row)	6
Nitrogen \times irrigation	2
Error(cell)	6
Variety	2
Variety \times nitrogen	4
Error(subrow)	18
Irrigation \times variety	2
Irrigation \times nitrogen \times variety	4
Error(subcell)	30

error term is obtained by pooling the *block \times variety \times herbicide*, *block \times variety \times nitrogen* and the *block \times variety \times herbicide \times nitrogen* interactions, providing 18 degrees of freedom for the cell error. The subplot error is obtained by computing the *block \times seeding rate* interaction for each combination of *herbicide*, *nitrogen*, and *variety* and then pooling these interactions across the levels of *herbicide*, *nitrogen*, and *variety*, providing 72 degrees of freedom for the residual error.

		Variety		
		1	2	3
Herbicide and nitrogen rates		H1N1	S1 S2 S3	S1 S2 S3
		H1N2	S1 S2 S3	S1 S2 S3
		H2N1	S1 S2 S3	S1 S2 S3
Seeding rates, S1, S2, and S3		H2N2	S1 S2 S3	S1 S2 S3

FIGURE 25.7 Graphic of One of the blocks of the field levels of herbicide and nitrogen assigned to rows, levels of varieties assigned to columns, and seeding rates assigned to the cells.

TABLE 25.16

Analysis of Variance for the Strip-Split-Plot Design of Section 25.5

Source	df
Block	3
H	1
N	1
$N \times H$	1
Error(row)	9
V	2
Error(column)	6
$V \times H$	2
$V \times N$	2
$V \times H \times N$	2
Error(cell)	18
S	2
$H \times S$	2
$N \times S$	2
$N \times H \times S$	2
$V \times S$	4
$V \times H \times S$	4
$V \times N \times S$	4
$V \times H \times N \times S$	4
Error(subplot)	72

25.7 Split-Plot with Strip-Plot 4

This experiment involves studying the effect of two varieties, two seeding rates, two herbicides and two nitrogen rates on the yield of corn. The graphic in Figure 25.8 shows how one of four blocks of experimental units was treated. The levels of variety and seeding rate are assigned to the four big squares within each block. Within each square, the levels of herbicide and nitrogen form a strip-plot. The strip-plot part is a subplot of the big squares. If one ignores the levels of nitrogen and herbicide, the design for varieties and seeding rates is a two-way treatment structure in a randomized complete block design structure. The error term is computed as the treatment structure by block interaction and which is obtained by the pooling the $block \times variety$, $block \times seeding\ rate$ and the $block \times variety \times seeding\ rate$ interactions, providing a total of nine degrees of freedom. The big block error term can be obtained using $variety \times seeding \times block$ in SAS-Mixed, and the procedure will automatically do the pooling for you. The analysis of variance table for the big block part of the analysis is given in Table 25.17. For variety 1 and seeding rate 1, the levels of herbicide and nitrogen provide a strip-plot design and the analysis of variance are summarized in Table 25.18. The error terms are pooled across the four combinations of variety and seeding rate to provide 12 degrees of freedom for $Error(row/block)$ (error of rows within a block), $Error(column/block)$ (error for columns within a block), and $Error(cell/block)$ (error for cells within a block). The SAS-Mixed code and a complete analysis of variance table are displayed in Table 25.19. The

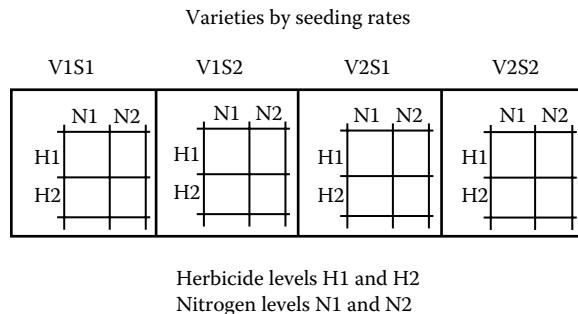


FIGURE 25.8 One of the blocks of the field layout with varieties and seeding rates as whole-plots and herbicide and nitrogen levels in a strip-plot as subplots.

TABLE 25.17

Big Block Analysis of Variance for Example in Section 25.7

Source	df
Block	3
Variety	1
Seeding rate	1
Variety \times seeding rate	1
Error(big block)	9

TABLE 25.18

Strip-Plot Analysis of Variance for the Herbicide and Nitrogen Data from Variety 1 and Seeding Rate 1 for the Example in Section 25.7

Source	df
Block	3
Herbicide	1
Error(row/block)	3
Nitrogen	1
Error(column/block)	3
Nitrogen \times herbicide	1
Error(cell/block)	3

error terms are computed as follows: $Error(\text{big squares}) = \text{Variety} \times \text{Seeding} \times \text{Block}$, $Error(\text{row}/\text{block}) = \text{Herbicide} \times \text{Block}(\text{Variety Seeding})$, $Error(\text{column}/\text{block}) = \text{Nitrogen} \times \text{Block}(\text{Variety Seeding})$, and $Error(\text{cell}/\text{block}) = \text{Nitrogen} \times \text{Herbicide} \times \text{Block}(\text{Variety Seeding})$.

25.8 Strip-Strip-Plot Design with Analysis via JMP7

The strip-strip-plot design structure occurs when the experimental units are arranged into three-dimensional rectangles with rows, columns and tiers. An experiment from the

TABLE 25.19

SAS-Mixed Code and Analysis of Variance for the Strip-Plot in Split-Plot Example of Section 25.7

```
Proc Mixed covtest CL;
Class Block Variety SeedingR Herbicide Nitrogen;
Model yield=Variety|SeedingR|Herbicide|Nitrogen/
      ddfm=KR;
Random Variety*SeedingR*Block Herbicide*Block
      (Variety SeedingR)
Nitrogen*Block(Variety SeedingR);
```

Source	df
<i>Block</i>	3
<i>Variety</i>	1
<i>Seeding rate</i>	1
<i>Variety × seeding rate</i>	1
<i>Error(big square)</i>	9
<i>Herbicide</i>	1
<i>Herbicide × variety</i>	1
<i>Herbicide × seeding rate</i>	1
<i>Herbicide × variety × seeding rate</i>	1
<i>Error(row/block)</i>	12
<i>Nitrogen</i>	1
<i>Nitrogen × variety</i>	1
<i>Nitrogen × seeding rate</i>	1
<i>Nitrogen × variety × seeding rate</i>	1
<i>Error(column/block)</i>	12
<i>Herbicide × nitrogen</i>	1
<i>Herbicide × nitrogen × variety</i>	1
<i>Herbicide × nitrogen × seeding rate</i>	1
<i>Herbicide × nitrogen × variety × seeding rate</i>	1
<i>Error(cell/block)</i>	12

semi-conductor industry is used to demonstrate the strip-strip-plot design structure. The experiment consists of evaluating three factors where each factor occurs at its own step in the process; that is, the experiment involves three steps. The first step is to add a layer of oxide to the surface of a silicon wafer (20 cm in diameter, 0.2 cm inches thick). The oxide layer is applied to the surface by putting the wafer into a furnace set to a specific temperature. Two levels of temperature were studied. The second step involves polishing or cleaning the surface of the wafer to smooth out any bumps that may have occurred, and two cleaning methods were evaluated. Finally, the third step involves washing off remaining surface particles that could cause interference with the circuitry being built. Two washing methods were included in the experiment. At each step four wafers are subjected together to the levels of the corresponding factor. The schematic in Figure 25.9 shows how eight wafers are moved through the three steps and how the wafers are grouped during each step. Figure 25.10 is a three-dimensional display showing that the levels of clean are assigned to rows, the levels of wash are assigned to columns and the levels of temperature are assigned to tiers. The study was repeated four times, generating four blocks. The experimental units for the levels of clean are the rows within a block and the row error term is computed as the

clean \times *block* interaction. The experimental units for the levels of wash are the columns within a block and the column error term is computed as the *wash* \times *block* interaction. If one sums over the levels of temperature, the resulting structure is a strip-plot design with clean and wash as the two factors. The experimental units for the clean by wash interaction are the two wafers from the two temperatures and the error for the two temperature wafers is computed as the *clean* \times *wash* \times *block* interaction. The experimental units for the levels of temperature are the tiers of the three-dimensional rectangle. The tier error term is computed as the *temperature* \times *block* interaction. If one sums over the levels of clean, the result is a strip-plot design with temperature and wash. The experimental units for the temperature by wash interaction are the two wafers from the two types of clean, which is computed as the *temperature* \times *wash* \times *block* interaction. If one sums over the levels of wash, the resulting design is a strip-plot with temperature and clean as the two factors. The experimental units

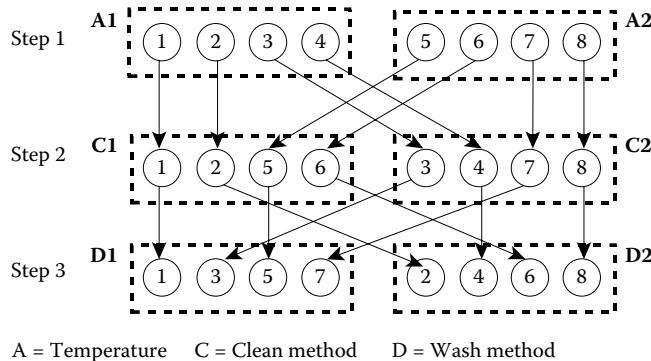


FIGURE 25.9 Three step process of applying the levels of temperature, cleaning, and washing.

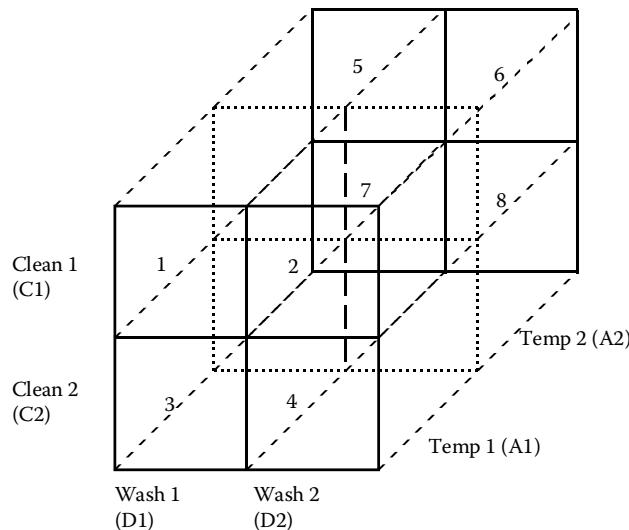


FIGURE 25.10 Three-dimensional rectangular showing the levels of clean are assigned to rows, the levels of wash are assigned to columns and the levels of temperature are assigned to tiers. The numbers correspond to the wafer numbers in Figure 25.9.

for the temperature by clean interaction are the two wafers from the two levels of wash and the error for the two wash wafers is computed as the $temperature \times clean \times block$ interaction. The experimental units for the temperature by clean by wash interaction are the individual wafers and the error for the wafer is computed as the $temperature \times clean \times wash \times block$ interaction. The response being measured is the average thickness of the resulting layer of oxide measured at nine locations on each wafer (in Angstroms) after the three steps. A model that can be used to describe the thickness data is

$$y_{ijkl} = \mu_{ijk} + b_l + r_{il} + c_{jl} + t_{kl} + w_{ijl}^{2W} + w_{ikl}^{2C} + w_{jkl}^{2T} + \varepsilon_{ijkl}, \quad i = 1, 2, \quad j = 1, 2, \quad k = 1, 2, \quad l = 1, 2, 3, 4$$

where μ_{ijk} denotes the mean thickness for level i of temperature, level j of clean and level k of wash and

$$\begin{aligned} b_l &\sim i.i.d. N(0, \sigma_{block}^2), \quad r_{il} \sim i.i.d. N(0, \sigma_{row}^2), \quad c_{jl} \sim i.i.d. N(0, \sigma_{column}^2) \\ t_{kl} &\sim i.i.d. N(0, \sigma_{tier}^2), \quad w_{ijl}^{2W} \sim i.i.d. N(0, \sigma_{2washwafer}^2) \\ w_{ikl}^{2C} &\sim i.i.d. N(0, \sigma_{2cleanwafer}^2), \quad w_{jkl}^{2T} \sim i.i.d. N(0, \sigma_{2tempwafer}^2) \end{aligned}$$

and

$$\varepsilon_{ijkl} \sim i.i.d. N(0, \sigma_{wafer}^2)$$

The analysis of variance is given in Table 25.20 where w , $2Tw$, $2Cw$, and $2Ww$ denote wafer, two temperature wafers, two clean wafers, and two wash wafers, respectively. Each factor effect has its own set of experimental units and corresponding error term.

TABLE 25.20

Analysis of Variance for the Strip-Strip-Plot Design of Section 25.8

Source	df	EMS
Block	3	$\sigma_w^2 + 2\sigma_{CTw}^2 + 2\sigma_{2Ww}^2 + 2\sigma_{2Tw}^2 + 4\sigma_{row}^2 + 4\sigma_{col}^2 + 4\sigma_{tier}^2 + 8\sigma_{blk}^2$
Temperature	1	$\sigma_w^2 + 2\sigma_{CTw}^2 + 2\sigma_{2Ww}^2 + 4\sigma_{tier}^2 + \varphi^2(T)$
Error(tier)	3	$\sigma_w^2 + 2\sigma_{CTw}^2 + 2\sigma_{2Ww}^2 + 4\sigma_{tier}^2$
Clean	1	$\sigma_w^2 + 2\sigma_{2Tw}^2 + 2\sigma_{2Ww}^2 + 4\sigma_{row}^2 + \varphi^2(C)$
Error(row)	3	$\sigma_w^2 + 2\sigma_{2Tw}^2 + 2\sigma_{2Ww}^2 + 4\sigma_{row}^2$
Wash	1	$\sigma_w^2 + 2\sigma_{2Tw}^2 + 2\sigma_{2Cw}^2 + 4\sigma_{col}^2 + \varphi^2(W)$
Error(column)	3	$\sigma_w^2 + 2\sigma_{2Tw}^2 + 2\sigma_{2Cw}^2 + 4\sigma_{col}^2$
Temperature \times wash	1	$\sigma_w^2 + 2\sigma_{2Cw}^2 + \varphi^2(T \times W)$
Error(two clean wafers)	3	$\sigma_w^2 + 2\sigma_{2Cw}^2$
Temperature \times clean	1	$\sigma_w^2 + 2\sigma_{2Ww}^2 + \varphi^2(T \times C)$
Error(two wash wafers)	3	$\sigma_w^2 + \sigma_{2Ww}^2$
Wash \times clean	1	$\sigma_w^2 + 2\sigma_{2Tw}^2 + \varphi^2(C \times W)$
Error(two temperature wafers)	3	$\sigma_w^2 + \sigma_{2Tw}^2$
Temperature \times clean \times wash	1	$\sigma_w^2 + \varphi^2(T \times C \times W)$
Error(wafer)	3	σ_w^2

Figure 25.11 contains the data set and Figure 25.12 is the JMP model specification screen. Each of the interactions involving *block* has been declared to be a random effect by using the Attributes button. The REML estimates of the variance components and the type III tests for the fixed effects are displayed in Figure 25.13. There is a significant *temperature* \times *clean* \times *wash* interaction and those cell means need to be evaluated. Least squares means for the three-way interaction and letters denoting significant differences are given in Figure 25.14. Another way to view the pairwise comparisons of the three-way interaction means is to plot all pairwise differences ranked from largest to smallest with confidence intervals about the

		block	Temp	Clean	Wash	thickness
	1	1	1	1	1	79.14
	2	1	1	1	2	94.12
	3	1	1	2	1	91.42
	4	1	1	2	2	115.33
	5	1	2	1	1	86.52
	6	1	2	1	2	113.49
	7	1	2	2	1	104.94
	8	1	2	2	2	128.89
	9	2	1	1	1	101.82
	10	2	1	1	2	104.70
	11	2	1	2	1	115.01
	12	2	1	2	2	129.77
	13	2	2	1	1	103.37
	14	2	2	1	2	112.21
	15	2	2	2	1	119.10
	16	2	2	2	2	132.85
	17	3	1	1	1	100.13
	18	3	1	1	2	104.97
	19	3	1	2	1	108.83
	20	3	1	2	2	124.41
	21	3	2	1	1	104.48
	22	3	2	1	2	116.58
All rows	32	23	3	2	2	116.24
Selected	1	24	3	2	2	122.74
Excluded	0	25	4	1	1	93.95
Hidden	0	26	4	1	1	93.73
Labelled	0	27	4	1	2	105.29
		28	4	1	2	120.06
		29	4	2	1	101.00
		30	4	2	1	109.55
		31	4	2	2	111.90
		32	4	2	2	120.09

FIGURE 25.11 Data set for three-step process for example in Section 25.8.

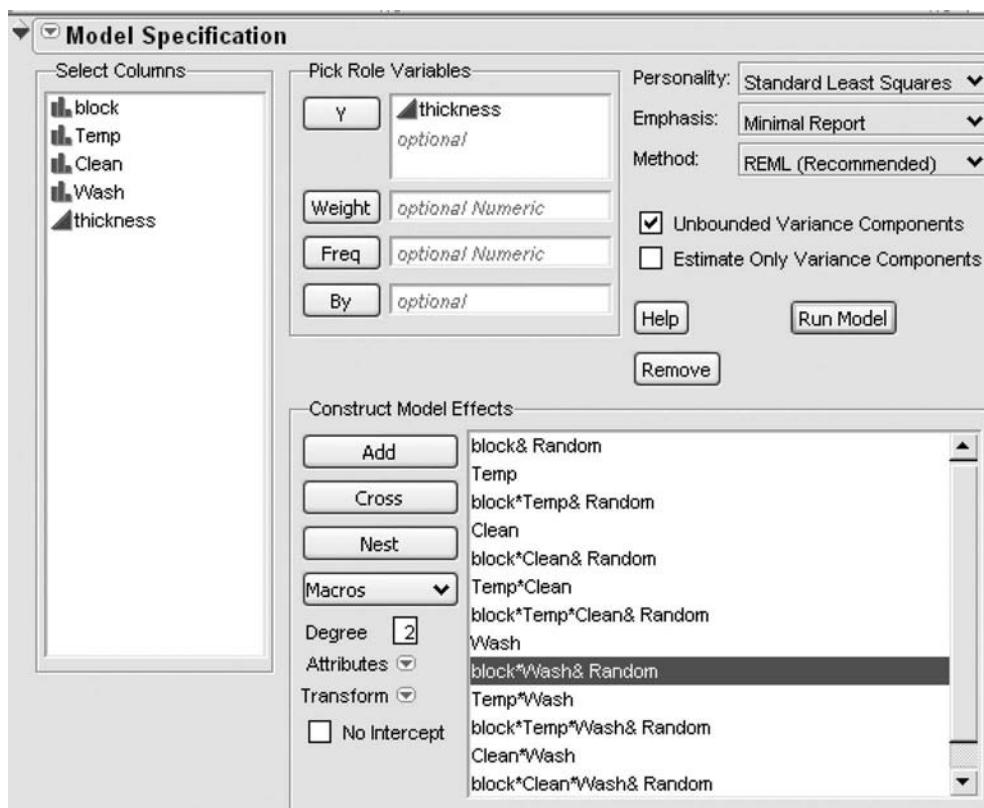


FIGURE 25.12 Fit model screen for three-step process of Section 25.8.

differences. Figure 24.15 is the display of the pairwise differences. The widths of the confidence intervals are not all equal as the variance of the difference depends on the type of comparison. The number of different variances is much larger for the strip-strip-plot than the number of different variances for the strip-plot displayed in Table 25.4.

25.9 Concluding Remarks

The strip-plot design is one of the basic design structures and the examples discussed in this chapter illustrate some of the complexities that can occur. Several examples are included so that the reader can develop a method for obtaining an appropriate analysis for complex designs. The strip-plot design structure was used to show how several sizes of experimental units can occur simply by the method of assigning the levels of the factors to the experimental units. The basic strip-plot was analyzed to demonstrate the complexity in computing the estimated standard errors for making various comparisons among the means. Obviously, as the design becomes more complex, there are many different standard errors that need to be evaluated, as demonstrated with the strip-plot and split-plot combination designs and the strip-strip-plot design. Sample size and power computations

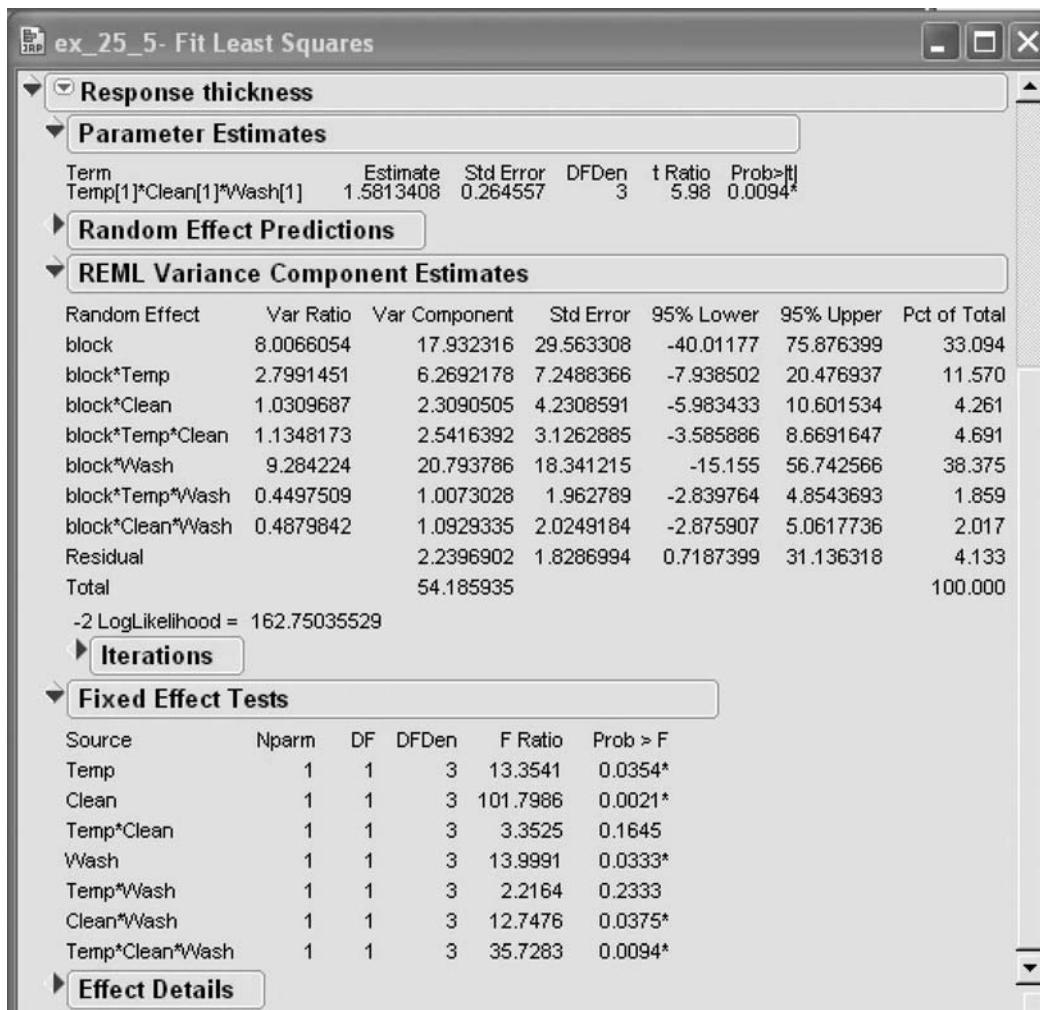


FIGURE 25.13 Estimates of the variance components and tests of the fixed effects for the three-step process.

Level	Least Sq Mean
2,2,2 A	126.14382
1,2,2 A B	122.39327
2,2,1 B C	113.04267
2,1,2 C D	112.95843
1,2,1 D E	105.13807
1,1,2 E F	99.37892
2,1,1 E F	98.84299
1,1,1 F	93.76016

Levels not connected by same letter are significantly different.

FIGURE 25.14 Three-way least square means with lines for the three-step process.

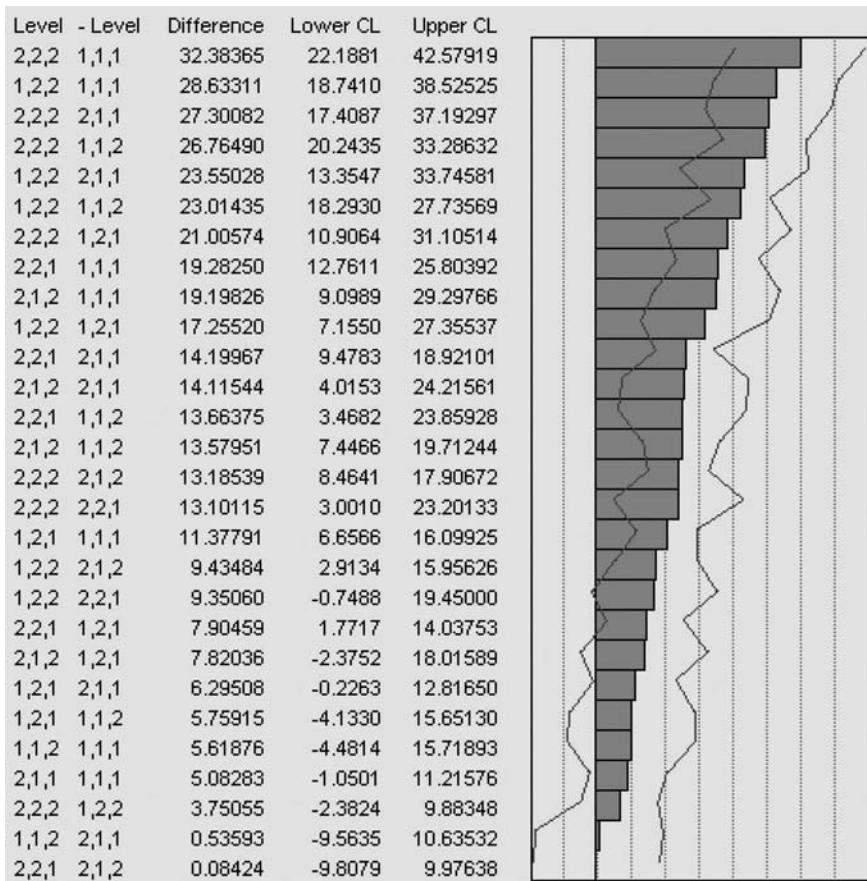


FIGURE 25.15 Graphical representation of the differences of the three-way least square means for the three-step process.

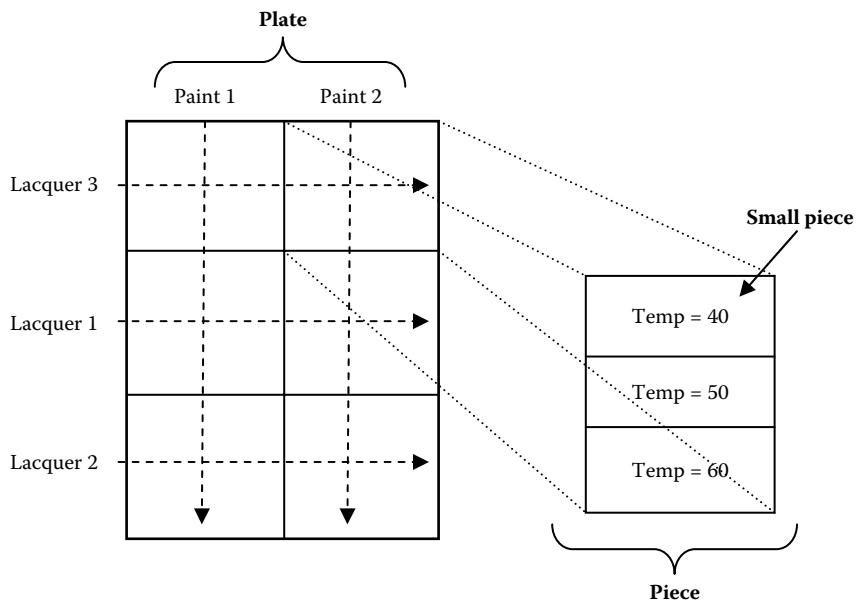
can be carried out using the method described in Chapter 24 and by using the estimated standard error(s) of the comparison(s) of interest.

25.10 Exercises

- 25.1 A cupcake baking experiment was carried out similar to the one described in Chapter 5 with the graphical representation in Figure 5.9. In this study there are four temperatures, four recipes, and five days. The four recipes are made up of two levels of fat and two levels of fiber. The data are in the following table where the response is the volume (cm^3) of the cupcake on day 1, day 2, ..., day 5.
- 1) Identify the experimental units and their design and treatment structures
 - 2) Write down a model to describe the data set.
 - 3) Fit the model to the data and carry out the necessary mean comparisons.

Temperature	Fat	Fiber	Day 1	Day 2	Day 3	Day 4	Day 5
325	H	H	37.6	33.0	31.8	36.4	31.1
325	H	L	35.8	35.8	32.7	36.4	35.1
325	L	H	30.7	28.2	27.1	32.1	25.7
325	L	L	36.1	26.8	29.0	32.6	28.5
340	H	H	44.1	40.1	32.2	38.9	41.4
340	H	L	42.2	40.2	35.9	39.8	44.8
340	L	H	37.6	35.2	31.3	34.4	34.4
340	L	L	43.6	33.3	32.3	37.2	41.6
360	H	H	45.8	39.8	36.3	42.5	41.6
360	H	L	42.2	45.1	42.6	40.9	45.8
360	L	H	35.8	40.6	32.4	38.4	32.8
360	L	L	44.2	35.1	39.1	40.3	38.9
400	H	H	48.0	43.6	34.8	41.2	39.8
400	H	L	45.5	47.1	44.7	42.7	48.2
400	L	H	39.6	42.7	36.6	41.8	35.8
400	L	L	45.9	36.9	38.3	41.7	38.6

- 4) Compute the standard errors for the six types of comparisons discussed in Section 25.2.
- 5) Determine the sample size (number of days) necessary do detect a difference of 3 units in the volumes of cupcakes made from two different recipes. Use type I and II error rates of 0.05 and 0.10 respectively.
- 6) Determine the sample size (number of days) necessary to detect a difference of 3 units in the volumes of two temperatures within the same recipe. Use type I and II error rates of 0.05 and 0.10 respectively.
- 7) Determine the power of the test for detecting a difference of 3 units in the volumes of two temperatures within the same recipe using 5 days of data. Use type I error rate of 0.05.
- 25.2 An engineer designing composite materials for aircraft designed an experiment to evaluate the effect of types of lacquer, paint, and temperature on the strength of a composite material. A plate of composite material was manufactured. The plate was divided into thirds row-wise and the three levels of lacquer were randomly assigned to the rows. The plate was divided in half column-wise and the two levels of paint were assigned to the columns. Finally each of four plates was cut into six pieces corresponding to the rows and columns. These six pieces were further cut into three pieces and the three levels of temperature were randomly assigned to the three small pieces within each of the larger pieces. The strength of each of the small pieces was measured by determining the amount of force required to break the piece while bending. The following is a diagram as to how one plate was treated.

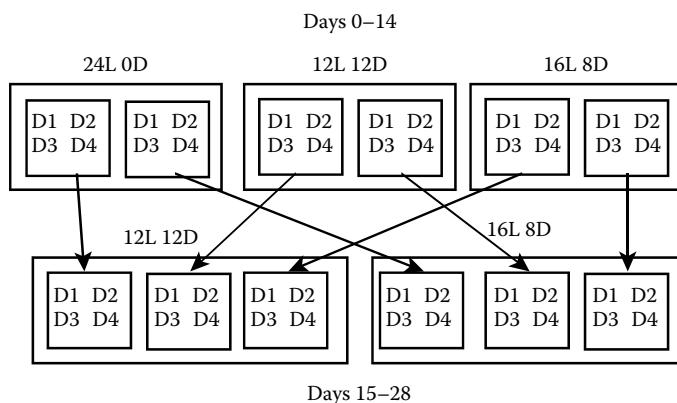


The data set follows.

LACQ	Paint	Temperature	Plate 1	Plate 2	Plate 3	Plate 4
1	1	40	143	180	186	201
1	1	50	145	173	198	191
1	1	60	146	187	197	202
1	2	40	201	192	249	234
1	2	50	209	187	243	240
1	2	60	207	185	253	227
2	1	40	180	175	175	223
2	1	50	182	175	186	238
2	1	60	185	182	179	236
2	2	40	254	290	300	308
2	2	50	267	286	303	295
2	2	60	287	293	306	299
3	1	40	199	196	226	256
3	1	50	219	212	228	260
3	1	60	216	210	238	267
3	2	40	301	309	324	348
3	2	50	310	322	319	359
3	2	60	313	333	343	371

- 1) Provide a description of the design structure by identifying each size of experimental unit and associated design structure.

- 2) Write out a model that can be used to describe this data.
- 3) Write out an analysis of variance table with the sources of variation and degrees of freedom. Identify the error terms for each size of experimental unit.
- 4) Fit the model to the data and make all necessary comparisons among the means.
- 5) Is there evidence of a linear or quadratic trend in temperature for any of the lacquer by paint combinations?
- 25.3 An animal scientist wanted to determine the effect of different types of light-dark cycles have on the ability of different diets to affect the growth of chickens from hatching to 28 days of age. There are four diets that are combinations of low- and high-oil corn with two sources of protein. During days 0–14 there are three light-dark cycles, 24 h light and 0 h dark, 12 h light and 12 h dark, and 16 h light and 8 h dark. During days 15–28 there are two light-dark cycles, 12 h light and 12 h dark, and 16 h light and 8 h dark. The experiment was repeated three times, providing three blocks and three replications of the treatment combinations. Twenty-four groups of five male chickens were formed. These 24 groups were divided into six sets of four cages and within a set the four diets were randomly assigned to the four cages. Two of the sets of four cages were randomly assigned to each of the three light-dark cycles for days 0 to 14. All cages of five chickens assigned to a light-dark cycle were put into one room; that is, there were eight cages (two sets of four cages) within a room at the same time, all subjected to the same light-dark cycle. At day 15, one of the cages from each of the first light-dark cycles was assigned to one of the two second light-dark cycles. The following graphic shows the flow of the pens through the two steps of light-dark cycles where the squares represent the sets of four pens and D1, D2, D3, and D4 represent the four diets.



The following table gives the data that was collected with the mean weight of the chickens in a pen. If the pen had four or five chickens at the end of the study, the pen was kept in the data set and the mean weight was recorded. If the pen

had three or fewer chickens at the end of the 28 days that data were not included in the data set. The experiment was repeated in three successive months, c_{ij} denotes the i th light-dark cycle in phase 1 and the j th light-dark cycle in the second phase.

Month	Oil	Protein	c11	c12	c21	c22	c31	c32
1	H	H	1.55	—	1.71	1.22	—	1.50
1	H	L	—	0.66	1.34	0.94	1.65	—
1	L	H	—	0.81	—	1.15	—	1.35
1	L	L	1.08	0.64	—	0.94	1.55	1.16
2	H	H	1.99	1.59	2.19	1.65	2.12	—
2	H	L	1.60	1.27	1.76	—	1.71	1.55
2	L	H	1.84	1.50	1.98	—	1.93	—
2	L	L	—	1.28	1.66	1.32	1.59	1.52
3	H	H	2.32	2.08	2.25	—	2.20	1.91
3	H	L	1.88	1.84	1.88	1.68	1.81	1.64
3	L	H	2.12	2.02	—	1.86	2.01	1.84
3	L	L	—	1.79	1.82	1.65	1.69	1.67

- 1) Identify the sizes of experimental units and then specify the corresponding design and treatment structures.
 - 2) Write out a model that can be used to describe this data.
 - 3) Write out an analysis of variance table with the sources of variation and degrees of freedom assuming there are no missing data. Identify the error terms for each size of experimental unit.
 - 4) Fit the model to the data and make all necessary comparisons among the means.
- 25.4 For the model in Section 25.4, determine the variances of the estimates for each of the following comparisons.
- 1) $\mu_{111} - \mu_{112}$
 - 2) $\mu_{111} - \mu_{121}$
 - 3) $\mu_{111} - \mu_{211}$
 - 4) $\bar{\mu}_{11\cdot} - \bar{\mu}_{12\cdot} - \bar{\mu}_{21\cdot} - \bar{\mu}_{22\cdot}$

26

Methods for Analyzing Repeated Measures Experiments

Like experiments using split-plot designs, experiments utilizing repeated measures designs have structures that involve more than one size of experimental unit. For example, a subject may be measured over time where time is one of the factors in the treatment structure of the experiment. By measuring the subject at several different times, the subject is essentially being “split” into parts (time intervals), and the response is measured on each part. The larger experimental unit is the subject or the collection of time intervals. The smaller unit is the interval of time during which the subject is exposed to a treatment or an interval just between time measurements.

Repeated measures designs differ from split-plot designs in that the levels of one or more factors cannot be randomly assigned to one or more of the sizes of experimental units in the experiment. In this case, the levels of time cannot be assigned at random to the time intervals, and thus analyzing a repeated measures experiment as though it was a split-plot experiment may not be valid. Because of this nonrandom assignment, the errors corresponding to the respective experimental units may have a covariance matrix that does not conform to the covariance matrix corresponding to experiments for which the usual split-plot analysis is valid. Analyzing a repeated measures experiment as though it was a split-plot experiment is often called a split-plot in time analysis.

In Section 26.1, the repeated measures models are described, and the assumptions necessary for the split-plot in time analysis of variance to be valid are given. Section 26.2 gives three examples that demonstrate the split-plot in time analysis of variance computations, including the computations of standard errors for making various comparisons between means. In Chapter 27, methods are presented for analyzing repeated measures experiments when the split-plot in time assumption is not satisfied. In addition, methods are presented that allow one to test whether or not the split-plot in time assumption is satisfied.

26.1 Model Specifications and Ideal Conditions

Repeated measures designs can be applied in numerous situations. Example 26.1, analyzed in the next section, investigates the effects of three drugs on heart rates, where each drug was administered to eight people. Each person's heart rate was then measured 5, 10, 15, and 20 min after administering the drug.

In the general model for a simple repeated measures experiment, n_i subjects are randomly assigned to treatment i , and each subject is measured at p time points. Table 26.1 illustrates the layout for a simple repeated measures experiment.

The larger size of experimental unit is the subject, and the smaller size experimental unit is the time interval when using the split-plot in time notation. A model that describes each measured response is similar to the split-plot model in a completely randomized design and is given by

$$y_{ijk} = \mu + \alpha_i + \delta_{ik} + \tau_j + (\alpha\tau)_{ij} + \varepsilon_{ijk} \quad (26.1)$$

or

$$y_{ijk} = \mu_{ij} + \delta_{ik} + \varepsilon_{ijk}$$

TABLE 26.1
Layout for a Simple Repeated Measures Experiment

TRT	Subject	TIME				
		1	2	3	...	p
1	1	—	—	—	...	—
	2	—	—	—	...	—

	n_1	—	—	—	...	—
	1	—	—	—	...	—
	2	—	—	—	...	—

t	n_2	—	—	—	...	—

	1	—	—	—	...	—
	2	—	—	—	...	—

	n_t	—	—	—	...	—

where $\mu + \alpha_i + \delta_{ik}$ is the subject part of the model and $\tau_j + (\alpha\tau)_{ij} + \varepsilon_{ijk}$ is the within subject (time interval) part of the model. The mean of treatment i at time j is

$$\mu_{ij} = \mu + \alpha_i + \tau_j + (\alpha\tau)_{ij}$$

The δ_{ik} represent the subject error components, and the ε_{ijk} represent within-subject (time interval) errors. The ideal conditions for a split-plot in time analysis are that

- 1) The δ_{ik} are independently and identically $N(0, \sigma_\delta^2)$.
- 2) The ε_{ijk} are independently and identically $N(0, \sigma_\varepsilon^2)$.
- 3) The δ_{ik} and the ε_{ijk} are all independent of one another.

Note that these are the same assumptions that were made for the analysis of a split-plot experiment in a completely randomized design (see Chapter 24). Such assumptions may not always be appropriate for a repeated measures design. However, the split-plot analysis is also a correct analysis under more general conditions. The more general conditions will be given in Chapter 27.

In Example 26.3, the attitudes among family members were studied in relation to whether the home was in a rural or an urban environment. The type of family required for the study consisted of three members, a son, a father, and a mother. Then 10 such families were randomly selected from an urban environment and seven such families were randomly selected from a rural environment. The attitude of each family member towards a moral issue was determined. Also of interest was whether the attitudes changed over time; hence, measurements were made on each person at the 0, 6, and 12 month time points. There are three sizes of experimental units in this design: the experimental unit corresponding to the treatment effect environment is the family; the experimental unit corresponding to the type of family member effect is the person; and the experimental unit for time is the six-month time interval. A model to describe the attitudes has three error components, one for each size of experimental unit in the experiment. One such model is

$$\begin{aligned} y_{ijkm} &= \mu + \alpha_i + f_{im} && \text{family experimental unit} \\ &+ \beta_j + (\alpha\beta)_{ij} + p_{ijm} && \text{person experimental unit} \\ &+ \tau_k + (\alpha\tau)_{ik} + (\beta\tau)_{jk} + (\alpha\beta\tau)_{ijk} + \varepsilon_{ijkm} && \text{time interval experimental unit} \end{aligned}$$

or

$$y_{ijkm} = \mu_{ijk} + f_{im} + p_{ijm} + \varepsilon_{ijkm} \quad (26.2)$$

In both of the above models $i = 1, 2; j = 1, 2, 3; k = 1, 2, 3; m = 1, 2, \dots, n_i$ where $n_1 = 10$ and $n_2 = 7$.

The error terms are f_{im} which represents family within environment error, p_{ijm} which represents person within family error, and ε_{ijkm} which represents the time within-person error. Note that this experiment has two sets of repeated measures, the three family members form one set of repeated measures as family member designation cannot be randomly assigned to family members, and time forms the second set of repeated measures. If both sets of repeated measures satisfy traditional split-plot type assumptions, then the ideal conditions on the error terms in the above models are

- 1) The f_{im} are independently identically distributed $N(0, \sigma_f^2)$.
- 2) The p_{ijm} are independently identically distributed $N(0, \sigma_p^2)$.
- 3) The ε_{ijkm} are independently identically distributed $N(0, \sigma_\varepsilon^2)$.
- 4) All the f_{im} , p_{ijm} , and ε_{ijkm} are independent of one another.

Chapter 27 will consider an analysis of the family member example that does not require that the ideal conditions be satisfied.

26.2 The Split-Plot in Time Analyses

The usual split-plot in time analyses of variance refers to the analyses discussed in Chapters 24 and 25 for split-plot and strip-plot designs. Such analyses are provided by most computer packages. If the ideal conditions are satisfied then these analyses provide valid F -tests. Three examples are used in this section to demonstrate analyses of repeated measures designs and to show how to determine estimates of interesting effects, their estimated standard errors (which are necessary to make various multiple comparisons), and provide methods to study contrasts between among various kinds of means. The examples used here are also used in Chapter 27 to demonstrate the techniques needed when the ideal conditions on the model error terms are not satisfied.

26.2.1 Example 26.1: Effect of Drugs on Heart Rate

An experiment involving t drugs was conducted to study each drug's effect on the heart rate of humans. After the drug was administered, each person's heart rate was measured every five minutes for a total of p times. At the start of the study n female human subjects were randomly assigned to each drug. A model similar to Equation 26.1 is used to describe the data. The model is

$$y_{ijk} = \mu + \alpha_i + \delta_{ik} + \tau_j + (\alpha\tau)_{ij} + \varepsilon_{ijk}, \quad i = 1, 2, \dots, t; \quad j = 1, 2, \dots, p; \quad k = 1, 2, \dots, n$$

The model has two error terms: δ_i represents a subject error component, and ε_{ijk} represents a time error component. The ideal conditions for a split-plot in time analysis is that

- 1) The δ_{ik} are independently and identically $N(0, \sigma_\delta^2)$.
- 2) The ε_{ijk} are independently and identically $N(0, \sigma_\varepsilon^2)$.
- 3) The δ_{ik} and the ε_{ijk} are all independent of one another.

Table 26.2 gives the split-plot in time analysis of variance table for Example 26.1.

In Table 26.2, Q_{Drug} , Q_{Time} , and $Q_{Drug \times Time}$ are noncentrality parameters measuring the *Drug* effect, *Time* effect, and the *Drug* \times *Time* interaction effect, respectively. These Q values are zero if and only if the corresponding effects are equal to zero.

To test $H_{01}: Q_{Drug} = 0$, one rejects H_{01} if

$$F = \frac{MS_{Drug}}{MS_{Error(Subject)}} > F_{\alpha, t-1, t(n-1)}$$

TABLE 26.2

Analysis of Variance Table for Example 26.1

Source of Variation	df	SS	EMS
Drug	$t - 1$	$np \sum_{i=1}^t (\bar{y}_{i..} - \bar{y}_{...})^2$	$\sigma_e^2 + p\sigma_\delta^2 + Q_{Drug}$
Error(Subject)	$t(n - 1)$	$p \sum_{i=1}^t \sum_{k=1}^n (\bar{y}_{i,k} - \bar{y}_{i..})^2$	$\sigma_e^2 + p\sigma_\delta^2$
Time	$p - 1$	$nd \sum_{j=1}^p (\bar{y}_{.,j} - \bar{y}_{...})^2$	$\sigma_e^2 + Q_{Time}$
Drug \times Time	$(t - 1)(p - 1)$	$n \sum_{i=1}^t \sum_{j=1}^p (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j} + \bar{y}_{...})^2$	$\sigma_e^2 + Q_{Drug \times Time}$
Error(Time)	$t(n - 1)(p - 1)$	$\sum_{i=1}^t \sum_{j=1}^p \sum_{k=1}^n (\bar{y}_{ijk} - \bar{y}_{ij.} - \bar{y}_{i.k} + \bar{y}_{...})^2$	σ_e^2

To test $H_{02}: Q_{Time} = 0$, one rejects H_{02} if

$$F = \frac{MSTime}{MSError(Time)} > F_{\alpha, p-1, t(p-1)(n-1)}$$

To test $H_{03}: Q_{Drug \times Time} = 0$, one rejects H_{03} if

$$F = \frac{MSDrug \times Time}{MSError(Time)} > F_{\alpha, (t-1)(p-1), t(p-1)(n-1)}$$

Similar to that which was done for a split-plot experiment, method of moment solutions for the two variance components are given by

$$\begin{aligned}\tilde{\sigma}_e^2 &= MSError(Time) \\ \tilde{\sigma}_\delta^2 &= \frac{MSError(Subject) - \tilde{\sigma}_e^2}{p}\end{aligned}$$

Then the estimates of the two variance components are taken as

$$\hat{\sigma}_e^2 = \tilde{\sigma}_e^2$$

and

$$\hat{\sigma}_\delta^2 = \begin{cases} \tilde{\sigma}_\delta^2 & \text{if } \tilde{\sigma}_\delta^2 > 0 \\ 0 & \text{if } \tilde{\sigma}_\delta^2 \leq 0 \end{cases}$$

When one finds significant effects in the data, one will generally want to compare various means with one another. If there is no *Drug* \times *Time* interaction, then one will often want to make comparisons between the *Drug* main effect means and the *Time* main effect means. If the interaction effect is significant, then one may wish to compare drugs with one another at each time point and/or times to one another for each drug. The estimators of μ_{ij} , $\bar{\mu}_{i\cdot}$, $\bar{\mu}_{\cdot j}$ and $\bar{\mu}_{\cdot \cdot}$ are $\hat{\mu}_{ij} = \bar{y}_{ij\cdot}$, $\hat{\mu}_{i\cdot} = \bar{y}_{i\cdot \cdot}$, $\hat{\mu}_{\cdot j} = \bar{y}_{\cdot j\cdot}$ and $\hat{\mu}_{\cdot \cdot} = \bar{y}_{\cdot \cdot \cdot}$, respectively. Since repeated measures designs involve more than one size of experimental unit and more than one error term, the variance of each comparison could involve different functions of the variances of the respective error terms and thus need to be determined. Table 26.3 gives the best estimates of various functions of the mean parameters in a repeated measures experiment along with their respective estimated standard errors.

Finally, the best estimate of $\mu_{ij} - \mu_{i'j} - \mu_{ij'} + \mu_{i'j'}$ is $\hat{\mu}_{ij} - \hat{\mu}_{i'j} - \hat{\mu}_{ij'} + \hat{\mu}_{i'j'}$ and its estimated standard error is $\sqrt{(4\hat{\sigma}_e^2/n)}$ when $i \neq i'$ and $j \neq j'$.

Inferences about each of the parameter functions above can be made using the ratio of a best estimate to its estimated standard error. Each ratio either has a *t*-distribution or can be approximated by a *t*-distribution. For those ratios whose estimated standard error involves $\hat{\sigma}_e$ only, the ratio has an exact *t*-distribution with $t(n-1)(p-1)$ degrees of freedom. For those ratios whose estimated standard error is a multiple of $\sqrt{(\hat{\sigma}_e^2 + p\hat{\sigma}_\delta^2)}$, the ratio also has an exact *t*-distribution but with $t(n-1)$ degrees of freedom. Finally, for those ratios whose estimated standard error is a multiple of $\sqrt{(\hat{\sigma}_e^2 + \hat{\sigma}_\delta^2)}$, the ratio has an approximate *t*-distribution whose degrees of freedom must be approximated by Satterthwaite's method.

In general, let $\phi = \sum_{i=1}^t \sum_{j=1}^p c_{ij}\mu_{ij}$ be any linear combination of the μ_{ij} , then the same linear combination of the $\hat{\mu}_{ij}$, $\hat{\phi} = \sum_{i=1}^t \sum_{j=1}^p c_{ij}\hat{\mu}_{ij}$, is its best estimate. The estimated variance of $\hat{\phi}$ is equal to $k[c_1\hat{\sigma}_1^2 + c_2\hat{\sigma}_2^2]$ for some k , c_1 , and c_2 where $v_i\hat{\sigma}_i^2/\sigma_i^2 \sim \text{independent } \chi^2(v_i)$ for $i = 1, 2$. Then $(\hat{\phi} - \phi) \widehat{s.e.}(\hat{\phi})$ is approximately $t(\hat{v})$ where

$$\hat{v} = \frac{(c_1\hat{\sigma}_1^2 + c_2\hat{\sigma}_2^2)^2}{(c_1^2\hat{\sigma}_1^2/v_1) + (c_2^2\hat{\sigma}_2^2/v_2)}$$

TABLE 26.3

Various Parameter Functions, Best Estimates, and Estimated Standard Errors

Parameter	Best Estimate	Estimated Standard Error
μ_{ij}	$\hat{\mu}_{ij}$	$\sqrt{(\hat{\sigma}_e^2 + \hat{\sigma}_\delta^2)/n}$
$\bar{\mu}_{i\cdot}$	$\hat{\mu}_{i\cdot}$	$\sqrt{(\hat{\sigma}_e^2 + p\hat{\sigma}_\delta^2)/pn}$
$\bar{\mu}_{\cdot j}$	$\hat{\mu}_{\cdot j}$	$\sqrt{(\hat{\sigma}_e^2 + \hat{\sigma}_\delta^2)/tn}$
$\bar{\mu}_{i\cdot} - \bar{\mu}_{i'\cdot}$	$\hat{\mu}_{i\cdot} - \hat{\mu}_{i'\cdot}$	$\sqrt{2(\hat{\sigma}_e^2 + p\hat{\sigma}_\delta^2)/pn}$
$\bar{\mu}_{\cdot j} - \bar{\mu}_{\cdot j'}$	$\hat{\mu}_{\cdot j} - \hat{\mu}_{\cdot j'}$	$\sqrt{2\hat{\sigma}_e^2/tn}$
$\mu_{ij} - \mu_{ij'}$	$\hat{\mu}_{ij} - \hat{\mu}_{ij'}$	$\sqrt{2\hat{\sigma}_e^2/n}$
$\mu_{ij} - \mu_{i'j}$	$\hat{\mu}_{ij} - \hat{\mu}_{i'j}$	$\sqrt{2(\hat{\sigma}_e^2 + \hat{\sigma}_\delta^2)/n}$

For the example being considered, $\hat{\sigma}_1^2$ is the $MSE_{error(Subject)}$ with $v_1 = t(n - 1)$ degrees of freedom and $\hat{\sigma}_2^2$ is the $MSE_{error(Time)}$ with $v_2 = t(n - 1)(p - 1)$ degrees of freedom. Furthermore, for those ratios whose estimated standard error is a multiple of $\hat{\sigma}_e^2 + \hat{\sigma}_\delta^2$ then $c_1 = 1/p$ and $c_2 = (p - 1)/p$.

Next consider a linear contrast of the drug main effect means, $\sum_{i=1}^t c_i \bar{\mu}_{i\cdot}$. The estimated standard error of the estimate of such a contrast is

$$\sqrt{\frac{(\hat{\sigma}_e^2 + p\hat{\sigma}_\delta^2) \sum_{i=1}^t c_i^2}{np}}$$

Thus

$$\frac{\sum_{i=1}^t c_i \hat{\mu}_{i\cdot} - \sum_{i=1}^t c_i \bar{\mu}_{i\cdot}}{\sqrt{\frac{(\hat{\sigma}_e^2 + p\hat{\sigma}_\delta^2) \sum_{i=1}^t c_i^2}{np}}} \sim t[t(n - 1)]$$

Consider a linear contrast in the time main effect means, $\sum_{j=1}^p d_j \bar{\mu}_{\cdot j}$. The estimated standard error of such a contrast is

$$\sqrt{\frac{\hat{\sigma}_e^2 \sum_{j=1}^p d_j^2}{nt}}$$

Thus

$$\frac{\sum_{j=1}^p d_j \hat{\mu}_{\cdot j} - \sum_{j=1}^p d_j \bar{\mu}_{\cdot j}}{\sqrt{\frac{\hat{\sigma}_e^2 \sum_{j=1}^p d_j^2}{nt}}} \sim t[t(n - 1)(p - 1)]$$

Such a contrast could be used to check for linear, quadratic, and similar trends over time.

If there is a $Drug \times Time$ interaction, then we need to compare the drugs at each time period and/or compare time periods for each drug. That is, one could consider contrasts in the μ_{ij} for each possibility for j , and/or contrasts in the μ_{ij} for each possibility for i . For the former, it can be shown that

$$\sqrt{\frac{\sum_{i=1}^t c_i \hat{\mu}_{ij} - \sum_{i=1}^t c_i \mu_{ij}}{\frac{(\hat{\sigma}_e^2 + \hat{\sigma}_\delta^2) \sum_{i=1}^t c_i^2}{n}}}$$

is approximately t with \hat{v} degrees of freedom for each j , and for the latter it can be shown that

$$\sqrt{\frac{\sum_{i=1}^t d_j \hat{\mu}_{ij} - \sum_{i=1}^t d_j \mu_{ij}}{\frac{\hat{\sigma}_e^2 \sum_{i=1}^t d_j^2}{n}}}$$

is exactly t with $t(n - 1)(p - 1)$ degrees of freedom for each i . These two results can be used to test hypotheses and set confidence intervals on within row or within column contrasts of the μ_{ij} .

The data in Table 26.4 are used to demonstrate the analyses described above. In this experiment there were three drugs, eight people per drug, and four time periods. The analysis of variance table is given in Table 26.5. There is a significant $Time \times Drug$ interaction; thus, we need to compare times with one another for each drug and drugs with one another at each time point.

TABLE 26.4
Heart Rate Data for Example 26.1

Person within Drug	Drug											
	AX23				BWW9				Control			
	T_1	T_2	T_3	T_4	T_1	T_2	T_3	T_4	T_1	T_2	T_3	T_4
1	72	86	81	77	85	86	83	80	69	73	72	74
2	78	83	88	81	82	86	80	84	66	62	67	73
3	71	82	81	75	71	78	70	75	84	90	88	87
4	72	83	83	69	83	88	79	81	80	81	77	72
5	66	79	77	66	86	85	76	76	72	72	69	70
6	74	83	84	77	85	82	83	80	65	62	65	61
7	62	73	78	70	79	83	80	81	75	69	69	68
8	69	75	76	70	83	84	78	81	71	70	65	63

Note: T_i denotes the i th time period.

TABLE 26.5
Analysis of Variance Table for Data Table 26.4

Source of Variation	df	SS	MS	F	EMS
Drug	2	1,333.00	666.5	5.99	$\sigma_e^2 + 4\sigma_\delta^2 + Q_{Drug}$
Error(Person)	21	2,337.91	111.33		$\sigma_1^2 = \sigma_e^2 + 4\sigma_\delta^2$
Time	3	289.61	96.54	12.96	$\sigma_e^2 + Q_{Time}$
Time \times Drug	6	527.42	87.90	11.80	$\sigma_e^2 + Q_{Drug \times Time}$
Error(Time)	63	469.22	7.45		$\sigma_2^2 = \sigma_e^2$

Note: Q denotes the respective noncentrality parameter.

To compare times with one another for each drug, the estimated standard error of the difference of the two means (see Table 26.3) is

$$\widehat{s.e.}(\hat{\mu}_{ij} - \hat{\mu}_{i'j}) = \sqrt{2\hat{\sigma}_e^2/n} = \sqrt{2(7.45)/8} = 1.365$$

and a 5% LSD value for comparing time means within each drug is

$$LSD_{0.05} = t_{0.025, 63} [\widehat{s.e.}(\hat{\mu}_{ij} - \hat{\mu}_{i'j})] = (2.00)(1.365) = 2.730$$

Comparisons of the time means within a drug are given in Table 26.6.

Since the levels of time are quantitative and equally spaced, orthogonal polynomials can be used to check for linear and quadratic trends in the response to each drug. The measure of the linear trend in time for the first drug is $\theta_{LT_1} = -3\mu_{11} - 1\mu_{12} + 1\mu_{13} + 3\mu_{14}$, its estimate is

$$\hat{\theta}_{LT_1} = -3(70.50) - 1(80.50) + 1(81.00) + 3(73.13) = 8.39$$

and the corresponding estimated standard error is

$$\widehat{s.e.}(\hat{\theta}_{LT_1}) = \sqrt{\frac{\hat{\sigma}_e^2 \sum_{i=1}^t d_j^2}{n}} = \sqrt{\frac{7.45(9+1+1+9)}{8}} = 4.316$$

The corresponding t -statistic is $t_c = 8.39/4.316 = 1.94$. The measure of the quadratic trend in time for the first drug is

$$\theta_{QT_1} = 1\mu_{11} - 1\mu_{12} - 1\mu_{13} + 1\mu_{14}$$

its estimate is

$$\hat{\theta}_{QT_1} = 1\hat{\mu}_{11} - 1\hat{\mu}_{12} - 1\hat{\mu}_{13} + 1\hat{\mu}_{14} = 70.50 - 80.50 - 81.00 + 73.13 = -17.87$$

and its estimated standard error is

$$\widehat{s.e.}(\hat{\theta}_{LT_1}) = \sqrt{\frac{\hat{\sigma}_e^2 \sum_{i=1}^t d_j^2}{n}} = \sqrt{\frac{7.45(1+1+1+1)}{8}} = 1.930$$

TABLE 26.6

Comparisons of Time Means at the Same Drug for the Data in Table 26.4

Time	Drug		
	AX23	BWW9	Control
1	70.50 (a)	81.75 (a, b)	72.75 (a)
2	80.50 (b)	84.00 (a)	72.38 (a)
3	81.00 (b)	78.63 (c)	71.50 (a)
4	73.13 (a)	79.75 (b, c)	71.00 (a)

Note: Means within a column with the same letter are not significantly different at the 5% significance level. $LSD_{0.05} = 2.730$.

The corresponding t -statistic is $t_c = -17.87/1.930 = 9.259$.

The linear and quadratic trends in time for all drugs are summarized in Table 26.7. Drug BWW9 shows a negative linear trend, and drug AX23 shows a strong quadratic trend. The graph in Figure 26.1 displays these relationships. To compare drugs to one another at each time point, the estimated standard error is $\hat{\sigma}_e(\hat{\mu}_{ij} - \hat{\mu}_{ij}) = \sqrt{[2(\hat{\sigma}_e^2 + \hat{\sigma}_\delta^2)/n]}$. To evaluate this quantity, one must find estimates of each of the variance components. One gets

$$\hat{\sigma}_e^2 = MSError(Time) = 7.45$$

and

$$\hat{\sigma}_\delta^2 = \frac{MSError(Subject) - \hat{\sigma}_e^2}{p} = \frac{111.33 - 7.45}{4} = 25.97$$

TABLE 26.7

Estimates of Linear and Quadratic Trends for the Data in Table 26.4 for Each Drug

Trend	Drug		
	AX23	BWW9	Control
Linear	8.39 (1.94)	-11.37 (-2.635)	-6.13 (-1.420)
Quadratic	-17.87 (9.259)	-1.13 (-0.585)	-0.13 (-0.067)

Note: The values in parentheses are the corresponding t statistics.

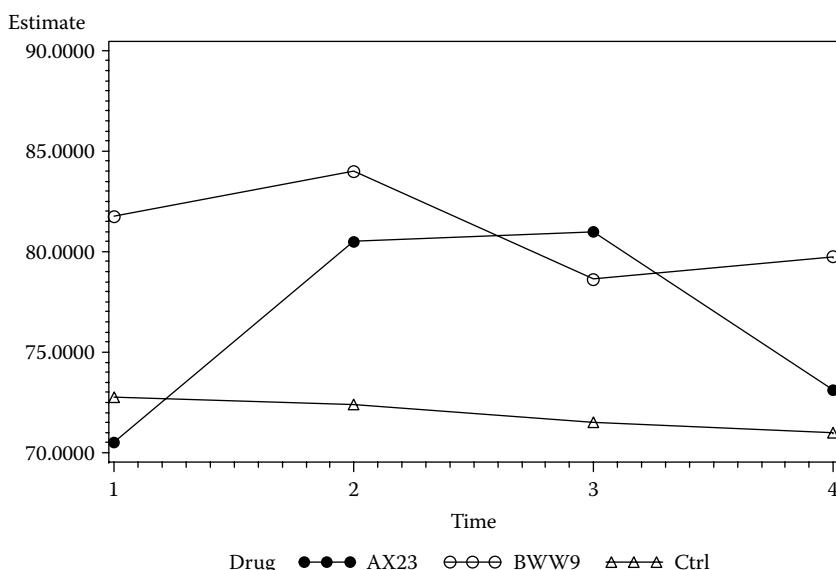


FIGURE 26.1 Means over time for each drug for the heart rate data.

Then

$$\widehat{s.e.}(\hat{\mu}_{ij} - \hat{\mu}_{i'j}) = \sqrt{2(\hat{\sigma}_e^2 + \hat{\sigma}_\delta^2)/n} = \sqrt{\frac{2(7.45 + 25.97)}{8}} = 2.891$$

The Satterthwaite estimated degrees of freedom for this estimated standard error is

$$\hat{v} = \frac{[c_1 \hat{\sigma}_1^2 + c_2 \hat{\sigma}_2^2]^2}{\frac{c_1^2 \hat{\sigma}_1^4}{V_1} + \frac{c_2^2 \hat{\sigma}_2^4}{V_2}} = \frac{\left[\frac{1}{4}(111.33) + \frac{3}{4}(7.45)\right]^2}{\frac{\frac{1}{16}(111.33)^2}{21} + \frac{\frac{9}{16}(7.45)^2}{63}} = \frac{(33.42)^2}{36.888 + 0.496} = 29.9$$

An approximate LSD at the 5% significance level to compare pairs of drug means to one another at each time point is

$$LSD_{0.05} = t_{0.025, 29.9} \times \widehat{s.e.}(\hat{\mu}_{ij} - \hat{\mu}_{i'j}) = (2.042)(2.891) = 5.903$$

Comparisons of the drugs to one another at each time point are given in Table 26.8.

Suppose drugs AX23 and BWW9 are experimental and we want to compare their average to the control at each time point. A relevant contrast for this comparison at time 1 is $\theta = \mu_{11} + \mu_{21} - 2\mu_{31}$. Its estimate is $\hat{\theta} = 70.5 + 81.75 - 2 \times 72.75 = 6.75$, and its estimated standard error is

$$\widehat{s.e.}(\hat{\theta}) = \sqrt{\frac{(\hat{\sigma}_e^2 + \hat{\sigma}_\delta^2) \sum_{i=1}^t c_i^2}{n}} = \sqrt{\frac{(7.45 + 25.97)(1+1+4)}{8}} = 5.006$$

The corresponding t -statistic is $t_c = 6.75/4.088 = 1.348$. We fail to reject $H_0: \theta = 0$ since $1.651 < t_{0.025, 29.9} = 2.042$.

The comparisons of the means of the two experimental drugs to the control at each time period are given in Table 26.9. The results show that the means of the drugs are significantly different at the 5% significance level for the last three time intervals, but they are not different at time 1.

TABLE 26.8

Comparisons of Time Means at the Same Drug for the Data in Table 26.4

Time	Drug		
	AX23	BWW9	Control
1	70.50 (a)	81.75 (b)	72.75 (a)
2	80.50 (a)	84.00 (a)	72.38 (b)
3	81.00 (a)	78.63 (a)	71.50 (b)
4	73.13 (a)	79.75 (b)	71.00 (a)

Note: Means within a row with the same letter are not significantly different at the 5% significance level. $LSD_{0.05} = 5.903$.

TABLE 26.9

Comparisons of the Means of Drugs AX23 and BWW9 to the Control at Each Time Point

Statistic	Time			
	1	2	3	4
$\hat{\theta}$	6.75	19.74	16.63	10.88
t_c	1.35	3.94*	3.32*	2.17*

Note: *Denotes significance at the 5% level, $t_{0.025, 29.9} = 2.042$.

26.2.2 Example 26.2: A Complex Comfort Experiment

An engineer had three environments in which to test two types of clothing. Since responses to an environment also differ between males and females, sex of person was included as a factor. Four people (two males and two females) were put into an environmental chamber (which was assigned to one of the three environments). One male and one female wore clothing type 1, and the other male and female wore clothing type 2. The comfort score of each person was recorded at the end of 1, 2, and 3h. The data for this experiment are shown in Table 26.10. There are three sizes of experimental units. The largest experimental unit is a chamber or, equivalently, a group of four people. The chamber experimental unit experimental design is a one-way treatment structure (environment is the treatment) in a completely randomized design structure with three replications at each level of the environment. The middle-sized experimental unit is a person. The experimental design for a person is a two-way treatment structure (sex \times clothing) in a randomized complete block design structure in nine blocks (each block contains four experimental units (people)). The smallest experimental unit is a 1 h time interval, which we will call hour. The experimental design for hour is a one-way treatment structure (time) in a randomized complete block design structure in 36 blocks [each block contains three experimental units (1 h time intervals)].

The model for this experiment (where the model is separated into parts corresponding to the three sizes of experimental units and assuming the hour measurements satisfy split-plot in time assumptions) is

$$\begin{aligned}
 Y_{ijkmn} &= \mu + E_i + r_{in} && \text{Chamber part} \\
 &+ S_j + C_k + (SC)_{jk} + (ES)_{ij} + (EC)_{ik} + (ESC)_{ijk} + p_{ijkn} && \text{Person part} \\
 &+ T_m + (ET)_{im} + (ST)_{jm} + (CT)_{km} + (SCT)_{jkm} + (EST)_{ijm} + (ECT)_{ikm} && \text{Hour part} \\
 &+ (ESCT)_{ikm} + \varepsilon_{ijkmn}
 \end{aligned}$$

where E denotes environment, S denotes sex, C denotes clothing type, T denotes time, r_{in} denotes the random chamber effect assumed to be distributed *i.i.d.* $N(0, \sigma_r^2)$, p_{ijkn} denotes the random person effect assumed to be distributed *i.i.d.* $N(0, \sigma_p^2)$, and ε_{ijkmn} denotes the measurement error for a given hour which is assumed to be *i.i.d.* $N(0, \sigma_e^2)$. In addition, all the r_{in} , p_{ijkn} , and ε_{ijkmn} are independently distributed error terms.

The analysis of variance table is given in Table 26.11 under the assumptions on the error terms given above. The F -statistics were computed by using the expected mean squares as a guide.

TABLE 26.10

Data for Comfort Experiment in Example 26.2

Replication	Sex	Clothing Type	Time	Environment		
				1	2	3
1	M	1	1	13.9001	10.2881	7.4205
1	M	1	2	7.5154	6.909	7.1752
1	M	1	3	10.9742	8.4138	7.1218
1	M	2	1	18.3941	13.8631	12.541
1	M	2	2	12.4151	10.1492	11.9157
1	M	2	3	15.2241	12.5372	12.2239
1	F	1	1	10.0149	6.1634	3.8293
1	F	1	2	3.7669	2.2837	3.5868
1	F	1	3	7.0326	4.0052	3.3004
1	F	2	1	16.4774	13.0291	11.0002
1	F	2	2	10.4104	9.7775	11.0282
1	F	2	3	13.1143	11.6576	10.5662
2	M	1	1	15.7185	11.9904	11.8158
2	M	1	2	9.717	8.4793	11.9721
2	M	1	3	12.508	9.8694	11.7187
2	M	2	1	19.7547	14.8587	16.4418
2	M	2	2	13.5293	11.0317	16.6355
2	M	2	3	16.5487	12.8317	16.6686
2	F	1	1	10.6902	6.7562	7.5707
2	F	1	2	4.8473	2.5634	7.3456
2	F	1	3	7.9829	4.7547	7.2404
2	F	2	1	17.1147	13.8977	13.5421
2	F	2	2	11.3858	9.6643	13.5672
2	F	2	3	14.1502	11.6034	14.024
3	M	1	1	14.9015	9.7589	7.2364
3	M	1	2	9.1825	6.1772	7.8304
3	M	1	3	11.5819	8.0785	7.4147
3	M	2	1	18.0402	13.5513	12.0689
3	M	2	2	12.1004	9.3052	12.5003
3	M	2	3	15.4893	11.5259	12.179
3	F	1	1	10.1944	4.5203	1.8330
3	F	1	2	4.1716	0.5913	1.6769
3	F	1	3	6.9688	2.8939	1.8065
3	F	2	1	16.0789	12.5057	9.4934
3	F	2	2	10.2357	7.7502	9.7000
3	F	2	3	12.4853	10.5226	10.0119

The error sums of squares in Table 26.11 were computed as

$$\begin{aligned}
 SS_{\text{ERROR}}(\text{PERSON}) = & \text{Replication} \times \text{Sex} \times \text{Clothing(} \\
 & \text{Environment)} \text{ SS} \\
 & + \text{Replication} \times \text{Sex(} \\
 & \text{Environment)} \text{ SS} \\
 & + \text{Replication} \times \text{Clothing(} \\
 & \text{Environment)} \text{ SS}
 \end{aligned}$$

TABLE 26.11

Analysis of Variance Table for Example 26.2

Source of Variation	df	SS	MS	F	Expected Mean Square
Environment	2	191.69	95.85	3.28	$\sigma_e^2 + 3\sigma_p^2 + 12\sigma_r^2 + Q_1$
Error(Chamber)	6	175.26	29.21		$\sigma_e^2 + 3\sigma_p^2 + 12\sigma_r^2$
Sex	1	289.46	289.46	501.89	$\sigma_e^2 + 3\sigma_p^2 + Q_2$
Clothing	1	806.11	806.11	1,397.70	$\sigma_e^2 + 3\sigma_p^2 + Q_3$
Sex \times Clothing	1	55.96	55.96	97.04	$\sigma_e^2 + 3\sigma_p^2 + Q_4$
Environment \times Sex	2	0.78	0.39	0.68	$\sigma_e^2 + 3\sigma_p^2 + Q_5$
Environment \times Clothing	2	4.31	2.16	3.73	$\sigma_e^2 + 3\sigma_p^2 + Q_6$
Environment \times Sex \times Clothing	2	4.41	2.21	3.82	$\sigma_e^2 + 3\sigma_p^2 + Q_7$
Error(Person)	18	10.38	0.58		$\sigma_e^2 + 3\sigma_p^2$
Time	2	194.6	97.3	1,672.24	$\sigma_e^2 + Q_8$
Time \times Environment	4	111.65	27.91	479.7	$\sigma_e^2 + Q_9$
Time \times Sex	2	0.08	0.04	0.67	$\sigma_e^2 + Q_{10}$
Time \times Clothing	2	0.08	0.04	0.71	$\sigma_e^2 + Q_{11}$
Time \times Sex \times Clothing	2	0.2	0.1	1.68	$\sigma_e^2 + Q_1^2$
Time \times Environment \times Sex	4	0.18	0.04	0.78	$\sigma_e^2 + Q_{13}$
Time \times Environment \times Clothing	4	0.26	0.06	1.14	$\sigma_e^2 + Q_{14}$
Time \times Environment \times Sex \times Clothing	4	0.17	0.04	0.71	$\sigma_e^2 + Q_{15}$
Error(Hour)	48	2.79	0.06		σ_e^2

Notes: All figures in the table are rounded to two significant figures, but the calculations were done in double precision. Q_i denotes the noncentrality parameter corresponding to the given effect.

and

$$SSERROR(CHAMBER) = Chamber(Environment) SS$$

The next step in the analysis is to make the needed comparisons. If one selects $\alpha = 0.01$ as the probability of a type I error, then there are two significant interactions, *Environment \times Time* and *Sex \times Clothing*. For the *Sex \times Clothing* interaction, one will want to compare the four *Sex \times Clothing* means. Since both treatments were applied to the same size of experimental unit (person), only one standard error needs to be computed.

Let

$$\begin{aligned} \mu_{ijkm} = & \mu + E_i + S_j + C_k + (SC)_{jk} + (ES)_{ij} + (EC)_{ik} + (ESC)_{ijk} + T_m + (ET)_{im} \\ & + (ST)_{jm} + (CT)_{km} + (SCT)_{jkm} + (EST)_{ijm} + (ECT)_{ikm} + (ESCT)_{ijkm} \end{aligned}$$

The differences $\bar{\mu}_{jk..} - \bar{\mu}_{j'k..}$ can be estimated by $\bar{Y}_{jk..} - \bar{Y}_{j'k..}$ and the estimated standard error of $\bar{Y}_{jk..} - \bar{Y}_{j'k..}$ is given by

$$\widehat{s.e.}(\bar{Y}_{jk..} - \bar{Y}_{j'k..}) = \sqrt{\frac{2MSERROR(PERSON)}{3 \cdot 3 \cdot 3}} = \sqrt{\frac{2 \cdot (0.58)}{27}} = 0.207$$

with 18 degrees of freedom. Fisher's LSD at the 1% significance level is

$$(t_{0.005,18}) \widehat{s.e.}(\bar{Y}_{jk..} - \bar{Y}_{j'k..}) = 2.878 \cdot 0.207 = 0.596$$

TABLE 26.12Sex \times Clothing Means with LSD for Example 26.2

Sex	Clothing	Mean Score
M	1	9.839
M	2	13.864
F	1	5.126
F	2	12.029

Note: $LSD_{0.01} = 0.595$. All means are based on 27 observations.

Table 26.12 has the *Sex \times Clothing* means, all of which are significantly different from one another.

The *Environment \times Time* interaction involves two different sizes of experimental units; thus, there are two types of comparisons. First, one can compare two time means for each environment, and one can check for linear and quadratic trends between the time means for each environment. Second, one can compare the environments to one another at each time point.

The standard error of a comparison of two time means at the same environment is

$$s.e.(\bar{Y}_{i..m.} - \bar{Y}_{i..m'.}) = s.e.[(\bar{r}_i - \bar{r}_i) + (\bar{p}_{i..} - \bar{p}_{i..}) + (\bar{\epsilon}_{i..m.} - \bar{\epsilon}_{i..m'.})] = s.e.[(\bar{\epsilon}_{i..m.} - \bar{\epsilon}_{i..m'.})] = \sqrt{\frac{2\sigma_e^2}{2 \cdot 2 \cdot 3}}$$

and the

$$\widehat{s.e.}(\bar{Y}_{i..m.} - \bar{Y}_{i..m'.}) = \sqrt{\frac{2\widehat{\sigma}_e^2}{2 \cdot 2 \cdot 3}} = \sqrt{\frac{2(0.06)}{12}} = 0.01$$

The 1% LSD value for comparing two times in the same environment is $t_{0.005,48}(0.01) = 2.682(0.01) = 0.268$.

Table 26.13 contains the *Environment \times Time* means and comparisons among times for each environment.

TABLE 26.13Environment \times Time Means for Comparing Time Means within Each Environment for Example 26.2

Time	Environment		
	1	2	3
1	15.11 (a)	10.93 (a)	9.57 (a)
2	9.11 (c)	7.07 (c)	9.57 (a)
3	12.01 (b)	9.06 (b)	9.52 (a)

Note: Within a given environment (column), time means with the same letter are not significantly different; $LSD_{0.01} = 0.268$.

Since the three levels of time are equally spaced, orthogonal polynomials can be used to measure linear and quadratic trends over time for each environment. The linear trend at environment 1 is measured by $\hat{\theta}_{LT_1} = -\bar{Y}_{1..1.} + 0\bar{Y}_{1..2.} + \bar{Y}_{1..3.} = -15.11 + 12.01 = -3.10$, and its estimated standard error is

$$s.e.(\hat{\theta}_{LT_1}) = \sqrt{\frac{\hat{\sigma}_e^2[(-1)^2 + (1)^2]}{12}} = \sqrt{\frac{(0.06)(2)}{12}} = 0.1$$

The corresponding t -statistic is $t_c = -3.10/0.1 = -31.0$ with 48 degrees of freedom. The quadratic trend across time points at environment 1 is measured by $\hat{\theta}_{QT_1} = 1\bar{Y}_{1..1.} - 2\bar{Y}_{1..2.} + 1\bar{Y}_{1..3.} = 15.11 - 2(9.11) + 12.01 = 8.90$. Its estimated standard error is given by

$$s.e.(\hat{\theta}_{QT_1}) = \sqrt{\frac{\hat{\sigma}_e^2[(1)^2 + (-2)^2 + (1)^2]}{12}} = \sqrt{\frac{(0.06)(6)}{12}} = 0.173$$

and the corresponding t -statistic is $t_c = 8.90/0.173 = -51.45$ with 48 degrees of freedom.

The linear and quadratic trends for each environment are given in Table 26.14. There are significant linear and quadratic trends for environments 1 and 2, but none for environment 3. Figure 26.2 displays the relationships.

The second type of $Environment \times Time$ comparison is to compare the different environments at the same time or at different times. The standard error of such a comparison is

$$\begin{aligned} s.e.(\bar{Y}_{i..m.} - \bar{Y}_{i'..m.}) &= s.e.[(\bar{r}_{i.} - \bar{r}_{i'.}) + (\bar{p}_{i..m.} - \bar{p}_{i'..m.}) + (\bar{\varepsilon}_{i..m.} - \bar{\varepsilon}_{i'..m.})] \\ &= \sqrt{\frac{2\sigma_r^2}{3} + \frac{2\sigma_p^2}{2 \cdot 2 \cdot 3} + \frac{2\sigma_e^2}{2 \cdot 2 \cdot 3}} = \sqrt{\frac{2}{12}(\sigma_e^2 + \sigma_p^2 + 4\sigma_r^2)} \end{aligned}$$

The quantity $\sigma_e^2 + \sigma_p^2 + 4\sigma_r^2$ can be estimated by $(1/3) MSERROR(CHAMBER) + (2/3) MSERROR(HOUR)$. Therefore the Satterthwaite estimated degrees of freedom is

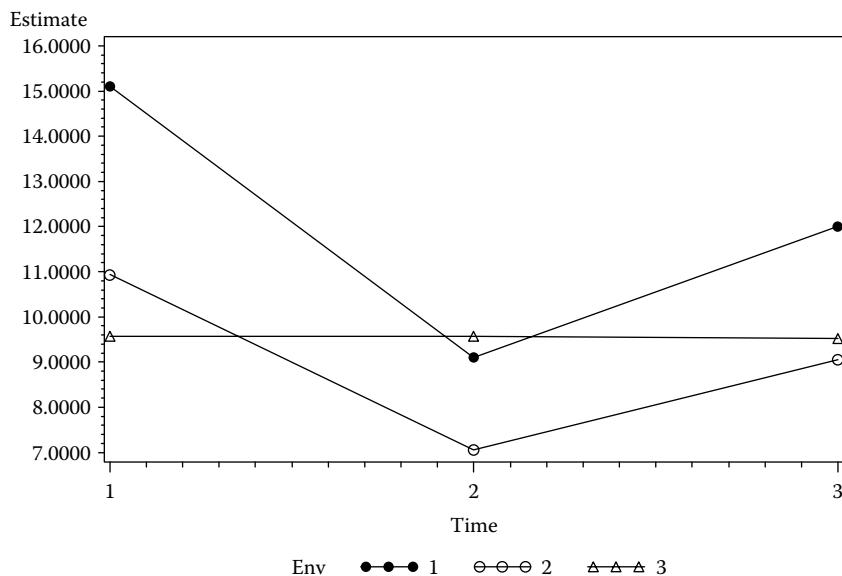
$$\hat{v} = \frac{\frac{[c_1\hat{\sigma}_1^2 + c_2\hat{\sigma}_2^2]^2}{c_1^2\hat{\sigma}_1^4} + \frac{[c_1\hat{\sigma}_1^2 + c_2\hat{\sigma}_2^2]^2}{c_2^2\hat{\sigma}_2^4}}{\frac{c_1^2\hat{\sigma}_1^4}{v_1} + \frac{c_2^2\hat{\sigma}_2^4}{v_2}} = \frac{\frac{[\frac{1}{3}(29.21) + \frac{2}{3}(0.06)]^2}{\frac{1}{9}(29.21)^2} + \frac{[\frac{1}{3}(29.21) + \frac{2}{3}(0.06)]^2}{\frac{4}{9}(0.06)^2}}{6} = \frac{(9.777)^2}{15.800 + 0.006} = 6.05$$

TABLE 26.14

Measures of Linear and Quadratic Trends for Each Environment
in Example 26.2

Trend	Environment		
	1	2	3
Linear	-3.10 (-31.0)	-1.87 (-18.70)	-0.05 (-0.5)
Quadratic	8.90 (51.45)	5.85 (33.81)	-0.05 (-0.289)

Note: t -values are in parentheses.

**FIGURE 26.2** Response over time at each environment for the comfort data.

Since

$$\sqrt{\frac{2\left[\frac{1}{3}MSERROR(Chamber) + \frac{2}{3}MSERROR(Hour)\right]}{12}} = \sqrt{\frac{2\left[\frac{1}{3}(29.21) + \frac{2}{3}(0.06)\right]}{12}} = 1.276$$

a 1% LSD for comparing environments within each time point is $LSD_{0.01} = t_{0.005,6}(1.276) = (3.707)(1.276) = 4.73$. Table 26.15 contains the *Environment* \times *Time* means with Fisher LSD comparisons made between different environments at the same times. This example illustrates why it is very important to use correct error terms when comparing within a set of interaction means: The two LSD values may be extremely different. In this example, the two values are 0.268 and 4.73. Thus, the $LSD_{0.01}$ value for comparing environments at the same level of time is more than 17 times larger than the $LSD_{0.01}$ value for comparing times within the same environment.

TABLE 26.15
Environment \times Time Means with Comparisons
between Environment Means at Each Time Point

Time	Environment		
	1	2	3
1	15.11 (a)	10.93 (ab)	9.57 (b)
2	9.11 (a)	7.07 (a)	9.57 (a)
3	12.01 (a)	9.06 (a)	9.52 (a)

Note: Within a given time (row), environment means with the same letter are not significantly different; $LSD_{0.01} = 4.73$.

To demonstrate how to use the information in Section 24.3 to compute standard errors and corresponding t -statistics for cases where more than two sizes of experimental units are used, we construct an LSD for comparing environments and sexes within the same time point. Suppose the comparison of interest is $\bar{\mu}_{11..} - \bar{\mu}_{22..}$. The best estimate of this is $\bar{Y}_{11..} - \bar{Y}_{22..} = 8.04$. Next note that the variance of this estimate is

$$\begin{aligned}\text{Var}(\bar{Y}_{11..} - \bar{Y}_{22..}) &= \text{Var}[(\bar{r}_1 - \bar{r}_2) + (\bar{p}_{11..} - \bar{p}_{22..}) + (\bar{\epsilon}_{11..} - \bar{\epsilon}_{22..})] \\ &= \frac{2\sigma_r^2}{3} + \frac{2\sigma_p^2}{2 \cdot 3} + \frac{2\sigma_\epsilon^2}{2 \cdot 3} = \frac{2}{6}(\sigma_\epsilon^2 + \sigma_p^2 + 2\sigma_r^2)\end{aligned}$$

This quantity inside the parentheses is estimated by

$$\begin{aligned}(1/6)\text{MSERROR(CHAMBER)} + (1/6)\text{MSERROR(PERSON)} \\ + (2/3)\text{MSERROR(HOUR)} = (1/6)(29.21) + (1/6)(0.58) + (2/3)(0.06) = 5.005\end{aligned}$$

so the estimated standard error of $\bar{Y}_{11..} - \bar{Y}_{22..}$ is $\widehat{s.e.}(\bar{Y}_{11..} - \bar{Y}_{22..}) = \sqrt{[2(5.005)/6]} = 1.2916$. The Satterthwaite estimated degrees of freedom associated with this estimated standard error is

$$\frac{\left[\left(\frac{1}{6}\right)(29.21) + \left(\frac{1}{6}\right)(0.58) + \left(\frac{2}{3}\right)(0.06)\right]^2}{\frac{\left(\frac{1}{6}\right)^2(29.21)^2}{6} + \frac{\left(\frac{1}{6}\right)^2(0.58)^2}{18} + \frac{\left(\frac{2}{3}\right)^2(0.06)^2}{48}} = \frac{(5.005)^2}{3.9501 + 0.0005 + 0.00003} = \frac{25.05}{3.9506} = 6.34$$

The methods used to compute the above two standard errors can be applied to other situations where a given comparison can be partitioned into the sum of components where each component is a comparison of only one size of experimental unit.

26.2.3 Example 26.3: Family Attitudes

The attitudes of families from rural and urban environments were measured every six months for three time periods. The data were obtained from seven rural families and 10 urban families, each family consisting of a son, father, and mother which are displayed in Table 26.16. The model used to describe the data was previously given in Equation 26.2. The analysis in this section assumes that the ideal conditions that were given in Section 26.1 are satisfied. Chapter 27 will consider more general analyses that do not require that the ideal conditions be satisfied.

The analysis of variance table corresponding to the model in Equation 26.2 is given in Table 26.17. If we operate at $\alpha=0.05$, there are four significant effects: *Area*, *Person*, *Time*, and *Area* \times *Time*.

Thus, there are three comparisons we wish to make: between persons, between times for each area, and between areas for each time.

The variance of a comparison between family members is

$$\begin{aligned}\text{Var}(\bar{y}_{.j..} - \bar{y}_{.j'..}) &= \text{Var}[(\bar{f}_{..} - \bar{f}_{..}) + (\bar{p}_{.j..} - \bar{p}_{.j'..}) + (\bar{\epsilon}_{.j..} - \bar{\epsilon}_{.j'..})] \\ &= \frac{2}{17}\sigma_p^2 + \frac{2}{3 \cdot 17}\sigma_\epsilon^2 = \frac{2}{51}(\sigma_\epsilon^2 + 3\sigma_p^2)\end{aligned}$$

TABLE 26.16

Data for Family Attitude Study of Example 26.3

Family	Person								
	Son			Father			Mother		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
<i>Urban</i>									
I	17	17	19	18	19	21	16	16	18
2	12	i4	15	19	19	21	16	16	18
3	8	10	11	16	18	19	11	12	12
4	5	7	7	12	12	13	13	14	14
S	2	5	6	12	14	14	14	16	18
6	9	11	11	16	17	18	14	15	16
7	8	9	9	19	20	20	15	16	18
8	13	14	16	16	17	18	18	18	20
9	11	12	13	13	16	17	7	8	10
10	19	20	20	13	15	15	11	12	12
<i>Rural</i>									
1	12	11	14	18	19	22	16	16	19
2	13	13	17	18	19	22	16	16	19
3	12	13	16	19	18	22	17	16	20
4	18	18	21	23	23	26	23	22	26
5	15	14	16	15	15	19	17	17	20
6	6	6	10	15	16	19	18	19	21
7	16	17	18	17	17	21	18	20	23

TABLE 26.17

Analysis of Variance Table for Family Attitude Data of Example 26.3

Source of Variation	df	SS	MS	F	EMS
Area	1	382.12	382.122	7.06	$\sigma_e^2 + 3\sigma_p^2 + 9\sigma_f^2 + Q_1$
Error(Family)	15	812.32	54.155		$\sigma_e^2 + 3\sigma_p^2 + 9\sigma_f^2$
Person	2	666.84	330.92	13.5	$\sigma_e^2 + 3\sigma_p^2 + Q_2$
Person \times Area	2	60.27	30.135	1.23	$\sigma_e^2 + 3\sigma_p^2 + Q_3$
Error(Person)	30	735.37	25.512		$\sigma_e^2 + 3\sigma_p^2$
Time	2	204.28	102.14	268.94	$\sigma_e^2 + Q_4$
Time \times Area	2	30.95	15.475	40.74	$\sigma_e^2 + Q_5$
Time \times Person	4	1.19	0.298	0.78	$\sigma_e^2 + Q_6$
Time \times Person \times Area	4	1.5	0.375	0.99	$\sigma_e^2 + Q_7$
Error(Time)	90	34.18	0.37		σ_e^2

Note: Q_i denotes the noncentrality parameter corresponding to the respective sums of squares.

The estimate of the standard error of the difference of two family member means is

$$\widehat{s.e.}(\bar{y}_{j..} - \bar{y}_{j'..}) = \sqrt{\frac{2}{51} MSError(Person)} = \sqrt{\frac{2}{51}(25.512)} = 0.9906$$

A 5% LSD value is $LSD_{0.05} = 2.042(0.99) = 2.022$. Table 26.18 has a summary of the pairwise comparisons. The variance of the difference of two time means within a specified area is

$$Var(\bar{y}_{i..k.} - \bar{y}_{i..k'.}) = Var[(\bar{f}_{i..} - \bar{f}_{i..}) + (\bar{p}_{i..} - \bar{p}_{i..}) + (\bar{\epsilon}_{i..k.} - \bar{\epsilon}_{i..k'.})] = \frac{2}{3n_i} \sigma_e^2$$

where n_i is the number of families in area i .

The estimate of the standard error to compare two times in the urban area is

$$\widehat{s.e.}(\bar{y}_{1..k.} - \bar{y}_{1..k'.}) = \sqrt{\frac{2}{3n_1} \hat{\sigma}_e^2} = \sqrt{\frac{2}{30}(0.370)} = 0.157$$

and the standard error to compare two times in a rural area is

$$\widehat{s.e.}(\bar{y}_{2..k.} - \bar{y}_{2..k'.}) = \sqrt{\frac{2}{3n_2} \hat{\sigma}_e^2} = \sqrt{\frac{2}{21}(0.370)} = 0.188$$

The corresponding 5% LSDs are $LSD_{0.05}(\text{Urban}) = 1.987(0.157) = 0.312$ and $LSD_{0.05}(\text{Rural}) = 1.987(0.188) = 0.374$. The multiple comparisons are summarized in Table 26.19.

TABLE 26.18

Comparison of Person Means for Example 26.3

Person	Mean
Son	12.67 (a)
Father	17.47 (b)
Mother	16.53 (b)

Note: Means within a column with the same letter are not significantly different. $LSD_{0.05} = 2.022$.

TABLE 26.19

Comparison of Time Means within Each Area for Example 26.3

Time	Area	
	Rural	Urban
1	16.33 (a)	13.10 (a)
2	16.38 (a)	14.30 (b)
3	19.61 (b)	15.30 (c)

Note: Means within a column with the same letter are not significantly different. $LSD_{0.05} = 0.374$ for Rural and $LSD_{0.05} = 0.312$ for Urban.

TABLE 26.20

Comparison of Area Means at Each Time for Example 26.3

Time	Area	
	Rural	Urban
1	16.33 (a)	13.10 (b)
2	16.38 (a)	14.30 (a)
3	19.61 (a)	15.30 (b)

Note: Means within a row with the same letter are not significantly different. $LSD_{0.05} = 2.59$.

Next the variance of the difference between urban and rural means at a given time point is

$$\begin{aligned}\text{Var}(\bar{y}_{1..k.} - \bar{y}_{2..k.}) &= \text{Var}[(\bar{f}_{1..} - \bar{f}_{2..}) + (\bar{p}_{1..} - \bar{p}_{2..}) + (\bar{\epsilon}_{1..k.} - \bar{\epsilon}_{2..k.})] \\ &= \sigma_f^2 \left(\frac{1}{10} + \frac{1}{7} \right) + \sigma_p^2 \left(\frac{1}{3 \cdot 10} + \frac{1}{3 \cdot 7} \right) + \sigma_\epsilon^2 \left(\frac{1}{3 \cdot 10} + \frac{1}{3 \cdot 7} \right) = \left(\frac{1}{30} + \frac{1}{21} \right) (\sigma_\epsilon^2 + \sigma_p^2 + 3\sigma_f^2)\end{aligned}$$

The function of the variance components in the above expression can be estimated by

$$(1/3)MSE_{\text{Error}}(\text{Family}) + (2/3)MSE_{\text{Error}}(\text{Time}) = (1/3)(54.155) + (2/3)(0.370) = 18.2983$$

Therefore, the estimated standard error is

$$\widehat{s.e.}(\bar{y}_{1..k.} - \bar{y}_{2..k.}) = \sqrt{\left(\frac{1}{30} + \frac{1}{21} \right) (18.2983)} = 1.217$$

and the Satterthwaite estimated degrees of freedom corresponding to this estimate is

$$\frac{\left[\left(\frac{1}{3} \right) (54.155) + \left(\frac{2}{3} \right) (0.370) \right]^2}{\frac{\left(\frac{1}{3} \right)^2 (54.155)^2}{15} + \frac{\left(\frac{2}{3} \right)^2 (0.370)^2}{90}} = \frac{(18.2983)^2}{21.7242 + 0.0007} = \frac{334.8278}{21.7249} = 15.41$$

Thus an approximate 5% LSD value for comparing urban to rural at a given time point is $LSD_{0.05} = 2.131(1.217) = 2.59$. The multiple comparisons are summarized in Table 26.20. The linear and quadratic trends of time in each area can be investigated using the method described in Example 26.1.

26.3 Data Analyses Using the SAS-Mixed Procedure

This section illustrates SAS®-Mixed analyses of each of the three examples described in the preceding section. Consider the data in Table 26.4. To reproduce the results shown in Section 26.2, one can use the SAS commands shown in Table 26.21.

TABLE 26.21

SAS-Mixed Code to Analyze the Data in Table 26.4

```

DATA HRT_RATE;
DATA; SET HRT_RATE; DROP HR1-HR4;
Time=1; HR=HR1; OUTPUT; Time=2; HR=HR2; OUTPUT; Time=3; HR=HR3; OUTPUT;
Time=4; HR=HR4; OUTPUT;
RUN;

PROC MIXED;
CLASSES Drug Time PERSON;
MODEL HR=Drug Time Drug*Time/DDFM=SATTERTH;
RANDOM PERSON(Drug);
LSMEANS Drug|Time/PDIFF;
ODS OUTPUT LSMEANS=LSMS;
RUN;

PROC PRINT DATA=LSMS;
RUN;
SYMBOL1 V=DOT COLOR=BLACK I=JOIN;
SYMBOL2 V=CIRCLE COLOR=BLACK I=JOIN;
SYMBOL3 V=TRIANGLE COLOR=BLACK I=JOIN;

PROC GPLOT DATA=LSMS; WHERE EFFECT='Drug*Time';
PLOT ESTIMATE*Time=Drug/VAXIS=70 TO 90 BY 5;
RUN;
QUIT;

```

TABLE 26.22Covariance Parameter Estimates for the Data
in Table 26.4*Covariance Parameter Estimates*

Covariance Parameter	Estimate
Person(Drug)	25.9702
Residual	7.4479

TABLE 26.23

Statistical Tests for *Drug*, *Time*, and *Drug* \times *Time* Effects for the Data in Table 26.4

Type III Tests of Fixed Effects				
Effect	Num df	Den df	F-Value	Pr > F
Drug	2	21	5.99	0.0088
Time	3	63	12.96	<0.0001
Drug \times Time	6	63	11.80	<0.0001

26.3.1 Example 26.1

One portion of the output that should be noted is the portion that provides estimates of the two variance components, σ^2_ϵ and σ^2_δ . The estimates are given in Table 26.22. One can see that $\hat{\sigma}^2_\delta = 25.97$ and $\hat{\sigma}^2_\epsilon = 7.45$.

Tests corresponding to the ANOVA in Table 26.5 are given in Table 26.23. Note that the MIXED procedure does not provide error sums of squares, but it does give the same *F*-statistics as given in Table 26.5.

Drug main effect means, *Time* main effect means, and the *Drug* \times *Time* two-way means are given in Table 26.24 and pairwise comparisons among various subsets of these means are given in Tables 26.25–26.27. Estimate options that can be used in the MIXED procedure to compute linear and quadratic contrasts in time for each drug are given in Table 26.28. The results of these options are shown in Table 26.29.

TABLE 26.24

Drug, *Time* and *Drug* \times *Time* Means for the Data in Table 26.4

Least Squares Means

Effect	Drug	Time	Estimate	Standard Error	df	t-Value	Pr > t
Drug	AX23		76.2813	1.8652	21	40.90	<0.0001
Drug	BWW9		81.0312	1.8652	21	43.44	<0.0001
Drug	CTRL		71.9062	1.8652	21	38.55	<0.0001
Time		1	75.0000	1.1800	29.9	63.56	<0.0001
Time		2	78.9583	1.1800	29.9	66.91	<0.0001
Time		3	77.0417	1.1800	29.9	65.29	<0.0001
Time		4	74.6250	1.1800	29.9	63.24	<0.0001
Drug \times Time	AX23	1	70.5000	2.0438	29.9	34.49	<0.0001
Drug \times Time	AX23	2	80.5000	2.0438	29.9	39.39	<0.0001
Drug \times Time	AX23	3	81.0000	2.0438	29.9	39.63	<0.0001
Drug \times Time	AX23	4	73.1250	2.0438	29.9	35.78	<0.0001
Drug \times Time	BWW9	1	81.7500	2.0438	29.9	40.00	<0.0001
Drug \times Time	BWW9	2	84.0000	2.0438	29.9	41.10	<0.0001
Drug \times Time	BWW9	3	78.6250	2.0438	29.9	38.47	<0.0001
Drug \times Time	BWW9	4	79.7500	2.0438	29.9	39.02	<0.0001
Drug \times Time	CTRL	1	72.7500	2.0438	29.9	35.59	<0.0001
Drug \times Time	CTRL	2	72.3750	2.0438	29.9	35.41	<0.0001
Drug \times Time	CTRL	3	71.5000	2.0438	29.9	34.98	<0.0001
Drug \times Time	CTRL	4	71.0000	2.0438	29.9	34.74	<0.0001

TABLE 26.25

Pairwise Comparisons among Drug Main Effect Means and Time Main Effect Means for the Data in Table 26.4

Differences of Least Squares Means

Effect	Drug	Time	Drug	Time	Estimate	Standard Error	df	t-Value	Pr > t
Drug	AX23		BWW9		-4.7500	2.6378	21	-1.80	0.0861
Drug	AX23		CTRL		4.3750	2.6378	21	1.66	0.1121
Drug	BWW9		CTRL		9.1250	2.6378	21	3.46	0.0023
Time		1		2	-3.9583	0.7878	63	-5.02	<0.0001
Time		1		3	-2.0417	0.7878	63	-2.59	0.0119
Time		1		4	0.3750	0.7878	63	0.48	0.6357
Time		2		3	1.9167	0.7878	63	2.43	0.0178
Time		2		4	4.3333	0.7878	63	5.50	<0.0001
Time		3		4	2.4167	0.7878	63	3.07	0.0032

TABLE 26.26

Comparisons of Time Means within Each Drug for the Data in Table 26.4

Differences of Least Squares Means

Effect	Drug	Time	Drug	Time	Estimate	Standard Error	df	t-Value	Pr > t
Drug × Time	AX23	1	AX23	2	-10.0000	1.3645	63	-7.33	<0.0001
Drug × Time	AX23	1	AX23	3	-10.5000	1.3645	63	-7.69	<0.0001
Drug × Time	AX23	1	AX23	4	-2.6250	1.3645	63	-1.92	0.0589
Drug × Time	AX23	2	AX23	3	-0.5000	1.3645	63	-0.37	0.7153
Drug × Time	AX23	2	AX23	4	7.3750	1.3645	63	5.40	<0.0001
Drug × Time	AX23	3	AX23	4	7.8750	1.3645	63	5.77	<0.0001
Drug × Time	BWW9	1	BWW9	2	-2.2500	1.3645	63	-1.65	0.1041
Drug × Time	BWW9	1	BWW9	3	3.1250	1.3645	63	2.29	0.0254
Drug × Time	BWW9	1	BWW9	4	2.0000	1.3645	63	1.47	0.1477
Drug × Time	BWW9	2	BWW9	3	5.3750	1.3645	63	3.94	0.0002
Drug × Time	BWW9	2	BWW9	4	4.2500	1.3645	63	3.11	0.0028
Drug × Time	BWW9	3	BWW9	4	-1.1250	1.3645	63	-0.82	0.4128
Drug × Time	CTRL	1	CTRL	2	0.3750	1.3645	63	0.27	0.7844
Drug × Time	CTRL	1	CTRL	3	1.2500	1.3645	63	0.92	0.3631
Drug × Time	CTRL	1	CTRL	4	1.7500	1.3645	63	1.28	0.2044
Drug × Time	CTRL	2	CTRL	3	0.8750	1.3645	63	0.64	0.5237
Drug × Time	CTRL	2	CTRL	4	1.3750	1.3645	63	1.01	0.3175
Drug × Time	CTRL	3	CTRL	4	0.5000	1.3645	63	0.37	0.7153

TABLE 26.27

Comparisons of Drug Means within Each Time for the Data in Table 26.4

Differences of Least Squares Means

Effect	Drug	Time	Drug	Time	Estimate	Standard Error	df	t-Value	Pr > t
Drug × Time	AX23	1	BWW9	1	-11.2500	2.8904	29.9	-3.89	0.0005
Drug × Time	AX23	1	CTRL	1	-2.2500	2.8904	29.9	-0.78	0.4424
Drug × Time	AX23	2	BWW9	2	-3.5000	2.8904	29.9	-1.21	0.2354
Drug × Time	AX23	2	CTRL	2	8.1250	2.8904	29.9	2.81	0.0086
Drug × Time	AX23	3	BWW9	3	2.3750	2.8904	29.9	0.82	0.4178
Drug × Time	AX23	3	CTRL	3	9.5000	2.8904	29.9	3.29	0.0026
Drug × Time	AX23	4	BWW9	4	-6.6250	2.8904	29.9	-2.29	0.0291
Drug × Time	AX23	4	CTRL	4	2.1250	2.8904	29.9	0.74	0.4680
Drug × Time	BWW9	1	CTRL	1	9.0000	2.8904	29.9	3.11	0.0041
Drug × Time	BWW9	2	CTRL	2	11.6250	2.8904	29.9	4.02	0.0004
Drug × Time	BWW9	3	CTRL	3	7.1250	2.8904	29.9	2.47	0.0197
Drug × Time	BWW9	4	CTRL	4	8.7500	2.8904	29.9	3.03	0.0050

TABLE 26.28

Estimate Options to Compute Linear and Quadratic Contrasts in Time for Each Drug for the Data in Table 26.4

```

ESTIMATE 'TIME LINEAR FOR AX23' TIME -3 -1 1 3 DRUG*TIME -3 -1 1 3
0 0 0 0 0 0 0;
ESTIMATE 'TIME LINEAR FOR BWW9' TIME -3 -1 1 3 DRUG*TIME 0 0 0 0
-3 -1 1 3 0 0 0 0;
ESTIMATE 'TIME LINEAR FOR CNTL' TIME -3 -1 1 3 DRUG*TIME 0 0 0 0
0 0 0 0 -3 -1 1 3;
ESTIMATE 'TIME QUAD FOR AX23' TIME 1 -1 -1 1 DRUG*TIME 1 -1 -1 1
0 0 0 0 0 0 0;
ESTIMATE 'TIME QUAD FOR BWW9' TIME 1 -1 -1 1 DRUG*TIME 0 0 0 0
1 -1 -1 1 0 0 0 0;
ESTIMATE 'TIME QUAD FOR CNTL' TIME 1 -1 -1 1 DRUG*TIME 0 0 0 0
0 0 0 0 1 -1 -1 1;

```

TABLE 26.29

Linear and Quadratic Contrasts in Time for Each Drug for the Data in Table 26.4

Estimates

Label	Estimate	Standard Error	df	t-Value	Pr > t
Time linear for AX23	8.3750	4.3151	63	1.94	0.0568
Time linear for BWW9	-11.3750	4.3151	63	-2.64	0.0105
Time linear for CNTL	-6.1250	4.3151	63	-1.42	0.1607
Time quad for AX23	-17.8750	1.9298	63	-9.26	<0.0001
Time quad for BWW9	-1.1250	1.9298	63	-0.58	0.5620
Time quad for CNTL	-0.1250	1.9298	63	-0.06	0.9486

TABLE 26.30

SAS Commands to Analyze the Complex Comfort Experiment

```

PROC MIXED DATA=COMFORT;
CLASSES ENV REP SEX CLO TIME;
MODEL SCORE= ENV|SEX|CLO|TIME/DDFM=SATTERTH ;
RANDOM REP(ENV) SEX*CLO*REP(ENV);
LSMEANS SEX*CLO ENV*TIME /PDIFF;
ESTIMATE 'LINEAR TIME FOR ENV 1' TIME -1 0 1 ENV*TIME -1 0 1 0 0 0 0 0 0;
ESTIMATE 'LINEAR TIME FOR ENV 2' TIME -1 0 1 ENV*TIME 0 0 0 -1 0 1 0 0 0;
ESTIMATE 'LINEAR TIME FOR ENV 3' TIME -1 0 1 ENV*TIME 0 0 0 0 0 0 -1 0 1;
ESTIMATE 'QUAD TIME FOR ENV 1' TIME 1 -2 1 ENV*TIME 1 -2 1 0 0 0 0 0 0;
ESTIMATE 'QUAD TIME FOR ENV 2' TIME 1 -2 1 ENV*TIME 0 0 0 1 -2 1 0 0 0;
ESTIMATE 'QUAD TIME FOR ENV 3' TIME 1 -2 1 ENV*TIME 0 0 0 0 0 0 1 -2 1;
ESTIMATE 'ET11 VERSUS ET22' ENV 1 -1 0 TIME 1 -1 0 ENV*TIME 1 0 0 0 -1 0 0 0 0;
ODS OUTPUT LSMEANS = LSMS DIFFS=LSMDIFFS;
RUN;
DATA LSMDIFFS1; SET LSMDIFFS; DROP SEX CLO _SEX _CLO;
PROC PRINT DATA=LSMDIFFS1; WHERE TIME=_TIME AND EFFECT='ENV*TIME';
PROC PRINT DATA=LSMDIFFS1; WHERE ENV=_ENV AND EFFECT='ENV*TIME';
DATA LSMDIFFS2; SET LSMDIFFS; DROP ENV TIME _ENV _TIME;
PROC PRINT DATA=LSMDIFFS2; WHERE EFFECT='SEX*CLO';
RUN;
SYMBOL1 V=DOT I=JOIN COLOR=BLACK;
SYMBOL2 V=CIRCLE I=JOIN COLOR=BLACK;
SYMBOL3 V=TRIANGLE I=JOIN COLOR=BLACK;
PROC GPLOT DATA=LSMS; WHERE EFFECT='ENV*TIME';
PLOT ESTIMATE*TIME=ENV;
RUN;

```

26.3.2 Example 26.2

This section provides MIXED analyses of the Complex Comfort Experiment. The data in Example 26.2 were analyzed using the SAS commands shown in Table 26.30. The estimates of each of the three variance components corresponding to chamber, person, and time are given in Table 26.31. The type III tests of fixed effects are given in Table 26.32. Linear and quadratic contrasts for time within each environment are given in Table 26.33. The table also contains a comparison of Environment 1 at Time 1 to Environment 2 at Time 2 to

TABLE 26.31

Estimates of the Chamber, Person, and Time Variance Components for the Comfort Study

Covariance Parameter Estimates

Covariance Parameter	Estimate
Rep(Env)	2.3861
Rep × Sex × Clo(Env)	0.1729
Residual	0.05819

TABLE 26.32

Type III Tests of Fixed Effects for the Complex Comfort Experiment

Type III Tests of Fixed Effects

Effect	Num df	Den df	F-value	Pr > F
Env	2	6	3.28	0.1090
Sex	1	18	501.89	<0.0001
Env × Sex	2	18	0.68	0.5201
Clo	1	18	1397.70	<0.0001
Env × Clo	2	18	3.73	0.0440
Sex × Clo	1	18	97.04	<0.0001
Env × Sex × Clo	2	18	3.82	0.0414
Time	2	48	1672.23	<0.0001
Env × Time	4	48	479.70	<0.0001
Sex × Time	2	48	0.67	0.5164
Env × Sex × Time	4	48	0.78	0.5449
Clo × Time	2	48	0.71	0.4954
Env × Clo × Time	4	48	1.14	0.3499
Sex × Clo × Time	2	48	1.68	0.1978
Env × Sex × Clo × Time	4	48	0.71	0.5896

TABLE 26.33

Linear and Quadratic Contrasts for Time within Each Environment for the Comfort Study

Estimates

Label	Estimate	Standard		t-Value	Pr > t
		Error	df		
Linear time for Env 1	-3.1016	0.09848	48	-31.50	<0.0001
Linear time for Env 2	-1.8741	0.09848	48	-19.03	<0.0001
Linear time for Env 3	-0.04309	0.09848	48	-0.44	0.6637
Quad time for Env 1	8.8988	0.1706	48	52.17	<0.0001
Quad time for Env 2	5.8761	0.1706	48	34.45	<0.0001
Quad time for Env 3	-0.06653	0.1706	48	-0.39	0.6982
ET11 vs ET22	8.0498	1.2764	6.05	6.31	0.0007

illustrate such a comparison should it be of interest to someone. Table 26.34 gives the *Sex × Clothing* means and the *Environment × Time* means. Tables 26.35–26.37 provide pairwise comparison between various subsets of these means.

26.3.3 Example 26.3

This section provides MIXED analyses of the family member experiment. The data in Example 26.3 were analyzed using the SAS commands shown in Table 26.38. The estimates of each of the three variance components corresponding to area, family member, and time are given in Table 26.39. The type III tests of fixed effects are given in Table 26.40. Table 26.41

TABLE 26.34

Sex \times Clothing Means and Environment \times Time Means for the Comfort Study

Least Squares Means									
Effect	Env	Sex	Clo	Time	Estimate	Standard Error	df	t-Value	Pr > t
Sex \times Clo		1	1		9.8396	0.5352	6.72	18.38	<0.0001
Sex \times Clo		1	2		13.8639	0.5352	6.72	25.90	<0.0001
Sex \times Clo		2	1		5.1256	0.5352	6.72	9.58	<0.0001
Sex \times Clo		2	2		12.0294	0.5352	6.72	22.47	<0.0001
Env \times Time	1			1	15.1066	0.9026	6.05	16.74	<0.0001
Env \times Time	1			2	9.1065	0.9026	6.05	10.09	<0.0001
Env \times Time	1			3	12.0050	0.9026	6.05	13.30	<0.0001
Env \times Time	2			1	10.9319	0.9026	6.05	12.11	<0.0001
Env \times Time	2			2	7.0568	0.9026	6.05	7.82	0.0002
Env \times Time	2			3	9.0578	0.9026	6.05	10.04	<0.0001
Env \times Time	3			1	9.5661	0.9026	6.05	10.60	<0.0001
Env \times Time	3			2	9.5778	0.9026	6.05	10.61	<0.0001
Env \times Time	3			3	9.5230	0.9026	6.05	10.55	<0.0001

TABLE 26.35

Pairwise Comparisons of the Sex \times Clothing Means for the Comfort Study

Effect	Sex	Clo	Sex	Clo	Estimate	Standard Error	df	t-Value	Pr t
Sex \times Clo	1	1	1	2	-4.0243	0.2067	18	-19.47	<0.0001
Sex \times Clo	1	1	2	1	4.7140	0.2067	18	22.81	<0.0001
Sex \times Clo	1	1	2	2	-2.1898	0.2067	18	-10.59	<0.0001
Sex \times Clo	1	2	2	1	8.7383	0.2067	18	42.28	<0.0001
Sex \times Clo	1	2	2	2	1.8345	0.2067	18	8.88	<0.0001
Sex \times Clo	2	1	2	2	-6.9038	0.2067	18	-33.40	<0.0001

TABLE 26.36

Pairwise Comparisons of the Times within Each Environment for the Comfort Study

Effect	Env	Time	Env	Time	Estimate	Standard Error	df	t-Value	Pr t
Env \times Time	1	1	1	2	6.0002	0.09848	48	60.93	<0.0001
Env \times Time	1	1	1	3	3.1016	0.09848	48	31.50	<0.0001
Env \times Time	1	2	1	3	-2.8986	0.09848	48	-29.43	<0.0001
Env \times Time	2	1	2	2	3.8751	0.09848	48	39.35	<0.0001
Env \times Time	2	1	2	3	1.8741	0.09848	48	19.03	<0.0001
Env \times Time	2	2	2	3	-2.0010	0.09848	48	-20.32	<0.0001
Env \times Time	3	1	3	2	-0.01172	0.09848	48	-0.12	0.9057
Env \times Time	3	1	3	3	0.04309	0.09848	48	0.44	0.6637
Env \times Time	3	2	3	3	0.05481	0.09848	48	0.56	0.5804

TABLE 26.37

Pairwise Comparisons of Times within Each Environment for the Comfort Study

Effect	Env	Time	Env	Time	Estimate	Standard Error	df	t-Value	Pr t
Env × Time	1	1	2	1	4.1747	1.2764	6.05	3.27	0.0168
Env × Time	1	1	3	1	5.5405	1.2764	6.05	4.34	0.0048
Env × Time	1	2	2	2	2.0496	1.2764	6.05	1.61	0.1591
Env × Time	1	2	3	2	-0.4714	1.2764	6.05	-0.37	0.7245
Env × Time	1	3	2	3	2.9472	1.2764	6.05	2.31	0.0600
Env × Time	1	3	3	3	2.4820	1.2764	6.05	1.94	0.0994
Env × Time	2	1	3	1	1.3658	1.2764	6.05	1.07	0.3254
Env × Time	2	2	3	2	-2.5210	1.2764	6.05	-1.98	0.0953
Env × Time	2	3	3	3	-0.4652	1.2764	6.05	-0.36	0.7279

TABLE 26.38

SAS Commands to Analyze the Family Attitudes Experiment

```

DATA AMD26_3;
  INPUT AREA $ FAM ATTITUD1-ATTITUD9;
  IF AREA='RURAL' THEN FAM=FAM+10;
  CARDS;
  URBAN  1 17 17 19 18 19 21 16 16 18
  URBAN  2 12 14 15 19 19 21 16 16 18
  URBAN  3 8 10 11 16 18 19 11 12 12
  URBAN  4 5 7 7 12 12 13 13 14 14
  URBAN  5 2 5 6 12 14 14 14 16 18
  URBAN  6 9 11 11 16 17 18 14 15 16
  URBAN  7 8 9 9 19 20 20 15 16 18
  URBAN  8 13 14 16 16 17 18 18 18 20
  URBAN  9 11 12 13 13 16 17 7 8 10
  URBAN 10 19 20 20 13 15 15 11 12 12
  RURAL  1 12 11 14 18 19 22 16 16 19
  RURAL  2 13 13 17 16 15 19 19 19 23
  RURAL  3 12 13 16 19 18 22 17 16 20
  RURAL  4 18 18 21 23 23 26 23 22 26
  RURAL  5 15 14 16 15 15 19 17 17 20
  RURAL  6 6 6 10 15 16 19 18 19 21
  RURAL  7 16 17 18 17 17 21 18 20 23
PROC PRINT;
  RUN;

DATA USUAL; SET AMD26_3; DROP ATTITUD1-ATTITUD9;
  FMEMB=1; TIME= 0; ATTITUD=ATTITUD1; OUTPUT;
  FMEMB=1; TIME= 6; ATTITUD=ATTITUD2; OUTPUT;
  FMEMB=1; TIME=12; ATTITUD=ATTITUD3; OUTPUT;
  FMEMB=2; TIME= 0; ATTITUD=ATTITUD4; OUTPUT;
  FMEMB=2; TIME= 6; ATTITUD=ATTITUD5; OUTPUT;
  FMEMB=2; TIME=12; ATTITUD=ATTITUD6; OUTPUT;
  FMEMB=3; TIME= 0; ATTITUD=ATTITUD7; OUTPUT;
  FMEMB=3; TIME= 6; ATTITUD=ATTITUD8; OUTPUT;
  FMEMB=3; TIME=12; ATTITUD=ATTITUD9; OUTPUT;
RUN;

```

Continued

TABLE 26.38 (continued)

```
PROC MIXED DATA=USUAL;
CLASSES AREA FMEMB TIME FAM;
MODEL ATTITUD=AREA|FMEMB|TIME/DDFM=SATTERTH;
RANDOM FAM(AREA) FMEMB*FAM(AREA);
LSMEANS AREA*TIME/PDIFF;
RUN;
QUIT;
```

TABLE 26.39

Covariance Parameter Estimates for the Family Attitudes Experiment

Covariance Parameter Estimates

Covariance Parameter	Estimate
Fam(Area)	3.2936
Fmemb × Fam(Area)	8.0442
Residual	0.3798

TABLE 26.40

Type III Tests of Fixed Effects for the Family Attitudes Experiment

Type III Tests of Fixed Effects

Effect	Num df	Den df	F-Value	Pr > F
Area	1	15	7.06	0.0180
Fmemb	2	30	13.47	<0.0001
Area × Fmemb	2	30	1.23	0.3068
Time	2	90	268.94	<0.0001
Area × Time	2	90	40.74	<0.0001
Fmemb × Time	4	90	0.78	0.5409
Area × Fmemb × Time	4	90	0.99	0.4190

TABLE 26.41

Area × Time Two-Way Means for the Family Attitudes Experiment

Least Squares Means

Effect	Area	Time	Estimate	Standard Error	df	t-Value	Pr > t
Area × Time	Rural	0	16.3333	0.9336	15.4	17.49	<0.0001
Area × Time	Rural	6	16.3810	0.9336	15.4	17.55	<0.0001
Area × Time	Rural	12	19.6190	0.9336	15.4	21.01	<0.0001
Area × Time	Urban	0	13.1000	0.7811	15.4	16.77	<0.0001
Area × Time	Urban	6	14.3000	0.7811	15.4	18.31	<0.0001
Area × Time	Urban	12	15.3000	0.7811	15.4	19.59	<0.0001

TABLE 26.42

Pairwise Comparisons among the Area \times Time Two-Way Means for the Family Attitudes Experiment

Differences of Least Squares Means

Effect	Area	Time	Area	Time	Estimate	Standard Error	df	t-Value	Pr > t
Area \times Time	Rural	0	Rural	6	-0.04762	0.1902	90	-0.25	0.8029
Area \times Time	Rural	0	Rural	12	-3.2857	0.1902	90	-17.28	<0.0001
Area \times Time	Rural	0	Urban	0	3.2333	1.2173	15.4	2.66	0.0177
Area \times Time	Rural	0	Urban	6	2.0333	1.2173	15.4	1.67	0.1150
Area \times Time	Rural	0	Urban	12	1.0333	1.2173	15.4	0.85	0.4089
Area \times Time	Rural	6	Rural	12	-3.2381	0.1902	90	-17.03	<0.0001
Area \times Time	Rural	6	Urban	0	3.2810	1.2173	15.4	2.70	0.0163
Area \times Time	Rural	6	Urban	6	2.0810	1.2173	15.4	1.71	0.1074
Area \times Time	Rural	6	Urban	12	1.0810	1.2173	15.4	0.89	0.3882
Area \times Time	Rural	12	Urban	0	6.5190	1.2173	15.4	5.36	<0.0001
Area \times Time	Rural	12	Urban	6	5.3190	1.2173	15.4	4.37	0.0005
Area \times Time	Rural	12	Urban	12	4.3190	1.2173	15.4	3.55	0.0028
Area \times Time	Urban	0	Urban	6	-1.2000	0.1591	90	-7.54	<0.0001
Area \times Time	Urban	0	Urban	12	-2.2000	0.1591	90	-13.83	<0.0001
Area \times Time	Urban	6	Urban	12	-1.0000	0.1591	90	-6.28	<0.0001

gives the *Area \times Time* means. Table 26.42 provides pairwise comparisons between various subsets of these means.

26.4 Concluding Remarks

The analysis of repeated measures designs was described for three examples involving repeated measures where split-plot in time assumptions hold. Included in the discussion were the models and assumptions for each example. Computational formulas for obtaining standard errors for multiple comparisons were given as well as methods for investigating various contrasts among the means. Analyses using the SAS-Mixed procedure were obtained in Section 26.3 and can be compared with the results given in Section 26.2.

26.5 Exercises

- 26.1 Phlebitis is an inflammation of a blood vein that can occur when intravenously administering drugs. The active drug was thought to be the main contributing factor to inflammation, although the solution used as a vehicle to carry the drug throughout the blood stream could be a possible contributor. Investigators wanted to be able to detect, if possible, the onset of phlebitis, and a study was

TABLE 26.43

Differences in Temperature between the Treated Ear and the Untreated Ear

Rabbit	Treatment	Time			
		0	30	60	90
1	Amiodarone	-0.3	-0.2	1.2	3.1
2	Amiodarone	-0.5	2.2	3.3	3.7
3	Amiodarone	-1.1	2.4	2.2	2.7
4	Amiodarone	1.0	1.7	2.1	2.5
5	Amiodarone	-0.3	0.8	0.6	0.9
6	Vehicle	-1.1	-2.2	0.2	0.3
7	Vehicle	-1.4	-0.2	-0.5	-0.1
8	Vehicle	-0.1	-0.1	-0.5	-0.3
9	Vehicle	-0.2	0.1	-0.2	0.4
10	Vehicle	-0.1	-0.2	0.7	-0.3
11	Saline	-1.8	0.2	0.1	0.6
12	Saline	-0.5	0.0	1.0	0.5
13	Saline	-1.0	-0.3	-2.1	0.6
14	Saline	0.4	0.4	-0.7	-0.3
15	Saline	-0.5	0.9	-0.4	-0.3

designed to explore mechanisms for early detection of phlebitis during amiodarone therapy. They believed that a change in tissue temperature near the intravenous administration is an early signal of impending inflammation. To test their belief and to see if amiodarone had an affect, one of three intravenous treatments was administered to one of the ears of a rabbit. The surface area temperature of the treated ear was measured at various time points and compared with the surface area temperature of the rabbit's other (untreated) ear. The treatments were amiodarone with a vehicle solution to carry the drug, the vehicle solution only, and a saline solution. Five rabbits were randomly assigned to each of the three treatments. The difference in the temperature between the treated ear and the untreated ear is used as the response. The data are given in Table 26.43. Give an analysis of these data assuming that the repeated measures satisfy the split-plot-in-time assumption. Please be sure to address the following questions:

- 1) Is there a significant $TIME \times TRT$ interaction? Why or why not?
 - 2) Are there differences between the three treatments? Are there time differences? Explain your answers.
 - 3) Which treatment seems to have the greatest affect on ear temperature difference?
 - 4) The investigators chose to use the difference in temperature between the treated ear and the untreated ear as the variable to analyze. Some statisticians might have suggested that they use the untreated ear temperature as a covariate and then perform an analysis of covariance. You might wish to discuss such an alternative during one of your class sessions.
- 26.2 A study was conducted to evaluate the effect of six diets on cholesterol levels of adult males. The diets are constructed from the six combinations of two Fiber

levels (Low, High) and three Fat levels (Low, Med, and High). Thirty-six persons were randomly assigned to the six diets, six persons per diet. Each person was on her assigned diet for eight months. Each person's cholesterol level was determined every two months (Chol1 = 2 months, Chol2 = 4 months, Chol3 = 6 months, and Chol4 = 8 months). Provide a complete analysis of this data assuming that the split-plot-in-time assumptions hold in order to evaluate the effect of diet and time on cholesterol level. The data are given in Table 26.44.

TABLE 26.44
Cholesterol Levels over Time

Subject	Fat	Fiber	Chol1	Chol2	Chol3	Chol4
1	Low	Low	175		134	138
2	Low	Low	192	169	142	137
3	Low	Low	153	132	115	114
4	Low	Low	204	184	162	164
5	Low	Low	194	173	149	151
6	Low	Low	224	194	164	170
7	Low	High	163	145	132	129
8	Low	High	204	183	163	166
9	Low	High	170	148	129	128
10	Low	High	182	165		
11	Low	High	186	160	140	139
12	Low	High	181	169	150	148
13	Med	Low	223		191	191
14	Med	Low	217	213	204	
15	Med	Low	236	219	205	206
16	Med	Low	201	182	165	168
17	Med	Low	220	205	198	205
18	Med	Low	191	189	183	188
19	Med	High	183	183	168	174
20	Med	High	218		192	183
21	Med	High	200	189	182	181
22	Med	High	210	193	182	
23	Med	High	208	202	192	192
24	Med	High	211	193	185	185
25	High	Low	237	243	244	246
26	High	Low	242	237	231	227
27	High	Low	224	217	219	229
28	High	Low	217	225	237	238
29	High	Low	225	226	227	228
30	High	Low	229	223	226	226
31	High	High	193	202	206	197
32	High	High	189	193	189	185
33	High	High	231	222	217	209
34	High	High	237	233	226	224
35	High	High	210	208	208	206
36	High	High	214	205	206	199

TABLE 26.45
Cheese Quality Measures

Refrigerator	Position	Moisture	Y
1	Top	Low	97
1	Top	High	99
1	Middle	Low	101
1	Middle	High	105
1	Bottom	Low	105
1	Bottom	High	110
2	Top	Low	107
2	Top	High	107
2	Mid	Low	108
2	Mid	High	111
2	Bottom	Low	120
2	Bottom	High	125
3	Top	Low	103
3	Top	High	92
3	Mid	Low	103
3	Mid	High	93
3	Bottom	Low	107
3	Bottom	High	106
4	Top	Low	91
4	Top	High	99
4	Mid	Low	93
4	Mid	High	106
4	Bottom	Low	93
4	Bottom	High	
5	Top	Low	97
5	Top	High	103
5	Mid	Low	93
5	Mid	High	
5	Bottom	Low	95
5	Bottom	High	111
6	Top	Low	96
6	Top	High	94
6	Mid	Low	101
6	Mid	High	106
6	Bottom	Low	107
6	Bottom	High	115
7	Top	Low	101
7	Top	High	100
7	Mid	Low	97
7	Mid	High	98
7	Bottom	Low	100
7	Bottom	High	105
8	Top	Low	93
8	Top	High	
8	Mid	Low	95
8	Mid	High	
8	Bottom	Low	96
8	Bottom	High	

- 26.3 An experiment involves the curing of cheese in refrigerators where each refrigerator is divided into left and right compartments and into top, middle, and bottom sections. The right and left sides are randomly assigned a level of moisture (low or high). There was a belief that the position in the refrigerator might have slightly different temperatures as heat rises, so the refrigerator was divided into thirds (top, middle, bottom), and one package of cheese was placed in each section of the refrigerator; that is, there are six packages of cheese in each refrigerator with one package in each section. This process was repeated for eight different refrigerators. The data are given in Table 26.45.
- 1) Identify each size of experimental unit.
 - 2) Describe the treatment and design structures for each size of experimental unit.
 - 3) Write out a model that can be used to describe the data structure.
 - 4) Carry out analysis of this data set and make any necessary comparisons.

27

Analysis of Repeated Measures Experiments When the Ideal Conditions Are Not Satisfied

Repeated measures designs involve one or more steps where a researcher cannot randomly assign the levels of one or more of the factors of to an experimental unit. The use of time as a factor is the most common experimental situation where one cannot use randomization. For example, when data are collected on the same experimental unit at several time points, one cannot randomize the order of the time points. Time 1 must be first, time 2 must be second, and so on. This nonrandom assignment of the repeated measures factor influences the variances and covariances between the experimental units, and the ideal conditions described in Chapter 26 may not be valid. This chapter presents strategies for analyzing data from repeated measures experiments when the ideal conditions given in Chapter 26 do not hold. In addition, procedures are described that allow one to check to see whether the ideal conditions described in Chapter 26 are satisfied.

27.1 Introduction

Consider an experimental situation similar to that described in Table 26.1. Let y_{ijk} represent the observed response for subject k in treatment group i at time j and let

$$\mathbf{y}_{ik} = \begin{bmatrix} y_{i1k} \\ y_{i2k} \\ \vdots \\ y_{ipk} \end{bmatrix}$$

be the vector of responses for subject k in treatment group i .

A model that can be used to describe these data is

$$y_{ijk} = \mu + \alpha_i + \tau_j + \gamma_{ij} + \varepsilon_{ijk}^* \quad i = 1, 2, \dots, t; \quad j = 1, 2, \dots, p; \quad k = 1, 2, \dots, n_i$$

Let

$$\boldsymbol{\varepsilon}_{ik}^* = \begin{bmatrix} \varepsilon_{i1k}^* \\ \varepsilon_{i2k}^* \\ \vdots \\ \varepsilon_{ipk}^* \end{bmatrix}$$

be the vector of errors for subject k in treatment group i . Suppose that ε_{ik}^* are distributed independently and identically as p -variate multivariate normal distributions with mean, $\mathbf{0}$, and covariance matrix, $\boldsymbol{\Sigma}$. That is, $\varepsilon_{ik}^* \sim i.i.d. N(\mathbf{0}, \boldsymbol{\Sigma})$, $i = 1, 2, \dots, t$; $k = 1, 2, \dots, n_i$.

Let

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}$$

represent the covariance matrix of the vector of repeated measures.

Definition 27.1: If $\boldsymbol{\Sigma} = \lambda I_t + \boldsymbol{\eta} j' + j \boldsymbol{\eta}'$, where j is a $p \times 1$ vector of ones, and $\boldsymbol{\eta}$ is a $p \times 1$ vector of constants, the repeated measures are said to satisfy conditions known as the Huynh–Feldt (H–F) conditions (see Huynh and Feldt, 1970).

When Definition 27.1 holds, $\boldsymbol{\Sigma}$ has the form

$$\boldsymbol{\Sigma} = \begin{bmatrix} \lambda + 2\eta_1 & \eta_1 + \eta_2 & \cdots & \eta_1 + \eta_p \\ \eta_2 + \eta_1 & \lambda + 2\eta_2 & \cdots & \eta_2 + \eta_p \\ \vdots & \vdots & \ddots & \vdots \\ \eta_p + \eta_1 & \eta_p + \eta_2 & \cdots & \lambda + 2\eta_p \end{bmatrix}$$

A special case of the H–F conditions is when the repeated measures possess a compound symmetry covariance structure, then

$$\boldsymbol{\Sigma} = \sigma^2 \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{bmatrix}$$

for some σ^2 and ρ . Note that the ideal conditions described for the split-plot-in-time analysis discussed in Section 26.1 are a special case of the covariance matrix of the repeated measures possessing compound symmetry with $\sigma^2 = \sigma_\delta^2 + \sigma_\epsilon^2$ and $\rho = \sigma_\delta^2 / (\sigma_\delta^2 + \sigma_\epsilon^2)$. Compound symmetry structure is more general than the split-plot-in-time structure since ρ can be negative in the compound symmetry structure.

Remark: If the H–F conditions are satisfied, then many of the important questions involving time comparisons can be answered by analyzing the repeated measures experiment in the same way one analyzes a repeated measures experiment satisfying the ideal conditions given in Section 26.1, and if the repeated measures satisfy compound symmetry, then the split-plot-in-time methods of analysis given in Chapter 26 can be used. In particular, the split-plot-in-time tests for the *Time* main effect and the *Time* × *Trt* interaction effect can be shown to be statistically valid if and only if the repeated measures satisfy the H–F conditions given in Definition 27.1. Furthermore, contrasts that compare *Time* main effects and contrasts that compare *Time* effects within a specified value of the *Trt* variable are statistically valid when using the split-plot-in-time analysis. If one is interested in the two-way means or marginal means, the split-plot-in-time analysis gives the correct estimates of the means, but the standard errors will be incorrect.

Remark: If the repeated measures possess compound symmetry with $\rho > 0$, then all of the results given by a split-plot-in-time analysis will be correct. In this case, one can say that $\varepsilon_{ik}^* = \delta_{ik} + \varepsilon_{ijk}$ where $\delta_{ik} \sim i.i.d. N(0, \sigma_\delta^2)$, $\varepsilon_{ijk} \sim i.i.d. N(0, \sigma_\varepsilon^2)$, and the δ_{ik} and the ε_{ijk} are independent.

Question: What if the H–F conditions are not satisfied?

In the case where the H–F conditions are not satisfied, several methods of analysis can be considered. One approach that is always appropriate is to treat the vector of repeated measures as a multivariate response vector and use multivariate analysis of variance (MANOVA) methods. A second is to use the split-plot-in-time analysis, but adjust the *p*-values by adjusting the degrees of freedom corresponding to relevant effect mean squares. The third approach is to use a mixed model approach available in the SAS®-Mixed and the SAS®-Glimix procedures and model the covariance structure.

MANOVA methods are described in Section 27.2, adjusted degrees of freedom methods are discussed in Section 27.3, and Mixed model methods will be discussed in Section 27.4.

27.2 MANOVA Methods

This section considers using multivariate analysis of variance methods to analyze repeated measures experiments. These methods are always appropriate when

$$\varepsilon_{ik}^* \sim i.i.d. N(\mathbf{0}, \boldsymbol{\Sigma}), \quad i = 1, 2, \dots, t; \quad k = 1, 2, \dots, n_i$$

The MANOVA methods can only be applied to experiments where either all repeated measures on a given experimental unit are present or all are missing since any experimental unit that has a missing value for one or more of the repeated measures will be automatically deleted from the analysis by the statistical software being used. This is not a problem with the other two approaches. The number of experimental units assigned to each treatment need not be balanced. To get the most out of this section, one should be able to understand the matrix form of the model discussed in Chapter 6.

A general multivariate model for a repeated measures experiment that has a structure similar to that given in Table 26.1 is a generalization of the matrix form of the model defined in Equation 26.1. The multivariate model can be expressed as

$$\mathbf{Y} = \mathbf{XB} + \mathbf{E} \quad (27.1)$$

where \mathbf{Y} denotes all of the data measured in the experiment. Each row of the data matrix \mathbf{Y} corresponds to a particular experimental unit, and each column corresponds to one of the repeated measures. Thus \mathbf{Y} is an $N \times p$ matrix where $N = \sum_{i=1}^t n_i$. The matrix \mathbf{X} is an $N \times r$ design matrix assumed to be of rank t . Each column of \mathbf{B} is an $r \times 1$ vector of unknown parameters with each column corresponding to a particular repeated measure. The matrix \mathbf{E} is an $N \times p$ matrix of unobservable random errors. It is assumed that the rows of \mathbf{E} are independently distributed $N(\mathbf{0}, \Sigma)$. Thus, while the rows in \mathbf{E} are independent, the elements in a row may be correlated with one another and may have different variances.

For the multivariate model (27.1), one can test general hypotheses of the form

$$H_0: \mathbf{CBM} = \mathbf{0} \text{ vs } H_a: \mathbf{CBM} \neq \mathbf{0} \quad (27.2)$$

where \mathbf{C} is a $g \times r$ matrix of rank g and \mathbf{M} is a $p \times q$ matrix of rank q .

To test the hypothesis in Equation 27.2, one first needs the least squares estimates of the parameters in \mathbf{B} and an observed residual sum-of-squares and cross-products matrix. These are denoted by $\hat{\mathbf{B}}$ and $\hat{\mathbf{E}}$, respectively, and given by

$$\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \quad \text{and} \quad \hat{\mathbf{E}} = \mathbf{Y}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{Y} \quad (27.3)$$

A likelihood ratio test statistic for testing the hypothesis in Equation 27.2 is given by

$$\Lambda = \frac{|\mathbf{R}|}{|\mathbf{H} + \mathbf{R}|} \quad (27.4)$$

where

$$\mathbf{R} = \mathbf{M}'\hat{\mathbf{E}}\mathbf{M}, \quad \mathbf{H} = \mathbf{M}'\hat{\mathbf{B}}\mathbf{C}'\left[\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'\right]^{-1}\mathbf{C}\hat{\mathbf{B}}\mathbf{M}$$

and $|\mathbf{W}|$ denotes the determinant of the matrix \mathbf{W} .

The statistic Λ is called Wilks' likelihood ratio criterion (Morrison, 1976). The sampling distribution of Λ is quite complicated, but for most practical purposes, an approximate α -level test can be obtained by rejecting H_0 when

$$-\left(N - t - \frac{|q - g| + 1}{2}\right)\log_e(\Lambda) > \chi_{\alpha, qg}^2$$

A better approximation, to be used only when both q and g are greater than 2, is to reject H_0 when

$$F > F_{\alpha, qg, ab-c}$$

where

$$F = \frac{(1 - \Lambda^{1/b})(ab - c)}{qg\Lambda^{1/b}}$$

and where

$$a = N - t - \frac{|q - s| + 1}{2}$$

$$b = \left(\frac{q^2 s^2 - 4}{q^2 + s^2 - 5} \right)^{1/2} \quad (27.5)$$

$$c = \frac{qs - 2}{2}$$

and

$$s = \min(q, g)$$

Exact F tests for Equation 27.2 exist whenever $q = 1, 2$ or whenever $g = 1, 2$. These tests are as follows:

1) For $g = 1$ and any q , reject H_0 if

$$F = \left(\frac{1 - \Lambda}{\Lambda} \right) \left(\frac{N - t - q + 1}{q} \right) > F_{\alpha, q, N-t-q+1} \quad (27.6)$$

2) For $q = 1$ and any g , reject H_0 if

$$F = \left(\frac{1 - \Lambda}{\Lambda} \right) \left(\frac{N - t}{g} \right) > F_{\alpha, g, N-t} \quad (27.7)$$

3) For $g = 2$ and any $q > 1$, reject H_0 if

$$F = \left(\frac{1 - \sqrt{\Lambda}}{\sqrt{\Lambda}} \right) \left(\frac{N - t - q + 1}{q} \right) > F_{\alpha, 2q, 2(N-t-q+1)} \quad (27.8)$$

4) For $q = 2$ and any $g > 1$, reject H_0 if

$$F = \left(\frac{1 - \sqrt{\Lambda}}{\sqrt{\Lambda}} \right) \left(\frac{N - t - 1}{g} \right) > F_{\alpha, 2g, 2(N-t-1)} \quad (27.9)$$

One drawback to the multivariate method is that one must have $p < N - t$. When $p \geq N - t$, it is often possible to combine adjacent repeated measures into p^* new variables or to analyze only a size p^* subset of the repeated measures where $p^* < N - t$.

To illustrate the analysis described in this section, consider an experiment conducted to study the differences among four varieties of sorghum and five fertilizer levels on a leaf area index where the four varieties of sorghum are denoted by V_1 , V_2 , V_3 , and V_4 and the five fertilizer levels are denoted by 1, 2, 3, 4, and 5. Also suppose that these 20 variety \times fertilizer combinations were randomly assigned to 20 plots in a field. For this example, it is assumed that there is no interaction between fertilizer levels and varieties and that a basic two-way additive model can be used to analyze the data. Finally, assume that leaf area index measurements are made on each variety \times fertilizer plot at weekly intervals for five weeks beginning two weeks after emergence of the plant. The data obtained are given in Table 27.1.

TABLE 27.1
Leaf Area Index on Four Sorghum Varieties

Time						
Variety	Fertilizer	Week 1	Week 2	Week 3	Week 4	Week 5
V_1	1	5.00	4.84	4.02	3.75	3.13
	2	4.42	4.30	3.67	3.23	2.83
	3	4.42	4.10	3.46	3.09	2.82
	4	4.01	3.89	3.21	2.89	2.56
	5	3.36	3.10	2.67	2.47	2.16
V_2	1	5.82	5.60	5.05	4.72	4.46
	2	5.73	5.59	5.00	4.65	4.42
	3	5.31	5.19	4.86	4.44	4.22
	4	4.92	4.66	4.56	4.16	3.99
	5	3.96	3.86	3.50	3.13	2.95
V_3	1	5.65	5.97	5.27	5.07	4.52
	2	5.39	5.49	5.08	4.87	4.32
	3	5.15	5.28	4.93	4.67	4.15
	4	4.50	4.89	4.74	4.49	4.10
	5	3.75	3.74	3.55	3.28	3.00
V_4	1	5.86	5.60	5.37	5.00	4.37
	2	5.82	5.55	5.29	4.95	4.07
	3	5.26	5.06	4.76	4.48	3.94
	4	4.87	4.75	4.55	4.33	3.83
	5	3.96	3.76	3.56	3.18	2.96

For the data in Table 27.1, the data matrix is

$$\boldsymbol{Y} = \begin{bmatrix} 5.00 & 4.84 & 4.02 & 3.75 & 3.13 \\ 4.42 & 4.30 & 3.67 & 3.23 & 2.83 \\ 4.42 & 4.10 & 3.46 & 3.09 & 2.82 \\ 4.01 & 3.89 & 3.21 & 2.89 & 2.56 \\ 3.36 & 3.10 & 2.67 & 2.47 & 2.16 \\ 5.82 & 5.60 & 5.05 & 4.72 & 4.46 \\ 5.73 & 5.59 & 5.00 & 4.65 & 4.42 \\ 5.31 & 5.19 & 4.86 & 4.44 & 4.22 \\ 4.92 & 4.66 & 4.56 & 4.16 & 3.99 \\ 3.96 & 3.86 & 3.50 & 3.13 & 2.95 \\ 5.65 & 5.97 & 5.27 & 5.07 & 4.52 \\ 5.39 & 5.49 & 5.08 & 4.87 & 4.32 \\ 5.15 & 5.28 & 4.93 & 4.67 & 4.15 \\ 4.50 & 4.89 & 4.74 & 4.49 & 4.10 \\ 3.75 & 3.74 & 3.55 & 3.28 & 3.00 \\ 5.86 & 5.60 & 5.37 & 5.00 & 4.37 \\ 5.82 & 5.55 & 5.29 & 4.95 & 4.07 \\ 5.26 & 5.06 & 4.76 & 4.48 & 3.94 \\ 4.87 & 4.75 & 4.55 & 4.33 & 3.83 \\ 3.96 & 3.76 & 3.56 & 3.18 & 2.96 \end{bmatrix}$$

The matrix of parameters is given by

$$\boldsymbol{B} = \begin{bmatrix} \mu^{(1)} & \mu^{(2)} & \mu^{(3)} & \mu^{(4)} & \mu^{(5)} \\ \tau_1^{(1)} & \tau_1^{(2)} & \tau_1^{(3)} & \tau_1^{(4)} & \tau_1^{(5)} \\ \tau_2^{(1)} & \tau_2^{(2)} & \tau_2^{(3)} & \tau_2^{(4)} & \tau_2^{(5)} \\ \tau_3^{(1)} & \tau_3^{(2)} & \tau_3^{(3)} & \tau_3^{(4)} & \tau_3^{(5)} \\ \tau_4^{(1)} & \tau_4^{(2)} & \tau_4^{(3)} & \tau_4^{(4)} & \tau_4^{(5)} \\ \beta_1^{(1)} & \beta_1^{(2)} & \beta_1^{(3)} & \beta_1^{(4)} & \beta_1^{(5)} \\ \beta_2^{(1)} & \beta_2^{(2)} & \beta_2^{(3)} & \beta_2^{(4)} & \beta_2^{(5)} \\ \beta_3^{(1)} & \beta_3^{(2)} & \beta_3^{(3)} & \beta_3^{(4)} & \beta_3^{(5)} \\ \beta_4^{(1)} & \beta_4^{(2)} & \beta_4^{(3)} & \beta_4^{(4)} & \beta_4^{(5)} \\ \beta_5^{(1)} & \beta_5^{(2)} & \beta_5^{(3)} & \beta_5^{(4)} & \beta_5^{(5)} \end{bmatrix}$$

where the τ correspond to the different varieties and the β correspond to the different levels of fertilizer, The design matrix is

$$X = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Note that each column of B represents the parameters needed for a two-way additive model for the j th response column of the data matrix, Y , $j = 1, 2, 3, 4, 5$. The design matrix X is a 20×10 matrix and has rank equal to 8; thus $N = 20$, $p = 10$, and $t = 8$. The value of \hat{B} given by Equation 27.3 is

$$\hat{B} = \begin{bmatrix} 3.350 & 3.283 & 3.003 & 2.788 & 2.150 \\ 0.222 & 0.106 & -0.198 & -0.260 & -0.312 \\ 1.128 & 1.040 & 0.990 & 0.874 & 0.996 \\ 0.868 & 1.134 & 1.110 & 1.130 & 1.006 \\ 1.134 & 1.004 & 1.102 & 1.042 & 0.822 \\ 1.395 & 1.398 & 1.173 & 1.115 & 0.982 \\ 1.152 & 1.128 & 1.006 & 0.940 & 0.772 \\ 0.847 & 0.803 & 0.748 & 0.685 & 0.645 \\ 0.387 & 0.443 & 0.511 & 0.483 & 0.482 \\ -0.430 & -0.489 & -0.434 & -0.470 & -0.370 \end{bmatrix}$$

and the value of \hat{E} in Equation 27.3 is

$$\hat{E} = \begin{bmatrix} 0.237 & 0.171 & 0.162 & 0.228 & 0.129 \\ 0.171 & 0.247 & 0.163 & 0.231 & 0.135 \\ 0.162 & 0.163 & 0.268 & 0.303 & 0.184 \\ 0.228 & 0.231 & 0.303 & 0.392 & 0.241 \\ 0.129 & 0.135 & 0.184 & 0.241 & 0.247 \end{bmatrix}$$

The test for equal variety main effect means is obtained from Equation 27.4 by taking

$$C = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad M = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Then $\Lambda = 0.04345$ with $g = 3$ and $q = 1$. Since $q = 1$, one can use Equation 27.7, and get

$$F = \frac{1 - 0.04345}{0.04345} \times \frac{12}{3} = 88.06$$

with 3 and 12 degrees of freedom. The observed significance level $\hat{\alpha}$ is less than 0.0001.

The test for equal time main effect means is obtained from Equation 27.4 by taking

$$C = \begin{bmatrix} 1 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{bmatrix} \quad \text{and} \quad M = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

One gets $\Lambda = 0.04345$ with $g = 1$ and $q = 4$. Since $g = 1$, one can use Equation 27.6, and get

$$F = \frac{1 - 0.00502}{0.00502} \times \frac{9}{4} = 445.96$$

with 4 and 9 degrees of freedom. The observed significance level $\hat{\alpha}$ is less than 0.0001.

The test for Variety \times Time interaction is obtained from Equation 27.6 by taking

$$C = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad M = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

One gets $\Lambda = 0.01426$ with $g = 3$ and $q = 4$. Since both g and q are greater than 1, one can use Equation 27.5 with $s = 3$, $a = 11$, $b = 2.646$, and $c = 5$. Then

$$F = \frac{(1 - 0.01426^{1/2.646})(11 \times 2.646 - 5)}{4 \times 3 \times 0.01426^{1/2.646}} = \frac{(1 - 0.2006)(24.1060)}{12(0.2006)} = 8.00$$

with 12 and 24.1 degrees of freedom. The observed significance level $\hat{\alpha}$ is less than 0.0001.

Note that the test for equal variety means is the same as that obtained by performing the split-plot-in-time analysis. It is the only test of the three that is the same as the corresponding tests obtained by a split-plot-in-time analysis. Test statistics for the *fertilizer* main effect ($F = 94.36$) and the *fertilizer* \times *time* interaction effect ($F = 1.91$) can be obtained in a similar manner.

Many special functions of the parameters in \mathbf{B} are estimable, and inferences can be made on those estimable functions. Most of the interesting linear functions of the parameters in \mathbf{B} can be written in the form $\mathbf{c}'\mathbf{B}\mathbf{m}$ where \mathbf{c} is an $r \times 1$ vector and \mathbf{m} is a $p \times 1$ vector. As in Chapter 6, $\mathbf{c}'\mathbf{B}\mathbf{m}$ is estimable if and only if there exists a vector \mathbf{u} such that $\mathbf{X}'\mathbf{X}\mathbf{u} = \mathbf{c}$. There are no restrictions on \mathbf{m} .

The best estimate of $\mathbf{c}'\mathbf{B}\mathbf{m}$ is $\mathbf{c}'\hat{\mathbf{B}}\mathbf{m}$ where $\hat{\mathbf{B}}$ is given in Equation 27.3. The estimated standard error of $\mathbf{c}'\hat{\mathbf{B}}\mathbf{m}$ is given by

$$\widehat{s.e.}(\mathbf{c}'\hat{\mathbf{B}}\mathbf{m}) = \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c} \cdot \frac{\mathbf{m}'\hat{\mathbf{E}}\mathbf{m}}{N-t}}$$

and its corresponding degrees of freedom are $N-t$. Thus a $(1-\alpha)100\%$ confidence interval for $\mathbf{c}'\mathbf{B}\mathbf{m}$ is given by

$$\mathbf{c}'\hat{\mathbf{B}}\mathbf{m} \pm t_{\alpha/2, N-t} \left[\widehat{s.e.}(\mathbf{c}'\hat{\mathbf{B}}\mathbf{m}) \right] \quad (27.10)$$

and a t -statistic for testing $H_0: \mathbf{c}'\mathbf{B}\mathbf{m} = a_0$ is given by

$$t = \frac{\mathbf{c}'\hat{\mathbf{B}}\mathbf{m} - a_0}{\widehat{s.e.}(\mathbf{c}'\hat{\mathbf{B}}\mathbf{m})} \quad (27.11)$$

If $t > t_{\alpha/2, N-t}$, then H_0 is rejected.

As an example, consider estimating the V_1 marginal mean. For this marginal mean,

$$\mathbf{c}' = [1 \ 1 \ 0 \ 0 \ 0 \ 0.2 \ 0.2 \ 0.2 \ 0.2 \ 0.2] \quad \text{and} \quad \mathbf{m} = \begin{bmatrix} 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \end{bmatrix}$$

The value of $\mathbf{c}'\hat{\mathbf{B}}\mathbf{m}$ is 3.496, and its estimated standard error is 0.0594. A 95% confidence interval for the V_1 marginal mean is found from Equation 27.10 as $3.496 \pm (2.179)(0.0594)$.

As a second example, consider estimating the Time 1 marginal mean for the data in Table 27.2. For this marginal mean,

$$c' = [1 \ 0.25 \ 0.25 \ 0.25 \ 0.25 \ 0.2 \ 0.2 \ 0.2 \ 0.2 \ 0.2] \text{ and } m = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

The value of $c'\hat{B}m$ is 4.858, and its estimated standard error is 0.0315.

As a third example, consider estimating the difference between the V_1 and V_2 marginal means. Here

$$c' = [0 \ 1 \ -1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0] \text{ and } m = \begin{bmatrix} 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \end{bmatrix}$$

The value of $c'\hat{B}m$ is -1.094, and its estimated standard error is 0.0840. A t -statistic for comparing these two marginal means is $t = -1.094/0.0840 = 13.02$ and its observed significance level is $\hat{\alpha} < 0.0001$.

Many of the inference results given above can be obtained from the SAS®-GLM procedure when one uses a MANOVA option along with its $M =$ option. To illustrate, the data in Table 27.1 are reanalyzed using the SAS commands given in Table 27.2. The first MANOVA option is used to obtain \hat{E} , and the second MANOVA option is used to obtain tests that compare *Variety* main effect means and *Fertilizer* main effect means [note that $M = (0.2 \ 0.2 \ 0.2 \ 0.2 \ 0.2)$ tells the MANOVA option to average across the five repeated time measures]. Tests for *Fertilizer* \times *Time* interaction and *Variety* \times *Time* interaction are obtained from the third MANOVA option where M is a 4×5 matrix of time contrasts.

TABLE 27.2

SAS-GLM Code to Analyze the Data in Table 27.1 Using MANOVA

```

PROC GLM DATA=LAI;
  CLASSES FERTILIZER VARIETY;
  MODEL LAI1-LAI5=FERTILIZER VARIETY/NOUNI;
  MANOVA /PRINT=;
  MANOVA H=FERTILIZER VARIETY
    M=(.2 .2 .2 .2 .2);
  MANOVA H=FERTILIZER VARIETY M=(1 -1 0 0 0,
                                1 0 -1 0 0,
                                1 0 0 -1 0,
                                1 0 0 0 -1);
  CONTRAST 'V1-V2' VARIETY 1 -1 0 0;
RUN;

```

Table 27.3 gives the value \hat{E} obtained from the MANOVA analysis in Table 27.2. Table 27.4 gives the MANOVA tests for *Fertilizer* and *Variety* main effects, and Table 27.5 gives the MANOVA interaction tests.

MANOVA and CONTRAST options can both be used to test hypotheses of the form $H_0: c'Bm = 0$. For example, by adding the following three statements to the SAS commands in Table 27.2

```
CONTRAST 'V1-V2' VARIETY 1 -1 0 0;
MANOVA M=(.2 .2 .2 .2 .2) / PRINTE;
MANOVA M=(1 0 0 0 0) / PRINTE;
```

one can obtain a Wilks statistic that compares the *Variety* 1 and 2 main effect means to one another from the first MANOVA statement, and a Wilks statistic that compares *Variety* 1 to *Variety* 2 at *Time* 1 from the second MANOVA statement. By including the PRINTE option in each of the above MANOVA commands, one also gets values of $m'\hat{E}m$. Table 27.6 gives a test statistic ($F = 169.4$) that compares the *Variety* 1 main effect mean to the *Variety* 2 main effect mean. Note that 169.4 is the square of the *t*-statistic given earlier in this section. Also note that $m'\hat{E}m = 0.2115$ when $m' = [0.2 \ 0.2 \ 0.2 \ 0.2 \ 0.2]$.

TABLE 27.3

Error Sums of Squares and Cross-Products Matrix

E = Error SSCP Matrix

	LAI1	LAI2	LAI3	LAI4	LAI5
LAI1	0.23699	0.171285	0.162175	0.22799	0.129105
LAI2	0.171285	0.24678	0.16292	0.231465	0.135435
LAI3	0.162175	0.16292	0.26763	0.30261	0.184435
LAI4	0.22799	0.231465	0.30261	0.39232	0.24145
LAI5	0.129105	0.135435	0.184435	0.24145	0.24683

TABLE 27.4

Tests for Fertilizer and Variety Main Effects

M Matrix Describing Transformed Variables

	LAI1	LAI2	LAI3	LAI4	LAI5
MVAR1	0.2	0.2	0.2	0.2	0.2

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall Fertilizer Effect
on the Variables Defined by the *M* Matrix Transformation *H* = Type III SSCP Matrix for Fertilizer
E = Error SSCP Matrix *S* = 1 *M* = 1 *N* = 5

Statistic	Value	F-Value	Num df	Den df	Pr > F
Wilks's lambda	0.03081325	94.36	4	12	<0.0001

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall Variety Effect
on the Variables Defined by the *M* Matrix Transformation *H* = Type III SSCP Matrix for Variety
E = Error SSCP Matrix *S* = 1 *M* = 0.5 *N* = 5

Statistic	Value	F-Value	Num df	Den df	Pr > F
Wilks's lambda	0.04345320	88.05	3	12	<0.0001

TABLE 27.5

MANOVA Interaction Tests

M Matrix Describing Transformed Variables

	LAI1	LAI2	LAI3	LAI4	LAI5
MVAR1	1	-1	0	0	0
MVAR2	1	0	-1	0	0
MVAR3	1	0	0	-1	0
MVAR4	1	0	0	0	-1

MANOVA Test Criteria and F Approximations for the Hypothesis of No Overall Fertilizer Effect
on the Variables Defined by the M Matrix Transformation H = Type III SSCP Matrix for Fertilizer
E = Error SSCP Matrix S = 4 M = -0.5 N = 3.5

Statistic	Value	F-Value	Num df	Den df	Pr > F
Wilks's lambda	0.10535731	1.91	16	28.133	0.0640

MANOVA Test Criteria and F Approximations for the Hypothesis of No Overall Variety Effect
on the Variables Defined by the M Matrix Transformation H = Type III SSCP Matrix for Variety
E = Error SSCP Matrix S = 3 M = 0 N = 3.5

Statistic	Value	F-Value	Num df	Den df	Pr > F
Wilks's lambda	0.01425985	8.00	12	24.103	<0.0001

TABLE 27.6

Test Comparing the Variety 1 and Variety 2 Main Effect Means

M Matrix Describing Transformed Variables

	LAI1	LAI2	LAI3	LAI4	LAI5
MVAR1	0.2	0.2	0.2	0.2	0.2

E = Error SSCP Matrix

MVAR1	0.2115316
-------	-----------

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall V₁ – V₂ Effect
on the Variables Defined by the M Matrix Transformation H = Contrast SSCP Matrix for
V₁ – V₂ E = Error SSCP Matrix S = 1 M = -0.5 N = 5

Statistic	Value	F-Value	Num df	Den df	Pr > F
Wilks's lambda	0.06602890	169.74	1	12	<0.0001

27.3 p-Value Adjustment Methods

Recall that the analysis of a repeated measures experiment when the repeated measures satisfy the compound symmetry assumptions can be obtained by the split-plot-in-time methods described in Chapter 26. If one is only considering tests that involve differences

in the time factor, then the split-plot-in-time tests are also valid when the repeated measures satisfy the H-F conditions. A second approach to analyzing repeated measures experiments that have been shown to be more powerful than analyses based on the MANOVA approach described in Section 27.2. This approach is to perform the split-plot-in-time analysis with an adjustment to the numerator and denominator degrees of freedom of the test statistics should the H-F conditions not be satisfied, resulting in adjusted p -values.

Suppose

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}$$

is the covariance matrix of the repeated measures. Let

$$\theta = \frac{p^2(\bar{\sigma}_{ii} - \bar{\sigma}_{..})^2}{(p-1) \left[\sum_{i=1}^p \sum_{j=1}^p \sigma_{ij}^2 - 2p \sum_{i=1}^p \bar{\sigma}_{i.}^2 + p^2 \bar{\sigma}_{..}^2 \right]} \quad (27.12)$$

where

$$\bar{\sigma}_{ii} = \frac{1}{p} \sum_{j=1}^p \sigma_{jj}, \quad \bar{\sigma}_{i.} = \frac{1}{p} \sum_{j=1}^p \sigma_{ij}, \quad \text{and} \quad \bar{\sigma}_{..} = \frac{1}{p^2} \sum_{i=1}^p \sum_{j=1}^p \sigma_{ij}$$

Box (1954) proposed θ as a measure of how far $\boldsymbol{\Sigma}$ deviates from compound symmetry and he showed that $1/(p-1) < \theta < 1$. The smaller the value of θ , the further $\boldsymbol{\Sigma}$ is from compound symmetry. Suppose that F_{TIME} and $F_{TIME \times TRT}$ are the split-plot-in-time test statistics for *Time* main effect and *Time* \times *Trt* interaction, respectively, for a repeated measures scenario such as that described by Table 26.1. Box showed that, when $\boldsymbol{\Sigma}$ deviates from compound symmetry by θ , then

$$F_{TIME} \text{ is approximately distributed as } F[\theta(p-1), \theta(N-t)(p-1)] \quad (27.13)$$

when there is no time effect, and

$$F_{TIME \times TRT} \text{ is approximately distributed as } F[\theta(t-1)(p-1), \theta(N-t)(p-1)] \quad (27.14)$$

when there is no interaction effect.

Unfortunately, θ is not known since the σ_{ij} are not known. If the σ_{ij} can be estimated, then θ can be estimated.

Three estimates of θ have been proposed. The first of these is called Box's conservative estimate. Box (1954) suggested that a conservative approach is to take θ as its minimum possible value. That is, take $\theta = 1/(p-1)$. This is probably too conservative to be recommended.

A second possibility was proposed by Greenhouse and Geisser (1959). Let $\hat{Q} = C\hat{\Sigma}C'$ where C is any $(p-1) \times p$ matrix satisfying $Cj = 0$ and $CC' = I_{p-1}$ where j is a $p \times 1$ vector of 1's and I_{p-1} is the $(p-1) \times (p-1)$ identity matrix. Then

$$\hat{\theta} = \frac{\left(\sum_{i=1}^{p-1} \hat{q}_{ii} \right)^2}{(p-1) \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} \hat{q}_{ij}^2} = \frac{\left[\text{tr}(\hat{Q}) \right]^2}{(p-1) \text{tr}(\hat{Q}\hat{Q}')}}$$

is the Greenhouse and Geisser (G-G) estimate of θ where $\text{tr}(B)$ is the trace of the matrix B .

A third method of estimating θ was suggested by Huynh and Feldt (1976). Their approach estimates θ by

$$\tilde{\theta} = \frac{N(p-1)\hat{\theta} - 2}{(p-1)(N-r-(p-1)\hat{\theta})}$$

where N = the total sample size and $N-r$ = degrees of freedom for error in a split-plot-in-time analysis.

It can be noted that Box's correction is the most conservative of the three adjustment methods and that the H-F correction is the least conservative of the three. It should also be noted that, if any of the estimates of θ are greater than 1, then 1 is substituted for θ in Equations 27.13 and 27.14 when computing adjusted p -values because one would not want to increase the degrees of freedom for any effect.

It can be noted that if Σ satisfies the H-F conditions, then $C\Sigma C' = \lambda I$ for some λ . When this is true, one says that $C\Sigma C'$ satisfies a sphericity condition. A likelihood ratio test of $H_0: C\Sigma C' = \lambda I$ is available. The likelihood ratio test statistic is given by

$$\Lambda = \frac{|\hat{Q}|}{\left[\frac{1}{p-1} \text{tr}(\hat{Q}) \right]^{p-1}} \quad (27.15)$$

and one rejects H_0 if $-2 \log_e(\Lambda) > \chi_{\alpha, p(p-1)/2-1}^2$. There are many C matrices that satisfy $Cj = 0$ and $CC' = I_{p-1}$. It should be noted that the value of Λ does not depend on which possible C matrix might be selected.

As an example, consider the data the Table 27.1. For these data

$$\hat{E} = \begin{bmatrix} 0.237 & 0.171 & 0.162 & 0.228 & 0.129 \\ 0.171 & 0.247 & 0.163 & 0.231 & 0.135 \\ 0.162 & 0.163 & 0.268 & 0.303 & 0.184 \\ 0.228 & 0.231 & 0.303 & 0.392 & 0.241 \\ 0.129 & 0.135 & 0.184 & 0.241 & 0.247 \end{bmatrix}$$

and is based on 12 degrees of freedom.

Therefore

$$\hat{\Sigma} = \frac{1}{12} \hat{E} = \frac{1}{12} \begin{bmatrix} 0.237 & 0.171 & 0.162 & 0.228 & 0.129 \\ 0.171 & 0.247 & 0.163 & 0.231 & 0.135 \\ 0.162 & 0.163 & 0.268 & 0.303 & 0.184 \\ 0.228 & 0.231 & 0.303 & 0.392 & 0.241 \\ 0.129 & 0.135 & 0.184 & 0.241 & 0.247 \end{bmatrix}$$

$$= \begin{bmatrix} 0.0198 & 0.0143 & 0.0135 & 0.0190 & 0.0108 \\ 0.0143 & 0.0206 & 0.0136 & 0.0193 & 0.0113 \\ 0.0135 & 0.0136 & 0.0223 & 0.0253 & 0.0153 \\ 0.0190 & 0.0193 & 0.0253 & 0.0327 & 0.0201 \\ 0.0108 & 0.0113 & 0.0153 & 0.0201 & 0.0206 \end{bmatrix}$$

Taking

$$C = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & 0 & 0 \\ \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & -\frac{3}{\sqrt{12}} & 0 \\ \frac{1}{\sqrt{20}} & \frac{1}{\sqrt{20}} & \frac{1}{\sqrt{20}} & \frac{1}{\sqrt{20}} & -\frac{4}{\sqrt{20}} \end{bmatrix}$$

one gets

$$Q = C \hat{\Sigma} C' = \begin{bmatrix} 0.00592 & -0.00019 & -0.00003 & 0.00013 \\ -0.00019 & 0.00830 & 0.00399 & 0.00178 \\ -0.00003 & 0.00399 & 0.00486 & 0.00078 \\ 0.00013 & 0.00178 & 0.00078 & 0.00875 \end{bmatrix}$$

One then gets $\hat{\theta} = 0.795$ and $\tilde{\theta} = 1.747$. The value of the likelihood ratio test statistic for testing $H_0: C\Sigma C' = \lambda I$ is

$$\Lambda = \frac{|\hat{Q}|}{\left[\frac{1}{p-1} \text{tr}(\hat{Q}) \right]^{p-1}} = \frac{1.20895(10)^{-9}}{\left(\frac{0.027831}{4} \right)^4} = 0.5159$$

Thus $-2 \log_e(\Lambda) = 1.32$, which would be compared to $\chi^2_{0.05,9} = 16.919$. Thus $H_0: C\Sigma C' = \lambda I$ cannot be rejected for these data.

Suppose the data in Table 27.1 are analyzed using a split-plot-in-time analysis using SAS-GLM procedure. The SAS commands used are given in Table 27.7.

TABLE 27.7

SAS-GLM Code to Analyze the Data in Table 27.1 Using a Split-Plot-In-Time Analysis

```

DATA LAI2; SET LAI;
DROP LAI1-LAI5;
TIME=1; LAI=LAI1; OUTPUT;
TIME=2; LAI=LAI2; OUTPUT;
TIME=3; LAI=LAI3; OUTPUT;
TIME=4; LAI=LAI4; OUTPUT;
TIME=5; LAI=LAI5; OUTPUT;

ODS RTF FILE='C:\TEMP.RTF';
PROC GLM DATA=LAI2;
CLASSES FERTILIZER VARIETY TIME;
MODEL LAI=FERTILIZER VARIETY FERTILIZER*VARIETY TIME FERTILIZER*TIME
VARIETY*TIME;
RANDOM FERTILIZER*VARIETY/TEST;
RUN;

```

TABLE 27.8

Split-Plot-In-Time Tests for Main Effects

Source	df	Type III SS	Mean Square	F-Value	Pr > F
Fertilizer	4	33.267126	8.316781	94.36	<0.0001
Variety	3	23.282507	7.760836	88.05	<0.0001
Error	12	1.057658	0.088138		
Error: MS(Fertilizer × Variety)					

The tests for *Fertilizer* and *Variety* main effects are given in Table 27.8 and the tests for the *Time* main effect, as well as the *Time* \times *Fertilizer* and *Time* \times *Variety* interaction effects are given in Table 27.9 along with G–G adjusted degrees of freedom. The adjusted degrees of freedom were computed from the split-plot-in-time values by multiplying each of the original degrees of freedom by $\hat{\theta}$. The tests in Table 27.8 need no adjustment. Note that the *F* statistics in Table 27.8 agree with the MANOVA test statistics for *Variety* and *Fertilizer* main effects given earlier. Even with the adjusted degrees of freedom being used, all significance probabilities in Table 27.9 are still less than 0.0001.

The SAS-GLM procedure with its Repeated option can also be used to produce the analyses described in this section as well as some of the results of a MANOVA analysis. Table 27.10 gives the SAS commands that can be used for such an analysis of the data in Table 27.1.

Table 27.11, in the row labeled “Orthogonal components,” gives the test for $H_0: \mathbf{C}\Sigma\mathbf{C}' = \lambda I$. That is, this is the test of whether the H–F conditions are satisfied or not. Table 27.12 gives the MANOVA test for comparing *Time* main effects, Table 27.13 gives the MANOVA tests for *Time* \times *Variety* interaction and *Time* \times *Fertilizer* interaction, and Table 27.14 gives the tests

for *Variety* and *Fertilizer* main effects. Finally, Table 27.15 gives the split-plot-in-time analyses of *Time* main effects as well as the *Time* \times *Variety* and *Time* \times *Fertilizer* interactions. Three different sets of *p*-values are given. They are unadjusted split-plot-in-time *p*-values, Greenhouse–Geisser (G–G) adjusted *p*-values, and Huyhn–Feldt (H–F) adjusted *p*-values. The values of $\hat{\theta} = 0.7954$ and $\bar{\theta} = 1.7473$ are also given.

TABLE 27.9

Split-Plot-in-Time Tests with G–G Adjusted Degrees of Freedom

Source	<i>df</i> (Adj <i>df</i>)	Type III SS	Mean Square	<i>F</i> -Value	Pr > <i>F</i>
Time	4 (3.18)	20.478356	5.119589	738.20	<0.0001
Fertilizer \times time	16 (12.72)	0.674324	0.042145	6.08	<0.0001
Variety \times time	12 (9.54)	1.246268	0.103856	14.98	<0.0001
Error: MS(Error)	48 (38.16)	0.332892	0.006935		

TABLE 27.10

SAS-GLM Code to Analyze the Data in Table 27.1 Using the Repeated Option

```
PROC GLM DATA=LAI;
CLASSES FERTILIZER VARIETY;
MODEL LAI1-LAI5=FERTILIZER VARIETY/NOUNI;
REPEATED TIME 5 (1 2 3 4 5) /SUMMARY PRINT;
RUN;
```

TABLE 27.11Test for H_0 : $\mathbf{C}\Sigma\mathbf{C}' = \lambda I$ *Sphericity Tests*

Variables	<i>df</i>	Mauchly's Criterion	Chi-Square	Pr > Chi-Square
Transformed variates	9	0.1009828	23.883383	0.0045
Orthogonal components	9	0.5208938	6.7938446	0.6586

TABLE 27.12

MANOVA Test for Time Main Effect

*MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Time Effect**H* = Type III SSCP Matrix for Time *E* = Error SSCP Matrix*S* = 1 *M* = 1 *N* = 3.5

Statistic	Value	<i>F</i> -Value	Num <i>df</i>	Den <i>df</i>	Pr > <i>F</i>
Wilks's lambda	0.00502375	445.62	4	9	<0.0001

TABLE 27.13

MANOVA Test for Interactions between Time and Fertilizer and Variety

*MANOVA Test Criteria and F Approximations for the Hypothesis of No Time × Fertilizer Effect**H = Type III SSCP Matrix for Time × Fertilizer E = Error SSCP Matrix**S = 4 M = -0.5 N = 3.5*

Statistic	Value	F-Value	Num df	Den df	Pr > F
Wilks' lambda	0.10535731	1.91	16	28.133	0.0640
<i>MANOVA Test Criteria and F Approximations for the Hypothesis of No Time × Variety Effect</i>					
<i>H = Type III SSCP Matrix for Time × Variety E = Error SSCP Matrix</i>					
<i>S = 3 M = 0 N = 3.5</i>					
Statistic	Value	F-Value	Num df	Den df	Pr > F
Wilks's lambda	0.01425985	8.00	12	24.103	<0.0001

TABLE 27.14

Tests for Fertilizer and Variety Main Effects

Source	df	Type III SS	Mean Square	F-Value	Pr > F
Fertilizer	4	33.26712600	8.31678150	94.36	<0.0001
Variety	3	23.28250700	7.76083567	88.05	<0.0001
Error	12	1.05765800	0.08813817		

TABLE 27.15Split-Plot-In-Time Analysis with Adjusted *p*-Values

Source	df	Type III SS	Mean Square	F-Value	Pr > F	Adj Pr > F	
						G-G	H-F
Time	4	20.47835600	5.11958900	738.20	<0.0001	<0.0001	<0.0001
Time × fertilizer	16	0.67432400	0.04214525	6.08	<0.0001	<0.0001	<0.0001
Time × variety	12	1.24626800	0.10385567	14.98	<0.0001	<0.0001	<0.0001
Error (time)	48	0.33289200	0.00693525				
		Greenhouse–Geisser epsilon		0.7954			
		Huynh–Feldt epsilon		1.7473			

27.4 Mixed Model Methods

Consider once again the notation described in Section 27.1 where

$$\mathbf{y}_{ik} = \begin{bmatrix} y_{i1k} \\ y_{i2k} \\ \vdots \\ y_{ipk} \end{bmatrix}$$

is the vector of responses for subject k in treatment group i . Suppose that the error vector

$$\varepsilon_{ik}^* \sim \text{independent } N(\mathbf{0}, \Sigma_i), \quad i = 1, 2, \dots, t; \quad k = 1, 2, \dots, n_i \quad (27.16)$$

Let \mathbf{y} be a vector that contains all of the data vectors. That is,

$$\mathbf{y}' = [y'_{11} \quad y'_{12} \quad \dots \quad y'_{1n_1} \quad y'_{21} \quad y'_{22} \quad \dots \quad y'_{2n_2} \quad \dots \quad y'_{t1} \quad y'_{t2} \quad y'_{t2} \quad \dots \quad y'_{tn_t}]$$

Under the conditions in Equation 27.16,

$$V = \text{Cov}(\mathbf{y}) = \begin{bmatrix} \Sigma_1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \Sigma_1 & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \Sigma_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \Sigma_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \Sigma_2 & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ddots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \Sigma_2 & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \Sigma_t & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \Sigma_t & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \Sigma_t \end{bmatrix}$$

Suppose $\mathbf{y} \sim N(X\beta, V)$ or equivalently, $\mathbf{y} = X\beta + \varepsilon$ where $\varepsilon \sim N(\mathbf{0}, V)$. Suppose for now that V is known and that $\ell'\beta$ is estimable. In this case, it can be shown that the best estimate of $\ell'\beta$ is $\ell'\hat{\beta}_V$ where $\hat{\beta}_V = (X'V^{-1}X)^{-1}X'V^{-1}\mathbf{y}$. Furthermore, the standard error of $\ell'\hat{\beta}_V$ can be shown to be equal to $\ell'(X'V^{-1}X)^{-1}\ell$. The estimator $\ell'\hat{\beta}_V$ is called a generalized least squares estimator of $\ell'\beta$. Unfortunately, V is rarely known. But suppose V could be estimated by \hat{V} , then $\ell'\beta$ could be estimated by $\ell'\hat{\beta}_{\hat{V}}$ where $\hat{\beta}_{\hat{V}} = (X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}\mathbf{y}$, and an estimate of the standard error of $\ell'\hat{\beta}_{\hat{V}}$ could be taken as $\widehat{s.e.}(\ell'\hat{\beta}_{\hat{V}}) = \ell'(X'\hat{V}^{-1}X)^{-1}\ell$. The estimator $\ell'\hat{\beta}_{\hat{V}}$ is called an estimated generalized least squares estimator of $\ell'\beta$. An approximate t -statistic whose degrees of freedom must also be approximated that tests $\ell'\beta = 0$ is given by

$$t = \frac{\ell'\hat{\beta}_{\hat{V}}}{\ell'(X'\hat{V}^{-1}X)\ell} \quad (27.17)$$

Suppose one wants to test $H_0: H\beta = \mathbf{0}$ vs $H_a: H\beta \neq \mathbf{0}$ where H is a $q \times p$ matrix of rank q . An approximate F -statistic that can be used to test H_0 is given by

$$F = \frac{(H\hat{\beta}_{\hat{V}})' \left[H(X'\hat{V}^{-1}X)^{-1}H' \right]^{-1} (H\hat{\beta}_{\hat{V}})}{q} \quad (27.18)$$

The numerator degrees of freedom for this F -statistic is given by q and the denominator degrees of freedom must be approximated. Approximation methods are beyond the scope of this book, but two methods can be recommended. If there are no missing data values, a method known as Satterthwaite's method can be used. This is a generalization of Satterthwaite's method discussed in Chapter 2. See Giesbrecht and Burns (1985), McLean and Sanders (1988), and Fai and Cornelius (1996) for more details. When there are missing values, the method known as the Kenward–Roger's method should be used. See Kenward and Roger (1997) for more details.

Mixed-model procedures can be used to obtain estimates of the Σ_i from which V can be estimated. The methods of estimation that are used are likelihood methods based on the assumptions in Equation 27.16, that is, on y having a multivariate normal distribution. Such methods are beyond the scope of this book except that some general comments will be made later. The estimate of Σ_i depends on whether Σ_i has any structure. Various structures can be considered. The more popular structures include a compound symmetry structure, a Huynh–Feldt structure, an AR(1) structure, heterogeneous compound symmetry, and heterogeneous AR(1). Mixed-model programs can also handle an unstructured case where Σ_i has no structure at all. Once V is estimated, the mixed model procedures can compute $\ell' \hat{\beta}_V$ and its estimated standard error. Table 27.16 gives some of the more popular covariance structures for repeated measures. The structures in Table 27.16 assume that each of the t treatments has the same parameters in its covariance structure. That is, it assumes that all of the Σ_i are equal. It is also possible to estimate covariance parameters separately for each treatment.

Summarizing, the repeated measures model is

$$y_{ijk} = \mu + \alpha_i + \tau_j + \gamma_{ij} + \varepsilon_{ijk}^* \quad i = 1, 2, \dots, t; j = 1, 2, \dots, p; k = 1, 2, \dots, n_i \quad (27.19)$$

where

$$\varepsilon_{ik}^* = \begin{bmatrix} \varepsilon_{i1k}^* \\ \varepsilon_{i2k}^* \\ \vdots \\ \varepsilon_{ipk}^* \end{bmatrix}$$

is the vector of errors for subject k in treatment group i and where $\varepsilon_{ik}^* \sim \text{independent } N(\mathbf{0}, \Sigma_i)$, $i = 1, 2, \dots, t; k = 1, 2, \dots, n_i$. Additional covariance structures can be obtained by adding a random subject component to the model in Equation 27.19. Such a model would be given by

$$y_{ijk} = \mu + \alpha_i + \delta_{ik} + \tau_j + \gamma_{ij} + \varepsilon_{ijk}^* \quad i = 1, 2, \dots, t; j = 1, 2, \dots, p; k = 1, 2, \dots, n_i \quad (27.20)$$

where the $\delta_{ik} \sim \text{i.i.d. } N(\mathbf{0}, \sigma_\delta^2)$, $i = 1, 2, \dots, t; k = 1, 2, \dots, n_i$.

Adding a random subject component to the model changes the covariance structure of the repeated measures to

$$\text{Cov}(y_{ik}) = \text{Cov} \begin{pmatrix} y_{i1k} \\ y_{i2k} \\ \vdots \\ y_{ipk} \end{pmatrix} = \sigma_\delta^2 J_p + \Sigma_i$$

TABLE 27.16

Some Popular Covariance Structures for Repeated Measures Experiments

Compound symmetry structure (two parameters)	$\Sigma_i = \sigma^2 \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{bmatrix}$ for $i = 1, 2, \dots, t$
Hyuhn–Feldt structure ($p + 1$ parameters)	$\Sigma_i = \begin{bmatrix} \sigma^2 + 2\eta_1 & \eta_1 + \eta_2 & \cdots & \eta_1 + \eta_p \\ \eta_2 + \eta_1 & \sigma^2 + 2\eta_2 & \cdots & \eta_2 + \eta_p \\ \vdots & \vdots & \ddots & \vdots \\ \eta_p + \eta_1 & \eta_p + \eta_2 & \cdots & \sigma^2 + 2\eta_p \end{bmatrix}$ for $i = 1, 2, \dots, t$
AR(1) structure (two parameters)	$\Sigma_i = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{p-1} \\ \rho & 1 & \rho & \cdots & \rho^{p-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{p-1} & \rho^{p-2} & \rho^{p-3} & \cdots & 1 \end{bmatrix}$ for $i = 1, 2, \dots, t$
Unstructured [$p(p + 1)/2$ parameters]	$\Sigma_i = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}$ for $i = 1, 2, \dots, t$
Heterogeneous compound symmetry structure ($p + 1$ parameters)	$\Sigma_i = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & \cdots & \rho\sigma_1\sigma_p \\ \rho\sigma_2\sigma_1 & \sigma_2^2 & \cdots & \rho\sigma_2\sigma_p \\ \vdots & \vdots & \ddots & \vdots \\ \rho\sigma_p\sigma_1 & \rho\sigma_p\sigma_2 & \cdots & \sigma_p^2 \end{bmatrix}$ for $i = 1, 2, \dots, t$
Heterogeneous AR(1) structure ($p + 1$ parameters)	$\Sigma_i = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & \rho^2\sigma_1\sigma_3 & \cdots & \rho^{p-1}\sigma_1\sigma_p \\ \rho\sigma_2\sigma_1 & \sigma_2^2 & \rho\sigma_2\sigma_3 & \cdots & \rho^{p-2}\sigma_2\sigma_p \\ \rho^2\sigma_3\sigma_1 & \rho\sigma_3\sigma_2 & \sigma_3^2 & \cdots & \rho^{p-3}\sigma_3\sigma_p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{p-1}\sigma_p\sigma_1 & \rho^{p-2}\sigma_p\sigma_2 & \rho^{p-3}\sigma_p\sigma_3 & \cdots & \sigma_p^2 \end{bmatrix}$ for $i = 1, 2, \dots, t$

where J_p is a $p \times p$ matrix of 1s. Adding a random subject component has no effect when Σ_i satisfies compound symmetry or when Σ_i has the unstructured form, but adding a random subject component could be very useful for many other covariance structures.

Two methods that have been used to fit models such as those in Equations 27.16 and 27.17 are discussed. One is to find the maximum likelihood estimates (ML) of the model parameters, and the other finds restricted maximum likelihood estimates (REML) of the model parameters. Both of these methods are numerically intensive and fitting models by these methods is beyond the scope of this book. Nevertheless, some of the issues involved with these two methods are discussed next.

27.4.1 Maximum Likelihood Method

One can form a likelihood function based on the distribution of y , and consider the likelihood as a function of the parameters in the covariance matrix V and the fixed effect parameters in β . Then one finds values of these parameters that maximize the likelihood function over the parameter space. This method is numerically intensive, the resulting estimators may not give rise to unbiased estimates of parameter functions that may be of interest such as $P'\beta$, distributional properties of the estimators of parameter functions may not be known except asymptotically, solving the likelihood equations requires an iterative process that may or may not converge and, even when it converges, it may converge at a local maxima rather than a global maximum. The greatest disadvantage of the ML method is that it tends to underestimate the variance–covariance parameters, which results in estimated standard errors of estimators of fixed effects that are too small. This leads to type I error rates that are much higher than desired, and confidence intervals that do not achieve the desired confidence levels. See Chapter 22 for a discussion of a two-way mixed model.

27.4.2 Restricted Maximum Likelihood Method

The restricted maximum likelihood method is also numerically intensive and distributional properties of the estimates are not known except asymptotically. However, it is preferred over the ML method because the resulting estimators of parameter functions of interest have less bias, and the REML method does not underestimate the variance–covariance parameters nearly as much as the ML method. This leads to estimated standard errors of estimators of fixed effects that are more appropriate which leads to type I error rates that are more desirable, and confidence intervals that tend to be closer to the desired confidence levels.

Consider the matrix form of the model, $y = X\beta + \epsilon$ where $\epsilon \sim N(\mathbf{0}, V)$. Let L be a full row rank matrix that satisfies $LX = \mathbf{0}$ and such that $\text{rank}(L) = n - \text{rank}(X)$ where $n =$ the dimension of y .

Let $y^* = Ly$. Then $y^* \sim N(\mathbf{0}, LVL')$. The likelihood function formed from y^* depends only on the variance–covariance parameters. The REML estimates of the variance–covariance parameters are the values of the parameters that maximize the restricted likelihood function based on the distribution of y^* . Once estimates of the variance–covariance parameters are found, V can be estimated, and then test statistics can be computed from Equations 27.17 and 27.18.

As a strategy for using mixed model methods, it is generally recommended that one determine an appropriate structure for the variance–covariance matrix of the repeated measures using REML, and once a covariance matrix structure is obtained, then one can consider inferences about estimable functions of the fixed effect parameters. See Chapter 22 for a discussion of REML for a two-way mixed model. Some of the possibilities for choosing a covariance matrix structure for the repeated measures will be considered next.

Statistical software such as the SAS-Mixed procedure compute several statistics that provide useful information that can help one choose a covariance structure for the repeated measures. The first such statistic was suggested by Akaike (1974) and is known as Akaike's Information Criterion (AIC). A second statistic was suggested by Schwarz (1978) and is known as Schwarz's Bayesian Criterion (BIC). A third possibility was given by Hurvich and Tsai (1989) and is generally denoted AICC. Each of these criterion is a function of $-2\log_e(\hat{L})$ where \hat{L} is the maximum of the restricted maximum likelihood function. The criteria differ only in the way that $-2\log_e(\hat{L})$ is penalized as the number of parameters in

the covariance structure increases. For each of these three criteria, smaller is considered to be better. Let d = the number of covariance parameters in the structure being considered, N = the number of subjects, and N^* = total number of observation minus the rank(X). Then each of these criteria is defined by:

$$\begin{aligned} \text{AIC} &= -2 \log_e(\hat{L}) + 2d \\ \text{BIC} &= -2 \log_e(\hat{L}) + d[\log_e(N)] \end{aligned} \quad (27.21)$$

and

$$\text{AICC} = -2 \log_e(\hat{L}) + 2d[N^*/(N^* - d - 1)]$$

When samples sizes are relatively large, one can also compute a likelihood ratio test (LRT) statistic that can be used to compare two covariance structures for the repeated measures whenever the structure under the null hypothesis is a special case of the structure under the alternative. For example, LRT statistics can be obtained that would compare a compound symmetry structure with an H-F structure or with an unstructured structure, and a LRT statistic could be obtained that would compare an H-F structure or an AR(1) structure with an unstructured structure, but one cannot get an LRT statistic to compare a compound symmetry structure or an H-F structure to an AR(1) structure since neither one is a special case of the other. In order to compare a compound symmetry structure or an H-F structure with an AR(1) structure, one would need to rely upon the criteria in Equation 27.21.

As an example, consider once again the Drug data in Table 26.4. Various analyses of these data will be obtained using the SAS-Mixed procedure. The basic SAS commands that can be used are shown in Table 27.17. The basic model is a two-way model with effects *Drug*, *Time*, and *Drug* \times *Time* interaction. The Repeated statement indicates that *Time* is a repeated factor, the *Type* = CS option indicates that the model being fitted assumes the covariance matrix of the repeated measures has compound symmetry, and the *Subject* = Person option indicates that Person is a variable in the data set that identifies those observations which make up the repeated measures for a particular subject. Note that for these data the Person variable takes on the unique values of 1–24, one value for each subject. *Caution:* If one had let the Person variable have values 1–8 for each of the three drug groups, then one would have had to use *Subject* = Person(*Drug*) in order to identify each subject correctly. Finally, the *R* option is used so that the covariance matrix for *Person* = 1 will be provided in the output. Note that this analysis assumes that all 24 subjects have the same covariance matrix, Σ , so only the first one needs be printed. If one desired the covariance matrices for subjects 1, 9, and 17, then one would use *R* = 1, 9, 17 instead of just *R*.

The approach that is used in this example is to refit models assuming various covariance structures in order to select the covariance structure that will be used when considering inferences on the *Drug* and *Time* factors. Consequently, the data in this example was reanalyzed using each of the Repeated statements given in Table 27.18. In each analysis, the statement “ODS RTF SELECT R;” was used to get each of the *R* matrices placed in the RTF output file, and the “ODS OUTPUT FITSTATISTICS = FIT1;” statement was used to create a data set that contains the values of various fit statistics for each analysis. The fit statistics obtained are AIC, AICC, and BIC as well as the value $-2 \log_e(\hat{L})$ for each analysis.

TABLE 27.17

SAS-Mixed Code to Analyze the Data in Table 26.4 Using the Repeated Option

```

DATA HRT_RATE;
INPUT DRUG $ PERSON HR1-HR4 @@;
CARDS;
A 1 72 86 81 77 B 2 85 86 83 80 C 3 69 73 72 74
A 4 78 83 88 81 B 5 82 86 80 84 C 6 66 62 67 73
A 7 71 82 81 75 B 8 71 78 70 75 C 9 84 90 88 87
A 10 72 83 83 69 B 11 83 88 79 81 C 12 80 81 77 72
A 13 66 79 77 66 B 14 86 85 76 76 C 15 72 72 69 70
A 16 74 83 84 77 B 17 85 82 83 80 C 18 65 62 65 61
A 19 62 73 78 70 B 20 79 83 80 81 C 21 75 69 69 68
A 22 69 75 76 70 B 23 83 84 78 81 C 24 71 70 65 63
;
DATA HR; SET HRT_RATE; DROP HR1-HR4;
TIME=1; HR=HR1; OUTPUT; TIME=2; HR=HR2; OUTPUT; TIME=3; HR=HR3;
OUTPUT;
TIME=4; HR=HR4; OUTPUT;

TITLE Analyses of Drug Data using the MIXED Prodedure';
ODS RTF FILE='TEMP.RTF';

PROC MIXED DATA=HR;
CLASSES DRUG TIME PERSON;
MODEL HR = DRUG TIME DRUG*TIME;
REPEATED TIME/TYPE=CS SUBJECT=PERSON R;
ODS RTF SELECT R;
ODS OUTPUT FITSTATISTICS=FIT1;
TITLE2 'Repeated Measures Analysis - Assuming Compound Symmetry';
RUN;

DATA FIT1; SET FIT1; TYPE='CS' ';
RUN;

ODS RTF CLOSE;

```

TABLE 27.18

Covariance Structures Considered with Different Repeated Options

REPEATED TIME/ TYPE =CS	SUBJECT =PERSON R;
REPEATED TIME/ TYPE =HF	SUBJECT =PERSON R;
REPEATED TIME/ TYPE =UN	SUBJECT =PERSON R;
REPEATED TIME/ TYPE =AR(1)	SUBJECT =PERSON R;
REPEATED TIME/ TYPE =CSH	SUBJECT =PERSON R;
REPEATED TIME/ TYPE =ARH(1)	SUBJECT =PERSON R;
REPEATED TIME/ TYPE =CS	SUBJECT =PERSON R= 1,9,17 GROUP =DRUG;
REPEATED TIME/ TYPE =HF	SUBJECT =PERSON R= 1,9,17 GROUP =DRUG;
REPEATED TIME/ TYPE =UN	SUBJECT =PERSON R= 1,9,17 GROUP =DRUG;
REPEATED TIME/ TYPE =AR(1)	SUBJECT =PERSON R= 1,9,17 GROUP =DRUG;

The first six of the Repeated statements in Table 27.18 correspond to the covariance structures given in Table 27.16. The last four Repeated statements include a Group = Drug option, which causes a different covariance matrix to be fitted for each Drug. The first person in each drug group is person numbers 1, 9, and 17, and the R = 1, 9, 17 option places those estimated covariance matrices in the output.

Table 27.19 gives the estimated covariance matrix of the repeated measures when assuming compound symmetry. Note that $\hat{\sigma}^2 = 33.4182$ and

$$\hat{\rho} = \frac{25.9702}{33.4182} = 0.777$$

Table 27.20 gives the estimated covariance matrix assuming the H-F conditions. Here

$$\hat{\sigma}^2 + 2\hat{\eta}_1 = 28.5296$$

$$\hat{\sigma}^2 + 2\hat{\eta}_2 = 40.2442$$

and

$$\hat{\eta}_1 + \hat{\eta}_2 = 26.9390$$

Solving these simultaneously for $\hat{\sigma}^2$, $\hat{\eta}_1$, and $\hat{\eta}_2$ gives

$$\hat{\sigma}^2 = 7.4479, \quad \hat{\eta}_1 = 10.5409, \quad \text{and} \quad \hat{\eta}_2 = 16.3982$$

TABLE 27.19
 Σ for Compound Symmetry (CS)

Estimated R Matrix for Person 1

Row	Col1	Col2	Col3	Col4
1	33.4182	25.9702	25.9702	25.9702
2	25.9702	33.4182	25.9702	25.9702
3	25.9702	25.9702	33.4182	25.9702
4	25.9702	25.9702	25.9702	33.4182

TABLE 27.20
 Σ for H-F Conditions

Estimated R Matrix for Person 1

Row	Col1	Col2	Col3	Col4
1	28.5296	26.9390	25.8070	22.1733
2	26.9390	40.2442	31.6643	28.0307
3	25.8070	31.6643	37.9802	26.8987
4	22.1733	28.0307	26.8987	30.7130

From these, and the other two diagonal elements of $\hat{\Sigma}$, one can get

$$\hat{\eta}_3 = 15.2671, \quad \text{and} \quad \hat{\eta}_4 = 11.6326$$

Table 27.21 gives the estimated covariance matrix assuming an unstructured covariance matrix. This is the same estimate that one gets from a SAS-GLM analysis when one includes the MANOVA/PRINTE; option.

Table 27.22 gives the estimated covariance matrix assuming an AR(1) structure. Here $\hat{\sigma}^2 = 32.4945$ and $\hat{\rho} = 26.7309/32.4945 = 0.8226$.

Table 27.23 gives the estimated covariance matrix assuming a heterogeneous compound symmetry structure. In this case, one can get

$$\hat{\sigma}_1^2 = 31.1038,$$

$$\hat{\sigma}_2^2 = 38.6218,$$

$$\hat{\sigma}_3^2 = 29.3752,$$

$$\hat{\sigma}_4^2 = 34.7840$$

and

$$\hat{\rho} = \frac{27.0585}{\sqrt{31.1038 \cdot 38.6218}} = 0.7807$$

TABLE 27.21
 $\hat{\Sigma}$ for Unstructured Case

Estimated R Matrix for Person 1

Row	Col1	Col2	Col3	Col4
1	30.5238	28.6548	25.4881	20.0952
2	28.6548	39.2321	29.3095	25.5476
3	25.4881	29.3095	31.2321	26.7262
4	20.0952	25.5476	26.7262	32.6845

TABLE 27.22
 $\hat{\Sigma}$ for an AR(1) Structure

Estimated R Matrix for Person 1

Row	Col1	Col2	Col3	Col4
1	32.4945	26.7309	21.9896	18.0893
2	26.7309	32.4945	26.7309	21.9896
3	21.9896	26.7309	32.4945	26.7309
4	18.0893	21.9896	26.7309	32.4945

TABLE 27.23

$\hat{\Sigma}$ for a Compound Symmetry Structure with Heterogeneous Variances

Estimated R Matrix for Person 1

Row	Col1	Col2	Col3	Col4
1	31.1038	27.0585	23.5982	25.6790
2	27.0585	38.6218	26.2959	28.6146
3	23.5982	26.2959	29.3752	24.9552
4	25.6790	28.6146	24.9552	34.7840

TABLE 27.24

$\hat{\Sigma}$ for an AR(1) Structure with Heterogeneous Variances

Estimated R Matrix for Person 1

Row	Col1	Col2	Col3	Col4
1	30.7872	29.0118	21.4829	18.3212
2	29.0118	39.3259	29.1204	24.8346
3	21.4829	29.1204	31.0182	26.4531
4	18.3212	24.8346	26.4531	32.4516

Table 27.24 shows the results from a heterogeneous AR(1) structure. Here

$$\hat{\sigma}_1^2 = 30.7872$$

$$\hat{\sigma}_2^2 = 39.3259$$

$$\hat{\sigma}_3^2 = 31.0182$$

$$\hat{\sigma}_4^2 = 32.4256$$

and

$$\hat{\rho} = \frac{29.0118}{\sqrt{30.7872 \cdot 39.3256}} = 0.8338$$

Tables 27.25–27.28 give the estimated covariance matrices when assuming a different covariance matrix for each Drug. Table 27.25 is for compound symmetry structures, Table 27.26 is for H–F structures, Table 27.27 is for the unstructured structures, and Table 27.28 is for AR(1) structures. Finding estimates of each of the individual covariance parameter is left for the reader to do.

Table 27.29 shows the values of the AIC for each of the covariance structures considered above. One can see that the minimum value for AIC is 488.603 and this minimum occurs for the AR(1) structure. Thus under AIC, the AR(1) structure would be the best covariance structure to choose.

TABLE 27.25

Three Estimated Covariance Matrices for Each Drug under Compound Symmetry

Estimated R Matrix for Person 1

Row	Col1	Col2	Col3	Col4
1	21.5313	15.8333	15.8333	15.8333
2	15.8333	21.5313	15.8333	15.8333
3	15.8333	15.8333	21.5313	15.8333
4	15.8333	15.8333	15.8333	21.5313

Estimated R Matrix for Person 9

Row	Col1	Col2	Col3	Col4
1	63.9063	53.2679	53.2679	53.2679
2	53.2679	63.9063	53.2679	53.2679
3	53.2679	53.2679	63.9063	53.2679
4	53.2679	53.2679	53.2679	63.9063

Estimated R Matrix for Person 17

Row	Col1	Col2	Col3	Col4
1	14.8170	8.8095	8.8095	8.8095
2	8.8095	14.8170	8.8095	8.8095
3	8.8095	8.8095	14.8170	8.8095
4	8.8095	8.8095	8.8095	14.8170

TABLE 27.26

Three Estimated Covariance Matrices for Each Drug under H-F Conditions

Estimated R Matrix for Person 1

Row	Col1	Col2	Col3	Col4
1	31.2012	20.4843	18.0424	22.7955
2	20.4843	21.1632	13.0234	17.7765
3	18.0424	13.0234	16.2794	15.3346
4	22.7955	17.7765	15.3346	25.7857

Estimated R Matrix for Person 9

Row	Col1	Col2	Col3	Col4
1	44.8443	63.7376	48.8587	46.8783
2	63.7376	103.91	78.3905	76.4101
3	48.8587	78.3905	74.1499	61.5311
4	46.8783	76.4101	61.5311	70.1891

Estimated R Matrix for Person 17

Row	Col1	Col2	Col3	Col4
1	16.6470	8.1668	11.6811	6.7731
2	8.1668	11.7014	9.2083	4.3003
3	11.6811	9.2083	18.7300	7.8146
4	6.7731	4.3003	7.8146	8.9140

TABLE 27.27

Three Estimated Covariance Matrices for Each Drug under Unstructured Conditions

Estimated R Matrix for Person 1

Row	Col1	Col2	Col3	Col4
1	24.0000	17.0000	16.7143	19.5000
2	17.0000	20.0000	12.2857	13.5000
3	16.7143	12.2857	16.0000	16.0000
4	19.5000	13.5000	16.0000	26.1250

Estimated R Matrix for Person 9

Row	Col1	Col2	Col3	Col4
1	43.9286	57.9643	44.5714	35.4286
2	57.9643	88.2679	68.2143	57.8571
3	44.5714	68.2143	60.0000	55.5714
4	35.4286	57.8571	55.5714	63.4286

Estimated R Matrix for Person 17

Row	Col1	Col2	Col3	Col4
1	23.6429	11.0000	15.1786	5.3571
2	11.0000	9.4286	7.4286	5.2857
3	15.1786	7.4286	17.6964	8.6071
4	5.3571	5.2857	8.6071	8.5000

TABLE 27.28

Three Estimated Covariance Matrices for Each Drug under AR(1) Conditions

Estimated R Matrix for Person 1

Row	Col1	Col2	Col3	Col4
1	23.0054	17.6563	13.5509	10.4001
2	17.6563	23.0054	17.6563	13.5509
3	13.5509	17.6563	23.0054	17.6563
4	10.4001	13.5509	17.6563	23.0054

Estimated R Matrix for Person 9

Row	Col1	Col2	Col3	Col4
1	57.3490	50.6923	44.8082	39.6071
2	50.6923	57.3490	50.6923	44.8082
3	44.8082	50.6923	57.3490	50.6923
4	39.6071	44.8082	50.6923	57.3490

Estimated R Matrix for Person 17

Row	Col1	Col2	Col3	Col4
1	15.1400	9.6852	6.1957	3.9634
2	9.6852	15.1400	9.6852	6.1957
3	6.1957	9.6852	15.1400	9.6852
4	3.9634	6.1957	9.6852	15.1400

TABLE 27.29

AIC for Each Covariance Structure

Type	Criterion	Fitstat
CS	AIC (smaller is better)	492.797
HF	AIC (smaller is better)	497.034
UN	AIC (smaller is better)	497.372
AR(1)	AIC (smaller is better)	488.603
CSH	AIC (smaller is better)	497.514
ARH(1)	AIC (smaller is better)	492.767
CS, Heterogenous Groups	AIC (smaller is better)	492.735
HF, Heterogenous Groups	AIC (smaller is better)	500.601
UN, Heterogenous Groups	AIC (smaller is better)	511.261
AR(1), Heterogenous Group	AIC (smaller is better)	490.948

TABLE 27.30

AICC for Each Covariance Structure

Type	Criterion	Fitstat
CS	AICC (smaller is better)	492.945
HF	AICC (smaller is better)	497.804
UN	AICC (smaller is better)	500.386
AR(1)	AICC (smaller is better)	488.751
CSH	AICC (smaller is better)	498.283
ARH(1)	AICC (smaller is better)	493.537
CS, Heterogenous Groups	AICC (smaller is better)	493.826
HF, Heterogenous Groups	AICC (smaller is better)	507.660
UN, Heterogenous Groups	AICC (smaller is better)	546.355
AR(1), Heterogenous Group	AICC (smaller is better)	492.039

Table 27.30 shows the values of the AICC for each of the covariance structures considered above. One can see that the minimum value for AICC is 488.751 and this minimum also occurs for the AR(1) structure. Thus under AICC, the AR(1) structure would be the best covariance structure to choose.

Table 27.31 shows the values of the BIC for each of the covariance structures considered above. One can see that the minimum value for BIC is 490.959 and this minimum occurs for the AR(1) structure. Thus under BIC, the AR(1) structure would be the best covariance structure to choose. So under each of the three criteria, AIC, AICC, and BIC, the preferred covariance structure for the repeated measures is AR(1).

Table 27.32 gives the values of $-2\log_e(\hat{L})$ for each covariance structure considered in this example. Suppose one wants to use a likelihood ratio test to compare the compound symmetry structure to the H-F structure. This can be done since compound symmetry is a special case of the H-F structure. That is, suppose we want to test $H_0: \Sigma$ has CS structure vs $H_a: \Sigma$ has H-F structure. The value of the LRT statistic is the difference in their respective $-2\log_e(\hat{L})$ values. Under H_0 , this result is an approximate chi-square test statistic with degrees equal to the difference in the number of parameters in each covariance structure. The CS structures has two parameters when and the H-F structure has five parameters

TABLE 27.31

BIC for Each Covariance Structure

Type	Criterion	Fitstat
CS	BIC (smaller is better)	495.153
HF	BIC (smaller is better)	502.925
UN	BIC (smaller is better)	509.153
AR(1)	BIC (smaller is better)	490.959
CSH	BIC (smaller is better)	503.404
ARH(1)	BIC (smaller is better)	498.658
CS, Heterogenous Groups	BIC (smaller is better)	499.803
HF, Heterogenous Groups	BIC (smaller is better)	518.272
UN, Heterogenous Groups	BIC (smaller is better)	546.602
AR(1), Heterogenous Group	BIC (smaller is better)	498.016

TABLE 27.32Values of $-2 \log_e(\hat{L})$ for Each Covariance Structure

Type	Criterion	Fitstat
CS	$-2 \text{ Res Log Likelihood}$	488.797
HF	$-2 \text{ Res Log Likelihood}$	487.034
UN	$-2 \text{ Res Log Likelihood}$	477.372
AR(1)	$-2 \text{ Res Log Likelihood}$	484.603
CSH	$-2 \text{ Res Log Likelihood}$	487.514
ARH(1)	$-2 \text{ Res Log Likelihood}$	482.767
CS, Heterogenous Groups	$-2 \text{ Res Log Likelihood}$	480.735
HF, Heterogenous Groups	$-2 \text{ Res Log Likelihood}$	470.601
UN, Heterogenous Groups	$-2 \text{ Res Log Likelihood}$	451.261
AR(1), Heterogenous Group	$-2 \text{ Res Log Likelihood}$	478.948

when $p = 4$. Thus the chi-square test statistic is $\chi^2 = 488.797 - 487.034 = 1.763$ with $5 - 2 = 3$ degrees of freedom. The resulting observed significance level is $\hat{\alpha} = 0.6230$, and H_0 cannot be rejected.

As a second example using a LRT, consider comparing the AR(1) structure to the ARH(1) structure. Then number of parameters in these two structures are 2 and 5, respectively, when $p = 4$. Here $\chi^2 = 484.603 - 482.767 = 1.836$ with 3 degrees of freedom. The chi-square critical point is $\chi^2_{0.05,3} = 7.815$ and since $1.836 < 7.815$, one cannot reject the AR(1) structure in favor of the ARH(1) structure.

As a third example, consider comparing the AR(1) structure to UN structure. The number of parameters in the AR(1) structure is 2, and the number of parameters in the UN structure is $p(p + 1)/2 = 10$ when $p = 4$. Here $\chi^2 = 484.603 - 477.372 = 7.231$ with $10 - 2 = 8$ degrees of freedom. The resulting observed significance level is $\hat{\alpha} = 0.5119$, and one cannot reject an AR(1) structure in favor of the UN structure.

The three tests above are possible because one covariance structure is a special case of the other. One cannot give a LRT that compares AR(1) to CS since neither one is a special case of the other. Likewise one can not get a LRT that compares ARH(1) to CSH.

Before considering tests about the fixed effects, suppose one fits a model where the covariance structure is AR(1), but a random component corresponding to subject is added to the model. That is, the model in Equation 27.20 is fitted where it is assumed the repeated measures satisfy an AR(1) structure. This is equivalent to choosing a covariance structure for the repeated measures that is equal to

$$\Sigma = \sigma_{\delta}^2 \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} + \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}$$

Table 27.33 gives the SAS-Mixed commands used to fit a model with AR(1) structure on the repeated measures and with a random subject effect, and Table 27.34 gives the values of the fit statistic for the analysis using the commands in Table 27.33. Note that the values of AIC, AICC, and BIC are all larger than they were for an AR(1) structure model without a random subject effect. Thus for these drug data, the addition of a random subject term cannot be recommended. In addition, the LRT statistic comparing the model with a random subject term to the model without a random subject term is $\chi^2 = 484.6 - 483.7 = 0.9$. Adding a random subject term to the model increases the number of parameters by 1.

TABLE 27.33

SAS-Mixed Commands for Adding a Random Subject Effect to a Repeated Measures Experiment

```
PROC MIXED DATA=HR MAXITER=200 ;
  CLASSES DRUG TIME PERSON;
  MODEL HR=DRUG TIME DRUG*TIME;
  RANDOM PERSON;
  REPEATED TIME/TYPE=AR(1) SUBJECT=PERSON R;
  TITLE2 'Repeated Measures Analysis - Assuming Autoregressive
  Errors';
  ODS RTF SELECT R FITSTATISTICS COVPARMS;
  RUN;
```

TABLE 27.34

Fit Statistics for a Repeated Measures Model with a Random Subject Effect

Fit Statistics	
-2 Res Log Likelihood	483.7
AIC (smaller is better)	489.7
AICC (smaller is better)	490.0
BIC (smaller is better)	493.2

Therefore $\chi^2 = 0.9$ would be compared with a chi-square critical point with 1 degree of freedom. This critical point is $\chi^2_{0.05,1} = 3.84$, and adding a random subject term to the model does not give a significantly better fit.

Next the data are analyzed to consider inferences about the Drug and Time effects. The SAS-Mixed commands are shown in Table 27.35. The values of the test statistics given by Equation 27.18 are shown in Table 27.36 for the *Drug* and *Time* main effects and the *Drug* \times *Time* interaction. It is noted that the *Drug* \times *Time* interaction effect is highly significant. Hence the two-way *Drug* \times *Time* means should be compared with one another within each drug group and within each time level. The least squares means are shown in Table 27.37 and pairwise comparisons for drugs within each time level are given in Table 27.38 and pairwise comparisons among time levels for each drug are given in Table 27.39.

TABLE 27.35

Final SAS-Mixed Commands for Analyzing Fixed Effects

```

PROC MIXED DATA=HR MAXITER=200 ;
  CLASSES DRUG TIME PERSON;
  MODEL HR=DRUG TIME DRUG*TIME/DDFM=KR;
  REPEATED TIME/TYPE=AR(1) SUBJECT=PERSON;
  TITLE2 'Repeated Measures Analysis - Assuming Autoregressive
  Errors';
  LSMEANS DRUG TIME DRUG*TIME/PDIFF ADJUST=TUKEY;
  ODS OUTPUT DIFFS=DIFFS;
  RUN;
PROC PRINT DATA=DIFFS; WHERE EFFECT='DRUG*TIME' AND DRUG=_DRUG;
  TITLE 'TIME DIFFERENCES FOR EACH DRUG';
  RUN;

PROC PRINT DATA=DIFFS; WHERE EFFECT='DRUG*TIME' AND TIME=_TIME;
  TITLE 'DRUG DIFFERENCES AT EACH LEVEL OF TIME';
  RUN;

```

TABLE 27.36

Type III Tests on Fixed Effects

Type III Tests of Fixed Effects

Effect	Num df	Den df	F-Value	Pr > F
Drug	2	22.8	6.52	0.0058
Time	3	63.1	15.43	<0.0001
Drug \times time	6	63.3	12.71	<0.0001

27.5 Summary

This chapter considered three alternative methods that can be used to analyze repeated measures experiments when the split-plot-in-time analysis is not appropriate. The

TABLE 27.37

Least Squares Means for Each Effect in the Model

Least Squares Means

Effect	Drug	Time	Estimate	Standard Error	df	t-Value	Pr > t
Drug	A		76.2813	1.7870	22.8	42.69	<0.0001
Drug	B		81.0313	1.7870	22.8	45.34	<0.0001
Drug	C		71.9062	1.7870	22.8	40.24	<0.0001
Time		1	75.0000	1.1636	33.7	64.46	<0.0001
Time		2	78.9583	1.1636	33.7	67.86	<0.0001
Time		3	77.0417	1.1636	33.7	66.21	<0.0001
Time		4	74.6250	1.1636	33.7	64.13	<0.0001
Drug × time	A	1	70.5000	2.0154	33.7	34.98	<0.0001
Drug × time	A	2	80.5000	2.0154	33.7	39.94	<0.0001
Drug × time	A	3	81.0000	2.0154	33.7	40.19	<0.0001
Drug × time	A	4	73.1250	2.0154	33.7	36.28	<0.0001
Drug × time	B	1	81.7500	2.0154	33.7	40.56	<0.0001
Drug × time	B	2	84.0000	2.0154	33.7	41.68	<0.0001
Drug × time	B	3	78.6250	2.0154	33.7	39.01	<0.0001
Drug × time	B	4	79.7500	2.0154	33.7	39.57	<0.0001
Drug × time	C	1	72.7500	2.0154	33.7	36.10	<0.0001
Drug × time	C	2	72.3750	2.0154	33.7	35.91	<0.0001
Drug × time	C	3	71.5000	2.0154	33.7	35.48	<0.0001
Drug × time	C	4	71.0000	2.0154	33.7	35.23	<0.0001

TABLE 27.38

Comparisons between Drug Means for Each Time Level with Tukey–Kramer Adjusted Significance Levels

Effect	Drug	Time	Drug	Time	Estimate	Standard Error	df	t-Value	Pr t	Adjustment		Adj p
										Tukey–Kramer	0.0100	
Drug × time	A	1	B	1	-11.2500	2.8502	33.7	-3.95	0.0004	Tukey–Kramer	0.0100	
Drug × time	A	1	C	1	-2.2500	2.8502	33.7	-0.79	0.4354	Tukey–Kramer	0.9997	
Drug × time	A	2	B	2	-3.5000	2.8502	33.7	-1.23	0.2280	Tukey–Kramer	0.9846	
Drug × time	A	2	C	2	8.1250	2.8502	33.7	2.85	0.0074	Tukey–Kramer	0.1846	
Drug × time	A	3	B	3	2.3750	2.8502	33.7	0.83	0.4106	Tukey–Kramer	0.9995	
Drug × time	A	3	C	3	9.5000	2.8502	33.7	3.33	0.0021	Tukey–Kramer	0.0586	
Drug × time	A	4	B	4	-6.6250	2.8502	33.7	-2.32	0.0263	Tukey–Kramer	0.4710	
Drug × time	A	4	C	4	2.1250	2.8502	33.7	0.75	0.4611	Tukey–Kramer	0.9998	
Drug × time	B	1	C	1	9.0000	2.8502	33.7	3.16	0.0033	Tukey–Kramer	0.0915	
Drug × time	B	2	C	2	11.6250	2.8502	33.7	4.08	0.0003	Tukey–Kramer	0.0066	
Drug × time	B	3	C	3	7.1250	2.8502	33.7	2.50	0.0175	Tukey–Kramer	0.3591	
Drug × time	B	4	C	4	8.7500	2.8502	33.7	3.07	0.0042	Tukey–Kramer	0.1129	

TABLE 27.39

Comparisons between Time Means for Each Drug with Tukey–Kramer Adjusted Significance Levels

Effect	Drug	Time	Drug	Time	Standard						
					Estimate	Error	df	t-Value	Pr t	Adjustment	Adj p
Drug × time	A	1	A	2	-10.0000	1.2315	62.7	-8.12	<0.0001	Tukey–Kramer	<0.0001
Drug × time	A	1	A	3	-10.5000	1.6645	72.1	-6.31	<0.0001	Tukey–Kramer	<0.0001
Drug × time	A	1	A	4	-2.6250	1.9503	79.3	-1.35	0.1822	Tukey–Kramer	0.9693
Drug × time	A	2	A	3	-0.5000	1.2315	62.7	-0.41	0.6861	Tukey–Kramer	1.0000
Drug × time	A	2	A	4	7.3750	1.6645	72.1	4.43	<0.0001	Tukey–Kramer	0.0021
Drug × time	A	3	A	4	7.8750	1.2315	62.7	6.39	<0.0001	Tukey–Kramer	<0.0001
Drug × time	B	1	B	2	-2.2500	1.2315	62.7	-1.83	0.0725	Tukey–Kramer	0.7970
Drug × time	B	1	B	3	3.1250	1.6645	72.1	1.88	0.0645	Tukey–Kramer	0.7681
Drug × time	B	1	B	4	2.0000	1.9503	79.3	1.03	0.3083	Tukey–Kramer	0.9965
Drug × time	B	2	B	3	5.3750	1.2315	62.7	4.36	<0.0001	Tukey–Kramer	0.0026
Drug × time	B	2	B	4	4.2500	1.6645	72.1	2.55	0.0128	Tukey–Kramer	0.3278
Drug × time	B	3	B	4	-1.1250	1.2315	62.7	-0.91	0.3645	Tukey–Kramer	0.9987
Drug × time	C	1	C	2	0.3750	1.2315	62.7	0.30	0.7617	Tukey–Kramer	1.0000
Drug × time	C	1	C	3	1.2500	1.6645	72.1	0.75	0.4551	Tukey–Kramer	0.9998
Drug × time	C	1	C	4	1.7500	1.9503	79.3	0.90	0.3723	Tukey–Kramer	0.9989
Drug × time	C	2	C	3	0.8750	1.2315	62.7	0.71	0.4800	Tukey–Kramer	0.9999
Drug × time	C	2	C	4	1.3750	1.6645	72.1	0.83	0.4115	Tukey–Kramer	0.9995
Drug × time	C	3	C	4	0.5000	1.2315	62.7	0.41	0.6861	Tukey–Kramer	1.0000

methods considered were a MANOVA approach, an adjusted p -value approach, and a mixed model approach. Which approach one chooses to use may depend on the statistical software one has available to use. When one has mixed model software available, then this approach is likely the easiest one to use. If one has trouble getting convergence with mixed model procedures, then one can consider the MANOVA approach if one has statistical software that allows one to use the MANOVA methods. If software is not available for either of these two approaches, then one can consider using the adjusted p -value approach.

27.6 Exercises

27.1 Consider the rabbit data described in Exercise 26.1.

- 1) Are the Hyuhn–Feldt conditions satisfied for these data? Justify your answer.
- 2) What is the G–G estimate of Box's θ defined in Equation 27.12? What is the H–F estimate of θ ?

Answer the following questions by treating the repeated measures as a multivariate response vector.

- 3) Is there a significant $Time \times Trt$ interaction? Why or why not?

- 4) Are there differences between the three treatments? Are there time differences? Explain your answers.
 - 5) Which treatment seems to have the greatest affect on ear temperature difference?
- 27.2 Consider the rabbit data described in Exercise 26.1.
- 1) Based on the AIC criterion, which of the following covariance structures would be selected: compound symmetry, heterogeneous compound symmetry, AR(1), heterogeneous AR(1), or unstructured.
 - 2) Based on the BIC criterion, which of the following covariance structures would be selected: compound symmetry, heterogeneous compound symmetry, AR(1), heterogeneous AR(1), or unstructured.
 - 3) Based on the AICC criterion which of the following covariance structures would be selected: compound symmetry, heterogeneous compound symmetry, AR(1), heterogeneous AR(1), or unstructured.
 - 4) Using likelihood ratio tests compare:
 - a) Compound symmetry to heterogenous compound symmetry.
 - b) Compound symmetry to unstructured.
 - c) Heterogeneous AR(1) to AR(1).
- 27.3 Consider the rabbit data described in Exercise 26.1. Answer questions 3–5 in Exercise 27.1 using a mixed model procedure assuming the covariance structure that you would recommend for these data.
- 27.4 What covariance structure for the repeated measures would you recommend for the data in Exercise 26.3? Justify your answer.

28

Case Studies: Complex Examples Having Repeated Measures

This chapter considers several examples and their statistical analyses using mixed model procedures. The examples considered include split-plot designs with repeated measures, repeated measures nested within repeated measures, and a multilocation study.

28.1 Complex Comfort Experiment

An engineer had three environments in which to test two types of clothing. Since responses to an environment also differ between males and females, sex of person was included as a factor. Four people (two males and two females) were put into an environmental chamber (which was assigned one of the three environments). One male and one female wore clothing type 1, and the other male and female wore clothing type 2. The comfort score of each person was recorded at the end of 1, 2, and 3 hours. The data for this experiment are shown in Table 28.1.

There are three sizes of experimental units. The largest experimental unit is a chamber or, equivalently, a group of four people. The chamber experimental unit experimental design is a one-way treatment structure (*Environment*) in a completely randomized design structure with three replications at each level of the environment. The middle-sized experimental unit is a person. The experimental design for a person is a two-way treatment structure (*Sex* \times *Clothing*) in a randomized complete block design structure in nine blocks (each block contains four experimental units, people). The smallest experimental unit is a 1 h time interval (*Hour*), which is a repeated measure. The experimental design for *Hour* is a one-way treatment structure (*Time*) in a randomized complete block design structure in 36 blocks (each block contains three experimental units (1 h time intervals).

TABLE 28.1

Data for Complex Comfort Experiment

<i>Env</i>	<i>Rep</i>	<i>Sex</i>	<i>Clo</i>	Score 1	Score 2	Score 3
1	1	1	1	13.9001	7.5154	10.9742
1	1	1	2	18.3941	12.4152	15.2242
1	1	2	1	10.0149	3.7669	7.0326
1	1	2	2	16.4774	10.4104	13.1143
1	2	1	1	15.7185	9.7170	12.5080
1	2	1	2	19.7547	13.5293	16.5487
1	2	2	1	10.6902	4.8473	7.9829
1	2	2	2	17.1147	11.3858	14.1503
1	3	1	1	14.9015	9.1825	11.5819
1	3	1	2	18.0402	12.1005	15.4893
1	3	2	1	10.1944	4.1716	6.9688
1	3	2	2	16.0789	10.2357	12.4853
2	1	1	1	10.2881	6.9090	8.4138
2	1	1	2	13.8631	10.1492	12.5372
2	1	2	1	6.1634	2.2837	4.0052
2	1	2	2	13.0291	9.7775	11.6576
2	2	1	1	11.9904	8.4793	9.8695
2	2	1	2	14.8587	11.0317	12.8317
2	2	2	1	6.7562	2.5634	4.7547
2	2	2	2	13.8977	9.6643	11.6034
2	3	1	1	9.7589	6.1772	8.0785
2	3	1	2	13.5513	9.3052	11.5259
2	3	2	1	4.5203	0.5913	2.8939
2	3	2	2	12.5057	7.7502	10.5226
3	1	1	1	7.4205	7.1752	7.1218
3	1	1	2	12.5410	11.9157	12.2239
3	1	2	1	3.8293	3.5868	3.3004
3	1	2	2	11.0002	11.0282	10.5662
3	2	1	1	11.8158	11.9721	11.7187
3	2	1	2	16.4418	16.6355	16.6686
3	2	2	1	7.5707	7.3456	7.2404
3	2	2	2	13.5421	13.5672	14.0240
3	3	1	1	7.2364	7.8304	7.4147
3	3	1	2	12.0690	12.5003	12.1790
3	3	2	1	1.8330	1.6769	1.8065
3	3	2	2	9.4934	9.7000	10.0119

The model for this experiment (where the model is separated into three parts corresponding to the three sizes of experimental units and where the repeated measures are treated as a split-plot-in-time factor) is

$$y_{ijklm} = [\mu + E_i + \eta_{il}] + [S_j + C_k + (SC)_{jk} + (ES)_{ij} + (EC)_{ik} + (ESC)_{ijk} + \delta_{ijkl}] + [T_m + (ET)_{im} + (ST)_{jm} + (CT)_{km} + (SCT)_{jkm} + (EST)_{ijm} + (ECT)_{ikm} + (ESCT)_{ijkm} + \epsilon_{ijklm}] \quad (28.1)$$

where E_i denotes the i th environment, S_j denotes the j th sex, C_k denotes the k th clothing, and T_m denotes the m th time point, $\eta_{i\ell}$ denotes a random chamber effect with the assumption that $\eta_{i\ell} \sim i.i.d. N(0, \sigma_\eta^2)$, $\delta_{ijk\ell}$ denotes a random person effect with the assumption that $\delta_{ijk\ell} \sim N(0, \sigma_\delta^2)$, and $\varepsilon_{ijkm\ell}$ denotes the random measurement error for a given hour with the assumption that $\varepsilon_{ijkm\ell} \sim i.i.d. N(0, \sigma_\varepsilon^2)$. It is also assumed that all random effects are independently distributed. Also note that the first bracket of Equation 28.1 is the Chamber part of the model, the second bracket is the Person part of the model, and the third bracket is the Hour part of the model. If the split-plot-in-time assumption was not appropriate, then the model would be

$$\begin{aligned} y_{ijkm\ell} = & \mu + E_i + \eta_{i\ell} + S_j + C_k + (SC)_{jk} + (ES)_{ij} + (EC)_{ik} + (ESC)_{ijk} + T_m + (ET)_{im} + (ST)_{jm} \\ & + (CT)_{km} + (SCT)_{jkm} + (EST)_{ijm} + (ECT)_{ikm} + (ESCT)_{ijkm} + \varepsilon_{ijkm\ell} \end{aligned} \quad (28.2)$$

where it is assumed that $\eta_{i\ell} \sim i.i.d. N(0, \sigma_\eta^2)$ and

$$\varepsilon_{ijk\ell} = \begin{bmatrix} \varepsilon_{ijk1\ell} \\ \varepsilon_{ijk2\ell} \\ \varepsilon_{ijk3\ell} \end{bmatrix} \sim N_3 \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} \right) \sim N_3(0, \Sigma) \text{ (say)}$$

and the analysis would depend on the structure assumed for Σ . Note that the model in Equation 28.2 is the same as the model in (28.1) except that the random effect corresponding to Person, $\delta_{ijk\ell}$, has been removed from the model. One might want to include Person as a random effect for some Σ such as when Σ has an AR(1) structure.

Should one wish to perform a split-plot-in-time analysis, the basic SAS®-Mixed Code that can be used is shown in Table 28.2. The results of such an analysis will be left for the reader to pursue. Such an analysis assumes the experimental design is a split-split-plot design with Chamber as the whole plot experimental unit, Person as the subplot experimental unit, and Time measurement as sub-sub-plot experimental unit.

The approach that will be taken here is to treat the Time factor as a repeated measure, and then consider various covariance structures for the repeated measures. The design then would be described as a split-plot design with repeated measures on the sub-plot factor (Person). The basic SAS-Mixed code that one would use when assuming compound symmetry for the covariance structure of the repeated measures is shown in Table 28.3. Note that the combination of $Sex \times Clo \times Rep \times Env$ is unique for each person, so this effect can be used to identify a person on whom the repeated measures occur.

TABLE 28.2

SAS-Mixed Code to Analyze the Data in Table 28.1 Treating the Repeated Measures Factor as a Split-Plot-In-Time Factor

```

PROC MIXED DATA=COMFORT;
  TITLE 'SPLIT-PLOT-IN-TIME ANALYSIS';
  CLASSES ENV REP SEX CLO TIME;
  MODEL SCORE=ENV|SEX|CLO|TIME/DDFM=SATTERTH;
  RANDOM REP(ENV) SEX*CLO*REP(ENV);
  LSMEANS ENV|SEX|CLO|TIME/PDIFF;
  RUN;

```

TABLE 28.3

SAS-Mixed Code to Analyze the Data in Table 28.1 Using the Repeated Option with a Compound Symmetry Covariance Structure

```
PROC MIXED DATA=COMFORT;
  TITLE 'ANALYSIS ASSUMES A SPLIT-PLOT DESIGN WITH REPEATED';
  TITLE2 'MEASURES ON THE SUBPLOT FACTOR';
  CLASSES ENV REP SEX CLO TIME;
  MODEL SCORE=ENV|SEX|CLO|TIME/DDFM=SATTERTH;
  RANDOM REP(ENV);
  REPEATED TIME/SUBJECT=SEX*CLO*REP*ENV TYPE=CS;
  LSMEANS ENV|SEX|CLO|TIME/PDIFF;
RUN;
```

Four structures for the covariance matrix of the repeated measures are considered for these data. The covariance structures considered are: compound symmetry, AR(1), an unstructured covariance matrix, and person was also included as a random effect with a covariance structure for the repeated measures being AR(1). For the first three of these, one just needs to change the “Type =” statement in the Repeated option in Table 28.3. To also include Person as a random factor one would use the random and repeated options given below:

```
RANDOM REP(ENV) SEX*CLO*REP*ENV;
REPEATED TIME/SUBJECT=SEX*CLO*REP*ENV TYPE=AR(1);
```

The resulting fit statistics for each of the four covariance structures being considered are given in Table 28.4.

TABLE 28.4

Fit Statistics for Four Covariance Structures

Description	Value	Type
-2 Res Log Likelihood	117.7	CS
-2 Res Log Likelihood	119.4	Ar(1)
-2 Res Log Likelihood	107.5	UN
-2 Res Log Likelihood	117.6	Ar(1) with random person error
AIC (smaller is better)	123.7	CS
AIC (smaller is better)	125.4	Ar(1)
AIC (smaller is better)	121.5	UN
AIC (smaller is better)	125.6	Ar(1) with random person error
AICC (smaller is better)	124.1	CS
AICC (smaller is better)	125.8	Ar(1)
AICC (smaller is better)	123.3	UN
$\chi^2 = 117.7 - 107.5 = 10.2$ with 4 degrees of freedom AICC (smaller is better)	126.2	Ar(1) with random person error
BIC (smaller is better)	124.3	CS
BIC (smaller is better)	126.0	Ar(1)
BIC (smaller is better)	122.9	UN
BIC (smaller is better)	126.4	Ar(1) with random person error

TABLE 28.5

SAS-Mixed Code to Analyze the Data in Table 28.1 Using the Repeated Option with a Unstructured Covariance Structure

```

PROC MIXED DATA=COMFORT;
  TITLE 'ANALYSIS ASSUMES A SPLIT-PLOT DESIGN WITH REPEATED';
  TITLE2 'MEASURES ON THE SUBPLOT FACTOR';

  CLASSES ENV REP SEX CLO TIME;
  MODEL SCORE=ENV|SEX|CLO|TIME/DDFM=SATTERTH;
  RANDOM REP (ENV) ;
  REPEATED TIME/SUBJECT=SEX*CLO*REP*ENV TYPE=UN;
  LSMEANS ENV|SEX|CLO|TIME/PDIFF;
  RUN;

```

An examination of the AIC, AICC, and BIC fit statistics in Table 28.4 reveals that each has their minimum for the unstructured covariance matrix assumption. The chi-square statistic for comparing compound symmetry to unstructured is $\chi^2 = 117.7 - 107.5 = 10.2$ with $6 - 2 = 4$ degrees of freedom. The corresponding observed significance level is $\hat{\alpha} = 0.0372$. Thus unstructured is significantly better than compound symmetry. The rest of the analysis of this example will assume an unstructured covariance matrix for the repeated measures. The SAS-Mixed code that produces an analysis of fixed effects table is given in Table 28.5. The covariance parameter estimates for the analysis described in Table 28.5 are shown in Table 28.6. From Table 28.6, one can see that $\hat{\sigma}_\eta^2 = 2.4047$ and

$$\boldsymbol{\Sigma} = \begin{bmatrix} 0.2157 & 0.2256 & 0.1248 \\ 0.2256 & 0.3525 & 0.1894 \\ 0.1248 & 0.1894 & 0.1462 \end{bmatrix}$$

The analysis of fixed effects table is given in Table 28.7. An examination of Table 28.7 reveals that the significant effects ($\hat{\alpha} < 0.05$) are: *Sex, Clo, Env × Clo, Sex × Clo, Env × Sex × Clo, Time, and Env × Time*. Since the three-way interaction, *Env × Sex × Clo* is significant, one should examine interesting pairwise comparisons among these three-way means. The only interaction involving *Time* that is significant is the *Env × Time* interaction, so one should also examine pairwise comparisons among these two-way means. In addition to the

TABLE 28.6

Covariance Parameter Estimates for the Analysis in Table 28.5

Covariance Parameter Estimates

Covariance Parameter	Subject	Estimate
Rep(Env)		2.4047
UN(1,1)	Env × Rep × Sex × Clo	0.2157
UN(2,1)	Env × Rep × Sex × Clo	0.2256
UN(2,2)	Env × Rep × Sex × Clo	0.3525
UN(3,1)	Env × Rep × Sex × Clo	0.1248
UN(3,2)	Env × Rep × Sex × Clo	0.1894
UN(3,3)	Env × Rep × Sex × Clo	0.1462

TABLE 28.7

Analysis of Fixed Effects

Type III Tests of Fixed Effects

Effect	Num <i>df</i>	Den <i>df</i>	F-Value	Pr > F
<i>Env</i>	2	6.11	3.25	0.1091
<i>Sex</i>	1	18	484.00	<0.0001
<i>Env</i> × <i>Sex</i>	2	18	0.65	0.5320
<i>Clo</i>	1	18	1347.89	<0.0001
<i>Env</i> × <i>Clo</i>	2	18	3.60	0.0484
<i>Sex</i> × <i>Clo</i>	1	18	93.59	<0.0001
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	2	18	3.69	0.0455
<i>Time</i>	2	24	1667.76	<0.0001
<i>Env</i> × <i>Time</i>	4	24	478.22	<0.0001
<i>Sex</i> × <i>Time</i>	2	24	0.66	0.5248
<i>Env</i> × <i>Sex</i> × <i>Time</i>	4	24	0.75	0.5666
<i>Clo</i> × <i>Time</i>	2	24	0.70	0.5059
<i>Env</i> × <i>Clo</i> × <i>Time</i>	4	24	1.12	0.3688
<i>Sex</i> × <i>Clo</i> × <i>Time</i>	2	24	1.62	0.2197
<i>Env</i> × <i>Sex</i> × <i>Clo</i> × <i>Time</i>	4	24	0.73	0.5811

pairwise comparisons described above, the analysis also considers linear and quadratic comparisons among the time means for each environment. This is done for illustration purposes here, and readers will need to determine for themselves whether such comparisons are of interest to them for a study similar to this one. The additional SAS-Mixed statements one will need to append to those in Table 28.5 are shown in Table 28.8. The results from these additional statements are shown in Tables 28.9–28.16.

Table 28.9 contains linear and quadratic contrasts in time averaged over the four combinations of *Sex* × *Clothing* and *Reps* for each environment. An examination of Table 28.9 reveals that both of the linear and quadratic contrasts are statistically significant for environments 1 and 2, and that neither is significant for environment 3.

Table 28.10 contains the *Environment* × *Sex* × *Clothing* least squares means. These means are averaged over *reps* and *time*. Table 28.11 contains the *Environment* × *Time* least squares means, and these means are averages over *Reps*, *Clothing*, and *Sex*. There are 12 *Environment* × *Sex* × *Clothing* means. If one were to make all pairwise comparisons among these twelve means, it would require 66 comparisons. However, many of these pairwise comparisons are not interesting to an experimenter. The ones that would be of interest would generally be those that compare possibilities for one of the three factors with one another for each combination of the other two factors. That is, one would usually want to compare clothing types for each combination of *Environment* × *Sex* (six comparisons), sexes for each combination of *Environment* × *Clothing* (six comparisons), and environments for each combination of *Sex* × *Clothing* (12 comparisons). Thus there are a total of 24 comparisons out of the 66 possible that are likely to be interesting to an experimenter. The results of these comparisons are shown in Tables 28.12–28.14, respectively.

An examination of the comparisons in Table 28.12 reveals that clothing types are significantly different for each combination of *Environment* × *Sex*, and an examination of Table 28.13 indicates that males are significantly different from females for each *Environment* × *Clothing*

TABLE 28.8

Additional SAS-Mixed Statements to Obtain Results for Comparisons of Interest

```

LSMEANS ENV*SEX*CLO ENV*TIME/PDIFF;
ODS LISTING EXCLUDE LSMEANS DIFFS;
ODS OUTPUT LSMEANS=LSMS DIFFS=LSMDIFFS;
ODS RTF EXCLUDE LSMEANS DIFFS;
ESTIMATE 'LINEAR TIME FOR ENV 1' TIME -1 0 1 ENV*TIME -1 0 1 0 0 0 0 0 0;
ESTIMATE 'LINEAR TIME FOR ENV 2' TIME -1 0 1 ENV*TIME 0 0 0 -1 0 1 0 0 0;
ESTIMATE 'LINEAR TIME FOR ENV 3' TIME -1 0 1 ENV*TIME 0 0 0 0 0 -1 0 1;
ESTIMATE 'QUAD TIME FOR ENV 1' TIME 1 -2 1 ENV*TIME 1 -2 1 0 0 0 0 0;
ESTIMATE 'QUAD TIME FOR ENV 2' TIME 1 -2 1 ENV*TIME 0 0 0 1 -2 1 0 0 0;
ESTIMATE 'QUAD TIME FOR ENV 3' TIME 1 -2 1 ENV*TIME 0 0 0 0 0 1 -2 1;
RUN;

PROC PRINT DATA=LSMS; WHERE EFFECT='ENV*SEX*CLO';
TITLE4 'ENV*SEX*CLO LEAST SQUARES MEANS';
RUN;

PROC PRINT DATA=LSMS; WHERE EFFECT='ENV*TIME';
TITLE4 'ENV*TIME LEAST SQUARES MEANS';
RUN;

PROC PRINT DATA=LSMDIFFS; WHERE EFFECT='ENV*SEX*CLO' AND ENV=_ENV AND SEX=_SEX;
TITLE4 'CLOTHING DIFFERENCES FOR EACH COMBINATON OF ENVIRONMENT*SEX';
RUN;

PROC PRINT DATA=LSMDIFFS; WHERE EFFECT='ENV*SEX*CLO' AND ENV=_ENV AND CLO=_CLO;
TITLE4 'SEX DIFFERENCES FOR EACH COMBINATON OF ENVIRONMENT*CLOTHING';
RUN;

PROC PRINT DATA=LSMDIFFS; WHERE EFFECT='ENV*SEX*CLO' AND SEX=_SEX AND CLO=_CLO;
TITLE4 'ENVIRONMENT DIFFERENCES FOR EACH COMBINATON OF CLOTHING*SEX';
RUN;

PROC PRINT DATA=LSMDIFFS; WHERE EFFECT='ENV*TIME' AND TIME=_TIME;
TITLE4 'ENVIRONMENT DIFFERENCES FOR EACH HOUR';
RUN;

PROC PRINT DATA=LSMDIFFS; WHERE EFFECT='ENV*TIME' AND ENV=_ENV;
TITLE4 'TIME DIFFERENCES FOR EACH ENVIRONMENT';
RUN;

```

TABLE 28.9

Linear and Quadratic Contrasts in Time for Each Environment

Estimates

Label	Estimate	Standard Error	df	t-Value	Pr > t
Linear time for Env 1	-3.1016	0.09673	24	-32.06	<0.0001
Linear time for Env 2	-1.8741	0.09673	24	-19.37	<0.0001
Linear time for Env 3	-0.04309	0.09673	24	-0.45	0.6600
Quad time for Env 1	8.8988	0.1735	24	51.28	<0.0001
Quad time for Env 2	5.8761	0.1735	24	33.86	<0.0001
Quad time for Env 3	-0.06653	0.1735	24	-0.38	0.7048

TABLE 28.10

Three-Way Least Squares Means for Environment, Sex, and Clothing

Effect	<i>Env</i>	<i>Sex</i>	<i>Clo</i>	Estimate	Standard Error	<i>df</i>	<i>t</i> -Value	<i>Pr</i> > <i>t</i>
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	1	1	1	11.7777	0.9317	6.86	12.64	<0.0001
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	1	1	2	15.7218	0.9317	6.86	16.87	<0.0001
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	1	2	1	7.2966	0.9317	6.86	7.83	0.0001
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	1	2	2	13.4947	0.9317	6.86	14.48	<0.0001
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	2	1	1	8.8849	0.9317	6.86	9.54	<0.0001
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	2	1	2	12.1838	0.9317	6.86	13.08	<0.0001
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	2	2	1	3.8369	0.9317	6.86	4.12	0.0047
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	2	2	2	11.1565	0.9317	6.86	11.97	<0.0001
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	3	1	1	8.8562	0.9317	6.86	9.51	<0.0001
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	3	1	2	13.6861	0.9317	6.86	14.69	<0.0001
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	3	2	1	4.2433	0.9317	6.86	4.55	0.0028
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	3	2	2	11.4370	0.9317	6.86	12.28	<0.0001

TABLE 28.11

Two-Way Least Squares Means for Environment and Time

Least Squares Means									
Effect	<i>Env</i>	<i>Sex</i>	<i>Clo</i>	<i>Time</i>	Estimate	Standard Error	<i>df</i>	<i>t</i> -Value	<i>Pr</i> > <i>t</i>
<i>Env</i> × <i>Time</i>	1			1	15.1066	0.9053	6.13	16.69	<0.0001
<i>Env</i> × <i>Time</i>	1			2	9.1065	0.9116	6.3	9.99	<0.0001
<i>Env</i> × <i>Time</i>	1			3	12.0050	0.9021	6.04	13.31	<0.0001
<i>Env</i> × <i>Time</i>	2			1	10.9319	0.9053	6.13	12.08	<0.0001
<i>Env</i> × <i>Time</i>	2			2	7.0568	0.9116	6.3	7.74	0.0002
<i>Env</i> × <i>Time</i>	2			3	9.0578	0.9021	6.04	10.04	<0.0001
<i>Env</i> × <i>Time</i>	3			1	9.5661	0.9053	6.13	10.57	<0.0001
<i>Env</i> × <i>Time</i>	3			2	9.5778	0.9116	6.3	10.51	<0.0001
<i>Env</i> × <i>Time</i>	3			3	9.5230	0.9021	6.04	10.56	<0.0001

TABLE 28.12

Comparisons between Clothing Types for Each Combination of Environment and Sex

Effect	<i>Env</i>	<i>Sex</i>	<i>Clo</i>	<i>Time</i>	<i>Env</i>	<i>Sex</i>	<i>Clo</i>	<i>Time</i>	Estimate	Standard Error	<i>df</i>	<i>t</i> -Value	<i>pt</i>
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	1	1	1	—	1	1	2	—	-3.9441	0.3646	18	-10.82	<0.0001
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	1	2	1	—	1	2	2	—	-6.1981	0.3646	18	-17.00	<0.0001
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	2	1	1	—	2	1	2	—	-3.2988	0.3646	18	-9.05	<0.0001
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	2	2	1	—	2	2	2	—	-7.3196	0.3646	18	-20.08	<0.0001
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	3	1	1	—	3	1	2	—	-4.8299	0.3646	18	-13.25	<0.0001
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	3	2	1	—	3	2	2	—	-7.1938	0.3646	18	-19.73	<0.0001

TABLE 28.13

Comparisons between Sexes for Each Combination of Environment and Clothing Type

Effect	<i>Env</i>	<i>Sex</i>	<i>Clo</i>	<i>Time</i>	<i>Env</i>	<i>Sex</i>	<i>Clo</i>	<i>Time</i>	Estimate	Standard Error	<i>df</i>	<i>t</i> -Value	<i>pt</i>
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	1	1	1	—	1	2	1	—	4.4811	0.3646	18	12.29	<0.0001
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	1	1	2	—	1	2	2	—	2.2270	0.3646	18	6.11	<0.0001
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	2	1	1	—	2	2	1	—	5.0481	0.3646	18	13.85	<0.0001
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	2	1	2	—	2	2	2	—	1.0273	0.3646	18	2.82	0.0114
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	3	1	1	—	3	2	1	—	4.6129	0.3646	18	12.65	<0.0001
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	3	1	2	—	3	2	2	—	2.2490	0.3646	18	6.17	<0.0001

TABLE 28.14

Comparisons between Environments for Each Combination of Sex and Clothing Type

Effect	<i>Env</i>	<i>Sex</i>	<i>Clo</i>	<i>Time</i>	<i>Env</i>	<i>Sex</i>	<i>Clo</i>	<i>Time</i>	Estimate	Standard Error	<i>df</i>	<i>t</i> -Value	<i>pt</i>
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	1	1	1	—	2	1	1	—	2.8927	1.3176	6.86	2.20	0.0649
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	1	1	1	—	3	1	1	—	2.9215	1.3176	6.86	2.22	0.0629
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	1	1	2	—	2	1	2	—	3.5380	1.3176	6.86	2.69	0.0319
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	1	1	2	—	3	1	2	—	2.0357	1.3176	6.86	1.55	0.1671
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	1	2	1	—	2	2	1	—	3.4597	1.3176	6.86	2.63	0.0347
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	1	2	1	—	3	2	1	—	3.0533	1.3176	6.86	2.32	0.0543
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	1	2	2	—	2	2	2	—	2.3383	1.3176	6.86	1.77	0.1201
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	1	2	2	—	3	2	2	—	2.0577	1.3176	6.86	1.56	0.1632
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	2	1	1	—	3	1	1	—	0.02878	1.3176	6.86	0.02	0.9832
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	2	1	2	—	3	1	2	—	-1.5023	1.3176	6.86	-1.14	0.2924
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	2	2	1	—	3	2	1	—	-0.4064	1.3176	6.86	-0.31	0.7669
<i>Env</i> × <i>Sex</i> × <i>Clo</i>	2	2	2	—	3	2	2	—	-0.2806	1.3176	6.86	-0.21	0.8376

combination. Finally, an examination of Table 28.14 reveals that environments 1 and 2 are significantly different for only the combinations where *Sex* = 1 and *Clo* = 2; and for *Sex* = 2, *Clo* = 1.

If one were to look at all possible pairwise comparisons among the nine combinations of the *Environment* × *Time*, it would require 36 comparisons. Only 18 of these would likely be of interest to an experimenter. Table 28.15 contains comparisons among the environment means for each time (these means are averages over reps, sex, and clothing), and Table 28.16 contains comparisons among the time means for each environment (these means are also averages over reps, sex, and clothing). An examination of Table 28.15 reveals that, at time 1, environment 1 is significantly different from both environments 2 and 3, and that environments 2 and 3 are not significantly different. At times 2 and 3, none of the three environments are significantly different from one another. An examination of Table 28.16 reveals that all time points are significantly different from one another for environments 1 and 2, and that no time points are significantly different from one another for environment 3.

A plot showing the time means for each comparison is shown in Figure 28.1. This plot was created by the SAS commands shown in Table 28.17. An examination of the plot illustrates the statements made about time comparisons for Table 28.16.

TABLE 28.15

Comparisons between Environments for Each Time

Effect	<i>Env</i>	<i>Sex</i>	<i>Clo</i>	<i>Time</i>	<i>Env</i>	<i>Sex</i>	<i>Clo</i>	<i>Time</i>	Estimate	Standard Error	<i>df</i>	<i>t</i> -Value	<i>pt</i>
<i>Env</i> × <i>Time</i>	1	—	—	1	2	—	—	1	4.1747	1.2803	6.13	3.26	0.0167
<i>Env</i> × <i>Time</i>	1	—	—	1	3	—	—	1	5.5405	1.2803	6.13	4.33	0.0047
<i>Env</i> × <i>Time</i>	1	—	—	2	2	—	—	2	2.0496	1.2891	6.3	1.59	0.1606
<i>Env</i> × <i>Time</i>	1	—	—	2	3	—	—	2	-0.4714	1.2891	6.3	-0.37	0.7266
<i>Env</i> × <i>Time</i>	1	—	—	3	2	—	—	3	2.9472	1.2757	6.04	2.31	0.0599
<i>Env</i> × <i>Time</i>	1	—	—	3	3	—	—	3	2.4820	1.2757	6.04	1.95	0.0993
<i>Env</i> × <i>Time</i>	2	—	—	1	3	—	—	1	1.3658	1.2803	6.13	1.07	0.3263
<i>Env</i> × <i>Time</i>	2	—	—	2	3	—	—	2	-2.5210	1.2891	6.3	-1.96	0.0960
<i>Env</i> × <i>Time</i>	2	—	—	3	3	—	—	3	-0.4652	1.2757	6.04	-0.36	0.7278

TABLE 28.16

Comparisons between Times for Each Environment

Effect	<i>Env</i>	<i>Sex</i>	<i>Clo</i>	<i>Time</i>	<i>Env</i>	<i>Sex</i>	<i>Clo</i>	<i>Time</i>	Estimate	Standard Error	<i>df</i>	<i>t</i> -Value	<i>pt</i>
<i>Env</i> × <i>Time</i>	1	—	—	1	1	—	—	2	6.0002	0.09872	24	60.78	<0.0001
<i>Env</i> × <i>Time</i>	1	—	—	1	1	—	—	3	3.1016	0.09673	24	32.06	<0.0001
<i>Env</i> × <i>Time</i>	1	—	—	2	1	—	—	3	-2.8986	0.09995	24	-29.00	<0.0001
<i>Env</i> × <i>Time</i>	2	—	—	1	2	—	—	2	3.8751	0.09872	24	39.25	<0.0001
<i>Env</i> × <i>Time</i>	2	—	—	1	2	—	—	3	1.8741	0.09673	24	19.37	<0.0001
<i>Env</i> × <i>Time</i>	2	—	—	2	2	—	—	3	-2.0010	0.09995	24	-20.02	<0.0001
<i>Env</i> × <i>Time</i>	3	—	—	1	3	—	—	2	-0.0117	0.09872	24	-0.12	0.9064
<i>Env</i> × <i>Time</i>	3	—	—	1	3	—	—	3	0.04309	0.09673	24	0.45	0.6600
<i>Env</i> × <i>Time</i>	3	—	—	2	3	—	—	3	0.05481	0.09995	24	0.55	0.5885

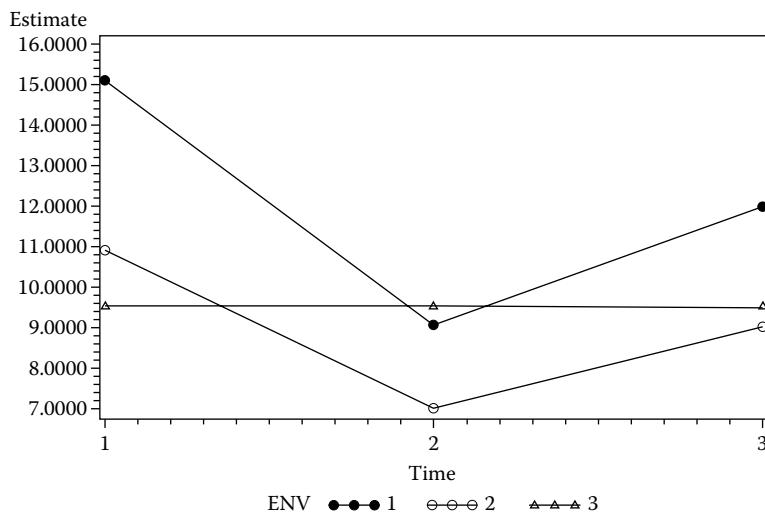
**FIGURE 28.1** Response over time for each environment.

TABLE 28.17

Additional SAS Statements to Obtain Plots for Time Comparisons within Each Environment

```

SYMBOL1 V=DOT I=JOIN COLOR=BLACK;
SYMBOL2 V=CIRCLE I=JOIN COLOR=BLACK;
SYMBOL3 V=TRIANGLE I=JOIN COLOR=BLACK;

PROC GPLOT DATA=LSMS; WHERE EFFECT='ENV*TIME';
  TITLE;
  PLOT ESTIMATE*TIME=ENV;
RUN;

```

28.2 Family Attitudes Experiment

The attitudes of families from rural and urban environments were measured every six months for three time periods. The data were obtained from seven rural families and 10 urban families with each family consisting of a son, a father, and a mother. The data are given in Table 28.18. Note that in this example there are two sets of repeated measures, the three time periods, and the three family members. Family member is a repeated measure

TABLE 28.18

Data for Family Attitude Study

Family	Family Member								
	Son			Father			Mother		
Family	T1	T2	T3	T1	T2	T3	T1	T2	T3
<i>Urban</i>									
1	17	17	19	18	19	21	16	16	18
2	12	14	15	19	19	21	16	16	18
3	8	10	11	16	18	19	11	12	12
4	5	7	7	12	12	13	13	14	14
5	2	5	6	12	14	14	14	16	18
6	9	11	11	16	17	18	14	15	16
7	8	9	9	19	20	20	15	16	18
8	13	14	16	16	17	18	18	18	20
9	11	12	13	13	16	17	7	8	10
10	19	20	20	13	15	15	11	12	12
<i>Rural</i>									
1	12	11	14	18	19	22	16	16	19
2	13	13	17	18	19	22	16	16	19
3	12	13	16	19	18	22	17	16	20
4	18	18	21	23	23	26	23	22	26
5	15	14	16	15	15	19	17	17	20
6	6	6	10	15	16	19	18	19	21
7	16	17	18	17	17	21	18	20	23

since one cannot randomly assign father, mother, and son to the three family members. This experiment can be described as a one-way experiment in a completely randomized design with two sets of repeated measures, family member, and time nested within family member.

The treatment structure for this experiment is a three-way treatment structure with factors area (A : rural vs urban), family member (M : son, father, mother), and time (T : 1, 2, 3). The design structure is a completely randomized design with two sets of repeated measures. One model that might represent this experiment is given by

$$y_{ijkl} = \mu + A_i + \eta_{il} + M_j + (AM)_{ij} + \delta_{ijl} + T_k + (AT)_{ik} + (MT)_{jk} + (AMT)_{ijk} + \varepsilon_{ijkl} \quad (28.3)$$

$i = 1, 2; j = 1, 2, 3; k = 1, 2, 3$

Under split-plot-in-time assumptions on both sets of repeated measures one would have η_{il} denoting a random family effect with the assumption that $\eta_{il} \sim i.i.d. N(0, \sigma_\eta^2)$, δ_{ijl} denoting a random person effect with the assumption that $\delta_{ijl} \sim N(0, \sigma_\delta^2)$, and ε_{ijkl} denoting the random measurement error for a given time point with the assumption that $\varepsilon_{ijkl} \sim i.i.d. N(0, \sigma_\varepsilon^2)$. The SAS-Mixed commands to perform an analysis under the above assumptions are shown in Table 28.19. However, it is up to the readers to perform this analysis should they wish to do so. The analysis using the commands in Table 28.19 would be similar to the analysis from a split-split-plot experiment. See Example 24.5 for such an example.

TABLE 28.19

SAS Commands to Analyze the Date in Table 28.18 Assuming That Both Sets of Repeated Measures Satisfy Split-Plot-In-Time Assumptions

```

DATA ATTITUDEEES;
  INPUT A $ F ATTITUDE1-ATTITUDE9;
  IF A='RURAL' THEN F=F+10;
  CARDS;
URBAN  1 17 17 19 18 19 21 16 16 18
URBAN  2 12 14 15 19 19 21 16 16 18
URBAN  3  8 10 11 16 18 19 11 12 12
URBAN  4  5  7  7 12 12 13 13 14 14
URBAN  5  2  5  6 12 14 14 14 16 18
URBAN  6  9 11 11 16 17 18 14 15 16
URBAN  7  8  9  9 19 20 20 15 16 18
URBAN  8 13 14 16 16 17 18 18 18 20
URBAN  9 11 12 13 13 16 17  7  8 10
URBAN 10 19 20 20 13 15 15 11 12 12
RURAL  1 12 11 14 18 19 22 16 16 19
RURAL  2 13 13 17 16 15 19 19 19 23
RURAL  3 12 13 16 19 18 22 17 16 20
RURAL  4 18 18 21 23 23 26 23 22 26
RURAL  5 15 14 16 15 15 19 17 17 20
RURAL  6  6  6 10 15 16 19 18 19 21
RURAL  7 16 17 18 17 17 21 18 20 23
PROC PRINT;
RUN;

```

Continued

TABLE 28.19 (continued)

```

DATA USUAL; SET ATTITUDEES; DROP ATTITUDE1-ATTITUDE9;
M='S'; T=0; ATTITUDE=ATTITUDE1; OUTPUT;
M='S'; T=6; ATTITUDE=ATTITUDE2; OUTPUT;
M='S'; T=12; ATTITUDE=ATTITUDE3; OUTPUT;
M='F'; T=0; ATTITUDE=ATTITUDE4; OUTPUT;
M='F'; T=6; ATTITUDE=ATTITUDE5; OUTPUT;
M='F'; T=12; ATTITUDE=ATTITUDE6; OUTPUT;
M='M'; T=0; ATTITUDE=ATTITUDE7; OUTPUT;
M='M'; T=6; ATTITUDE=ATTITUDE8; OUTPUT;
M='M'; T=12; ATTITUDE=ATTITUDE9; OUTPUT;
RUN;

PROC MIXED DATA=USUAL;
TITLE 'ANALYSIS ASSUMING SPLIT-PLOT-IN-TIME ASSUMPTIONS';
TITLE2 'FOR BOTH SETS OF REPEATED MEASURES';
CLASSES A M T F;
MODEL ATTITUDE=A|M|T/DDFM=BETWITHIN;
RANDOM F(A) M*F(A);
LSMEANS A*T/PDIFF;
RUN;

```

The analyses considered in this section will consider the model

$$y_{ijkl} = \mu + A_i + M_j + (AM)_{ij} + T_k + (AT)_{ik} + (MT)_{jk} + (AMT)_{ijk} + \varepsilon_{ijkl} \quad (28.4)$$

$$i = 1, 2; j = 1, 2, 3; k = 1, 2, 3$$

Let

$$\boldsymbol{\varepsilon}_{i\ell} = \begin{bmatrix} \varepsilon_{i11\ell} \\ \varepsilon_{i12\ell} \\ \varepsilon_{i13\ell} \\ \varepsilon_{i21\ell} \\ \varepsilon_{i22\ell} \\ \varepsilon_{i23\ell} \\ \varepsilon_{i31\ell} \\ \varepsilon_{i32\ell} \\ \varepsilon_{i33\ell} \end{bmatrix}$$

be the vector of errors for the ℓ th family in the i th area.

Assume that $\varepsilon_{i\ell} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$, and the statistical analysis will depend upon the structure assumed for $\boldsymbol{\Sigma}$. The covariance matrix $\boldsymbol{\Sigma}$ is a 9×9 matrix. Such a set-up ignores the nesting structure of the repeated measures and simply assumes that there are nine repeated

measures, and all of the structures that were considered in Chapter 27 would be possible here. Structures that take into account the nesting of the repeated measures assume that

$$\boldsymbol{\Sigma} = \text{Var} \begin{pmatrix} \boldsymbol{\varepsilon}_{i11\ell} \\ \boldsymbol{\varepsilon}_{i12\ell} \\ \boldsymbol{\varepsilon}_{i13\ell} \\ \boldsymbol{\varepsilon}_{i21\ell} \\ \boldsymbol{\varepsilon}_{i22\ell} \\ \boldsymbol{\varepsilon}_{i23\ell} \\ \boldsymbol{\varepsilon}_{i31\ell} \\ \boldsymbol{\varepsilon}_{i32\ell} \\ \boldsymbol{\varepsilon}_{i33\ell} \end{pmatrix} = \begin{bmatrix} \theta_{11}V & \theta_{12}V & \theta_{13}V \\ \theta_{21}V & \theta_{22}V & \theta_{23}V \\ \theta_{31}V & \theta_{32}V & \theta_{33}V \end{bmatrix} = \boldsymbol{\Theta} \otimes V \text{ (say)}$$

where $\boldsymbol{\Theta} \otimes V$ represents the direct product of the two variance covariance matrices

$$\boldsymbol{\Theta} = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \\ \theta_{31} & \theta_{32} & \theta_{33} \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{bmatrix}$$

The SAS-Mixed procedure will allow the user to select such a covariance structure when $\boldsymbol{\Theta}$ is unstructured and V has either compound symmetry, AR(1), or unstructured. Analyzing the data as a split-split plot experiment is equivalent to assuming that both $\boldsymbol{\Theta}$ and V satisfy the compound symmetry structure.

The data in Table 28.18 were analyzed under four different covariance structures with the first three being obtained by taking the direct product of a covariance matrix corresponding to family member and a covariance matrix corresponding to time. The structure considered for family member was unstructured, and the structures considered for time were compound symmetry, AR(1), and unstructured. The fourth covariance structure considered was an unstructured 9×9 covariance matrix that treats all combinations of family member and time as nine repeated measures. The resulting fit statistics are examined in order to determine the best covariance structure to consider when comparing the fixed effects of area, family member, and time and all possible interactions of these effects. The SAS-Mixed commands used are given in Table 28.20 and continued in Table 28.21. It might be noted that the option DDFM = BETWITHIN was used in all analyses rather than the DDFM = KR option. For the fit statistics that are obtained, it does not matter which of these two options are used. However, if one is going to consider tests on the fixed effects under a direct product covariance structure, using the DDFM = KR fails to give reasonable results since the 9×9 estimated covariance matrix is singular under the direct product structures. The results for the fit statistics are given in Table 28.22. In the two cases where V was assumed to have compound symmetry or AR(1), the Mixed routine did not produce any results because of obtaining an infinite likelihood value. This seems to have happened because of the default starting values for the covariance parameters that the Mixed procedure chose. To force the Mixed procedure to begin with different starting values, a PARMS option was included. The statement

```
PARMS 10, 1, 10, 1, 1, 10, .5;
```

TABLE 28.20
SAS Commands Used to Compare Covariance Structures

```

PROC MIXED DATA=USUAL MAXITER=500;
  TITLE 'ANALYSIS ASSUMING TYPE=UN@UN';
  CLASSES A M T F;
  MODEL ATTITUDE=A|M|T/DDFM=BETWITHIN;
  REPEATED M T/TYPE=UN@UN R SUBJECT=F;
  ODS LISTING EXCLUDE ALL;
  ODS OUTPUT FITSTATISTICS=FIT1;
  RUN;
DATA FIT1; SET FIT1; TYPE='UN@UN'
RUN;

PROC MIXED DATA=USUAL MAXITER=500;
  TITLE 'ANALYSIS ASSUMING TYPE=UN@CS';
  CLASSES A M T F;
  MODEL ATTITUDE=A|M|T/DDFM=BETWITHIN;
  REPEATED M T/TYPE=UN@CS R SUBJECT=F;
  PARMs 10, 1, 10, 1, 1, 10, .5;
  ODS LISTING EXCLUDE ALL;
  ODS OUTPUT FITSTATISTICS=FIT2;
  RUN;
DATA FIT2; SET FIT2; TYPE='UN@CS';
RUN;

PROC MIXED DATA=USUAL MAXITER=500;
  TITLE 'ANALYSIS ASSUMING TYPE=UN@AR(1)';
  CLASSES A M T F;
  MODEL ATTITUDE=A|M|T/DDFM=BETWITHIN;
  REPEATED M T/TYPE=UN@AR (1) R SUBJECT=F;
  PARMs 10, 1, 10, 1, 1, 10, .5;
  ODS LISTING EXCLUDE ALL;
  ODS OUTPUT FITSTATISTICS=FIT3;
  RUN;

DATA FIT3; SET FIT3; TYPE='UN@AR (1)';
RUN;

PROC MIXED DATA=USUAL MAXITER=500;
  TITLE 'ANALYSIS ASSUMING TYPE=UN 9X9 COVARIANCE MATRIX';
  CLASSES A M T F;
  MODEL ATTITUDE=A|M|T/DDFM=KR;
  REPEATED M*T/TYPE=UN R SUBJECT=F;
  ODS LISTING EXCLUDE ALL;
  ODS OUTPUT FITSTATISTICS=FIT4;
  RUN;

DATA FIT4; SET FIT4; TYPE='UN 9X9 ';
RUN;

```

selects starting values for the parameters in

$$\boldsymbol{\Theta} = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \\ \theta_{31} & \theta_{32} & \theta_{33} \end{bmatrix} \quad \text{and} \quad \mathbf{V} = \begin{bmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{bmatrix}$$

TABLE 28.21

SAS Commands Used to Compare Covariance Structures (Continued)

```

DATA FITSTATS; SET FIT1 FIT2 FIT3 FIT4;
PROC SORT; BY DESCRL;

ODS RTF FILE='C:\TEMP1.RTF';
PROC PRINT DATA=FITSTATS;
TITLE 'FIT STATISTICS';
ODS LISTING SELECT ALL;
ODS RTF SELECT ALL;
RUN;
ODS RTF CLOSE;

```

as $\theta_{11} = 10$, $\theta_{21} = 1$, $\theta_{22} = 10$, $\theta_{31} = 1$, $\theta_{32} = 1$, $\theta_{33} = 10$, and $\rho = 0.5$. Without loss of generality the diagonal elements in V can be taken to be equal to 1. The same starting values were also used for the case when V has AR(1) structure. In this case,

$$V = \begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix}$$

and again without loss of generality, the diagonal elements in V can be taken as equal to 1.

An examination of the AIC, AICC, and BIC values in Table 28.22 reveals that all three of these criteria achieve their respective minimums for the covariance structure $\Theta \otimes V$ where

TABLE 28.22

Fit Statistics for Family Attitudes Data

Description	Value	Type
-2 Res Log Likelihood	480.2	UN@UN
-2 Res Log Likelihood	482.6	UN@CS
-2 Res Log Likelihood	487.8	UN@AR(1)
-2 Res Log Likelihood	433.0	UN 9 × 9
AIC (smaller is better)	502.2	UN@UN
AIC (smaller is better)	496.6	UN@CS
AIC (smaller is better)	501.8	UN@AR(1)
AIC (smaller is better)	523.0	UN 9 × 9
AICC (smaller is better)	504.3	UN@UN
AICC (smaller is better)	497.5	UN@CS
AICC (smaller is better)	502.7	UN@AR(1)
AICC (smaller is better)	569.5	UN 9 × 9
BIC (smaller is better)	511.4	UN@UN
BIC (smaller is better)	502.5	UN@CS
BIC (smaller is better)	507.6	UN@AR(1)
BIC (smaller is better)	560.5	UN 9 × 9

Θ is unstructured and V has compound symmetry. This is the assumption that is used to consider inferences on the fixed effect parameters.

To obtain tests of the fixed effects the SAS-Mixed commands are given in Table 28.23. The data were sorted prior to this analysis since the SAS-Mixed procedure sorts the variables in the Classes statement alphabetically. Sorting the data helps to match the covariance parameter estimates with the appropriate family member. The estimates of the covariance parameters are shown in Table 28.24. Note that

$$\hat{\Theta} = \begin{bmatrix} \hat{\theta}_{FF} & \hat{\theta}_{FM} & \hat{\theta}_{FS} \\ \hat{\theta}_{MF} & \hat{\theta}_{MM} & \hat{\theta}_{MS} \\ \hat{\theta}_{SF} & \hat{\theta}_{SM} & \hat{\theta}_{SS} \end{bmatrix} = \begin{bmatrix} 8.1570 & 2.7308 & 1.2415 \\ 2.7308 & 9.7410 & 2.6035 \\ 1.2415 & 2.6035 & 14.7547 \end{bmatrix} \text{ and } \hat{\rho} = 0.9637$$

The subscripts in $\hat{\Theta}$ have been changed to correspond to father (F), mother (M), and son (S). The resulting estimate of the 9×9 covariance matrix $\Theta \otimes V$ is given in Table 28.25 and the tests on the fixed effects are shown in Table 28.26.

TABLE 28.23

SAS Commands Used to Obtain the Tests of Fixed Effects Table and Covariance Parameter Estimates

```
PROC SORT DATA=USUAL; BY F A M T;
RUN;

PROC MIXED DATA=USUAL MAXITER=500;
TITLE 'ANALYSIS ASSUMING TYPE=UN@CS';
CLASSES A M T F;
MODEL ATTITUDE=A|M|T/DDFM=BETWITHIN;
REPEATED M T/TYPE=UN@CS R SUBJECT=F;
PARMS 10, 1, 10, 1, 1, 10, .5;
ODS RTF SELECT ALL;
RUN;
```

TABLE 28.24

Covariance Parameter Estimates

Covariance Parameter Estimates

Covariance Parameter	Subject	Estimate
M UN(1,1)	F	8.1570
UN(2,1)	F	2.7308
UN(2,2)	F	9.7410
UN(3,1)	F	1.2415
UN(3,2)	F	2.6035
UN(3,3)	F	14.7547
T Corr	F	0.9637

TABLE 28.25

Estimated Covariance Matrix for Each Family

Estimated R Matrix for F 1

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7	Col8	Col9
1	8.1570	7.8606	7.8606	2.7308	2.6315	2.6315	1.2415	1.1964	1.1964
2	7.8606	8.1570	7.8606	2.6315	2.7308	2.6315	1.1964	1.2415	1.1964
3	7.8606	7.8606	8.1570	2.6315	2.6315	2.7308	1.1964	1.1964	1.2415
4	2.7308	2.6315	2.6315	9.7410	9.3870	9.3870	2.6035	2.5089	2.5089
5	2.6315	2.7308	2.6315	9.3870	9.7410	9.3870	2.5089	2.6035	2.5089
6	2.6315	2.6315	2.7308	9.3870	9.3870	9.7410	2.5089	2.5089	2.6035
7	1.2415	1.1964	1.1964	2.6035	2.5089	2.5089	14.7547	14.2186	14.2186
8	1.1964	1.2415	1.1964	2.5089	2.6035	2.5089	14.2186	14.7547	14.2186
9	1.1964	1.1964	1.2415	2.5089	2.5089	2.6035	14.2186	14.2186	14.7547

TABLE 28.26

Tests of Fixed Effects

Type III Tests of Fixed Effects

Effect	Num df	Den df	F-Value	Pr > F
A	1	15	8.55	0.0105
M	2	30	10.12	0.0004
A × M	2	30	1.56	0.2271
T	2	30	184.11	<0.0001
A × T	2	30	27.89	<0.0001
M × T	4	60	0.80	0.5271
A × M × T	4	60	1.23	0.3080

An examination of Table 28.26 reveals that A , M , T , and $A \times T$ are statistically significant. To explore these effects, one should look at the main effect means for family member (M) and the two-way means for area by time ($A \times T$).

The SAS-Mixed commands to compute the corresponding least squares means and to make pairwise comparisons among them are shown in Table 28.27. These commands should be appended to those in Table 28.23. The least squares means for M and $A \times T$ are shown in Tables 28.28 and 28.29, respectively. Interesting pairwise comparisons among these means are given in Tables 28.30–28.32. An examination of Tables 28.28 and 28.30 reveals that the son's mean is significantly smaller than both of the parent means, and that the father and mother means are not significantly different from one another.

An examination of Table 28.29 suggests that means are increasing over time for both rural families and urban families. However, an examination of Table 28.31 reveals that the 0 and 6 month means for rural families are not significantly different, but both are significantly smaller than the 12 month mean. For urban families all time means are significantly different from one another. An examination of Table 28.32 reveals that urban families have means that are significantly smaller than those of rural families at 0 and 12 months.

TABLE 28.27

Additional SAS Commands Used to Obtain Least Squares Means and Pairwise Comparison among Them

```

LSMEANS M A*T/PDIFF;
ODS RTF EXCLUDE ALL;
ODS LISTING EXCLUDE ALL;
ODS OUTPUT LSMEANS = LSMS DIFFS=LSMDIFFS;
RUN;

PROC PRINT DATA=LSMS; WHERE EFFECT='M';
ODS RTF SELECT ALL;
TITLE 'FAMILY MEMBER LEAST SQUARES MEANS';
RUN;

PROC PRINT DATA=LSMS; WHERE EFFECT='A*T';
TITLE 'AREA X TIME LEAST SQUARES MEANS';
RUN;

PROC PRINT DATA=LSMDIFFS; WHERE EFFECT='M';
TITLE4 'FAMILY MEMBER DIFFERENCES';
RUN;

PROC PRINT DATA=LSMDIFFS; WHERE EFFECT='A*T' AND A=_A;
TITLE4 'TIME DIFFERENCES FOR EACH AREA';
RUN;

PROC PRINT DATA=LSMDIFFS; WHERE EFFECT='A*T' AND T=_T;
TITLE4 'AREA DIFFERENCES FOR EACH TIME POINT';
RUN;

ODS RTF CLOSE;
QUIT;

```

TABLE 28.28

Family Member Least Squares Means

Effect	A	M	T	Estimate	Standard Error	df	t-Value	pt
M		F	—	17.6643	0.6952	30	25.41	<0.0001
M		M	—	16.9714	0.7597	30	22.34	<0.0001
M		S	—	12.8810	0.9349	30	13.78	<0.0001

TABLE 28.29

Two-Way Least Squares Mean for Area and Time Combinations

Effect	A	M	T	Estimate	Standard Error	df	t-Value	pt
$A \times T$	Rural		0	16.3333	0.8527	30	19.16	<0.0001
$A \times T$	Rural		6	16.3810	0.8527	30	19.21	<0.0001
$A \times T$	Rural		12	19.6190	0.8527	30	23.01	<0.0001
$A \times T$	Urban		0	13.1000	0.7134	30	18.36	<0.0001
$A \times T$	Urban		6	14.3000	0.7134	30	20.04	<0.0001

TABLE 28.30

Family Members Comparisons

Effect	A	M	T	A	M	T	Estimate	Standard Error	df	t-Value	pt
M		F	—	M	—	—	0.6929	0.8584	30	0.81	0.4259
M		F	—	S	—	—	4.7833	1.1001	30	4.35	0.0001
M		M	—	S	—	—	4.0905	1.0690	30	3.83	0.0006

TABLE 28.31

Time Comparisons within Each Area

Effect	A	M	T	A	M	T	Estimate	Standard Error	df	t-Value	pt
$A \times T$	Rural		0	Rural		6	-0.04762	0.2299	30	-0.21	0.8373
$A \times T$	Rural		0	Rural		12	-3.2857	0.2299	30	-14.29	<0.0001
$A \times T$	Rural		6	Rural		12	-3.2381	0.2299	30	-14.09	<0.0001
$A \times T$	Urban		0	Urban		6	-1.2000	0.1923	30	-6.24	<0.0001
$A \times T$	Urban		0	Urban		12	-2.2000	0.1923	30	-11.44	<0.0001
$A \times T$	Urban		6	Urban		12	-1.0000	0.1923	30	-5.20	<0.0001

TABLE 28.32

Area Comparisons within Each Time Level

Effect	A	M	T	A	M	T	Estimate	Standard Error	df	t-Value	pt
$A \times T$	Rural		0	Urban		0	3.2333	1.1118	30	2.91	0.0068
$A \times T$	Rural		6	Urban		6	2.0810	1.1118	30	1.87	0.0710
$A \times T$	Rural		12	Urban		12	4.3190	1.1118	30	3.88	0.0005

The difference between the rural family and urban family means at 6 months is not large enough to achieve significance at the 0.05 level.

28.3 Multilocation Experiment

This section considers an experiment involving three drugs and where each subject was measured repeatedly at three different time points. The data were collected by three different investigators (or in three different centers). The data are shown in Table 28.33. Note that there are quite a few missing data points, as indicated by the empty cells.

The analysis assumes that centers and subjects are random effects, and that drugs and times are fixed effects. If one assumes that the repeated measures satisfy the split-plot-in-time assumptions, the model for the data in Table 28.33 is

$$y_{ijkl} = \mu + \eta_i + D_j + \delta_{ijl} + T_k + (DT)_{jk} + \varepsilon_{ijkl} \quad (28.5)$$

TABLE 28.33

Multicenter Drug Experiment

Center	Drug	Subject	Y1	Y2	Y3	Center	Drug	Subject	Y1	Y2	Y3	Center	Drug	Subject	Y1	Y2	Y3
R	1	1	17			R	2	1	18	19	21	R	3	1	16	16	18
R	1	2	12	14	15	R	2	2	19			R	3	2	16	16	18
R	1	3	12	11	14	R	2	3	18	19		R	3	3	16	16	19
R	1	4	13	13	17	R	2	4	16	15	19	R	3	4	19	19	23
R	1	5	12	13		R	2	5	19	18		R	3	5	17	16	20
S	1	1	18	18	21	S	2	1	23	23	26	S	3	1	23		
S	1	2	15	14	16	S	2	2	15	15	19	S	3	2	17	17	20
S	1	3	6	6		S	2	3	15	16	19	S	3	3	18	19	21
S	1	4	16	17	18	S	2	4	17	17	21	S	3	4	18		
T	1	1	8	10	11	T	2	1	16	18		T	3	1	11		
T	1	2	5	7	7	T	2	2	12			T	3	2	13	14	14
T	1	3	2	5	6	T	2	3	12			T	3	3	14	16	18
T	1	4	9	11	11	T	2	4	16	17	18	T	3	4	14	15	16
T	1	5	8	9	9	T	2	5	19	20	20	T	3	5	15	16	18
T	1	6	13	14	16	T	2	6	16			T	3	6	18	18	20
T	1	7	11	12		T	2	7	13	16	17	T	3	7	7		
T	1	8	19	20		T	2	8	13	15		T	3	8	11		

where D_j denotes the j th drug, T_k denotes the k th time effect, and $(DT)_{jk}$ denotes the interaction effect between time and drug. In addition, η_i denotes a random center effect with the assumption that $\eta_i \sim i.i.d. N(0, \sigma_\eta^2)$, $\delta_{ij\ell}$ denotes a random subject effect for the ℓ th subject assigned the j th drug in the i th center. It is assumed that $\delta_{ij\ell} \sim N(0, \sigma_\delta^2)$. Finally, $\varepsilon_{ijk\ell}$ denotes the random error associated with the k th time for the ℓ th subject assigned the j th drug in the i th center with the assumption that $\varepsilon_{ijk\ell} \sim i.i.d. N(0, \sigma_\varepsilon^2)$. It is also assumed that all random effects are independently distributed.

If the split-plot-in-time assumption is not appropriate, then the model would be

$$y_{ijkl} = \mu + \eta_i + D_j + T_k + (DT)_{jk} + \varepsilon_{ijkl} \quad (28.6)$$

where it is assumed that $\eta_i \sim i.i.d. N(0, \sigma_\eta^2)$ and

$$\boldsymbol{\varepsilon}_{ij\ell} = \begin{bmatrix} \varepsilon_{ij1\ell} \\ \varepsilon_{ij2\ell} \\ \varepsilon_{ij3\ell} \end{bmatrix} \sim N_3 \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} \right) \sim N_3(\mathbf{0}, \boldsymbol{\Sigma}) \text{ (say)}$$

and the analysis would depend on the structure assumed for $\boldsymbol{\Sigma}$. Note that the model in Equation 28.6 is the same as the model in (28.1) except that the random effect corresponding to a subject, $\delta_{ij\ell}$, has been removed from the model. One might also want to include person as a random effect for some $\boldsymbol{\Sigma}$ such as when $\boldsymbol{\Sigma}$ has an AR(1) structure.

Four covariance structures for the repeated measures were considered: unstructured, compound symmetry, AR(1), and AR(1) plus a random subject effect. The BIC criterion was smallest for the unstructured assumption while AIC and AICC were both smallest for the AR(1) plus a random subject effect structure. For this latter structure the estimate of ρ was

TABLE 28.34

SAS Commands Used to Analyze the Data in Table 28.33

```

TITLE 'Ex. 28.3 A MULTI-CENTER TRIAL';
DATA ONEB;
  INPUT CENTER $ DRUG SUBJ Y1-Y3 @@;
  CARDS;
R 1 1 17 . . R 2 1 18 19 21 R 3 1 16 16 18
R 1 2 12 14 15 R 2 2 19 . . R 3 2 16 16 18
R 1 3 12 11 14 R 2 3 18 19 . R 3 3 16 16 19
R 1 4 13 13 17 R 2 4 16 15 19 R 3 4 19 19 23
R 1 5 12 13 . R 2 5 19 18 . R 3 5 17 16 20
S 1 1 18 18 21 S 2 1 23 23 26 S 3 1 23 .
S 1 2 15 14 16 S 2 2 15 15 19 S 3 2 17 17 20
S 1 3 6 6 . S 2 3 15 16 19 S 3 3 18 19 21
S 1 4 16 17 18 S 2 4 17 17 21 S 3 4 18 .
T 1 1 8 10 11 T 2 1 16 18 . T 3 1 11 .
T 1 2 5 7 7 T 2 2 12 . . T 3 2 13 14 14
T 1 3 2 5 6 T 2 3 12 . . T 3 3 14 16 18
T 1 4 9 11 11 T 2 4 16 17 18 T 3 4 14 15 16
T 1 5 8 9 9 T 2 5 19 20 20 T 3 5 15 16 18
T 1 6 13 14 16 T 2 6 16 . . T 3 6 18 18 20
T 1 7 11 12 . T 2 7 13 16 17 T 3 7 7 .
T 1 8 19 20 . T 2 8 13 15 . T 3 8 11 .

RUN;

DATA TWO; SET ONE; DROP Y1-Y3;
  IF CENTER='R' THEN SUBJ=100+10*DRUG+SUBJ;
  IF CENTER='S' THEN SUBJ=200+10*DRUG+SUBJ;
  IF CENTER='T' THEN SUBJ=300+10*DRUG+SUBJ;

  TIME=1; Y=Y1; OUTPUT;
  TIME=2; Y=Y2; OUTPUT;
  TIME=3; Y=Y3; OUTPUT;
RUN;
ODS RTF FILE='C:\TEMP.RTF';

PROC MIXED DATA=TWO;
  TITLE2 'ANALYSIS USING MIXED - ASSUMING AN UNSTRUCTURED';
  TITLE3 'COVARIANCE MATRIX FOR THE REPEATED MEASURES';
  CLASSES CENTER SUBJ DRUG TIME;
  MODEL Y=DRUG|TIME/DDFM=SATTERTH;
  LSMEANS DRUG|TIME/PDIFF;
  ODS LISTING EXCLUDE LSMEANS DIFFS;
  ODS OUTPUT LSMEANS=LSMS DIFFS=PDIFFS;
  ODS RTF EXCLUDE LSMEANS DIFFS;
  RANDOM CENTER;
  REPEATED TIME/SUBJECT=SUBJ TYPE=UN R=2;
RUN;

PROC PRINT DATA=LSMS; WHERE EFFECT='DRUG';
PROC PRINT DATA=PDIFFS; WHERE EFFECT='DRUG';
PROC PRINT DATA=LSMS; WHERE EFFECT='TIME';
PROC PRINT DATA=PDIFFS; WHERE EFFECT='TIME';
RUN;

ODS RTF CLOSE;

```

-0.3141, which does not seem to make sense from a philosophical point of view as this would mean that the correlation between the first and second time points is negative, but the correlation between the first and third time points is positive, which does not seem reasonable. Consequently, inferences on the fixed effects will be obtained by assuming the repeated measures have an unstructured covariance structure.

The SAS-Mixed commands used to obtain inferences on the fixed effects are given in Table 28.34. The estimates of the covariance parameters are shown in Table 28.35, and hypothesis tests for the fixed effects are shown in Table 28.36. An examination of Table 28.36 reveals that the Drug and Time main effects are both significant and the Drug \times Time interaction effect is not significant. Consequently, comparisons can be made among the drug and time main effect means. Table 28.37 gives the drug main effect means and Table 28.38 gives pairwise comparisons among the drug main effect means. An examination of these two tables reveals that the drug 1 mean is significantly smaller than both the drug 2

TABLE 28.35

Estimates of Covariance Parameters

Covariance Parameter Estimates

Covariance Parameter	Subject	Estimate
CENTER		4.1996
UN(1,1)	SUBJ	10.9733
UN(2,1)	SUBJ	10.1984
UN(2,2)	SUBJ	10.5365
UN(3,1)	SUBJ	9.9781
UN(3,2)	SUBJ	9.3742
UN(3,3)	SUBJ	9.9515

TABLE 28.36

Hypothesis Tests of Fixed Effects

Type III Tests of Fixed Effects

Effect	Num df	Den df	F-Value	Pr > F
Drug	2	44.2	11.35	0.0001
Time	2	34.6	134.49	<0.0001
Drug \times Time	4	34.4	1.04	0.4022

TABLE 28.37

Drug Least Squares Means

Effect	Drug	Time	Estimate	Standard Error	df	t-Value	pt
Drug	1	—	13.1182	1.4176	2.44	9.25	0.0059
Drug	2	—	18.0637	1.4187	2.45	12.73	0.0027
Drug	3	—	16.9923	1.4186	2.45	11.98	0.0031

TABLE 28.38

Pairwise Comparison between Drug Main Effect Means

Effect	Drug	Time	Drug	Time	Estimate	Standard Error	df	t-Value	pt
Drug	1	—	2	—	-4.9454	1.0929	44.1	-4.53	<0.0001
Drug	1	—	3	—	-3.8740	1.0921	44	-3.55	0.0009
Drug	2	—	3	—	1.0714	1.0943	44.4	0.98	0.3328

TABLE 28.39

Time Least Squares Means

Effect	Drug	Time	Estimate	Standard Error	df	t-Value	pt
Time	—	1	14.8982	1.2767	1.6	11.67	0.0153
Time	—	2	15.6157	1.2753	1.57	12.24	0.0152
Time	—	3	17.6602	1.2722	1.59	13.88	0.0119

TABLE 28.40

Pairwise Comparisons between Time Main Effect Means

Effect	Drug	Time	Drug	Time	Estimate	Standard Error	df	t-Value	pt
Time	—	1	—	2	-0.7175	0.1651	38.4	-4.34	<0.0001
Time	—	1	—	3	-2.7619	0.1698	30.8	-16.26	<0.0001
Time	—	2	—	3	-2.0445	0.2199	33.8	-9.30	<0.0001

TABLE 28.41

Data from a Soil Study Experiment

IRR 1														
REP1														
Dur 1				Dur 2			Dur 1				Dur 2			
Depth	Site A	Site B	Site C	Site A	Site B	Site C	Depth	Site A	Site B	Site C	Site A	Site B	Site C	
1	5.91	5.98	5.94	6.73	6.44	6.22	1	5.94	5.83	5.91	6.35	6.16	6.26	
2	5.94	6.05	6.18	6.82	6.84	6.64	2	6.10	5.96	5.97	6.42	6.52	6.20	
3	7.40	7.31	7.44	8.12	7.89	7.96	3	7.55	7.15	7.13	7.96	7.63	7.34	
4	7.14	7.14	7.20	7.84	7.87	7.77	4	7.03	7.60	7.56	7.16	7.38	6.83	
5	7.62	7.62	7.67	8.56	8.54	8.65	5	7.59	7.60	7.59	8.62	8.61	8.45	
6	7.61	7.63	7.62	8.60	8.59	8.29	6	7.59	7.61	7.51	8.65	8.51	8.49	
IRR 2														
REP1														
Dur 1				Dur 2			Dur 1				Dur 2			
Depth	Site A	Site B	Site C	Site A	Site B	Site C	Depth	Site A	Site B	Site C	Site A	Site B	Site C	
1	5.81	5.86	5.59	6.62	6.48	6.14	1	5.71	5.59	5.19	6.61	6.51	6.00	
2	6.06	5.90	5.84	6.24	6.28	6.25	2	5.69	5.46	5.30	6.31	6.39	6.32	
3	6.98	6.76	7.00	8.47	8.33	8.63	3	7.38	7.73	7.67	7.81	7.78	7.80	
4	6.19	6.14	6.78	7.77	7.66	8.05	4	6.99	7.03	7.42	7.63	7.46	7.09	
5	7.42	7.48	7.55	8.40	8.36	8.34	5	7.62	6.75	7.55	8.19	8.46	8.50	
6	7.54	7.55	7.31	8.54	8.46	8.44	6	7.53	7.54	7.54	8.46	8.59	8.66	

mean and the drug 3 mean. In addition, there is no significant difference between the drug 2 and drug 3 means. Table 28.39 gives the time main effect means and Table 28.40 gives pairwise comparisons between the time means. The time means increase over time, and all pairwise comparisons are statistically significant.

28.4 Exercises

- 28.1 An experiment consists of eight columns of soil in tubes that are 20 cm in diameter and 100 cm long. Four of the columns were assigned to a type of irrigation (Irr1 and Irr2), and two columns within each type of irrigation were assigned an irrigation cycle each day [one watering for one hour each day (Dur 1) and watering three times a day for 20 min each watering (Dur 2)]. The columns were exposed to the sun and were orientated so that each column had a southern exposure. At the end of the study, three core soil samples were taken from each tube—one core was taken from the south side, one from the middle and one from the north side of the column. The value of pH of the soil was measured at six depths (0, 20, 40, 60, 80, and 100 cm) within each core sample. The data are shown in Table 28.41.
- 1) Determine the experimental units used in this study assuming all repeated measures in this experiment satisfy split-plot-in-time type assumptions.
 - 2) Determine the design and treatment structure for each size of experimental unit.
 - 3) Write out an appropriate model and key out the analysis of variance table.
 - 4) Carry out an appropriate repeated measures analysis for the following data set assuming all repeated measures in this experiment satisfy split-plot-in-time type assumptions.
 - 5) Consider several other kinds of structures for the repeated measures in this experiment. What covariance structure would you select for these data? Why?
 - 6) Give a complete analysis of the data assuming the covariance structure selected in 5.

29

Analysis of Crossover Designs

Crossover designs are used to compare treatments that are administered to an experimental unit (such as an animal or a person) in a sequence. That is, each experimental unit is subjected to each treatment in a predetermined sequence. The objective of crossover designs is to eliminate between-experimental unit variation in comparing treatments by observing treatments applied to the same experimental unit.

Although crossover designs eliminate between experimental unit variation from treatment comparisons, other problems may arise in the form of carryover or residual effects. Carryover effects occur, say, when treatment A is given first and its effect has not worn off by the time treatment B is applied. If this lingering effect of A interferes with the response of the subject to treatment B (either positively or negatively), then there is a residual effect of treatment A on the response to treatment B.

The crossover design model must contain a sequence effect, a time effect, a treatment effect, carryover effects, an experimental unit error term, and a time interval error term. The first section discusses a general model and its assumptions, and the last two sections discuss crossover designs for two treatments in two periods and more than two periods, respectively.

In the general crossover design, t treatments are compared where each treatment is observed on each of the experimental units; that is, the treatments are applied in a specified sequence to the experimental units. The experimenter constructs s sequences of the t treatments and randomly assigns experimental units to the i th sequence. Table 29.1 contains a set of sequences of three treatments that could be applied to experimental units. The assignment of sequences to subjects (a possible experimental unit) means that the subject is the experimental unit for sequences; the assignment of a treatment to a time interval means that the time interval is the experimental unit for treatments.

29.1 Definitions, Assumptions, and Models

A model to describe the response of an observation from the ℓ th animal assigned to the j th sequence during the i th time period is

$$y_{ijkl} = \mu + S_i + \delta_{i\ell} + P_j + T_k + \varepsilon_{ijkl} \quad (29.1)$$
$$i = 1, 2, \dots, s, \quad j = 1, 2, \dots, p, \quad k = 1, 2, \dots, t, \quad \text{and} \quad \ell = 1, 2, \dots, n_j$$

TABLE 29.1

Possible Set of Sequences for Applying Three Treatments (A, B, and C) to Experimental Units

Sequence	Time/Period		
	1	2	3
1	A	B	C
2	A	C	B
3	B	A	C
4	B	C	A
5	C	A	B
6	C	B	A

In the above model S_i is the effect of the i th sequence, P_j is the j th period effect, and T_k is the effect of the k th treatment where the value of k is determined by the combination of i and j for the i th sequence and the j th period. Under ideal conditions, one would also have $\delta_{il} \sim i.i.d. N(0, \sigma_\delta^2)$, $\varepsilon_{ijkl} \sim i.i.d. N(0, \sigma_\varepsilon^2)$, and all δ_{il} and ε_{ijkl} independent of one another.

The ideal conditions on the error terms δ_{il} and ε_{ijkl} are important in developing an appropriate analysis. Since the experimental units are randomly assigned to the sequences, an appropriate assumption is that $\delta_{il} \sim i.i.d. N(0, \sigma_\delta^2)$. Since ε_{ijkl} is the error of the time interval and is, in some sense, a repeated measure error, it may be that these errors are not independently distributed. The ideal conditions given above are equivalent to assuming the covariance matrix of the time period errors satisfies a compound symmetry condition (see Section 27.1). In that case, the usual analysis of variance methods for independent δ_{il} and ε_{ijkl} can be used to analyze the observed data. The compound symmetry assumption is necessarily met for two-period designs and may be met for many three-period designs. A time series error structure may be more appropriate for crossover designs with three or more periods, and the methods discussed in Chapter 27 should be utilized.

The two-period/two-treatment crossover design is described in Section 29.2 and designs with more than two treatments or more than two periods are described in Sections 29.3 and 29.4.

29.2 Two Period/Two Treatment Designs

Consider a two period/two treatment crossover design with two sequences of treatments, AB and BA , where AB means that treatment A is assigned to an experimental unit in period 1 and treatment B is assigned to the same experimental unit in period 2. Likewise the BA sequence means that treatment B is assigned to an experimental unit in period 1 and treatment A is assigned to the same experimental unit in period 2. Also suppose that there are n_1 experimental units assigned to the AB sequence and that there are n_2 experimental units assigned to the BA sequence. Let the model for an observed response be

$$\begin{aligned} y_{ijkl} &= \mu + S_i + \delta_{il} + P_j + T_k + \varepsilon_{ijkl} \\ i &= AB, BA, \quad j = 1, 2, \quad k = A, B, \quad \text{and} \quad l = 1, 2, \dots, n_i \end{aligned} \tag{29.2}$$

where the value of the subscript k is determined by the sequence i and period j combination. Let

$$\mu_{ij} = \mu + S_i + P_j + T_k$$

be a model for the cell mean corresponding to the i th sequence and the j th period, and let $\hat{\mu}_{ij}$ be the observed mean of all of the observations in the (i, j) th cell. Then note that

$$\begin{aligned}\hat{\mu}_{11} &= \bar{y}_{11..} \text{ estimates } \mu + S_{AB} + P_1 + T_A \\ \hat{\mu}_{12} &= \bar{y}_{12..} \text{ estimates } \mu + S_{AB} + P_2 + T_B \\ \hat{\mu}_{21} &= \bar{y}_{21..} \text{ estimates } \mu + S_{BA} + P_1 + T_B\end{aligned}\quad (29.3)$$

and

$$\hat{\mu}_{22} = \bar{y}_{22..} \text{ estimates } \mu + S_{BA} + P_2 + T_A$$

From Equation 29.3, one can see that

$$\begin{aligned}\hat{\mu}_1 &= \frac{\bar{y}_{11..} + \bar{y}_{12..}}{2} \text{ estimates } \mu + S_{AB} + \bar{P}_+ + \frac{T_A + T_B}{2} \\ \hat{\mu}_2 &= \frac{\bar{y}_{21..} + \bar{y}_{22..}}{2} \text{ estimates } \mu + S_{BA} + \bar{P}_+ + \frac{T_A + T_B}{2} \\ \hat{\mu}_{.1} &= \frac{\bar{y}_{11..} + \bar{y}_{21..}}{2} \text{ estimates } \mu + \bar{S}_+ + P_1 + \frac{T_A + T_B}{2}\end{aligned}\quad (29.4)$$

and

$$\hat{\mu}_{.2} = \frac{\bar{y}_{12..} + \bar{y}_{22..}}{2} \text{ estimates } \mu + \bar{S}_+ + P_2 + \frac{T_A + T_B}{2}$$

Note also that the difference in the two sequence means, $\hat{\mu}_1 - \hat{\mu}_2$, estimates $S_{AB} - S_{BA}$, the difference in the two sequence parameters. The difference in the period 1 and period 2 means, $\hat{\mu}_{.1} - \hat{\mu}_{.2}$ estimates $P_1 - P_2$. Finally, to estimate the difference in the two treatments, $T_A - T_B$, one uses $(\hat{\mu}_{11} - \hat{\mu}_{12} - \hat{\mu}_{21} + \hat{\mu}_{22})/2$. Under the ideal conditions, one can show that

$$\begin{bmatrix} \hat{\mu}_{11} \\ \hat{\mu}_{12} \\ \hat{\mu}_{21} \\ \hat{\mu}_{22} \end{bmatrix} \sim N \left(\begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{21} \\ \mu_{22} \end{bmatrix}, \begin{bmatrix} \frac{\sigma_e^2 + \sigma_\delta^2}{n_1} & \frac{\sigma_\delta^2}{n_1} & 0 & 0 \\ \frac{\sigma_\delta^2}{n_1} & \frac{\sigma_e^2 + \sigma_\delta^2}{n_1} & 0 & 0 \\ 0 & 0 & \frac{\sigma_e^2 + \sigma_\delta^2}{n_2} & \frac{\sigma_\delta^2}{n_2} \\ 0 & 0 & \frac{\sigma_\delta^2}{n_2} & \frac{\sigma_e^2 + \sigma_\delta^2}{n_2} \end{bmatrix} \right) \quad (29.5)$$

From Equation 29.5, one can show that

$$\begin{aligned}\hat{\mu}_{1\cdot} - \hat{\mu}_{2\cdot} &\sim N\left[S_{AB} - S_{BA}, \left(\frac{\sigma_e^2 + 2\sigma_\delta^2}{2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right] \\ \hat{\mu}_{\cdot 1} - \hat{\mu}_{\cdot 2} &\sim N\left[P_1 - P_2, \frac{\sigma_e^2}{2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right]\end{aligned}$$

and that

$$\frac{\hat{\mu}_{11} - \hat{\mu}_{12} - \hat{\mu}_{21} + \hat{\mu}_{22}}{2} \sim N\left[T_A - T_B, \frac{\sigma_e^2}{2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right]$$

The form of an ANOVA table corresponding to the two period/two treatment crossover design being discussed in this section is shown in Table 29.2.

The F -statistic for testing $H_0: T_A = T_B$ is given by $F = TMS/WSEMS$ where TMS is the treatment mean square in Table 29.2 and $WSEMS$ is the within subject error mean square. One rejects H_0 if $F > F_{\alpha, n_1+n_2-2}$. A $(1 - \alpha)100\%$ confidence interval for $T_A - T_B$ is given by

$$\frac{\hat{\mu}_{11} - \hat{\mu}_{12} - \hat{\mu}_{21} + \hat{\mu}_{22}}{2} \pm t_{\alpha, v} \sqrt{\frac{\hat{\sigma}_e^2}{2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

where $v = n_1 + n_2 - 2$ and $\hat{\sigma}_e^2 = WSEMS$. As an example, consider data used by Grizzle (1965). The data are shown in Table 29.3. The data were analyzed with the SAS®-GLM procedure using the commands shown in Table 29.4.

Table 29.5 shows the model ANOVA and gives the value of the $WSEMS$ as $\hat{\sigma}_e^2 = 1.245$. Table 29.6 gives the test of $H_0: T_A = T_B$ along with the between-subject error mean square, $BSEMS = \hat{\sigma}_e^2 + 2\hat{\sigma}_\delta^2 = 1.001$. The observed significance level for H_0 is $\hat{\alpha} = 0.1165$. The treatment A and B least squares means are shown in Table 29.7 along with the observed significance level of $\hat{\alpha} = 0.1165$ comparing treatment A to treatment B . The expected mean squares for the rows in the ANOVA table are shown in Table 28.8. If one wants to find the estimates of the two-way Sequence \times Period means, one can use the commands in Table 29.9. The two-way means are shown in Table 29.10.

The data in Table 29.3 can also be analyzed with the SAS-Mixed procedure using the commands shown in Table 29.11. The results differ slightly from those obtained from the GLM analysis due to the fact that the between-subject error mean square is smaller than the

TABLE 29.2

ANOVA Table for a Two Period/Two Treatment Crossover Experiment

Source of Variation	Degrees of Freedom	Expected Mean Square
Sequence	1	$\sigma_e^2 + 2\sigma_\delta^2 + Q(\text{Sequence})$
Error (between subject)	$n_1 + n_2 - 2$	$\sigma_e^2 + 2\sigma_\delta^2$
Treatment	1	$\sigma_e^2 + Q(\text{Treatment})$
Period	1	$\sigma_e^2 + Q(\text{Period})$
Error (within subject)	$n_1 + n_2 - 2$	σ_e^2

TABLE 29.3
Grizzle's (1965) Data

Seq	Period	Trt	Person	Y
AB	1	A	11	0.2
AB	2	B	11	1.0
AB	1	A	12	0.0
AB	2	B	12	-0.7
AB	1	A	13	-0.8
AB	2	B	13	0.2
AB	1	A	14	0.6
AB	2	B	14	1.1
AB	1	A	15	0.3
AB	2	B	15	0.4
AB	1	A	16	1.5
AB	2	B	16	1.2
BA	1	B	21	1.3
BA	2	A	21	0.9
BA	1	B	22	-2.3
BA	2	A	22	1.0
BA	1	B	23	0.0
BA	2	A	23	0.6
BA	1	B	24	-0.8
BA	2	A	24	-0.3
BA	1	B	25	-0.4
BA	2	A	25	-1.0
BA	1	B	26	-2.9
BA	2	A	26	1.7
BA	1	B	27	-1.9
BA	2	A	27	-0.3
BA	1	B	28	-2.9
BA	2	A	28	0.9

within-subject error mean square in the GLM analysis. Consequently, the Mixed procedure estimates the variance component corresponding to subjects as $\hat{\sigma}_\delta^2 = 0$ whereas GLM would estimate the same variance component using the method of moments criteria as

$$\hat{\sigma}_\delta^2 = \frac{BSEMS - WSEMS}{2} = \frac{1.005 - 1.245}{2} = \frac{-0.240}{2} = -0.120$$

The results using the code in Table 29.11 are not given at this time, and the interested reader will need to run the code in Table 29.11.

The analyses described above assume that there is no carryover from the treatment given in the first period to the observations taken in the second period. Next the case where carryover exists is considered. When there is carryover, then one can assume that

$$\begin{aligned}\hat{\mu}_{11} &= \bar{y}_{11..} \text{ estimates } \mu + S_{AB} + P_1 + T_A \\ \hat{\mu}_{12} &= \bar{y}_{12..} \text{ estimates } \mu + S_{AB} + P_2 + T_B + \lambda_A \\ \hat{\mu}_{21} &= \bar{y}_{21..} \text{ estimates } \mu + S_{BA} + P_1 + T_B\end{aligned}\tag{29.6}$$

TABLE 29.4

SAS-GLM Code to Analyze the Data in Table 29.3

```

TITLE 'CRSOVER EXAMPLE 29.1 - A TWO PERIOD/TWO TREATMENT DESIGN';
DATA GRIZ;
INPUT SEQ $ PERIOD TRT $ PERSON Y;
LINES;
AB    1     A     11    0.2
AB    2     B     11    1.0
AB    1     A     12    0.0
AB    2     B     12   -0.7
AB    1     A     13   -0.8
AB    2     B     13    0.2
AB    1     A     14    0.6
AB    2     B     14    1.1
AB    1     A     15    0.3
AB    2     B     15    0.4
AB    1     A     16    1.5
AB    2     B     16    1.2
BA    1     B     21    1.3
BA    2     A     21    0.9
BA    1     B     22   -2.3
BA    2     A     22    1.0
BA    1     B     23    0.0
BA    2     A     23    0.6
BA    1     B     24   -0.8
BA    2     A     24   -0.3
BA    1     B     25   -0.4
BA    2     A     25   -1.0
BA    1     B     26   -2.9
BA    2     A     26    1.7
BA    1     B     27   -1.9
BA    2     A     27   -0.3
BA    1     B     28   -2.9
BA    2     A     28    0.9
;
PROC GLM;
  TITLE2 'STATISTICAL ANALYSIS USING SAS-GLM';
  CLASSES SEQ TRT PERIOD PERSON;
  MODEL Y=SEQ PERSON(SEQ) TRT PERIOD;
  LSMEANS TRT/PDIFF;
  RANDOM PERSON(SEQ);
RUN;

```

TABLE 29.5

Model ANOVA with the Within Subject Error Mean Square

Source	df	Sum of Squares	Mean Square	F-Value	Pr > F
Model	15	27.96583333	1.86438889	1.50	0.2435
Error	12	14.94416667	1.24534722		
Corrected total	27	42.91000000			

TABLE 29.6

Type III ANOVA Table for Data in Table 29.3

Source	df	Type III SS	Mean Square	F-Value	Pr > F
Seq	1	4.57333333	4.57333333	3.67	0.0794
Person(Seq)	12	12.00666667	1.00055556	0.80	0.6446
Trt	1	3.56297619	3.56297619	2.86	0.1165
Period	1	6.24297619	6.24297619	5.01	0.0449

TABLE 29.7

Treatment Main Effect Means

$H_0: \text{LSMean1} = \text{LSMean2}$		
Trt	Y LSMean	Pr > t
A	0.36875000	0.1165
B	-0.35208333	

TABLE 29.8

Table of Expected Mean Squares for the Data in Table 29.3

Source	Type III Expected Mean Square
Seq	$\text{Var}(\text{Error}) + 2 \text{Var}[\text{Person}(Seq)] + Q(\text{Seq})$
Person(Seq)	$\text{Var}(\text{Error}) + 2 \text{Var}[\text{Person}(Seq)]$
Trt	$\text{Var}(\text{Error}) + Q(\text{Trt})$
Period	$\text{Var}(\text{Error}) + Q(\text{Period})$

TABLE 29.9

SAS-GLM Code to Analyze the Data in Table 29.3

```

PROC GLM;
  TITLE2 'STATISTICAL ANALYSIS USING SAS-GLM';
  CLASSES SEQ TRT PERIOD PERSON;
  MODEL Y=SEQ PERSON(SEQ) PERIOD SEQ*PERIOD;
    LSMEANS SEQ*PERIOD;
    RANDOM PERSON(SEQ);
RUN;

```

TABLE 29.10Sequence \times Period Means

Seq	Period	Y LSMean
AB	1	0.30000000
AB	2	0.53333333
BA	1	-1.23750000
BA	2	0.43750000

TABLE 29.11

SAS-Mixed Code to Analyze the Data in Table 29.3

```

PROC MIXED;
  TITLE2 'STATISTICAL ANALYSIS USING SAS-MIXED';
  CLASSES SEQ TRT PERIOD PERSON;
  MODEL Y=SEQ PERIOD TRT;
    LSMEANS TRT/PDIFF;
    RANDOM PERSON(SEQ);
  RUN;
PROC MIXED;
  TITLE2 'STATISTICAL ANALYSIS USING SAS-MIXED';
  CLASSES SEQ TRT PERIOD PERSON;
  MODEL Y=SEQ PERIOD SEQ*PERIOD;
    LSMEANS SEQ*PERIOD;
    RANDOM PERSON(SEQ);
  RUN;

```

and

$$\hat{\mu}_{22} = \bar{y}_{22..} \text{ estimates } \mu + S_{BA} + P_2 + T_A + \lambda_B$$

where λ_A is a parameter for the sequence AB corresponding to carryover from treatment A given in period 1 into period 2, and λ_B is a parameter for the sequence BA corresponding to carryover from treatment B given in period 1 into period 2.

From Equation 29.6, one can see that

$$\begin{aligned}
 \hat{\mu}_1 &= \frac{\bar{y}_{11..} + \bar{y}_{12..}}{2} \text{ estimates } \mu + S_{AB} + \bar{P}_+ + \frac{T_A + T_B}{2} + \frac{\lambda_A}{2} \\
 \hat{\mu}_2 &= \frac{\bar{y}_{21..} + \bar{y}_{22..}}{2} \text{ estimates } \mu + S_{BA} + \bar{P}_+ + \frac{T_A + T_B}{2} + \frac{\lambda_B}{2} \\
 \hat{\mu}_{.1} &= \frac{\bar{y}_{11..} + \bar{y}_{21..}}{2} \text{ estimates } \mu + \bar{S}_+ + P_1 + \frac{T_A + T_B}{2}
 \end{aligned} \tag{29.7}$$

and

$$\hat{\mu}_{.2} = \frac{\bar{y}_{12..} + \bar{y}_{22..}}{2} \text{ estimates } \mu + \bar{S}_+ + P_2 + \frac{T_A + T_B}{2} + \frac{\lambda_A + \lambda_B}{2}$$

Carryover is said to exist whenever $\lambda_A \neq \lambda_B$. That is, in the no carryover case λ_A and λ_B do not actually have to be equal to zero, they just have to be equal to one another. In the case where $\lambda_A = \lambda_B$, λ_A and λ_B are confounded with period effects, and one does not need to include the parameters in model (29.2). In the carryover case, the difference in the two sequence means, $\hat{\mu}_1 - \hat{\mu}_2$, estimates $S_{AB} - S_{BA} + [(\lambda_A - \lambda_B)/2]$. The difference in the period 1 and period 2 means, $\hat{\mu}_{.1} - \hat{\mu}_{.2}$, estimates $P_1 - P_2 - [(\lambda_A + \lambda_B)/2]$. Finally, $(\hat{\mu}_{11} - \hat{\mu}_{12} - \hat{\mu}_{21} + \hat{\mu}_{22})/2$ estimates $T_A - T_B + [(\lambda_B - \lambda_A)/2]$. In the case where there is carryover, the second period data does not help one estimate the difference in the two treatments. When there is carryover, one must estimate $T_A - T_B$ by $\hat{\mu}_{11} - \hat{\mu}_{21}$, and this estimator only depends on the

period 1 data. One might also note that, if only period 1 data are going to be used to estimate treatment effect, the model for period 1 can be simplified to

$$y_{k\ell} = \mu + T_k + \varepsilon_{k\ell}^*, \quad k = A, B, \text{ and } \ell = 1, 2, \dots, n_i \quad (29.8)$$

where $\varepsilon_{k\ell}^* = \delta_{k\ell} + \varepsilon_{k\ell}$. Furthermore,

$$\text{Var}(\hat{\mu}_{11} - \hat{\mu}_{21}) = (\sigma_\varepsilon^2 + \sigma_\delta^2) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$$

The form of an ANOVA table corresponding to a two period/two treatment crossover design when there is carryover is shown in Table 29.12. This table has the same form as Table 29.2 except for the expected mean square column.

If there is no carryover, then it should not matter which sequence a subject is assigned to. That is, philosophically, $S_{AB} - S_{BA}$ should be equal to zero. Thus, the F -statistic given by $F = \text{SeqMS/BSEMS}$ provides a test for whether there is carryover or not. If $F > F_{\alpha, 1, n_1+n_2-2}$, then one would conclude that there is a significant carryover effect. It should be noted that the test for carryover is a between-subject comparison, and as such, the test is not as powerful as a test that is based on a within-subject comparison. It is recommended that one should not use the two period/two treatment crossover design if one thinks that there might be carryover effects. In some instances, an experimenter might be able to include a so-called "wash-out" period between the two periods of the crossover design. A wash-out period would be a period of time that is long enough so that any residual effect of the treatment given in the first period would be eliminated or washed-out prior to applying the second treatment to an experimental unit. It is also extremely important to note that the problem of having carryover in a two period/two treatment crossover design is reduced and/or eliminated in some crossover designs that have more than two treatments and/or more than two periods. Such designs will be considered in the next section.

As an example of a two period/two treatment crossover experiment where there might be carryover, consider once again Grizzle's (1965) data that was shown in Table 29.3. The data will be analyzed using the SAS-Mixed procedure using the commands shown in Table 29.13.

The first set of Mixed commands is used to obtain a test for carryover and a test for treatment effects should there not be any significant carryover. These commands also give the treatment main effect means that one can use if there is no carryover. The second set of commands use a different model statement in order to get the Sequence \times Period two-way means, and the ESTIMATE option is included in order to obtain an estimate of treatment effect from period 1 data.

Table 29.14 gives tests on the fixed effects from the first set of Mixed commands in Table 29.13. Note that the SEQ effect is significant at the 0.0665 level, indicating that there is

TABLE 29.12

ANOVA Table for a Two Period/Two Treatment Crossover Experiment

Source of Variation	Degrees of Freedom	Expected Mean Square
Sequence	1	$\sigma_\varepsilon^2 + 2\sigma_\delta^2 + Q(\text{Sequence}, \text{Carryover})$
Error (between subject)	$n_1 + n_2 - 2$	$\sigma_\varepsilon^2 + 2\sigma_\delta^2$
Treatment	1	$\sigma_\varepsilon^2 + Q(\text{Treatment}, \text{Carryover})$
Period	1	$\sigma_\varepsilon^2 + Q(\text{Period}, \text{Carryover})$
Error (within subject)	$n_1 + n_2 - 2$	σ_ε^2

TABLE 29.13

SAS-Mixed Code to Analyze the Data in Table 29.3 when Carryover Exists

```

PROC MIXED;
  TITLE2 'STATISTICAL ANALYSIS USING SAS-MIXED';
  CLASSES SEQ TRT PERIOD PERSON;
  MODEL Y=SEQ PERIOD TRT;
    LSMEANS TRT/PDIFF;
    RANDOM PERSON(SEQ);
RUN;
PROC MIXED;
  TITLE2 'STATISTICAL ANALYSIS USING SAS-MIXED';
  CLASSES SEQ TRT PERIOD PERSON;
  MODEL Y=SEQ PERIOD SEQ*PERIOD;
    LSMEANS SEQ*PERIOD;
    ESTIMATE 'A-B FROM PERIOD 1' SEQ 1 -1 SEQ*PERIOD 1 0 -1 0;
    RANDOM PERSON(SEQ);
RUN;

```

TABLE 29.14

Tests on Fixed Effects

Type III Tests of Fixed Effects

Effect	Num df	Den df	F-Value	Pr > F
Seq	1	12	4.07	0.0665
Period	1	12	5.56	0.0362
Trt	1	12	3.17	0.1002

some evidence of carryover. Also note that the *Trt* effect is only significant at the 0.1002 level. The *Trt* main effect means are shown in Table 29.15, and a comparison between the two treatment means is also given in this table.

Table 29.16 gives the tests on fixed effects from the second set of Mixed commands in Table 29.13. Note that the test for *Seq* is the same as that given in Table 29.14. Also note that the test for *Seq × Per* is the same as the test for *TRT* in Table 29.14. Table 29.17 gives the *Seq × Per* means, and Table 29.18 gives a test that compares the two treatments from only the period 1 data. An examination of Table 29.18 reveals that the carryover effect evidently

TABLE 29.15

Treatment Means and a Test on the Difference of the Two Treatment Means

Least Squares Means

Effect	Trt	Estimate	Standard Error	df	t-Value	Pr > t
Trt	A	0.3688	0.2862	12	1.29	0.2218
Trt	B	-0.3521	0.2862	12	-1.23	0.2421

Differences of Least Squares Means

Effect	Trt	Trt	Estimate	Standard Error	df	t-Value	Pr > t
Trt	A	B	0.7208	0.4047	12	1.78	0.1002

TABLE 29.16

Additional Tests on Fixed Effects

Type III Tests of Fixed Effects

Effect	Num df	Den df	F-Value	Pr > F
Seq	1	12	4.07	0.0665
Period	1	12	5.56	0.0362
Seq × Period	1	12	3.17	0.1002

TABLE 29.17

Sequence × Period Two-Way Means

Least Squares Means

Effect	Seq	Period	Estimate	Standard Error	df	t-Value	Pr > t
Seq × Period	AB	1	0.3000	0.4326	12	0.69	0.5012
Seq × Period	AB	2	0.5333	0.4326	12	1.23	0.2413
Seq × Period	BA	1	-1.2375	0.3747	12	-3.30	0.0063
Seq × Period	BA	2	0.4375	0.3747	12	1.17	0.2656

TABLE 29.18

Treatment A vs Treatment B from Period 1 Estimates

Estimates

Label	Estimate	Standard Error	df	t-Value	Pr > t
A-B from Period 1	1.5375	0.5723	12	2.69	0.0198

masked part of the treatment effect in the analysis given in Table 29.7, where no carryover was assumed as the treatment effect was not significant there, while the comparison using only first-period data finds the treatments to be significantly different.

29.3 Crossover Designs with More than Two Periods

The statistical analysis of a crossover design that has more than two periods and/or more than two treatments when there is no carryover is straightforward. The model in Equation 29.1 is appropriate when the ideal conditions on between subject errors and within subject errors are satisfied. The SAS commands given in Tables 29.4, 29.9, and 29.11 can still be used to obtain the useful and interesting statistics. Since analyses when there is no carryover are straightforward, this section concentrates on situations where carryover exists.

The first case considered is the case where there are still two treatments, but involves sequences having three periods. One possibility is to use the two sequences *ABA* and *BAB*.

For the *ABA* sequence, an experimental unit would receive treatment *A* in period 1, treatment *B* in period 2, and treatment *A* again in period 3. The *BAB* sequence would be handled similarly. Another possibility is to use the two sequences *ABB* and *BAA*. A third possibility is to use all four of these sequences. That is, *ABA*, *BAB*, *ABB* and *BAA* would all be used. This section concentrates on a crossover design using the two sequences *ABA* and *BAB*. The other possibilities can be analyzed similarly.

Table 29.19 identifies the model parameters that can be associated with the sequence by period combinations of the *ABA* and *BAB* crossover design.

Suppose that n_1 subjects have been assigned to the *ABA* sequence and that n_2 subjects have been assigned to the *BAB* sequence. Let $y_{ij\ell}$ be the observed response from subject ℓ in sequence i and period j . Let $\hat{\mu}_{ij} = \bar{y}_{ij\ell}$, $i = 1, 2$; $j = 1, 2, 3$. Under the ideal conditions given for the error terms for the model in Equation 29.1, one can show that

$$\begin{bmatrix} \hat{\mu}_{11} \\ \hat{\mu}_{12} \\ \hat{\mu}_{13} \\ \hat{\mu}_{21} \\ \hat{\mu}_{22} \\ \hat{\mu}_{23} \end{bmatrix} \sim N \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \\ \mu_{21} \\ \mu_{22} \\ \mu_{23} \end{bmatrix}, \quad \left(\begin{array}{cccccc} \frac{\sigma_e^2 + \sigma_\delta^2}{n_1} & \frac{\sigma_\delta^2}{n_1} & \frac{\sigma_\delta^2}{n_1} & 0 & 0 & 0 \\ \frac{\sigma_\delta^2}{n_1} & \frac{\sigma_e^2 + \sigma_\delta^2}{n_1} & \frac{\sigma_\delta^2}{n_1} & 0 & 0 & 0 \\ \frac{\sigma_\delta^2}{n_1} & \frac{\sigma_\delta^2}{n_1} & \frac{\sigma_e^2 + \sigma_\delta^2}{n_1} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\sigma_e^2 + \sigma_\delta^2}{n_2} & \frac{\sigma_\delta^2}{n_2} & \frac{\sigma_\delta^2}{n_2} \\ 0 & 0 & 0 & \frac{\sigma_\delta^2}{n_2} & \frac{\sigma_e^2 + \sigma_\delta^2}{n_2} & \frac{\sigma_\delta^2}{n_2} \\ 0 & 0 & 0 & \frac{\sigma_\delta^2}{n_2} & \frac{\sigma_\delta^2}{n_2} & \frac{\sigma_e^2 + \sigma_\delta^2}{n_2} \end{array} \right) \quad (29.9)$$

Note that the contrast

$$\begin{aligned} \mu_{11} - \frac{1}{2}\mu_{12} - \frac{1}{2}\mu_{13} - \mu_{21} + \frac{1}{2}\mu_{22} + \frac{1}{2}\mu_{23} &= (\mu + S_{ABA} + P_1 - T_A) - \frac{1}{2}(\mu + S_{ABA} + P_2 + T_B + \lambda_A) \\ &\quad - \frac{1}{2}(\mu + S_{ABA} + P_3 + T_A + \lambda_B) - (\mu + S_{BAB} + P_1 + T_B) \\ &\quad + \frac{1}{2}(\mu + S_{BAB} + P_2 + T_A + \lambda_B) + \frac{1}{2}(\mu + S_{BAB} + P_3 + T_B + \lambda_A) \\ &= T_A - T_B \end{aligned}$$

TABLE 29.19

Cell Mean Parameters for a Three Period/Two Treatment Crossover Design

Sequence	Period 1	Period 2	Period 3
<i>ABA</i>	$\mu_{11} = \mu + S_{ABA} + P_1 + T_A$	$\mu_{12} = \mu + S_{ABA} + P_2 + T_B + \lambda_A$	$\mu_{13} = \mu + S_{ABA} + P_3 + T_A + \lambda_B$
<i>BAB</i>	$\mu_{21} = \mu + S_{BAB} + P_1 + T_B$	$\mu_{22} = \mu + S_{BAB} + P_2 + T_A + \lambda_B$	$\mu_{23} = \mu + S_{BAB} + P_3 + T_B + \lambda_A$

Thus even if carryover is present in the *ABA*, *BAB* crossover design, the direct difference in the two treatments can be estimated by

$$\hat{\mu}_{11} - \frac{1}{2}\hat{\mu}_{12} - \frac{1}{2}\hat{\mu}_{13} - \hat{\mu}_{21} + \frac{1}{2}\hat{\mu}_{22} + \frac{1}{2}\hat{\mu}_{23} \quad (29.10)$$

Also note that

$$\text{Var}\left(\hat{\mu}_{11} - \frac{1}{2}\hat{\mu}_{12} - \frac{1}{2}\hat{\mu}_{13} - \hat{\mu}_{21} + \frac{1}{2}\hat{\mu}_{22} + \frac{1}{2}\hat{\mu}_{23}\right) = \frac{3}{2}\sigma_e^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \quad (29.11)$$

so the contrast in Equation 29.10 is a within subject contrast whose variance depends only on σ_e^2 .

One question of interest might be how should treatment main effect means be defined? One possibility is to estimate the *A* main effect mean by averaging over all the cells in Table 29.19 that would receive treatment *A*. That is, the *A* mean would be $(\mu_{11} + \mu_{22} + \mu_{13})/3$ and the *B* mean would be $(\mu_{21} + \mu_{12} + \mu_{23})/3$. In terms of the effects model parameters defined in Table 29.19, these two functions are equal to

$$\mu + \frac{2}{3}S_{ABA} + \frac{1}{3}S_{BAB} + \bar{P}_. + T_A + \frac{1}{3}(\lambda_B + \lambda_A)$$

and

$$\mu + \frac{1}{3}S_{ABA} + \frac{2}{3}S_{BAB} + \bar{P}_. + T_B + \frac{1}{3}(\lambda_A + \lambda_B)$$

respectively. Such a definition would not make sense since the difference in such an *A* mean and *B* mean is equal to $\frac{1}{3}S_{ABA} - \frac{1}{3}S_{BAB} + T_A - T_B$ which is not equal to $T_A - T_B$. That is, the difference in such a defined *A* mean and *B* mean is aliased with sequence effect. A second definition of an *A* main effect mean is

$$\mu + \bar{S}_. + \bar{P}_. + T_A + \frac{1}{3}(\lambda_B + \lambda_A)$$

and a *B* main effect mean would be similarly defined by

$$\mu + \bar{S}_. + \bar{P}_. + T_B + \frac{1}{3}(\lambda_B + \lambda_A)$$

Such definitions are reasonable since the first is equal to the following linear function of the cell means in Table 29.19

$$\frac{2}{3}\mu_{11} - \frac{1}{12}\mu_{12} - \frac{1}{12}\mu_{13} - \frac{1}{3}\mu_{21} + \frac{5}{12}\mu_{22} + \frac{5}{12}\mu_{23}$$

and the second is equal to

$$-\frac{1}{3}\mu_{11} + \frac{5}{12}\mu_{12} + \frac{5}{12}\mu_{13} + \frac{2}{3}\mu_{21} - \frac{1}{12}\mu_{22} - \frac{1}{12}\mu_{23}$$

Furthermore, the difference in these two means is

$$\mu_{11} - \frac{1}{2}\mu_{12} - \frac{1}{2}\mu_{13} - \mu_{21} + \frac{1}{2}\mu_{22} + \frac{1}{2}\mu_{23} = T_A - T_B$$

Note that this is the same contrast in the cell means as that given in Equation 29.10.
A contrast in the cell means that measures carryover effect is

$$\mu_{11} - \mu_{13} - \mu_{21} + \mu_{23} = \lambda_A - \lambda_B \quad (29.12)$$

and the variance of the estimate of this contrast is

$$\text{Var}(\hat{\mu}_{11} - \hat{\mu}_{13} - \hat{\mu}_{21} + \hat{\mu}_{23}) = 2\sigma_e^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \quad (29.13)$$

and this contrast is also a within-subjects contrast.

To illustrate with an example, consider the data in Table 29.20 where there were three subjects assigned to the *ABA* sequence and another three subjects assigned to the *BAB* sequence. Thus $n_1 = n_2 = 3$. The SAS code that reads the data and computes the *Sequence* \times *Period* cell means is shown in Table 29.21. The cell means are shown in Table 29.22.

Thus the estimate of the *A* main effect mean is given by

$$\begin{aligned} & \frac{2}{3}\hat{\mu}_{11} - \frac{1}{12}\hat{\mu}_{12} - \frac{1}{12}\hat{\mu}_{13} - \frac{1}{3}\hat{\mu}_{21} + \frac{5}{12}\hat{\mu}_{22} + \frac{5}{12}\hat{\mu}_{23} \\ &= \frac{2}{3}(24.133) - \frac{1}{12}(26.533) - \frac{1}{12}(23.933) - \frac{1}{3}(26.367) + \frac{5}{12}(26.233) + \frac{5}{12}(24.833) = 24.372 \end{aligned}$$

and the estimate of the *B* main effect mean is given by

$$\begin{aligned} & -\frac{1}{3}\hat{\mu}_{11} + \frac{5}{12}\hat{\mu}_{12} + \frac{5}{12}\hat{\mu}_{13} + \frac{2}{3}\hat{\mu}_{21} - \frac{1}{12}\hat{\mu}_{22} - \frac{1}{12}\hat{\mu}_{23} \\ &= -\frac{1}{3}(24.133) + \frac{5}{12}(26.533) + \frac{5}{12}(23.933) + \frac{2}{3}(26.367) - \frac{1}{12}(26.233) - \frac{1}{12}(24.833) = 26.306 \end{aligned}$$

TABLE 29.20

Data for an *ABA/BAB* Crossover Experiment

Seq	Per	Trt	Person	Y
<i>ABA</i>	1	<i>A</i>	1	25.1
<i>ABA</i>	2	<i>B</i>	1	27.6
<i>ABA</i>	3	<i>A</i>	1	24.5
<i>ABA</i>	1	<i>A</i>	2	22.0
<i>ABA</i>	2	<i>B</i>	2	24.3
<i>ABA</i>	3	<i>A</i>	2	21.6
<i>ABA</i>	1	<i>A</i>	3	25.3
<i>ABA</i>	2	<i>B</i>	3	27.7
<i>ABA</i>	3	<i>A</i>	3	25.7
<i>BAB</i>	1	<i>B</i>	4	25.5
<i>BAB</i>	2	<i>A</i>	4	23.7
<i>BAB</i>	3	<i>B</i>	4	24.9
<i>BAB</i>	1	<i>B</i>	5	27.4
<i>BAB</i>	2	<i>A</i>	5	27.9
<i>BAB</i>	3	<i>B</i>	5	24.6
<i>BAB</i>	1	<i>B</i>	6	26.2
<i>BAB</i>	2	<i>A</i>	6	27.1
<i>BAB</i>	3	<i>B</i>	6	25.0

TABLE 29.21SAS Code to Obtain Seq \times Per Means for the Data in Table 29.20

```

TITLE 'CRSOVR EXAMPLE #29.2 AN ABA/BAB DESIGN';
DATA CRS;
  INPUT SEQ $ PER TRT $ PERSON Y;
  LINES;
ABA 1 A 1 25.1
ABA 2 B 1 27.6
ABA 3 A 1 24.5
ABA 1 A 2 22.0
ABA 2 B 2 24.3
ABA 3 A 2 21.6
ABA 1 A 3 25.3
ABA 2 B 3 27.7
ABA 3 A 3 25.7
BAB 1 B 4 25.5
BAB 2 A 4 23.7
BAB 3 B 4 24.9
BAB 1 B 5 27.4
BAB 2 A 5 27.9
BAB 3 B 5 24.6
BAB 1 B 6 26.2
BAB 2 A 6 27.1
BAB 3 B 6 25.0
;
PROC MEANS;
  CLASS SEQ PER;
  VAR Y;
RUN;

```

TABLE 29.22

Sequence by Period Cell Means

Seq	PER	N	Mean
ABA	1	3	24.1333333
	2	3	26.5333333
	3	3	23.9333333
BAB	1	3	26.3666667
	2	3	26.2333333
	3	3	24.8333333

The difference between the *A* mean and the *B* mean is

$$\hat{\tau}_A - \hat{\tau}_B = 24.372 - 26.306 = -1.934$$

One can use the SAS-Mixed procedure to estimate the treatment means, test for treatment differences, test for carryover, and estimate the variance components for two treatment/three period crossover designs. For example, in order to analyze the data in Table 29.20, one must define a new variable that identifies the treatment that was given in the previous

TABLE 29.23

SAS Code to Create a PRIORTRT Variable for the Data in Table 29.20

```

DATA CRS2; SET CRS;
IF SEQ='ABA' AND PER=2 THEN PRIORTRT='A';
ELSE IF SEQ='ABA' AND PER=3 THEN PRIORTRT='B';
ELSE IF SEQ='BAB' AND PER=2 THEN PRIORTRT='B';
ELSE IF SEQ='BAB' AND PER=3 THEN PRIORTRT='A';
ELSE PRIORTRT='O';
RUN;

PROC PRINT;
TITLE2 'A PRINT OF THE DATA WITH PRIORTRT VARIABLE INCLUDED';
RUN;

```

period. So that the first period data are not eliminated, this new variable has to also be defined for the period one data. For the *ABA* sequence, this new variable could be defined to take on the values *O*, *A*, and *B* for periods 1, 2, and 3, respectively. And for the *BAB* sequence, the new variable would take on values *O*, *B*, and *A*, respectively. The SAS code in Table 29.23 uses a new variable, called *PRIORTRT*, that has the properties given in the preceding sentences. A display of the data with this new variable defined is shown in Table 29.24. The SAS-Mixed commands that can be used to analyze the data in Table 29.20 are shown in Table 29.25. The resulting output is shown in Tables 29.26–29.29. Table 29.26 gives estimates of the variance components. From this table, one can see that $\hat{\sigma}_\delta^2 = 2.1378$ and $\hat{\sigma}_\epsilon^2 = 0.7872$. Table 29.27 gives tests for difference between treatment *A* and treatment *B*. The row of Table 29.27 labeled *Trt* is a test of $H_{01}: \tau_A = \tau_B$ and the row labeled *PRIORTRT* is

TABLE 29.24

Print of New Data with PRIORTRT Variable Defined

Obs	Seq	Per	Trt	Person	Y	PRIORTRT
1	ABA	1	A	1	25.1	O
2	ABA	2	B	1	27.6	A
3	ABA	3	A	1	24.5	B
4	ABA	1	A	2	22.0	O
5	ABA	2	B	2	24.3	A
6	ABA	3	A	2	21.6	B
7	ABA	1	A	3	25.3	O
8	ABA	2	B	3	27.7	A
9	ABA	3	A	3	25.7	B
10	BAB	1	B	4	25.5	O
11	BAB	2	A	4	23.7	B
12	BAB	3	B	4	24.9	A
13	BAB	1	B	5	27.4	O
14	BAB	2	A	5	27.9	B
15	BAB	3	B	5	24.6	A
16	BAB	1	B	6	26.2	O
17	BAB	2	A	6	27.1	B
18	BAB	3	B	6	25.0	A

TABLE 29.25

SAS Code to Analyze the Data in Table 29.24

```

PROC MIXED;
TITLE3 'AN ANALYSIS USING MIXED';
CLASSES SEQ PERSON TRT PER PRIORTRT;
MODEL Y=SEQ TRT PER PRIORTRT/DDFM=KR;
LSMEANS TRT/PDIFF;
ESTIMATE 'TRT DIFF' TRT 1 -1;
ESTIMATE 'CARRYOVER' PRIORTRT 1 -1;
CONTRAST 'PERIOD DIFF' PER 0 1 -1, PER 1 -.5 -.5 PRIORTRT -.5 -.5 1;
RANDOM PERSON(SEQ);
RUN;

```

TABLE 29.26

Estimates of the Variance Components

Covariance Parameter Estimates

Covariance Parameter	Estimate
Person(Seq)	2.1378
Residual	0.7872

TABLE 29.27

Tests on the Fixed Effects

Type III Tests of Fixed Effects

Effect	Num df	Den df	F-Value	Pr > F
Seq	1	4	0.05	0.8287
Trt	1	8	4.75	0.0610
Per	1	8	15.24	0.0045
PRIORTRT	1	8	1.69	0.2293

Contrasts

Label	Num df	Den df	F-Value	Pr > F
Period diff	2	8	7.67	0.0138

a test of H_{02} : $\lambda_A = \lambda_B$. Note that the row of Table 29.27 labeled PER has only one degree of freedom associated with it. This may seem strange since there are three periods. However, there is complete confounding between period 1 and λ_0 , so the row labeled PER compares only periods 2 and 3. That is, the row labeled PER tests H_{03} : $P_2 = P_3$. The CONTRAST option in Table 29.25 gives a test statistic that compares all three periods to one another; that is, the results of this CONTRAST option tests H_{03} : $P_1 + \lambda_0 = P_2 = P_3$. The results of this option are appended to the bottom of Table 29.27.

The estimates of the treatment main effect means are shown in Table 29.28. Note that the values given by the SAS-Mixed procedure are the same as the estimates that were computed from the Table 29.22 using the second definition for main effect means. The results from the two ESTIMATE statements in Table 29.25 are shown in Table 29.29. The first of

TABLE 29.28

Treatment Main Effect Means

Least Squares Means						
Effect	Trt	Estimate	Standard Error	df	t-Value	Pr > t
Trt	A	24.3722	0.7726	7.94	31.55	<0.0001
Trt	B	26.3056	0.7726	7.94	34.05	<0.0001

TABLE 29.29

Tests for Treatment and Carryover Effects from Estimate Options

Estimates						
Label	Estimate	Standard Error	df	t-Value	Pr > t	
Trt diff	-1.9333	0.8873	8	-2.18	0.0610	
Carryover	-1.3333	1.0245	8	-1.30	0.2293	

these estimates $\tau_A - \tau_B$, gives the estimate's estimated standard error and gives a *t*-statistic for testing $H_{01}: \tau_A = \tau_B$ as well as the test's observed significance level. The second ESTIMATE statement estimates $\lambda_A - \lambda_B$, gives the estimate's estimated standard error and gives a *t*-statistic for testing $H_{02}: \lambda_A = \lambda_B$ as well as the test's observed significance level. The reader should note that the observed significance levels in Table 29.29 for treatment difference and carryover difference are the same as the corresponding observed significance levels in Table 29.27. The estimated standard error for the estimated treatment difference can be computed by substituting $\hat{\sigma}_e^2$ for σ_e^2 in (29.11). One gets

$$\begin{aligned}\widehat{\text{Var}}\left(\hat{\mu}_{11} - \frac{1}{2}\hat{\mu}_{12} - \frac{1}{2}\hat{\mu}_{13} - \hat{\mu}_{21} + \frac{1}{2}\hat{\mu}_{22} + \frac{1}{2}\hat{\mu}_{23}\right) &= \frac{3}{2}\hat{\sigma}_e^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \\ &= \left(\frac{3}{2}\right)(0.7872)\left(\frac{1}{3} + \frac{1}{3}\right) = 0.7872\end{aligned}$$

And thus the estimated standard error of $\hat{\tau}_A - \hat{\tau}_B$ is $\sqrt{0.7872} = 0.8872$.

Similarly, one can get the estimated standard error of $\hat{\lambda}_A - \hat{\lambda}_B$ by substituting $\hat{\sigma}_e^2$ for σ_e^2 in Equation 29.13 and taking the square root of the result. The degrees of freedom associated with each of these estimated standard errors is $2(n_1 + n_2 - 2) = 8$.

29.4 Crossover Designs with More than Two Treatments

Next, consider the three treatment/three period crossover design in six sequences. The design sequences are *ABC*, *ACB*, *BAC*, *BCA*, *CAB*, and *CBA*. All six of these sequences are required in order to get a direct comparison of treatments that will not be aliased with carryover. Williams (1949) developed designs that are balanced with respect to carryover effects. A Williams crossover design that is balanced for carryover effects has the property that every treatment is followed by every other treatment exactly the same number of times.

The six-sequence design given above is a Williams design. In this six-sequence design, each treatment (letter) follows each other treatment twice. Note also that, while the three-sequence design with the sequences ABC , BCA , and CAB is a Latin square design, it is not a Williams design. In this design, A is followed by B twice, but A is never followed by C . Likewise, B is followed by C twice, but B is never followed by A ; and C is followed by A twice, but C is never followed by B . A Williams design for three treatments in three periods must use all six of the sequences ABC , ACB , BAC , BCA , CAB , and CBA .

Suppose one has four treatments and four periods. Clearly, the 24-sequence design that has all 24 possible orderings of the treatments A , B , C , and D will be a Williams design. Another Williams design for four treatments in four periods is given by the sequences $ABCD$, $BDAC$, $CADB$, and $DCBA$. Let t be the number of treatments and periods in the crossover design to be used. If t is an even number, then a Williams design can be constructed using a single special $t \times t$ Latin square design, and if t is odd, a Williams design can be constructed by using two special $t \times t$ Latin square designs.

Table 29.30 gives the cell mean parameters for the three period/three treatment Williams design. To show that one can estimate treatment differences when carryover is present, one can show that for the cell mean definitions in Table 29.30

$$(5\mu_{11} - 2\mu_{12} - 3\mu_{13} + 4\mu_{21} + 2\mu_{22} - 6\mu_{23} - 5\mu_{31} + 2\mu_{32} + 3\mu_{33} - 4\mu_{41} - 2\mu_{41} + 6\mu_{43} - \mu_{51} + 4\mu_{52} - 3\mu_{53} + \mu_{61} - 4\mu_{62} + 3\mu_{63})/24 = \tau_A - \tau_B \quad (29.14)$$

Also note that one can also compare carryover effects as

$$(\mu_{11} + 2\mu_{12} - 3\mu_{13} + 0\mu_{21} + 2\mu_{22} - 2\mu_{23} - \mu_{31} - 2\mu_{32} + 3\mu_{33} + 0\mu_{41} - 2\mu_{41} + 2\mu_{43} - \mu_{51} + 0\mu_{52} + \mu_{53} + \mu_{61} + 0\mu_{62} - \mu_{63})/8 = \lambda_A - \lambda_B \quad (29.15)$$

Similar functions of the cell mean parameters that simplify to $\tau_A - \tau_C$ and $\tau_B - \tau_C$ as well as for $\lambda_A - \lambda_C$ and $\lambda_B - \lambda_C$ can also be obtained.

As an example consider the data in Table 29.31. These data come from a three period/three treatment crossover design in six sequences. There were a few subjects that had missing data. A SAS data set was created using the commands in Table 29.32. These commands also create a carryover parameter that gives the value of the treatment in the previous period. This parameter has a value of "O" for first period data. The data were initially analyzed with SAS-Mixed using the commands in Table 29.33. Table 29.34 gives the estimates of the between-subject and within-subject variance components. That is $\hat{\sigma}_\delta^2 = 3.2278$ and $\hat{\sigma}_e^2 = 0.8934$.

TABLE 29.30

Cell Mean Parameters for a Three Period/Three Treatment Crossover Design in Six Sequences

Sequence	Period 1	Period 2	Period 3
ABC	$\mu_{11} = \mu + S_{ABC} + P_1 + T_A$	$\mu_{12} = \mu + S_{ABC} + P_2 + T_B + \lambda_A$	$\mu_{13} = \mu + S_{ABC} + P_3 + T_C + \lambda_B$
ACB	$\mu_{21} = \mu + S_{ACB} + P_1 + T_A$	$\mu_{22} = \mu + S_{ACB} + P_2 + T_C + \lambda_A$	$\mu_{23} = \mu + S_{ACB} + P_3 + T_B + \lambda_C$
BAC	$\mu_{31} = \mu + S_{BAC} + P_1 + T_B$	$\mu_{32} = \mu + S_{BAC} + P_2 + T_A + \lambda_B$	$\mu_{33} = \mu + S_{BAC} + P_3 + T_C + \lambda_A$
BCA	$\mu_{41} = \mu + S_{BCA} + P_1 + T_B$	$\mu_{42} = \mu + S_{BCA} + P_2 + T_C + \lambda_B$	$\mu_{43} = \mu + S_{BCA} + P_3 + T_A + \lambda_C$
CAB	$\mu_{51} = \mu + S_{CAB} + P_1 + T_C$	$\mu_{52} = \mu + S_{CAB} + P_2 + T_A + \lambda_C$	$\mu_{53} = \mu + S_{CAB} + P_3 + T_B + \lambda_A$
CBA	$\mu_{61} = \mu + S_{CBA} + P_1 + T_C$	$\mu_{62} = \mu + S_{CBA} + P_2 + T_B + \lambda_C$	$\mu_{63} = \mu + S_{CBA} + P_3 + T_A + \lambda_B$

TABLE 29.31

Data for a Three Period/Three Treatment Crossover Design

Seq	Subj	Y1	Y2	Y3	Seq	Subj	Y1	Y2	Y3
ABC	11	20.1	20.3	—	ACB	21	24.7	29.4	27.5
ABC	12	23.3	24.8	28.7	ACB	22	23.8	—	24.1
ABC	13	23.4	—	28.3	ACB	23	23.6	—	25.0
ABC	14	19.7	21.3	25.7	ACB	24	20.2	—	—
ABC	15	19.2	20.9	25.9	ACB	25	19.8	23.7	23.3
ABC	16	22.2	22.0	—	ACB	26	21.5	25.5	20.8
BAC	31	24.3	—	30.1	BCA	41	20.9	27.5	24.3
BAC	32	26.4	26.4	32.3	BCA	42	21.9	28.6	23.1
BAC	33	19.9	23.7	25.5	BCA	43	22.0	27.4	—
BAC	34	23.9	26.8	30.8	BCA	44	23.3	30.7	26.6
BAC	35	20.5	23.2	26.3	BCA	45	18.8	27.9	24.6
BAC	36	21.8	23.6	—	BCA	46	24.6	29.8	26.6
CAB	51	24.0	21.8	21.6	CBA	61	23.2	18.9	23.8
CAB	52	25.9	23.7	—	CBA	62	23.9	21.5	25.4
CAB	53	25.5	—	23.4	CBA	63	28.0	25.3	28.1
CAB	54	27.9	25.4	24.4	CBA	64	24.6	22.7	23.8
CAB	55	25.3	26.4	25.8	CBA	65	27.7	23.5	25.6
CAB	56	25.7	—	24.9	CBA	66	21.5	18.1	22.8

Table 29.35 gives the tests on the fixed effects. An examination of Table 29.35 shows that there is a large significant treatment effect ($p < 0.0001$) and a significant carryover effect ($p = 0.0306$). Tables 29.36 and 29.37 give the results of the LSMEANS option. It is interesting to note than neither the treatment nor the prior treatment least squares means are estimable. The reason for this is that the SAS-Mixed procedure generates functions of the model parameters that are not estimable.

For example, the SAS-Mixed procedure defines the treatment A least squares mean as

$$\mu + \frac{S_{ABC} + S_{ACB} + S_{BAC} + S_{BCA} + S_{CAB} + S_{CBA}}{6} + \frac{P_1 + P_2 + P_3}{3} + \tau_A + \frac{\lambda_A + \lambda_B + \lambda_C + \lambda_O}{4}$$

This function is not estimable because P_1 and λ_O are completely aliased with one another in the model, and consequently, they both must have the same multipliers. In the above expression, they do not as the multiplier on P_1 is equal to $\frac{1}{3}$ while the multiplier on λ_O is equal to $\frac{1}{4}$. A better definition of the treatment A least squares mean is

$$\mu + \frac{S_{ABC} + S_{ACB} + S_{BAC} + S_{BCA} + S_{CAB} + S_{CBA}}{6} + \frac{P_1 + P_2 + P_3}{3} + \tau_A + \frac{2\lambda_A + 2\lambda_B + 2\lambda_C + 3\lambda_O}{9} \quad (29.16)$$

In this definition, both P_1 and λ_O have a multiplier equal to $\frac{1}{3}$. Furthermore, since the sum of the multipliers on the four carryover parameters must add to one, the multipliers on λ_A , λ_B , and λ_C must each be equal to $\frac{2}{9}$. The function in Equation 29.16 is estimable. In a similar manner, one can define treatment B and treatment C least squares means, and pairwise differences among these means will not include carryover parameters.

TABLE 29.32

SAS Code to Analyze the Data in Table 29.31

```

DATA ONE; INPUT SEQ $ SUBJ @@;
DO PER=1 TO 3; INPUT Y @@; OUTPUT; END;
LINES;
ABC 11 20.1 20.3 . ACB 21 24.7 29.4 27.5
ABC 12 23.3 24.8 28.7 ACB 22 23.8 . 24.1
ABC 13 23.4 . 28.3 ACB 23 23.6 . 25.0
ABC 14 19.7 21.3 25.7 ACB 24 20.2 . .
ABC 15 19.2 20.9 25.9 ACB 25 19.8 23.7 23.3
ABC 16 22.2 22.0 . ACB 26 21.5 25.5 20.8
BAC 31 24.3 . 30.1 BCA 41 20.9 27.5 24.3
BAC 32 26.4 26.4 32.3 BCA 42 21.9 28.6 23.1
BAC 33 19.9 23.7 25.5 BCA 43 22.0 27.4 .
BAC 34 23.9 26.8 30.8 BCA 44 23.3 30.7 26.6
BAC 35 20.5 23.2 26.3 BCA 45 18.8 27.9 24.6
BAC 36 21.8 23.6 . BCA 46 24.6 29.8 26.6
CAB 51 24.0 21.8 21.6 CBA 61 23.2 18.9 23.8
CAB 52 25.9 23.7 . CBA 62 23.9 21.5 25.4
CAB 53 25.5 . 23.4 CBA 63 28.0 25.3 28.1
CAB 54 27.9 25.4 24.4 CBA 64 24.6 22.7 23.8
CAB 55 25.3 26.4 25.8 CBA 65 27.7 23.5 25.6
CAB 56 25.7 . 24.9 CBA 66 21.5 18.1 22.8
DATA TWO; SET ONE;
IF SEQ='ABC' AND PER=1 THEN TRT='A'; IF SEQ='ABC' AND PER=1 THEN PRIORTRT='O';
IF SEQ='ABC' AND PER=2 THEN TRT='B'; IF SEQ='ABC' AND PER=2 THEN PRIORTRT='A';
IF SEQ='ABC' AND PER=3 THEN TRT='C'; IF SEQ='ABC' AND PER=3 THEN PRIORTRT='B';
IF SEQ='ACB' AND PER=1 THEN TRT='A'; IF SEQ='ACB' AND PER=1 THEN PRIORTRT='O';
IF SEQ='ACB' AND PER=2 THEN TRT='C'; IF SEQ='ACB' AND PER=2 THEN PRIORTRT='A';
IF SEQ='ACB' AND PER=3 THEN TRT='B'; IF SEQ='ACB' AND PER=3 THEN PRIORTRT='C';
IF SEQ='BAC' AND PER=1 THEN TRT='B'; IF SEQ='BAC' AND PER=1 THEN PRIORTRT='O';
IF SEQ='BAC' AND PER=2 THEN TRT='A'; IF SEQ='BAC' AND PER=2 THEN PRIORTRT='B';
IF SEQ='BAC' AND PER=3 THEN TRT='C'; IF SEQ='BAC' AND PER=3 THEN PRIORTRT='A';
IF SEQ='BCA' AND PER=1 THEN TRT='B'; IF SEQ='BCA' AND PER=1 THEN PRIORTRT='O';
IF SEQ='BCA' AND PER=2 THEN TRT='C'; IF SEQ='BCA' AND PER=2 THEN PRIORTRT='B';
IF SEQ='BCA' AND PER=3 THEN TRT='A'; IF SEQ='BCA' AND PER=3 THEN PRIORTRT='C';
IF SEQ='CAB' AND PER=1 THEN TRT='C'; IF SEQ='CAB' AND PER=1 THEN PRIORTRT='O';
IF SEQ='CAB' AND PER=2 THEN TRT='A'; IF SEQ='CAB' AND PER=2 THEN PRIORTRT='C';
IF SEQ='CAB' AND PER=3 THEN TRT='B'; IF SEQ='CAB' AND PER=3 THEN PRIORTRT='A';
IF SEQ='CBA' AND PER=1 THEN TRT='C'; IF SEQ='CBA' AND PER=1 THEN PRIORTRT='O';
IF SEQ='CBA' AND PER=2 THEN TRT='B'; IF SEQ='CBA' AND PER=2 THEN PRIORTRT='C';
IF SEQ='CBA' AND PER=3 THEN TRT='A'; IF SEQ='CBA' AND PER=3 THEN PRIORTRT='B';
RUN;

```

TABLE 29.33

SAS Code to Analyze the Data in Table 29.31

```

PROC MIXED;
CLASSES SUBJ SEQ PER TRT PRIORTRT;
MODEL Y=SEQ PER TRT PRIORTRT/DDFM=KR;
RANDOM SUBJ(SEQ);
LSMEANS TRT PRIORTRT/PDIFF;
RUN;

```

TABLE 29.34

Estimates of the Variance Components

Covariance Parameter Estimates

Covariance Parameter	Estimate
Subj(Seq)	3.2278
Residual	0.8934

TABLE 29.35

Test on the Fixed Effects

Type III Tests of Fixed Effects

Effect	Num df	Den df	F-Value	Pr > F
Seq	5	31	1.07	0.3942
Per	1	54.7	11.93	0.0011
Trt	2	54.3	112.07	<0.0001
PRIORTRT	2	54.6	3.72	0.0306

TABLE 29.36

Treatment Least Squares Means

Least Squares Means

Effect	Trt	PRIORTRT	Estimate	Standard Error	df	t-Value	Pr > t
Trt	A		Nonestimable	—	—	—	—
Trt	B		Nonestimable	—	—	—	—
Trt	C		Nonestimable	—	—	—	—
PRIORTRT	A		Nonestimable	—	—	—	—
PRIORTRT	B		Nonestimable	—	—	—	—
PRIORTRT	C		Nonestimable	—	—	—	—
PRIORTRT	O		Nonestimable	—	—	—	—

Even though the treatment least squares means were nonestimable in Table 29.36, the pairwise differences are estimable, and the pairwise comparisons between the treatment and carryover parameters are given in Table 29.37. From Table 29.37, one can see that

$$\hat{\tau}_A - \hat{\tau}_B = 0.8195, \quad \hat{\tau}_A - \hat{\tau}_C = -3.0861, \quad \text{and} \quad \hat{\tau}_B - \hat{\tau}_C = -3.9056$$

The corresponding observed significance levels are $p = 0.0041$, $p < 0.0001$, and $p < 0.0001$. Thus all treatments are significantly different from one another.

Also from Table 29.37, one can see that

$$\hat{\lambda}_A - \hat{\lambda}_B = -0.7706, \quad \hat{\lambda}_A - \hat{\lambda}_C = 0.1842, \quad \text{and} \quad \hat{\lambda}_B - \hat{\lambda}_C = 0.9548$$

The corresponding observed significance levels are $p = 0.0557$, $p = 0.6380$, and $p = 0.0118$. Thus the carryover effects from treatments A and C are similar to one another, and both are significantly different from the carryover from treatment B.

TABLE 29.37

Pairwise Comparisons among Treatments and Carryover Parameters

Differences of Least Squares Means

Effect	Trt	PRIORTRT	Trt	PRIORTRT	Estimate	Standard Error	df	t-Value	Pr > t
Trt	A		B		0.8195	0.2737	54.3	2.99	0.0041
Trt	A		C		-3.0861	0.2738	54.5	-11.27	<0.0001
Trt	B		C		-3.9056	0.2753	54.2	-14.19	<0.0001
PRIORTRT		A		B	-0.7706	0.3942	54.8	-1.95	0.0557
PRIORTRT		A		C	0.1842	0.3893	54.8	0.47	0.6380
PRIORTRT		A		O	Nonestimable	—	—	—	—
PRIORTRT		B		C	0.9548	0.3665	54.3	2.61	0.0118
PRIORTRT		B		O	Nonestimable	—	—	—	—
PRIORTRT		C		O	Nonestimable	—	—	—	—

TABLE 29.38

Additional SAS Code to Compute Estimable Least Squares Means for Treatments

```

ESTIMATE 'LSM A' INTERCEPT 18 SEQ 3 3 3 3 3 PER 6 6 6 TRT 18 0 0
PRIORTRT 4 4 4 6/DIVISOR=18;
ESTIMATE 'LSM B' INTERCEPT 18 SEQ 3 3 3 3 3 PER 6 6 6 TRT 0 18 0
PRIORTRT 4 4 4 6/DIVISOR=18;
ESTIMATE 'LSM C' INTERCEPT 18 SEQ 3 3 3 3 3 PER 6 6 6 TRT 0 0 18
PRIORTRT 4 4 4 6/DIVISOR=18;
ESTIMATE 'A-B' TRT 1 -1 0;
ESTIMATE 'A-C' TRT 1 0 -1;
ESTIMATE 'B-C' TRT 0 1 -1;
RUN;

```

TABLE 29.39

Results from the Commands Given in Table 29.38

Estimates

Label	Estimate	Standard Error	df	t-Value	Pr > t
LSM A	23.7217	0.3525	45.3	67.29	<0.0001
LSM B	22.9022	0.3511	44.7	65.23	<0.0001
LSM C	26.8078	0.3567	46.9	75.16	<0.0001
A-B	0.8195	0.2737	54.3	2.99	0.0041
A-C	-3.0861	0.2738	54.5	-11.27	<0.0001
B-C	-3.9056	0.2753	54.2	-14.19	<0.0001
CARRY_A-CARRY_B	-0.7706	0.3942	54.8	-1.95	0.0557
CARRY_A-CARRY_C	0.1842	0.3893	54.8	0.47	0.6380
CARRY_B-CARRY_C	0.9548	0.3665	54.3	2.61	0.0118

Table 29.38 gives additional SAS code that computes treatment least squares means using definitions similar to that for treatment A given in (29.16). The results from the SAS commands in Table 29.38 are given in Table 29.39. An examination of Table 29.39 reveals that treatment C has the largest mean, followed in order by treatments A and B. Also note

TABLE 29.40

Pain Score

Subject	Sequence	Response 1	Response 2
1	AB	49.6	54.0
2	AB	45.5	58.6
3	AB	45.1	45.1
4	AB	46.5	40.7
5	AB	26.1	63.9
6	AB	45.6	61.6
7	AB	43.0	46.4
8	AB	51.2	56.7
9	AB	32.0	66.7
10	AB	55.3	33.3
11	AB	43.9	65.2
12	AB	41.4	46.6
21	BA	48.7	49.4
22	BA	40.8	26.9
23	BA	47.6	48.6
24	BA	60.5	30.3
25	BA	38.7	44.6
26	BA	49.7	53.1
27	BA	65.5	39.9
28	BA	51.5	43.4
29	BA	38.6	43.1
30	BA	48.8	30.4
31	BA	39.4	48.8
32	BA	44.6	33.9
33	BA	60.1	60.5
34	BA	44.8	24.7
35	BA	33.8	48.1
36	BA	48.2	46.1

TABLE 29.41

Milk Yields

Cow	Sequence	Period 1	Period 2	Period 3
1	ABC	38	25	15
2	BCA	109	86	39
3	CAB	124	72	27
4	ACB	86	76	46
5	BAC	75	35	34
6	CBA	101	63	1

TABLE 29.42
Data from a 13 Treatment/4 Period Crossover Design

Period	Steer	Trt	ACE Average (m 100 m ⁻¹)	PROP Average (m 100 m ⁻¹)	ISOBUT Average (m 100 m ⁻¹)	BUT Average (m 100 m ⁻¹)	ISOVAL Average (m 100 m ⁻¹)	VAL Average (m 100 m ⁻¹)	LAC Average (m 100 m ⁻¹)	VFA (m Ml ⁻¹)	Total
1	5048	HPLS	72.48	13.13	1.02	8.75	1.90	1.57	1.16	46.20	
2	5048	HPHD	45.53	9.52	0.72	12.96	0.96	1.09	29.22	64.99	
3	5048	LPHF	74.63	15.99	0.69	7.22	0.86	0.61	0.00	66.79	
4	5048	HPLF	72.19	13.50	1.23	8.08	1.96	1.57	1.46	67.33	
1	5106	HPLD	65.87	11.95	0.46	13.39	0.73	1.01	6.60	69.23	
2	5106	LPHF	74.53	14.02	0.53	9.44	0.82	0.48	0.18	54.88	
3	5106	LPHS	71.41	14.47	0.61	12.17	0.76	0.48	0.10	62.44	
4	5106	HPLS	71.62	12.17	1.02	10.02	1.54	1.30	2.33	76.18	
1	5107	HPHF	74.13	14.38	0.69	8.54	1.19	1.07	0.00	62.40	
2	5107	HPLF	73.44	14.15	0.60	9.54	1.24	1.02	0.00	57.12	
3	5107	HPLS	73.03	13.49	0.64	10.83	0.97	1.05	0.00	78.54	
4	5107	LPLD	62.06	11.93	0.35	13.12	0.51	0.43	11.60	67.97	
1	5113	LPHD	55.11	12.10	0.43	14.27	0.49	0.95	16.65	86.30	
2	5113	LPLD	60.39	10.55	0.37	11.94	0.55	0.51	15.69	58.55	
3	5113	HPHF	73.75	14.43	0.68	8.86	1.06	1.16	0.07	72.07	
4	5113	LPHS	74.04	13.23	0.52	9.66	0.79	0.66	1.10	63.03	
1	5182	LPLF	77.43	13.89	0.73	6.52	0.85	0.58	0.00	70.99	
2	5182	LPHS	72.09	15.80	0.66	10.15	1.00	0.30	0.00	54.08	
3	5182	LPHD	58.96	13.25	0.46	12.11	0.59	1.12	13.51	80.11	
4	5182	HPLD	61.52	12.00	0.45	10.97	0.84	0.88	13.33	75.59	
1	5186	LPHS	73.11	13.51	0.65	11.30	0.78	0.66	0.00	57.25	
2	5186	IPLS	72.71	13.04	0.52	11.57	0.79	1.25	0.13	52.49	
3	5186	IPLD	63.84	12.24	0.40	15.24	0.53	0.46	7.29	71.56	
4	5186	LPHF	75.16	12.85	0.48	9.56	0.77	0.49	0.69	58.99	
1	5189	HPLF	72.60	15.09	0.77	8.58	1.37	1.33	0.26	53.08	

Continued

TABLE 29.42 (continued)

Period	Steer	Trt	ACE Average (m 100 m ⁻¹)	PROP Average (m 100 m ⁻¹)	ISOBUT Average (m 100 m ⁻¹)	BUT Average (m 100 m ⁻¹)	ISOVAL Average (m 100 m ⁻¹)	VAL Average (m 100 m ⁻¹)	LAC Average (m 100 m ⁻¹)	VFA (m Ml ⁻¹)	Total
2	5189	HPHS	70.68	16.91	0.71	9.37	1.27	1.06	0.00	51.84	
3	5189	HPHD	46.38	10.97	0.40	10.14	0.47	0.82	30.81	88.38	
4	5189	NC	77.24	12.73	0.43	8.93	0.33	0.35	0.00	57.78	
1	5236	LPHF	76.45	14.15	0.68	6.93	0.78	0.58	0.42	63.11	
2	5236	IPLF	74.90	13.88	0.59	8.76	0.94	0.56	0.36	54.99	
3	5236	IPLS	74.37	13.25	0.61	9.47	0.74	1.56	0.00	72.69	
4	5236	HPHD	64.81	13.95	0.63	14.49	0.74	0.98	4.42	68.50	
1	5260	HPHS	69.16	15.36	0.88	11.17	1.42	1.39	0.61	64.12	
2	5260	HPLS	68.82	15.16	0.84	11.42	1.42	1.33	1.01	74.31	
3	5260	HPLD	61.44	12.96	0.49	15.92	0.78	1.17	7.23	83.77	
4	5260	HPHF	71.78	14.29	0.75	9.64	1.20	1.02	1.31	80.55	
1	5262	NC	77.63	11.12	0.72	8.51	0.42	0.36	1.25	66.44	
2	5262	HPHF	74.85	13.23	0.54	9.48	1.04	0.87	0.00	59.93	
3	5262	HPHS	70.30	16.49	0.85	9.83	1.36	1.18	0.00	96.30	
4	5262	LPHD	67.53	12.86	0.43	16.57	0.60	0.72	1.31	72.58	
1	5266	IPLD	51.23	28.06	0.40	8.55	0.26	0.59	10.92	53.90	
2	5266	NC	74.12	14.03	0.63	9.77	0.52	0.46	0.47	39.97	
3	5266	HPLF	71.00	15.40	0.62	10.48	1.04	1.21	0.25	67.79	
4	5266	IPLS	60.37	11.38	0.38	14.65	0.59	0.69	11.94	78.90	
1	5267	HPHD	55.06	10.89	0.67	22.59	0.70	1.00	9.10	80.74	
2	5267	HPLD	59.39	11.68	0.74	17.81	1.16	1.51	7.71	64.35	
3	5267	IPLF	74.70	14.55	0.77	8.41	0.97	0.60	0.00	72.21	
4	5267	HPHS	70.95	14.84	0.82	9.36	1.37	1.21	1.45	71.21	
1	5310	IPLS	75.14	11.44	0.60	10.20	0.78	1.03	0.81	57.12	
2	5310	LPHD	47.21	10.04	0.38	11.00	0.58	1.26	29.55	75.71	
3	5310	NC	77.42	11.59	0.51	9.68	0.41	0.39	0.00	63.77	
4	5310	IPLF	72.85	13.62	0.67	10.09	0.91	0.57	1.28	73.64	

that the pairwise comparisons between the treatment means and carryover effects given in Table 29.39 are exactly the same as the corresponding comparisons in Table 29.37.

29.5 Summary

This chapter considered crossover designs where each subject receives a series of treatments over time. The particular series of treatments that a specific subject receives is called a sequence. When using crossover designs one needs to consider the issue of carryover. If one suspects that carryover might be present, one can include a washout period prior to giving the next treatment in a sequence in order to minimize as much as possible the effects of carryover. When carryover exists in a two period/two treatment crossover design, then one cannot use the data from the second period to estimate treatment differences. However, when one is able to use more than two periods and/or more than two treatments, then estimates of treatment differences can be obtained whether carryover exists or not.

29.6 Exercises

- 29.1 An experiment was conducted to study a toothpaste's ability to desensitize sensitive teeth. In Table 29.40, A represents a desensitizing toothpaste and B represents a control toothpaste. Twelve subjects were assigned to the sequence AB and another 16 subjects were assigned to the sequence BA . Response 1 corresponds to a measure of pain on a 0–100 scale for the toothpaste received in period 1, and response 2 corresponds to the toothpaste received in period 2. Answer each of the following questions:
- 1) Does there appear to be any carryover effects? Why or why not?
 - 2) Assuming no carryover, find a 95% confidence interval for the difference between the desensitizing toothpaste and the control toothpaste.
 - 3) Assuming carryover exists, find a 95% confidence interval for the difference between the desensitizing toothpaste and the control toothpaste.
- 29.2 Table 29.41 contains data on milk yields for cows being fed one of three different diets over time. The data are taken from Cochran and Cox (1957). The diets are denoted by A , B , and C .
- 1) Do there appear to be any carryover effects? Why or why not?
 - 2) Compare the three diets with one another assuming a no carryover model.
 - 3) Compare the three diets with one another assuming a model with carryover.
- 29.3 Table 29.42 contains data from a 13 treatment/four period crossover design. There were 13 sequences of diet treatments with each sequence covering four periods. One cow was assigned to each sequence. Answer the following questions for the dependent variables ACE average and PROP average.
- 1) Do there appear to be any carryover effects? Why or why not?
 - 2) Compare the 13 diets with one another assuming a no carryover model.
 - 3) Compare the 13 diets with one another assuming a model with carryover.

30

Analysis of Nested Designs

Nested effects can occur in either the design structure or the treatment structure or both of a designed experiment. For nesting to occur in the design structure, there must be more than one size of experimental unit where a small experimental unit is nested within a larger one. Split-plot, repeated measures and hierarchical designs are examples where there is nesting in the design structure. Nesting in the treatment structure can occur when there are two or more factors. These factors may be all fixed effects, all random effects, or a mixture of both. Thus an experiment with nesting in the treatment structure can be modeled as a fixed, random, or mixed model. The concept of nested factors in the design structure was introduced in Chapter 5. This chapter presents some examples that demonstrate model construction, parameter estimation, and hypothesis testing. Nested designs with nesting in the design structure are often referred to as hierarchical designs.

30.1 Definitions, Assumptions, and Models

In the treatment structure, the levels of factor B are nested within the levels of factor A if each level of B occurs with only one level of factor A . The following examples demonstrate nesting in the treatment structure.

30.1.1 Example 30.1: Companies and Insecticides

Four chemical companies produce certain insecticides. Company A produces three such products, companies B and C produce two such products each, and company D produces four such products. No company produces a product exactly like that of another. The treatment structure is two-way with company as one factor and product as the other. Such a treatment structure is shown in Table 30.1, where each level of product occurs only once within each level of company. Thus the levels of product are nested within the levels of company.

TABLE 30.1

Treatment Structure for the Companies and Insecticide Example

Company	Products										
	1	2	3	4	5	6	7	8	9	10	11
A	X	X	X								
B				X	X						
C						X	X				
D							X	X	X	X	

Note: An "X" denotes that particular product is from the corresponding company.

The levels of both factors in the treatment structure are fixed effects. To conduct the experiment, a box of soil containing live bluegrass plants and 400 mosquitoes were put into each of 33 glass containers. Three glass containers were then randomly assigned to each product. The glass containers were treated with the product, and after 4 h the number of live mosquitoes was counted.

The design structure for this experiment is completely randomized. A model that can be used to describe the number of live mosquitoes from each container is

$$y_{ijk} = \mu + \gamma_i + \rho_{j(i)} + \varepsilon_{ijk}, \quad i = 1, 2, 3, 4, \quad j = 2, 3, \text{ or } 4, \quad k = 1, 2, 3 \quad (30.1)$$

where y_{ijk} is the observed number of mosquitoes from the k th replication of the j th product of the i th company, μ is the overall mean, γ_i is the effect of the i th company, $\rho_{j(i)}$ is the effect of the j th product in company i , and $\varepsilon_{ijk} \sim i.i.d. N(0, \sigma^2_\varepsilon)$ denotes the error associated with measuring y_{ijk} . The model has only one size of experimental unit and thus only one error term. The design of the experiment is a two-way nested treatment structure (both factors are fixed effects) in a completely randomized design structure. The parameters of the model to be estimated are σ^2_ε and estimable functions of μ , γ_i , and $\rho_{j(i)}$. The data for this example are in Table 30.2, and the analysis is discussed in Sections 30.2 and 30.3.

TABLE 30.2

Data for Example 30.1

Company	Products										
	1	2	3	4	5	6	7	8	9	10	11
A	151	118	131								
	135	132	137								
	137	135	121								
B				140	151						
				152	132						
				133	139						
C					96	84					
					108	87					
					94	82					
D						79	67	90	82		
						74	78	81	89		
						73	63	96	98		

30.1.2 Example 30.2: Comfort Experiment Revisited

The comfort experiment in Example 5.10 is an example of nesting in the design structure. The large experimental unit is an environmental chamber, and the small experimental unit is a person. The model is

$$y_{ijk} = \mu_{ik} + c_{j(i)} + p_{m(ijk)} \quad i = 65, 70, 75, \quad j = 1, 2, \dots, 9, \quad k = M, F, \quad m = 1, 2 \quad (30.2)$$

where μ_{ik} denotes the mean of the i th temperature and the k th gender, $c_{j(i)}$ denotes the random effect of the j th chamber assigned to the i th temperature where it is assumed that $c_{j(i)} \sim N(0, \sigma_{\text{Chamber}}^2)$ and $p_{m(ijk)}$ denotes the random effect of the m th person of the k th gender assigned to the j th chamber of the i th temperature where it is assumed that $p_{m(ijk)} \sim N(0, \sigma_{\text{Person}}^2)$. The terms $c_{j(i)}$ and $p_{m(ijk)}$ denote the errors of observing temperature i on chamber j (chambers are nested within temperature) and person m of the k th gender within the j th chamber of the i th temperature (persons are nested within chambers). Since there are two sizes of experimental units, the analysis has two levels and two error terms. The parameters that need to be estimated are μ_{ik} , $i = 65, 70, 75$, $k = M, F$, $\sigma_{\text{Chamber}}^2$, and σ_{Person}^2 . The analysis of this example is discussed in Sections 30.2 and 30.3.

30.1.3 Example 30.3: Coffee Price Example Revisited

The coffee price example in Chapter 18 is a design of a sample survey with a three-way treatment structure, where the levels of store are nested within the levels of city, which are nested within the levels of state (all three factors are random effects). The model used to describe the variability in the coffee prices in Chapter 18 was

$$y_{ijk} = \mu + s_i + c_{j(i)} + a_{k(ij)}, \quad i = 1, 2, \dots, r, \quad j = 1, 2, \dots, t_i, \quad k = 1, 2, \dots, n_{ij}$$

where μ denotes the average price of coffee in the United States, $s_i \sim \text{i.i.d. } N(0, \sigma_{\text{State}}^2)$, $c_{j(i)} \sim \text{i.i.d. } N(0, \sigma_{\text{City}}^2)$, and $a_{k(ij)} \sim \text{i.i.d. } N(0, \sigma_{\text{Store}}^2)$ denotes the random state, city and store effects, respectively. The parameters of interest are the variance components σ_{State}^2 , σ_{City}^2 , σ_{Store}^2 , and the overall mean, μ . The coffee price example is an example of a multistage sampling experiment, a very common application of nested designs. The analysis of this example will be discussed in the next two sections. Table 30.3 contains coffee prices for a small study that is used to demonstrate parameter estimation, confidence interval estimation, and hypothesis testing.

The above examples demonstrate the nesting that can occur in the treatment structure, design structure, or both. The treatment structure can involve random and/or fixed effects, and the design structure can involve several sizes of experimental units. Thus, the analysis of such designs involves using the techniques of fixed effects models, random effects models, and mixed effects models as discussed in the next two sections.

30.2 Parameter Estimation

A model involving nesting falls into one of the classes of models already discussed. That is, if there are fixed effects, the means need to be estimated; if there are random effects,

TABLE 30.3

Coffee Prices in U.S. Cents per Pound for Example 30.3

State	City	Store 1	Store 2	Store 3	Store 4	Store 5	Store 6
1	1	158	146	158	—	—	—
1	2	140	142	144	140	152	—
1	3	156	162	148	156	150	164
2	1	162	162	166	162	166	170
2	2	166	156	168	162	162	168
3	1	142	150	154	156	—	—
3	2	128	132	132	134	126	132
3	3	164	164	146	164	162	—
3	4	140	146	142	138	146	142
4	1	148	138	144	154	150	146
4	2	140	138	150	136	148	—
5	1	156	152	144	148	152	—
5	2	124	148	140	136	—	—
5	3	134	148	144	144	142	—
5	4	148	148	148	158	—	—

variance components need to be estimated; and if there are several sizes of experimental units, the analysis involves more than one error term. Thus the design may be quite simple or quite complex. The examples in this chapter are used to demonstrate the application of the previously discussed techniques to the analysis of designs involving nesting.

30.2.1 Example 30.1: Continuation

The estimable functions of model (30.1) involve linear combinations of $\mu + \gamma_i + \rho_{j(i)}$. For example, contrasts of the $\rho_{j(i)}$ for each i are estimable where $\rho_{1(i)} - \rho_{2(i)}$ is a contrast of the products within the same company where in this case the contrast compares products 1 and 2 from the i th company. Contrasts of $\mu + \gamma_i + \bar{\rho}_{(i)}$ are used to compare companies averaged across each companies' products where $\gamma_1 + \bar{\rho}_{(1)} - \gamma_2 - \bar{\rho}_{(2)}$ compares company 1 to company 2. The contrast $\gamma_1 + \rho_{1(1)} - \gamma_2 - \rho_{1(2)}$ compares product 1 of company 1 to product 1 of company 2. The estimates of these contrasts are

$$\begin{aligned}\widehat{\rho_{1(i)}} - \widehat{\rho_{2(i)}} &= \bar{y}_{i1\cdot} - \bar{y}_{i2\cdot} \\ \widehat{\gamma_1 + \bar{\rho}_{(1)}} - \widehat{\gamma_2 - \bar{\rho}_{(2)}} &= \bar{y}_{1..} - \bar{y}_{2..}\end{aligned}$$

and

$$\widehat{\gamma_1 + \rho_{1(1)}} - \widehat{\gamma_2 - \rho_{1(2)}} = \bar{y}_{11\cdot} - \bar{y}_{21\cdot}$$

respectively.

Model (30.1) can also be expressed as

$$y_{ijk} = \mu_{j(i)} + \varepsilon_{ijk}, \quad i = 1, 2, 3, 4, \quad j = 2, 3, \text{ or } 4, \quad k = 1, 2, 3$$

TABLE 30.4

Estimates of the Means for Example 30.1

Company A or 1			Company B or 2		Company C or 3		Company D or 4			
$\hat{\mu}_{1(1)}$	$\hat{\mu}_{2(1)}$	$\hat{\mu}_{3(1)}$	$\hat{\mu}_{1(2)}$	$\hat{\mu}_{2(2)}$	$\hat{\mu}_{1(3)}$	$\hat{\mu}_{2(3)}$	$\hat{\mu}_{1(4)}$	$\hat{\mu}_{2(4)}$	$\hat{\mu}_{3(4)}$	$\hat{\mu}_{4(4)}$
141.0	128.3	129.7	141.7	140.7	99.3	84.3	75.3	69.3	89.0	89.7

where $\mu_{j(i)}$ is the mean of product j of company i . The estimators of the $\mu_{j(i)}$ are the $\bar{y}_{ij.}$ Any linear combination including contrasts of the $\mu_{j(i)}$ is estimable, but one has to be careful in interpreting selected contrasts. The estimator of σ_e^2 comes from pooling the variances of the observations within the $j(i)$ combinations as

$$\hat{\sigma}_e^2 = \frac{1}{N-q} \sum_{i=1}^m \sum_{j=1}^{p_i} \sum_{k=1}^{n_{j(i)}} (y_{ijk} - \bar{y}_{ij.})^2$$

where there are $n_{j(i)}$ observations in the $j(i)$ cell, there are q cells with $n_{j(i)} > 0$, m is the number of companies, p_i is the number of products from the i th company, $n_{j(i)}$ is the number of glass containers assigned to the j th product of the i th company and $N = \sum_{j(i)} n_{j(i)} =$ total number of observations.

For this example,

$$\hat{\sigma}_e^2 = \frac{1}{33-11} \sum_{i=1}^4 \sum_{j=1}^{p_i} \sum_{k=1}^3 (y_{ijk} - \bar{y}_{ij.})^2 = 60.818$$

and the estimates of the sample means are in Table 30.4.

Multiple comparisons can be made between the $\mu_{j(i)}$. Other hypotheses can be tested about the $\mu_{j(i)}$ such as comparing companies or products within a company. For example, one could compare company B to company D by considering

$$\bar{\mu}_{(2)} = \bar{\mu}_{(4)} \quad \text{where } \bar{\mu}_{(2)} = \frac{\mu_{1(2)} + \mu_{2(2)}}{2} \quad \text{and} \quad \bar{\mu}_{(4)} = \frac{\mu_{1(4)} + \mu_{2(4)} + \mu_{3(4)} + \mu_{4(4)}}{4}$$

that is, the comparison would be between the mean of the two products from company B and the mean of the four products from company D . The researcher will need to decide whether such a comparison would be of interest. Hypothesis testing for this nested treatment structure is discussed in Section 30.3.

If there are unequal numbers of observations in the $j(i)$ cells, then the techniques for analyzing unbalanced models can be used to obtain estimates of the $\mu_{j(i)}$. The estimates of the population marginal means provide estimates of the $\hat{\mu}_{j(i)}$ and the estimate of $\hat{\sigma}_e^2$ is obtained from pooling the variances across the treatment combinations which can be obtained from an analysis of variance.

30.2.2 Example 30.2: Continuation

Data for the comfort study are given in Table 30.5. The nesting occurs in the design structure with person nested within chamber and chamber nested within temperature. The analysis

TABLE 30.5

Data for Example 30.2 Where Values Are Comfort Scores, Where 1 = Cold, 8 = Comfortable, and 15 = Hot

Temperature	Chamber 1		Chamber 2		Chamber 3	
	Male	Female	Male	Female	Male	Female
65	5	4	5	4	4	2
	1	2	5	5	1	3
Chamber 4		Chamber 5		Chamber 6		
70	8	8	6	3	5	7
	10	7	8	8	8	8
Chamber 7		Chamber 8		Chamber 9		
75	12	8	8	7	6	6
	11	13	8	8	6	7

TABLE 30.6

Analysis of Variance Table for Comfort Study of Example 30.2

Source	df	MS	EMS
Temperature	2	79.19	$\sigma_{\text{Person}}^2 + 4\sigma_{\text{Chamber}}^2 + \varphi^2(\text{Temp})$
Gender	1	3.36	$\sigma_{\text{Person}}^2 + \varphi^2(\text{Gender})$
Temperature \times Gender	2	7.86	$\sigma_{\text{Person}}^2 + \varphi^2(\text{Temp} \times \text{Gender})$
Chamber(Temperature)	6	11.08	$\sigma_{\text{Person}}^2 + 4\sigma_{\text{Chamber}}^2$
Person(Chamber)	24	1.65	σ_{Person}^2

of variance table with expected mean squares is shown in Table 30.6. The method-of-moments estimates of the two components of variance (which are also REML and MINQUE0 estimates) are $\hat{\sigma}_{\text{Person}}^2 = 1.65$ and $\hat{\sigma}_{\text{Chamber}}^2 = 2.36$ and the maximum likelihood estimates are $\hat{\sigma}_{\text{Person}}^2 = 1.47$ and $\hat{\sigma}_{\text{Chamber}}^2 = 1.48$. The methods used for split-plot experiments (described in Chapter 24) can be used to compare the μ_{ij} . If the design is unbalanced, then a mixed model analysis using REML to estimate the variance components would be appropriate.

30.2.3 Example 30.3: Continuation

The coffee price study involves a random effects model whose parameters of interest are μ , σ_{State}^2 , σ_{City}^2 , and σ_{Store}^2 . A method of maximum likelihood would provide an analysis and estimators of the parameters, as would the MINQUE0 technique. A method-of-moments analysis employing type I sums of squares has typically been used for this type of multi-stage sampling design. The results of the REML method are preferred. Table 30.7 contains the type I sums of squares, their expected mean squares, and the method of moment estimators for a general study. Table 30.8 contain the type I analysis for the data in Table 30.3. Estimates of the variance components using method-of-moments from type I sums of squares, REML, MIVQUE0, and ML are in Table 30.9. The 95% confidence intervals for the variance components were computed using the chi-square distribution with the stated

TABLE 30.7

Analysis of Variance Table with Type I Sums of Squares and Method of Moments Estimators for Example 30.3

Source	df	SS	EMS
State	$r - 1$	$\sum_{i=1}^r n_{i..} \bar{y}_{i..}^2 - n_{..} \bar{y}_{..}^2$	$\sigma_{\text{Store}}^2 + \frac{a_1 - a_1}{r - 1} \sigma_{\text{City}}^2 + \frac{n_{..} - a_2}{r - 1} \sigma_{\text{State}}^2$
City(State)	$\sum_{i=1}^r (t_i - 1)$	$\sum_{i=1}^r \sum_{j=1}^{t_i} n_{ij} \bar{y}_{ij..}^2 - \sum_{i=1}^r n_{i..} \bar{y}_{i..}^2$	$\sigma_{\text{Store}}^2 + \frac{n_{..} - a_1}{t_{..} - r} \sigma_{\text{City}}^2$
Store(City State)	$\sum_{i=1}^r \sum_{j=1}^{t_i} (n_{ij} - 1)$	$\sum_{i=1}^r \sum_{j=1}^{t_i} \sum_{k=1}^{n_{ij}} y_{ijk}^2 - \sum_{i=1}^r \sum_{j=1}^{t_i} n_{ij} \bar{y}_{ij..}^2$	σ_{Store}^2

Source: Searle (1987).

$$a_1 = \sum_{i=1}^r \left(\sum_{j=1}^{t_i} \frac{n_{ij}^2}{n_{i..}} \right)^2, \quad a_2 = \sum_{i=1}^r \frac{n_{i..}^2}{n_{..}}, \quad a_3 = \sum_{i=1}^r \sum_{j=1}^{t_i} \frac{n_{ij}^2}{n_{..}}$$

$$\hat{\sigma}_{\text{Person}}^2 = \frac{\text{MSSTORE}(City, State)}{n_{..} - t_{..}}, \quad \hat{\sigma}_{\text{City}}^2 = \frac{\text{MSCITY}(STATE) - \hat{\sigma}_{\text{Person}}^2}{(n_{..} - a_1)/(t_{..} - r)}$$

$$\hat{\sigma}_{\text{State}}^2 = \frac{\text{MSSTATE} - \hat{\sigma}_{\text{Person}}^2 - [(a_1 - a_3)/(r - 1)] \hat{\sigma}_{\text{City}}^2}{(n_{..} - a_2)/(r - 1)}$$

TABLE 30.8

Type I Sums of Squares, Mean Squares, Expected Mean Squares and Test Statistics for the Coffee Price Data

Source	df	SS	MS	EMS	Error Term	Error df	F-Value	Pr F	
State	4	3658.67	914.67	Var(Residual) + 5.3113	1.0726 MS[City(State)] - 0.0726 MS(Residual)	—	9.9	2.40	0.1204
				Var[City(State)] + 14.967 Var(State)					
City(State)	10	3579.02	357.90	Var(Residual) + 4.9518	MS(Residual)	61.0	11.58	<0.0001	Var[City(State)]
Residual	61	1885.47	30.91	Var(Residual)	—	—	—	—	—

number of degrees of freedom. Wald's method was used for the intervals provided by the method of moments, but they were recomputed using the chi-square distribution with degrees of freedom computed as $2(Z\text{-value})^2$. The estimates of the mean coffee price from each of the methods of estimating the variance components are given in Table 30.10.

30.3 Testing Hypotheses and Confidence Interval Construction

For balanced designs with nesting in one or both of the structures, the methods for balanced random effects models, balanced mixed effects models, and balanced split-plot-repeated

TABLE 30.9

Estimates of the Variance Components for the Coffee Data Using Four Methods

Method	Covariance Parameter	Estimate	Standard Error	df	Lower	Upper
Type I	State	35.6	56.8	0.8	6.4	225322.5
Type I	City(State)	66.0	33.6	7.7	29.8	249.7
Type I	Residual	30.9	5.6	61	22.3	45.6
REML	State	28.5	47.3	0.7	4.9	361333.6
REML	City(State)	67.2	34.7	7.5	30.0	261.2
REML	Residual	30.9	5.6	61	22.3	45.6
ML	State	12.3	37.2	0.2	1.3	1.169×10^{15}
ML	City(State)	70.8	39.3	6.5	30.2	315.9
ML	Residual	30.9	5.6	61	22.3	45.5
MINQUE0	State	21.3	58.2	0.3	2.4	4.545×10^{12}
MINQUE0	City(State)	88.3	60.0	4.3	32.7	645.5
MINQUE0	Residual	21.3	3.9	60	15.4	31.6

TABLE 30.10

Estimate of the Mean Coffee Price Using Four Methods of Estimating the Variance Components

Method	Estimate	Standard Error	df
Type I	149.7	3.4940	4
REML	149.6	3.2935	4
ML	149.3	2.7822	4
MINQUE0	149.4	3.2649	4

measures models can be used to construct confidence intervals and make comparisons between parameters. For unbalanced nested designs, the REML method of estimating the variance components provides a good method to construct confidence intervals using the Satterthwaite approximation. In this case, the approximate degrees of freedom are computed as $2(Z\text{-value})^2$. Confidence intervals constructed using type 1 SS estimates, REML estimates, ML estimates, and MINQUE0 estimates are displayed in Table 30.9 for the coffee price data. The method of moments can be effectively used to provide tests of hypotheses concerning the variance components. Table 30.8 uses the expected means squares to test $H_0: \sigma_{\text{State}}^2 = 0$ vs $H_a: \sigma_{\text{State}}^2 > 0$ and $H_0: \sigma_{\text{City}}^2 = 0$ vs $H_a: \sigma_{\text{City}}^2 > 0$ providing F -tests of 2.40 and 11.58, respectively.

When the nesting is in the treatment structure and involves fixed effect factors, the methods of analyzing fixed effects models are used. The major change is in the types of hypotheses that can be tested with the analysis of variance table as demonstrated with the data of Example 30.1.

30.3.1 Example 30.1: Continuation

There are two ways to write the model for the insecticide data

$$y_{ijk} = \mu_{j(i)} + \varepsilon_{ijk} \quad i = 1, 2, 3, 4, \quad j = 2, 3, \text{ or } 4, \quad k = 1, 2, 3 \quad (30.3)$$

TABLE 30.11

Analysis of Variance Table for Example 30.1 Based on Model (30.3)

Source	df	SS	MS	EMS	F-Value	Pr F
Product(Company)	10	24201.0	2420.1	Var(Residual) + Q[Product(Company)]	40.09	<0.0001
Residual	22	1328.00	60.36	Var(Residual)	—	—
$SSPRODUCT = \sum_{i=1}^4 \sum_{j=1}^{m_i} n_{j(i)} \bar{y}_{ij..}^2 - n_{..} \bar{y}_{...}^2$ and $SSERROR = \sum_{i=1}^4 \sum_{j=1}^{m_i} \sum_{k=1}^{n_{j(i)}} (y_{ijk} - \bar{y}_{ij..})^2 = 22\hat{\sigma}_e^2$						

or

$$y_{ijk} = \mu + \gamma_i + \rho_{j(i)} + \varepsilon_{ijk} \quad i = 1, 2, 3, 4, \quad j = 2, 3, \text{ or } 4, \quad k = 1, 2, 3 \quad (30.4)$$

Table 30.11 contains the analysis of variance table for model (30.3). The hypothesis being tested by the sum of squares due to products is

$$\mu_{1(1)} = \mu_{2(1)} = \mu_{3(1)} = \mu_{1(2)} = \mu_{2(2)} = \mu_{1(3)} = \mu_{2(3)} = \mu_{1(4)} = \mu_{2(4)} = \mu_{3(4)} = \mu_{4(4)}$$

Table 30.12 contains the analysis of variance table for model (30.4). The sum of squares due to products in Table 30.11 has been partitioned into the sum of squares due to company and the sum of squares due to products nested within companies. Since the two-way treatment structure is nested (products are nested within companies), there is no measure for interaction between the levels of product and the levels of company. The sums of squares due to companies tests the hypothesis $\bar{\mu}_{..(1)} = \bar{\mu}_{..(2)} = \bar{\mu}_{..(3)} = \bar{\mu}_{..(4)}$ where $\bar{\mu}_{..(i)} = (1/m_i) \sum_{j=1}^{m_i} \mu_{j(i)} = \mu + \gamma_i + \bar{\rho}_{..(i)}$.

The sum of squares due to products within companies tests the hypothesis that

$$\mu_{1(1)} = \mu_{2(1)} = \mu_{3(1)}, \quad \mu_{1(2)} = \mu_{2(2)} = \mu_{1(3)} = \mu_{2(3)}, \quad \text{and} \quad \mu_{1(4)} = \mu_{2(4)} = \mu_{3(4)} = \mu_{4(4)}$$

in terms of the means model parameters, or

$$\rho_{1(1)} = \rho_{2(1)} = \rho_{3(1)}, \quad \rho_{1(2)} = \rho_{2(2)} = \rho_{1(3)} = \rho_{2(3)}, \quad \text{and} \quad \rho_{1(4)} = \rho_{2(4)} = \rho_{3(4)} = \rho_{4(4)}$$

in terms of the effects model parameters.

TABLE 30.12

Analysis of Variance Table for Example 30.1 Based on Model (30.4)

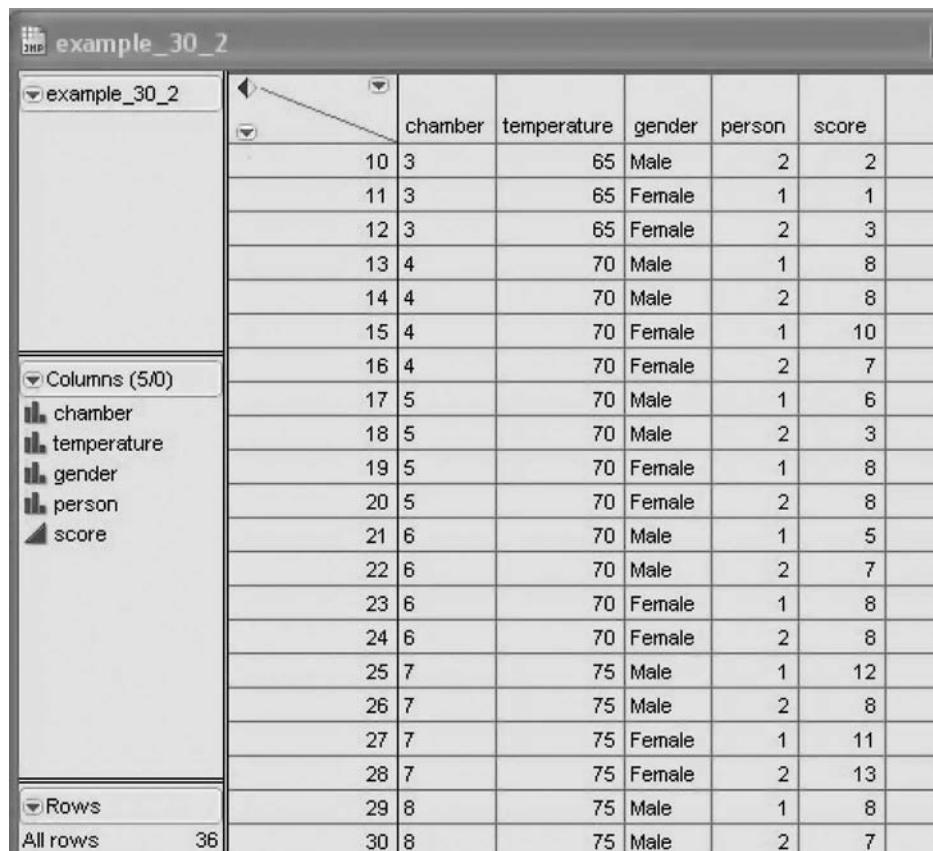
Source	df	SS	MS	EMS	F-Value	Pr F
Company	3	22649.6	7549.9	Var(Residual) + Q[Company, Product(Company)]	125.07	<0.0001
Product(Company)	7	1551.33	221.62	Var(Residual) + Q[Product(Company)]	3.67	0.0089
Residual	22	1328.00	60.36	Var(Residual)	—	—
$SSCOMPANY = \sum_{i=1}^4 n_{..(i)} \bar{y}_{..i}^2 - n_{..} \bar{y}_{...}^2$ $SSPRODUCT(Company) = \sum_{i=1}^4 \sum_{j=1}^{m_i} n_{j(i)} \bar{y}_{ij..}^2 - \sum_{i=1}^4 n_{..(i)} \bar{y}_{..i}^2$						
$SSERROR = \sum_{i=1}^4 \sum_{j=1}^{m_i} \sum_{k=1}^{n_{j(i)}} (y_{ijk} - \bar{y}_{ij..})^2 = 22\hat{\sigma}_e^2$						

Any of the appropriate multiple comparisons procedures (see Chapter 3) can be used where one may want to 1) compare the $\bar{\mu}_{(i)}$, 2) compare the $\mu_{j(i)}$, 3) compare the $\mu_{j(i)}$ for each i , or 4) compare any combination of the above. For unbalanced models, the type III or IV sums of squares are appropriate for testing the various hypotheses and the estimates of the population marginal means provide estimates of the $\bar{\mu}_{(i)}$ and of the $\mu_{j(i)}$.

For other nested models, the expected mean squares in the analysis of variance table can be used as guidelines for constructing proper F tests about the variance components and the mixed model “type III tests for fixed effects” can be used to test various hypotheses about the fixed effects. The methods in Chapter 22 and 23 are to be used in this context.

30.4 Analysis Using JMP

JMP® software is used to provide analyses for Examples 30.2 and 30.3. Figure 30.1 contains a partial listing of the data set where chamber, temperature, gender, and person are declared to be nominal variables. The fit model screen in Figure 30.2 includes *temperature*, *gender* and *temperature* \times *gender* as fixed effects and *chamber*(*temperature*) as a random effect. The *person* error term is the residual of the model. Figure 30.3 contains the REML estimates



	chamber	temperature	gender	person	score
10	3	65	Male	2	2
11	3	65	Female	1	1
12	3	65	Female	2	3
13	4	70	Male	1	8
14	4	70	Male	2	8
15	4	70	Female	1	10
16	4	70	Female	2	7
17	5	70	Male	1	6
18	5	70	Male	2	3
19	5	70	Female	1	8
20	5	70	Female	2	8
21	6	70	Male	1	5
22	6	70	Male	2	7
23	6	70	Female	1	8
24	6	70	Female	2	8
25	7	75	Male	1	12
26	7	75	Male	2	8
27	7	75	Female	1	11
28	7	75	Female	2	13
29	8	75	Male	1	8
30	8	75	Male	2	7
All rows	36				

FIGURE 30.1 JMP screen of comfort data for Example 30.2.

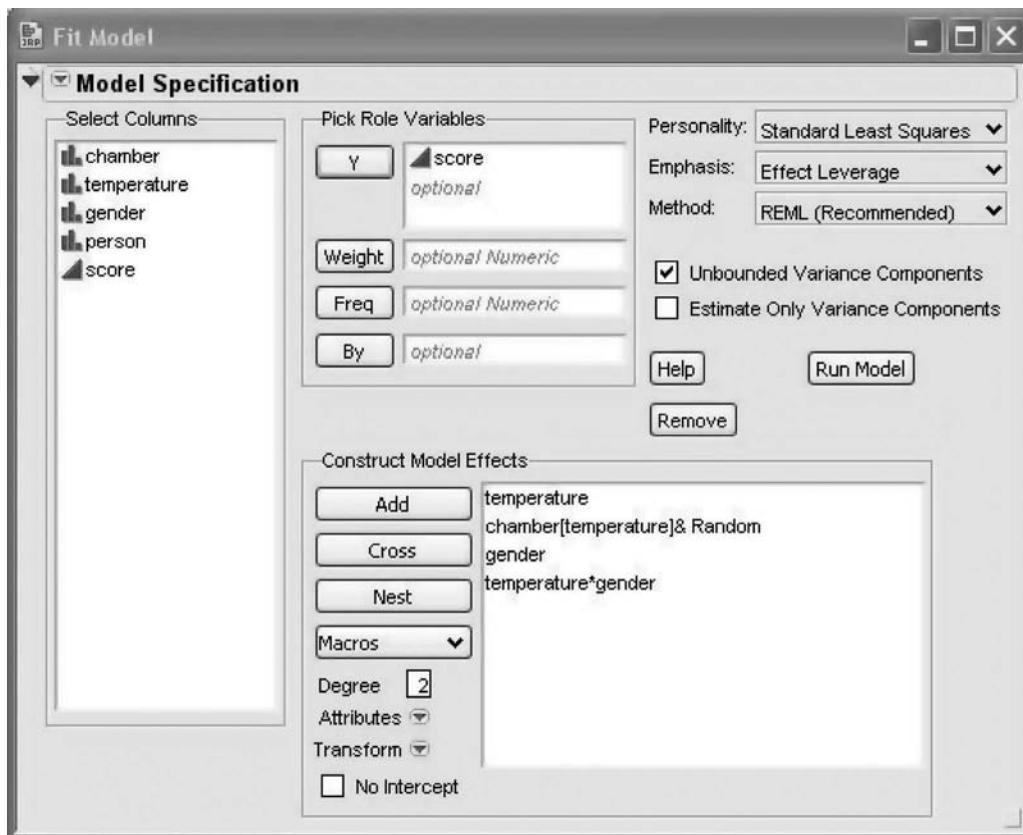


FIGURE 30.2 JMP fit model screen for comfort data of Example 30.2.

REML Variance Component Estimates							
Random Effect	Var Ratio	Var Component	Std Error	95% Lower	95% Upper	Pct of Total	
chamber[temperature]	1.4264706	2.3576389	1.604182	-0.786558	5.5018356	58.788	
Residual		1.6527778	0.4771158	1.0076869	3.198628	41.212	
Total		4.0104167				100.000	
$-2 \text{ LogLikelihood} = 130.93183289$							
Iterations							
Fixed Effect Tests							
Source	Nparm	DF	DFDen	F Ratio	Prob > F		
temperature	2	2	6	7.1454	0.0259*		
gender	1	1	24	2.0336	0.1667		
temperature*gender	2	2	24	4.7563	0.0182*		

FIGURE 30.3 JMP AOV table for random and fixed effects for Example 30.2.

of the variance components with a Wald confidence interval for the chamber error and a chi-square confidence interval for the residual or person variation and the type III tests for the fixed effects. The least squares means with Tukey multiple comparisons are displayed in Figure 30.4. The data screen in Figure 30.5 contains a partial listing of the coffee prices for Example 30.3. The fit model screen in Figure 30.6 has *state* and *city(state)* as

Level	Least Sq Mean
75,Female A	8.8333333
70,Female A B	8.1666667
75, Male A B	7.8333333
70, Male A B	6.1666667
65, Male A B	4.0000000
65, Female B	2.8333333

Levels not connected by same letter are significantly different.

FIGURE 30.4 Least squares means from JMP with Tukey multiple comparison for Example 30.2.

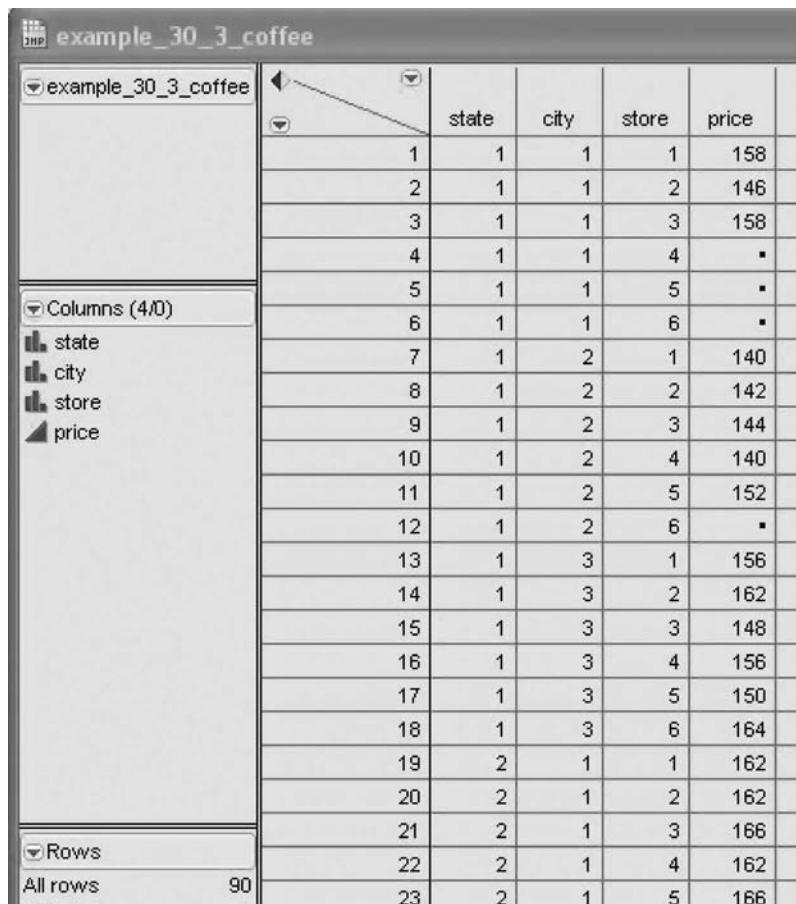


FIGURE 30.5 JMP data screen for coffee prices for Example 30.3.

random effects. The store to store variation is measured by the residual. Figure 30.7 contains the estimates of the variance components as well as the estimate of the overall mean. Wald confidence intervals are provided for the state and *city(state)* variance components and a chi-square confidence interval is provided for the residual or store to store variation. The estimate of the standard error of the overall mean is 3.359 while the estimate of the standard error of the overall mean from SAS®-Mixed using REML is 3.2935 (see Table 30.10).

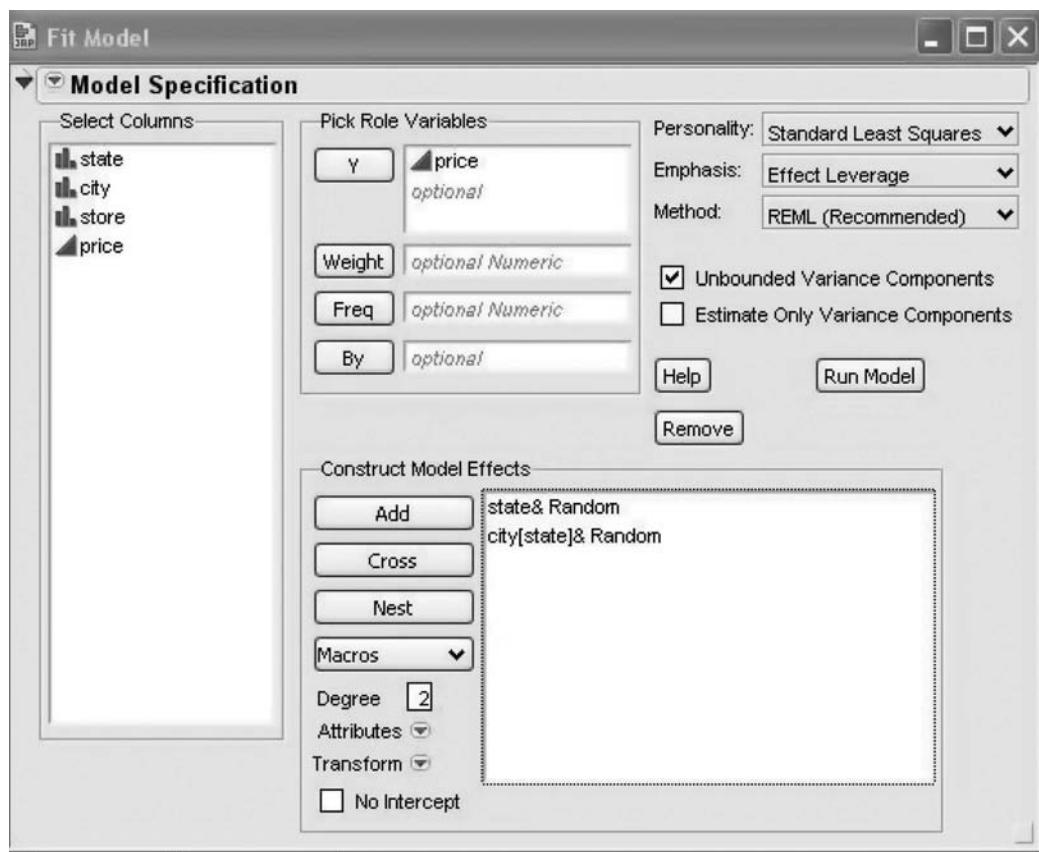


FIGURE 30.6 JMP fit model screen for Example 30.3.

Parameter Estimates						
Term	Estimate	Std Error	DFDen	t Ratio	Prob> t	
Intercept	149.59566	3.359306	2.934	44.53	<.0001*	

Random Effect Predictions						
REML Variance Component Estimates						
Random Effect	Var Ratio	Var Component	Std Error	95% Lower	95% Upper	Pct of Total
state	0.9253692	28.566337	47.346341	-64.23249	121.36516	22.561
city[state]	2.1762524	67.181359	34.706183	-0.84276	135.20548	53.058
Residual		30.870206	5.582685	22.299592	45.567205	24.381
Total		126.6179				100.000

FIGURE 30.7 REML estimates of the variance components for Example 30.3 from JMP.

30.5 Concluding Remarks

The concept of nesting was defined in this chapter and examples were used to demonstrate that nesting can occur in the treatment structure, the design structure or both. Nested models can involve fixed effects, random effects or mixed effects models, and analyses of the examples demonstrate some of the issues involved.

30.6 Exercises

- 30.1 *Types of salt:* A food scientist wanted to determine if the type and source of salt when mixed into instant mashed potatoes affect flavor profiles. Three types of salt, kosher, sea and table salt, were selected with four sources of sea salt (four different parts of the world, denoted by brands 1, 2, 3, and 4), three brands of kosher salt (brands 5, 6, and 7), and two brands of table salt (brands 7 and 8). Thus the study involves nine brands of salt, but the brands come from three different types. On a given day, a taste panel session was held where six of the salts were evaluated for salt intensity by five trained taste panelists. The median of the five panel members' evaluations (0–100) was used as the response for each salt tasted. The following table contains the session, brand (b_i) and the corresponding response (r_i). The order the panel tasted the products in was randomized within each session. Construct a model to describe this data, compare the salt type means and compare the brand means. Estimate the variance components and provide confidence intervals. Carry out a multiple comparison method to compare the type means (sea, kosher, and table) and the brand means within a type.

Data for Exercise 30.1

Session	b1	r1	b2	r2	b3	r3	b4	r4	b5	r5	b6	r6
1	1	54	2	57	4	68	5	49	7	49	8	46
2	2	67	3	73	5	57	6	62	8	55	9	66
3	1	62	3	66	4	64	6	60	7	53	9	58
4	1	72	2	63	5	57	6	57	7	60	9	60
5	1	63	3	60	4	79	5	58	8	61	9	66
6	2	72	3	67	4	66	6	49	7	52	8	50
7	1	61	3	71	5	59	6	64	7	66	8	67
8	1	60	2	64	4	69	6	57	8	41	9	64
9	2	59	3	66	4	59	5	49	7	53	9	47
10	4	73	5	64	6	65	7	59	8	62	9	63
11	1	56	2	52	3	55	4	62	5	54	6	56
12	1	66	2	72	3	72	7	71	8	59	9	60

- 30.2 *Coffee prices and types of stores:* The coffee example of Exercise 30.3 was modified: within each city of a state, three types of stores were included in the sampling. A random sample of each of three types of stores (large chain store, locally owned store, and convenience store) was selected from the stores of that type within

each city. Construct a model to describe the following data, and obtain estimates of the store type means as well as the variance components. Construct confidence intervals about the variance components and carry out a multiple comparison among the store type means.

Data for Exercise 30.2

State	City	Type	Store 1	Store 2	Store 3	Store 4	Store 5	Store 6
1	1	Chain	142	128	144	144	126	—
1	1	Convenience	152	154	134	144	140	150
1	1	Local	154	128	150	—	—	—
1	2	Chain	136	100	126	—	—	—
1	2	Convenience	124	134	132	136	124	116
1	2	Local	140	134	140	154	—	—
1	3	Chain	146	152	138	130	154	—
1	3	Convenience	148	160	150	152	—	—
1	3	Local	142	154	138	134	—	—
2	1	Chain	162	150	146	146	—	—
2	1	Convenience	162	168	148	158	—	—
2	1	Local	146	146	156	158	164	152
2	2	Chain	160	158	156	154	160	—
2	2	Convenience	162	182	170	166	—	—
2	2	Local	154	158	180	172	—	—
2	3	Chain	164	174	180	144	166	136
2	3	Convenience	164	162	166	—	—	—
2	3	Local	162	190	162	156	172	168
2	4	Chain	126	140	122	126	142	124
2	4	Convenience	148	144	148	—	—	—
2	4	Local	126	128	124	116	—	—
2	5	Chain	148	134	140	146	134	124
2	5	Convenience	132	130	134	170	—	—
2	5	Local	136	130	118	172	148	—
3	1	Chain	110	120	118	134	116	—
3	1	Convenience	124	138	114	134	126	—
3	1	Local	118	120	118	124	116	—
3	2	Chain	130	154	144	156	152	—
3	2	Convenience	166	132	166	160	—	—
3	2	Local	148	156	144	150	—	—
3	3	Chain	120	130	136	126	—	—
3	3	Convenience	152	134	136	136	144	—
3	3	Local	132	148	126	134	134	—
4	1	Chain	146	170	154	150	—	—
4	1	Convenience	154	148	148	142	—	—
4	1	Local	146	146	140	158	132	—
4	2	Chain	126	154	130	142	126	120
4	2	Convenience	130	138	144	138	140	—
4	2	Local	138	126	132	—	—	—

Appendix

TABLE A.1

 Percentage Points of the Maximum *F*-Ratio

		Upper 5% Points										
<i>v</i>	<i>k</i>	2	3	4	5	6	7	8	9	10	11	12
2		39.0	87.5	142	202	266	333	403	475	550	626	704
3		15.4	27.8	39.2	50.7	62.0	72.9	83.5	93.9	104	114	124
4		9.60	15.5	20.6	25.2	29.5	33.6	37.5	41.1	44.6	48.0	51.4
5		7.15	10.8	13.7	16.3	18.7	20.8	22.9	24.7	26.5	28.2	29.9
6		5.82	8.38	10.4	12.1	13.7	15.0	16.3	17.5	18.6	19.7	20.7
7		4.99	6.94	8.44	9.70	10.8	11.8	12.7	13.5	14.3	15.1	15.8
8		4.43	6.00	7.18	8.12	9.03	9.78	10.5	11.1	11.7	12.2	12.7
9		4.03	5.34	6.31	7.11	7.80	8.41	8.95	9.45	9.91	10.3	10.7
10		3.72	4.85	5.67	6.34	6.92	7.42	7.87	8.28	8.66	9.01	9.34
12		3.28	4.16	4.79	5.30	5.72	6.09	6.42	6.72	7.00	7.25	7.48
15		2.86	3.54	4.01	4.37	4.68	4.95	5.19	5.40	5.59	5.77	5.93
20		2.46	2.95	3.29	3.54	3.76	3.94	4.10	4.24	4.37	4.49	4.59
30		2.07	2.40	2.61	2.78	2.91	3.02	3.12	3.21	3.29	3.36	3.39
60		1.67	1.85	1.96	2.04	2.11	2.17	2.22	2.26	2.30	2.33	2.36
∞		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		Upper 1% Points										
<i>v</i>	<i>k</i>	2	3	4	5	6	7	8	9	10	11	12
2		199	448	729	1036	1362	1705	2063	2432	2813	3204	3605
3		47.5	85	120	151	184	21(6)	24(9)	28(1)	31(0)	33(7)	36(1)
4		23.2	37	49	59	69	79	89	97	106	113	120
5		14.9	22	28	33	38	42	46	50	54	57	60
6		11.1	15.5	19.1	22	25	27	30	32	34	36	37
7		8.89	12.1	14.5	16.5	18.4	20	22	23	24	26	27

Continued

TABLE A.1 (continued)

		Upper 1% Points										
<i>v</i>	<i>k</i>	2	3	4	5	6	7	8	9	10	11	12
8		7.50	9.9	11.7	13.2	14.5	15.8	16.9	17.9	18.9	19.8	21
9		6.54	8.5	9.9	11.1	12.1	13.1	13.9	14.7	15.3	16.0	16.6
10		5.85	7.4	8.6	9.6	10.4	11.1	11.8	12.4	12.9	13.4	13.9
12		4.91	6.1	6.9	7.6	8.2	8.7	9.1	9.5	9.9	10.2	10.6
15		4.07	4.9	5.5	6.0	6.4	6.7	7.1	7.3	7.5	7.8	8.0
20		3.32	3.8	4.3	4.6	4.9	5.1	5.3	5.5	5.6	5.8	5.9
30		2.63	3.0	3.3	3.4	3.6	3.7	3.8	3.9	4.0	4.1	4.2
60		1.96	2.2	2.3	2.4	2.4	2.5	2.5	2.6	2.6	2.7	2.7
∞		1.00	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Source: Beyer, W. H., ed., *Handbook of Tables for Probability and Statistics*, Second Edition, The Chemical Rubber Co., 1968. With permission.

TABLE A.2

Table of Bonferroni Critical Points

		Values of <i>c</i> for $1 - \alpha = .95$ $\int_{-\infty}^c f^{(v)}(t) dt = 1 - \frac{.05}{2m}$											
<i>m</i>	<i>v</i>	5	7	10	12	15	20	24	30	40	60	120	∞
2		3.17	2.84	2.64	2.56	2.49	2.42	2.39	2.36	2.33	2.30	2.27	2.24
3		3.54	3.13	2.87	2.78	2.69	2.61	2.58	2.54	2.50	2.47	2.43	2.39
4		3.81	3.34	3.04	2.94	2.84	2.75	2.70	2.66	2.62	2.58	2.54	2.50
5		4.04	3.50	3.17	3.06	2.95	2.85	2.80	2.75	2.71	2.66	2.62	2.58
6		4.22	3.64	3.28	3.15	3.04	2.93	2.88	2.83	2.78	2.73	2.68	2.64
7		4.38	3.76	3.37	3.24	3.11	3.00	2.94	2.89	2.84	2.79	2.74	2.69
8		4.53	3.86	3.45	3.31	3.18	3.06	3.00	2.94	2.89	2.84	2.79	2.74
9		4.66	3.95	3.52	3.37	3.24	3.11	3.05	2.99	2.93	2.88	2.83	2.77
10		4.78	4.03	3.58	3.43	3.29	3.16	3.09	3.03	2.97	2.92	2.86	2.81
15		5.25	4.36	3.83	3.65	3.48	3.33	3.26	3.19	3.12	3.06	2.99	2.94
20		5.60	4.59	4.01	3.80	3.62	3.46	3.38	3.30	3.23	3.16	3.09	3.02
25		5.89	4.78	4.15	3.93	3.74	3.55	3.47	3.39	3.31	3.24	3.16	3.09
30		6.15	4.95	4.27	4.04	3.82	3.63	3.54	3.46	3.38	3.30	3.22	3.15
35		6.36	5.09	4.37	4.13	3.90	3.70	3.61	3.52	3.43	3.34	3.27	3.19
40		6.56	5.21	4.45	4.20	3.97	3.76	3.66	3.57	3.48	3.39	3.31	3.23
45		6.70	5.31	4.53	4.26	4.02	3.80	3.70	3.61	3.51	3.42	3.34	3.26
50		6.86	5.40	4.59	4.32	4.07	3.85	3.74	3.65	3.55	3.46	3.37	3.29
100		8.00	6.08	5.06	4.73	4.42	4.15	4.04	3.90	3.79	3.69	3.58	3.48
250		9.68	7.06	5.70	5.27	4.90	4.56	4.4*	4.2*	4.1*	3.97	3.83	3.72

Continued

TABLE A.2 (continued)

$m \setminus v$	Values of c for $1 - \alpha = .99$ $\int_{-\infty}^{\infty} f^{(v)}(t) dt = -\frac{.01}{2m}$											
	5	7	10	12	15	20	24	30	40	60	120	∞
2	4.78	4.03	3.58	3.43	3.29	3.16	3.09	3.03	2.97	2.92	2.86	2.81
3	5.25	4.36	3.83	3.65	3.48	3.33	3.26	3.19	3.12	3.06	2.99	2.94
4	5.60	4.59	4.01	3.80	3.62	3.46	3.38	3.30	3.23	3.16	3.09	3.02
5	5.89	4.78	4.15	3.93	3.74	3.55	3.47	3.39	3.31	3.24	3.16	3.09
6	6.15	4.95	4.27	4.04	3.82	3.63	3.54	3.46	3.38	3.30	3.22	3.15
7	6.36	5.09	4.37	4.13	3.90	3.70	3.61	3.52	3.43	3.34	3.27	3.19
8	6.56	5.21	4.45	4.20	3.97	3.76	3.66	3.57	3.48	3.39	3.31	3.23
9	6.70	5.31	4.53	4.26	4.02	3.80	3.70	3.61	3.51	3.42	3.34	3.26
10	6.86	5.40	4.59	4.32	4.07	3.85	3.74	3.65	3.55	3.46	3.37	3.29
15	7.51	5.79	4.86	4.56	4.29	4.03	3.91	3.80	3.70	3.59	3.50	3.40
20	8.00	6.08	5.06	4.73	4.42	4.15	4.04	3.90	3.79	3.69	3.58	3.48
25	8.37	6.30	5.20	4.86	4.53	4.25	4.1*	3.98	3.88	3.76	3.64	3.54
30	8.68	6.49	5.33	4.95	4.61	4.33	4.2*	4.13	3.93	3.81	3.69	3.59
35	8.95	6.67	5.44	5.04	4.71	4.39	4.3*	4.26	3.97	3.84	3.73	3.63
40	9.19	6.83	5.52	5.12	4.78	4.46	4.3*	4.1*	4.01	3.89	3.77	3.66
45	9.41	6.93	5.60	5.20	4.84	4.52	4.3*	4.2*	4.1*	3.93	3.80	3.69
50	9.68	7.06	5.70	5.27	4.90	4.56	4.4*	4.2*	4.1*	3.97	3.83	3.72
100	11.04	7.80	6.20	5.70	5.20	4.80	4.7*	4.4*	4.5*	4.00	3.89	
250	13.26	8.83	6.9*	6.3*	5.8*	5.2*	5.0*	4.9*	4.8*			4.11

*Obtained by graphical interpolation.

Source: Dunn, O. J., *Journal of the American Statistical Association*, 56: 52–64, 1961. With permission.

TABLE A.3

Probability Points for Multivariate t -Distribution

Entries are $t_{\alpha/2;q,m}$ where $P[\max |T_i| \leq t_{\alpha/2;q,m}] = 1 - \alpha$ and $t = [T_1, T_2, \dots, T_q]$ is distributed $S(t; q, m; I)$

$m \setminus q$	1	2	3	4	5	6	8	10	12	15	20
	1 - $\alpha = .90$										
3	2.353	2.989	3.369	3.637	3.844	4.011	4.272	4.471	4.631	4.823	5.066
4	2.132	2.662	2.976	3.197	3.368	3.506	3.722	3.887	4.020	4.180	4.383
5	2.015	2.491	2.769	2.965	3.116	3.239	3.430	3.576	3.694	3.837	4.018
6	1.943	2.385	2.642	2.822	2.961	3.074	3.249	3.384	3.493	3.624	3.790
7	1.895	2.314	2.556	2.725	2.856	2.962	3.127	3.253	3.355	3.478	3.635
8	1.860	2.262	2.494	2.656	2.780	2.881	3.038	3.158	3.255	3.373	3.522
9	1.833	2.224	2.447	2.603	2.723	2.819	2.970	3.086	3.179	3.292	3.436
10	1.813	2.193	2.410	2.562	2.678	2.771	2.918	3.029	3.120	3.229	3.360
11	1.796	2.169	2.381	2.529	2.642	2.733	2.875	2.984	3.072	3.178	3.313
12	1.782	2.149	2.357	2.501	2.612	2.701	2.840	2.946	3.032	3.136	3.268
15	1.753	2.107	2.305	2.443	2.548	2.633	2.765	2.865	2.947	3.045	3.170

Continued

TABLE A.3 (continued)

Entries are $t_{\alpha/2;q,m}$ where $P[\max |T_i| \leq t_{\alpha/2;q,m}] = 1 - \alpha$ and $\mathbf{t} = [T_1, T_2, \dots, T_q]$ is distributed $S(t;q,m;I)$

$m \setminus q$	1	2	3	4	5	6	8	10	12	15	20
20	1.725	2.065	2.255	2.386	2.486	2.567	2.691	2.786	2.863	2.956	3.073
25	1.708	2.041	2.226	2.353	2.450	2.528	2.648	2.740	2.814	2.903	3.016
30	1.697	2.025	2.207	2.331	2.426	2.502	2.620	2.709	2.781	2.868	2.978
40	1.684	2.006	2.183	2.305	2.397	2.470	2.585	2.671	2.741	2.825	2.931
60	1.671	1.986	2.160	2.278	2.368	2.439	2.550	2.634	2.701	2.782	2.884
$1 - \alpha = .95$											
3	3.183	3.960	4.430	4.764	5.023	5.233	5.562	5.812	6.015	6.259	6.567
4	2.777	3.382	3.745	4.003	4.203	4.366	4.621	4.817	4.975	5.166	5.409
5	2.571	3.091	3.399	3.619	3.789	3.928	4.145	4.312	4.447	4.611	4.819
6	2.447	2.916	3.193	3.389	3.541	3.664	3.858	4.008	4.129	4.275	4.462
7	2.365	2.800	3.056	3.236	3.376	3.489	3.668	3.805	3.916	4.051	4.223
8	2.306	2.718	2.958	3.128	3.258	3.365	3.532	3.660	3.764	3.891	4.052
9	2.262	2.657	2.885	3.046	3.171	3.272	3.430	3.552	3.651	3.770	3.923
10	2.228	2.609	2.829	2.984	3.103	3.199	3.351	3.468	3.562	3.677	3.823
11	2.201	2.571	2.784	2.933	3.048	3.142	3.288	3.400	3.491	3.602	3.743
12	2.179	2.540	2.747	2.892	3.004	3.095	3.236	3.345	3.433	3.541	3.677
15	2.132	2.474	2.669	2.805	2.910	2.994	3.126	3.227	3.309	3.409	3.536
20	12.086	2.411	2.594	2.722	2.819	2.898	3.020	3.114	3.190	3.282	3.399
25	2.060	2.374	2.551	2.673	2.766	2.842	2.959	3.048	3.121	3.208	3.320
30	2.042	2.350	2.522	2.641	2.732	2.805	2.918	3.005	3.075	3.160	3.267
40	2.021	2.321	2.488	2.603	2.690	2.760	2.869	2.952	3.019	3.100	3.203
60	2.000	2.292	2.454	2.564	2.649	2.716	2.821	2.900	2.964	3.041	3.139
$1 - \alpha = .99$											
3	5.841	7.127	7.914	8.479	8.919	9.277	9.838	10.269	10.616	11.034	11.559
4	4.604	5.462	5.985	6.362	6.656	6.897	7.274	7.565	7.801	8.087	8.451
5	4.032	4.700	5.106	5.398	5.625	5.812	6.106	6.333	6.519	6.744	7.050
6	3.707	4.271	4.611	4.855	5.046	5.202	5.449	5.640	5.796	5.985	6.250
7	3.500	3.998	4.296	4.510	4.677	4.814	5.031	5.198	5.335	5.502	5.716
8	3.355	3.809	4.080	4.273	4.424	4.547	4.742	4.894	5.017	5.168	5.361
9	3.250	3.672	3.922	4.100	4.239	4.353	4.532	4.672	4.785	4.924	5.103
10	3.169	3.567	3.801	3.969	4.098	4.205	4.373	4.503	4.609	4.739	4.905
11	3.106	3.485	3.707	3.865	3.988	4.087	4.247	4.370	4.470	4.593	4.750
12	3.055	3.418	3.631	3.782	3.899	3.995	4.146	4.263	4.359	4.475	4.625
15	2.947	3.279	3.472	3.608	3.714	3.800	3.935	4.040	4.125	4.229	4.363
20	2.845	3.149	3.323	3.446	3.541	3.617	3.738	3.831	3.907	3.999	4.117
25	2.788	3.075	3.239	3.354	3.442	3.514	3.626	3.713	3.783	3.869	3.978
30	2.750	3.027	3.185	3.295	3.379	3.448	3.555	3.637	3.704	3.785	3.889
40	2.705	2.969	3.119	3.223	3.303	3.367	3.468	3.545	3.607	3.683	3.780
60	2.660	2.913	3.055	3.154	3.229	3.290	3.384	3.456	3.515	3.586	3.676

Source: Hahn, G. J. and Hendrickson, R. W., *Biometrika*, Vol. 58. With permission.

TABLE A.4
Percentage Points of the Studentized Range

		The Entries are $q_{\text{Student},n}$ where $P[Q < q_{\text{Student},n}] = .95$																		
<i>n</i>	<i>m</i>	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	17.97	26.98	32.82	37.08	40.41	43.12	45.40	47.36	49.07	50.59	51.96	53.20	54.33	55.36	56.32	57.22	58.04	58.83	59.56	
2	6.08	8.33	9.80	10.88	11.74	12.44	13.03	13.54	13.99	14.39	14.75	15.08	15.38	15.65	15.91	16.14	16.37	16.57	16.77	
3	4.50	5.91	6.82	7.50	8.04	8.48	8.85	9.18	9.46	9.72	9.95	10.15	10.35	10.52	10.69	10.84	10.98	11.11	11.24	
4	3.93	5.04	5.76	6.29	6.71	7.05	7.35	7.60	7.83	8.03	8.21	8.37	8.52	8.66	8.79	8.91	9.03	9.13	9.23	
5	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17	7.32	7.47	7.60	7.72	7.83	7.93	8.03	8.12	8.21	
6	3.46	4.34	4.90	5.30	5.63	5.90	6.12	6.32	6.49	6.65	6.79	6.92	7.03	7.14	7.24	7.34	7.43	7.51	7.59	
7	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16	6.30	6.43	6.55	6.66	6.76	6.85	6.94	7.02	7.10	7.17	
8	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05	6.18	6.29	6.39	6.48	6.57	6.65	6.73	6.80	6.87	
9	3.20	3.95	4.41	4.76	5.02	5.24	5.43	5.59	5.74	5.87	5.98	6.09	6.19	6.28	6.36	6.44	6.51	6.58	6.64	
10	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60	5.72	5.83	5.93	6.03	6.11	6.19	6.27	6.34	6.40	6.47	
11	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49	5.61	5.71	5.81	5.90	5.98	6.06	6.13	6.20	6.27	6.33	
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.39	5.51	5.61	5.71	5.80	5.88	5.95	6.02	6.09	6.15	6.21	
13	3.06	3.73	4.15	4.45	4.69	4.88	5.05	5.19	5.32	5.43	5.53	5.63	5.71	5.79	5.86	5.93	5.99	6.05	6.11	
14	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25	5.36	5.46	5.55	5.64	5.71	5.79	5.85	5.91	5.97	6.03	
15	3.01	3.67	4.08	4.37	4.59	4.78	4.94	5.08	5.20	5.31	5.40	5.49	5.57	5.65	5.72	5.78	5.85	5.90	5.96	
16	3.00	3.65	4.05	4.33	4.56	4.74	4.90	5.03	5.15	5.26	5.35	5.44	5.52	5.59	5.66	5.73	5.79	5.84	5.90	
17	2.98	3.63	4.02	4.30	4.52	4.70	4.86	4.99	5.11	5.21	5.31	5.39	5.47	5.54	5.61	5.67	5.73	5.79	5.84	
18	2.97	3.61	4.00	4.28	4.49	4.67	4.82	4.96	5.07	5.17	5.27	5.35	5.43	5.50	5.57	5.63	5.69	5.74	5.79	
19	2.96	3.59	3.98	4.25	4.47	4.65	4.79	4.92	5.04	5.14	5.23	5.31	5.39	5.46	5.53	5.59	5.65	5.70	5.75	
20	2.95	3.58	3.96	4.23	4.45	4.62	4.77	4.90	5.01	5.11	5.20	5.28	5.36	5.43	5.49	5.55	5.61	5.66	5.71	
24	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81	4.92	5.01	5.10	5.18	5.25	5.32	5.38	5.44	5.49	5.55	5.59	
30	2.89	3.49	3.85	4.10	4.30	4.46	4.60	4.72	4.82	4.92	5.00	5.08	5.15	5.21	5.27	5.33	5.38	5.43	5.47	
40	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.73	4.82	4.90	4.98	5.04	5.11	5.16	5.22	5.27	5.31	5.36	
60	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65	4.73	4.81	4.88	4.94	5.00	5.06	5.11	5.15	5.20	5.24	
120	2.80	3.36	3.68	3.92	4.10	4.24	4.36	4.47	4.56	4.64	4.71	4.78	4.84	4.90	4.95	5.00	5.04	5.09	5.13	
∞	2.77	3.31	3.63	3.89	4.03	4.17	4.29	4.39	4.47	4.55	4.62	4.68	4.74	4.80	4.85	4.89	4.93	4.97	5.01	

Continued

TABLE A.4 (continued)

		The Entries are $\pi_{1,05m,n}$ where $P(Q < \pi_{1,05m,n}) = .99$																		
$n \backslash m$	m	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	90.03	135.0	164.3	185.6	202.2	215.8	227.	237.0	245.6	253.2	260.0	266.2	271.8	277.0	281.8	286.3	290.4	294.3	298.0	
2	14.04	19.02	22.29	24.72	26.63	28.20	29.53	30.68	31.69	32.59	33.40	34.13	34.81	35.43	36.00	36.53	37.03	37.50	37.95	
3	8.26	10.62	12.17	13.33	14.24	15.00	15.64	16.20	16.69	17.13	17.53	17.89	18.22	18.52	18.81	19.07	19.32	19.55	19.77	
4	6.51	8.12	9.17	9.96	10.58	11.10	11.55	11.93	12.27	12.57	12.84	13.09	13.32	13.53	13.73	13.91	14.08	14.24	14.40	
5	5.70	6.98	7.80	8.42	8.91	9.32	9.67	9.97	10.24	10.48	10.70	10.89	11.08	11.24	11.40	11.55	11.68	11.81	11.93	
6	5.24	6.33	7.03	7.56	7.97	8.32	8.61	8.87	9.10	9.30	9.48	9.65	9.81	9.95	10.08	10.21	10.32	10.43	10.54	
7	4.95	5.92	6.54	7.01	7.37	7.68	7.94	8.17	8.36	8.55	8.71	8.86	9.00	9.12	9.24	9.35	9.46	9.55	9.65	
8	4.75	5.64	6.20	6.62	6.96	7.24	7.47	7.68	7.86	8.03	8.18	8.31	8.44	8.55	8.66	8.76	8.85	8.94	9.03	
9	4.60	5.43	5.96	6.35	6.66	6.91	7.13	7.33	7.49	7.65	7.78	7.91	8.03	8.13	8.23	8.33	8.41	8.49	8.57	
10	4.48	5.27	5.77	6.14	6.43	6.67	6.87	7.05	7.21	7.36	7.49	7.60	7.71	7.81	7.91	7.99	8.08	8.15	8.23	
11	4.39	5.15	5.62	5.97	6.25	6.48	6.67	6.84	6.99	7.13	7.25	7.36	7.46	7.56	7.65	7.73	7.81	7.88	7.95	
12	4.32	5.05	5.50	5.84	6.10	6.32	6.51	6.67	6.81	6.94	7.06	7.17	7.26	7.36	7.44	7.52	7.59	7.66	7.73	
13	4.26	4.96	5.40	5.73	5.98	6.19	6.37	6.53	6.67	6.79	6.90	7.01	7.10	7.19	7.27	7.35	7.42	7.48	7.55	
14	4.21	4.89	5.32	5.63	5.88	6.08	6.26	6.41	6.54	6.66	6.77	6.87	6.96	7.05	7.13	7.20	7.27	7.33	7.39	
15	4.17	4.84	5.25	5.56	5.80	5.99	6.16	6.31	6.44	6.55	6.66	6.76	6.84	6.93	7.00	7.07	7.14	7.20	7.26	
16	4.13	4.79	5.19	5.49	5.72	5.92	6.08	6.22	6.35	6.46	6.56	6.66	6.74	6.82	6.90	6.97	7.03	7.09	7.15	
17	4.10	4.74	5.14	5.43	5.66	5.85	6.01	6.15	6.27	6.38	6.48	6.57	6.66	6.73	6.81	6.87	6.94	7.00	7.05	
18	4.07	4.70	5.09	5.38	5.60	5.79	5.94	6.08	6.20	6.31	6.41	6.50	6.58	6.65	6.73	6.79	6.85	6.91	6.97	
19	4.05	4.67	5.05	5.33	5.55	5.73	5.89	6.02	6.14	6.25	6.34	6.43	6.51	6.58	6.65	6.72	6.78	6.84	6.89	
20	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09	6.19	6.28	6.37	6.45	6.52	6.59	6.65	6.71	6.77	6.82	
24	3.96	4.55	4.91	5.17	5.37	5.54	5.69	5.81	5.92	6.02	6.11	6.19	6.26	6.33	6.39	6.45	6.51	6.56	6.61	
30	3.89	4.45	4.80	5.05	5.24	5.40	5.54	5.65	5.76	5.85	5.93	6.01	6.08	6.14	6.20	6.26	6.31	6.36	6.41	
40	3.82	4.37	4.70	4.93	5.11	5.26	5.39	5.50	5.60	5.69	5.76	5.83	5.90	5.96	6.02	6.07	6.12	6.16	6.21	
60	3.76	4.28	4.59	4.82	4.99	5.13	5.25	5.36	5.45	5.53	5.60	5.67	5.73	5.78	5.84	5.89	5.93	5.97	6.01	
120	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30	5.37	5.44	5.50	5.56	5.61	5.66	5.71	5.75	5.79	5.83	
∞	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16	5.23	5.29	5.35	5.40	5.45	5.49	5.54	5.57	5.61	5.65	

Source: Pearson, E. S. and Hartley, H. O., *Biometrika Tables for Statisticians*, Vol. 1, Third Edition, Table 29, published by the Biometrika Trustees, Cambridge University Press, London, 1966. With permission.

TABLE A.5

Percentage Points of the Duncan New Multiple Range Test

Error df	α	r = Number of Ordered Steps between Means													
		2	3	4	5	6	7	8	9	10	12	14	16	18	20
1	.05	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0
	.01	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0
2	.05	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09
	.01	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0
3	.05	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50
	.01	8.26	8.5	8.6	8.7	8.8	8.9	8.9	9.0	9.0	9.0	9.1	9.2	9.3	9.3
4	.05	3.93	4.01	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02
	.01	6.51	6.8	6.9	7.0	7.1	7.1	7.2	7.2	7.3	7.3	7.4	7.4	7.5	7.5
5	.05	3.64	3.74	3.79	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83
	.01	5.70	5.96	6.11	6.18	6.26	6.33	6.40	6.44	6.5	6.6	6.6	6.7	6.7	6.8
6	.05	3.46	3.58	3.64	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68
	.01	5.24	5.51	5.65	5.73	5.81	5.88	5.95	6.00	6.0	6.1	6.2	6.2	6.3	6.3
7	.05	3.35	3.47	3.54	3.58	3.60	3.61	3.61	3.61	3.61	3.61	3.61	3.61	3.61	3.61
	.01	4.95	5.22	5.37	5.45	5.53	5.61	5.69	5.73	5.8	5.8	5.9	5.9	6.0	6.0
8	.05	3.26	3.39	3.47	3.52	3.55	3.56	3.56	3.56	3.56	3.56	3.56	3.56	3.56	3.56
	.01	4.74	5.00	5.14	5.23	5.32	5.40	5.47	5.51	5.5	5.6	5.7	5.7	5.8	5.8
9	.05	3.20	3.34	3.41	3.47	3.50	3.52	3.52	3.52	3.52	3.52	3.52	3.52	3.52	3.52
	.01	4.60	4.86	4.99	5.08	5.17	5.25	5.32	5.36	5.4	5.5	5.5	5.6	5.7	5.7
10	.05	3.15	3.30	3.37	3.43	3.46	3.47	3.47	3.47	3.47	3.47	3.47	3.47	3.47	3.48
	.01	4.48	4.73	4.88	4.96	5.06	5.13	5.20	5.24	5.28	5.36	5.42	5.48	5.54	5.55
11	.05	3.11	3.27	3.35	3.39	3.43	3.44	3.45	3.46	3.46	3.46	3.46	3.46	3.47	3.48
	.01	4.39	4.63	4.77	4.86	4.94	5.01	5.06	5.12	5.15	5.24	5.28	5.34	5.38	5.39
12	.05	3.08	3.23	3.33	3.36	3.40	3.42	3.44	3.44	3.46	3.46	3.46	3.46	3.47	3.48
	.01	4.32	4.55	4.68	4.76	4.84	4.92	4.96	5.02	5.07	5.13	5.17	5.22	5.23	5.26
13	.05	3.06	3.21	3.30	3.35	3.38	3.41	3.42	3.44	3.45	3.45	3.46	3.46	3.47	3.47
	.01	4.26	4.48	4.62	4.69	4.74	4.84	4.88	4.94	4.98	5.04	5.08	5.13	5.14	5.15
14	.05	3.03	3.18	3.27	3.33	3.37	3.39	3.41	3.42	3.44	3.45	3.46	3.46	3.47	3.47
	.01	4.21	4.42	4.55	4.63	4.70	4.78	4.83	4.87	4.91	4.96	5.00	5.04	5.06	5.07
15	.05	3.01	3.16	3.25	3.31	3.36	3.38	3.40	3.42	3.43	3.44	3.45	3.46	3.47	3.47
	.01	4.17	4.37	4.50	4.58	4.64	4.72	4.77	4.81	4.84	4.90	4.94	4.97	4.99	5.00
16	.05	3.00	3.15	3.23	3.30	3.34	3.37	3.39	3.41	3.43	3.44	3.45	3.46	3.47	3.47
	.01	4.13	4.34	4.45	4.54	4.60	4.67	4.72	4.76	4.79	4.84	4.88	4.91	4.93	4.94
17	.05	2.98	3.13	3.22	3.28	3.33	3.36	3.38	3.40	3.42	3.43	3.44	3.45	3.46	3.47
	.01	4.10	4.30	4.41	4.50	4.56	4.63	4.68	4.72	4.75	4.80	4.83	4.86	4.88	4.89
18	.05	2.97	3.12	3.21	3.27	3.32	3.35	3.37	3.39	3.41	3.43	3.45	3.46	3.47	3.47
	.01	4.07	4.27	4.38	4.46	4.53	4.59	4.64	4.68	4.71	4.76	4.79	4.82	4.84	4.85
19	.05	2.96	3.11	3.19	3.26	3.31	3.35	3.37	3.39	3.41	3.43	3.44	3.46	3.47	3.47
	.01	4.05	4.24	4.35	4.43	4.50	4.56	4.61	4.64	4.67	4.72	4.76	4.79	4.81	4.82
20	.05	2.95	3.10	3.18	3.25	3.30	3.34	3.36	3.38	3.40	3.43	3.44	3.46	3.46	3.47
	.01	4.02	4.22	4.33	4.40	4.47	4.53	4.58	4.61	4.65	4.69	4.73	4.76	4.78	4.79
22	.05	2.93	3.08	3.17	3.24	3.29	3.32	3.35	3.37	3.39	3.42	3.44	3.45	3.46	3.47
	.01	3.99	4.17	4.28	4.36	4.42	4.48	4.53	4.57	4.60	4.65	4.68	4.71	4.74	4.75

Continued

TABLE A.5 (continued)

Error df	α	<i>r</i> = Number of Ordered Steps between Means													
		2	3	4	5	6	7	8	9	10	12	14	16	18	20
24	.05	2.92	3.07	3.15	3.22	3.28	3.31	3.34	3.37	3.38	3.41	3.44	3.45	3.46	3.47
	.01	3.96	4.14	4.24	4.33	4.39	4.44	4.49	4.53	4.57	4.62	4.64	4.67	4.70	4.72
26	.05	2.91	3.06	3.14	3.21	3.27	3.30	3.34	3.36	3.38	3.41	3.43	3.45	3.46	3.47
	.01	3.93	4.11	4.21	4.30	4.36	4.41	4.46	4.50	4.53	4.58	4.62	4.65	4.67	4.69
28	.05	2.90	3.04	3.13	3.20	3.26	3.30	3.33	3.35	3.37	3.40	3.43	3.45	3.46	3.47
	.01	3.91	3.08	4.18	4.28	4.34	4.39	4.43	4.47	4.51	4.56	4.60	4.62	4.65	4.67
30	.05	2.89	3.04	3.12	3.20	3.25	3.29	3.32	3.35	3.37	3.40	3.43	3.44	3.46	3.47
	.01	3.89	4.06	4.16	4.22	4.32	4.36	4.41	4.45	4.48	4.54	4.58	4.61	4.63	4.65
40	.05	2.86	3.01	3.10	3.17	3.22	3.27	3.30	3.33	3.35	3.39	3.42	3.44	3.46	3.47
	.01	3.82	3.99	4.10	4.17	4.24	4.30	4.34	4.37	4.41	4.46	4.51	4.54	4.57	4.59
60	.05	2.83	2.98	3.08	3.14	3.20	3.24	3.28	3.31	3.33	3.37	3.40	3.43	3.45	3.47
	.01	3.76	3.92	4.03	4.12	4.17	4.23	4.27	4.31	4.34	4.39	4.44	4.47	4.50	4.53
100	.05	2.80	2.95	3.05	3.12	3.18	3.22	3.26	3.29	3.32	3.36	3.40	3.42	3.45	3.47
	.01	3.71	3.86	3.93	4.06	4.11	4.17	4.21	4.25	4.29	4.35	4.38	4.42	4.45	4.48
∞	.05	2.77	2.92	3.02	3.09	3.15	3.19	3.23	3.26	3.29	3.34	3.38	3.41	3.44	3.47
	.01	3.64	3.80	3.90	3.98	4.04	4.09	4.14	4.17	4.20	4.26	4.31	4.34	4.38	4.41

Source: Duncan, D. B., *Biometrics*, 11: 1–42, 1955. With permission.

References

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Autom. Control*, AC-19, 716–723.
- Anderson, V. L. and McLean, R. A. (1974). *Design of Experiments: A Realistic Approach*. Marcel Dekker, New York, NY.
- Bancroft, T. A. (1968). *Topics in Intermediate Statistical Methods*, Vol. I. The Iowa State University Press, Ames, IA.
- Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. *Proc. R. Soc., Ser. A*, 160, 268–282.
- Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. B., Methodological*, 57, 289–300.
- Beyer, W. H., ed. (1966). *Handbook of Tables of Probability and Statistics*. The Chemical Rubber Co., Cleveland, OH.
- Beyer, W. H. (1968). *CRC Handbook of Tables for Probability and Statistics*. The Chemical Rubber Co., Cleveland, OH.
- Boardman, T. J. (1974). Confidence intervals for variance components—A comparative Monte Carlo study. *Biometrics*, 30, 251–262.
- Box, G. E. P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems. *Ann. Math. Stat.*, 25, 290–302.
- Box, G. E. P., Hunter, W. G., and Hunter, J. S. (1978). *Statistics for Experimenters*. John Wiley and Sons, New York, NY.
- Brown, K. G. (1976). Asymptotic behavior of MINQUE-type estimators of variance components. *Ann. Stat.*, 4, 746–754.
- Brown, W. A. and Forsythe, A. B. (1974). Robust tests for equality of variances. *J. Am. Stat. Assoc.*, 69, 364–367.
- Burdick, R. K. and Graybill, F. A. (1992). *Confidence Intervals on Variance Components*. Marcel Dekker, New York, NY.
- Cobb, G. W. (1997). *Introduction to Design and Analysis of Experiments*. Springer, New York, NY.
- Cochran, W. G., and Cox, G. M. 1957. *Experimental Design*. 2nd ed. John Wiley and Sons, New York, NY.
- Conover, W. J., Jounson, M. E., and Johnson, M. M. (1981). A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data. *Technometrics*, 23, 351–361.
- Corbeil, R. R. and Searle, S. R. (1976). A comparison of variance component estimators. *Biometrics*, 32, 779–791.
- Cornell, J. A. (1990). *Experiments with Mixtures: Designs, Models and Analysis of Mixture Data*. John Wiley and Sons, New York, NY.

- Davies, O. L. 1954. *Design and Analysis of Industrial Experiments*. Oliver and Boyd, London, UK.
- Edwards, D. and Berry, J. J. (1987). The efficiency of simulation-based multiple comparisons. *Biometrics*, 43, 913–928.
- Eisenhert, C. (1947). Assumptions underlying analysis of variance. *Biometrics*, 3, 1–22.
- Fai, A. H. T. and Cornelius, P. L. (1996). Approximate F-tests of multiple degree of freedom hypotheses in generalized least squares analyses of unbalanced split-plot experiments. *J. Stat. Comput. Simul.*, 54, 363–378.
- Federer, W. T. (1995). *Experimental Design*. Macmillan, New York, NY.
- Fisher, R. A. (1949). *The Design and Analysis of Experiments*. Oliver and Boyd, Edinburgh, UK.
- Giesbrecht, F. G. and Burns, J. C. (1985). Two-stage analysis based on a mixed model: Large-sample asymptotic theory and small-samples simulation results. *Biometrics*, 41, 477–486.
- Graybill, F. A. (1976). *Theory and Application of the Linear Model*. North Scituate, MA.
- Graybill, F. A. and Wang, C. M. (1980). Confidence intervals on nonnegative linear combinations of variances. *J. Am. Stat. Assoc.*, 75, 869–873.
- Greenhouse, S. W., and Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24, 95–112.
- Grizzle, J. E. (1965). The two-period change-over design and its use in clinical trials. *Biometrics*, 21, 467–80.
- Hartley, H. O. (1950). The maximum F-ratio as a short-cut test for heterogeneity of variances. *Biometrika*, 37, 308–312.
- Hartley, H. O. (1967). Expectations, variances and covariances of ANOVA mean squares by “synthesis.” *Biometrics*, 23, 105–114.
- Hartley, H. O. and Rao, J. N. K. (1967). Maximum likelihood estimation for the mixed analysis of variance model. *Biometrika*, 54, 93–108.
- Hayter, A. J. (1984). A proof of the conjecture that the Tukey–Kramer method is conservative. *Ann. Stat.*, 12, 61–75.
- Hemmerle, W. J. and Hartley, H. O. (1973). Computing maximum likelihood estimates for the mixed A.O.V. model using the \$W\$ transformation. *Technometrics*, 15, 819–831.
- Henderson, C. R. (1984). ANOVA, MIVQUE, REML, and ML algorithms for estimation of variances and covariances. *Statistics: An Appraisal*, pp. 257–280. David, H. A. and David, H. T. (eds.), Iowa State University, Ames, IA.
- Henderson, C. R. (1984). *Applications of Linear Models in Animal Breeding*. University of Guelph, Ontario, Canada.
- Hicks, Charles R. (1993). *Fundamental Concepts in the Design of Experiments*, 4th ed. W.B. Saunders Co., Philadelphia, PA.
- Holland, B. S. and Copenhaver, M. D. (1987). An improved sequentially rejective Bonferroni test procedure. *Biometrics*, 43, 417–424.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Stat.*, 6, 65–70.
- Hurvich, Clifford M. and Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika*, 76, 297–307.
- Huynh, H., and Feldt, L. S. (1970). Conditions under which mean square ratios in repeated measures designs have exact F-distributions. *J. Am. Stat. Assoc.*, 65, 1582–1589.
- John, P. W. M. (1971). *Statistical Design and Analysis of Experiments*. John Wiley and Sons, New York, NY.
- Johnson, D. E. (1976). Some new multiple comparison procedures for the two-way AOV model with interaction. *Biometrics*, 32, 929–934.
- Johnson, D. E. and Graybill, F. A. (1972). An analysis of a two-way model with interaction and no replication. *J. Am. Stat. Assoc.*, 67, 862–868.
- Kackar, A. N. and Harville, D. A. (1984). Approximations for standard errors of estimators of fixed and random effects in mixed linear models. *J. Am. Stat. Assoc.*, 86, 557–568.
- Kempthorne, O. (1952). *Design and Analysis of Experiments*. John Wiley and Sons, New York, NY.
- Kendall, M. G. and Stuart, A. (1952). *The Advanced Theory of Statistics*, Vol. 1. Hafner, New York, NY.
- Kendall, M. G. and Stuart, A. (1973). *The Advanced Theory of Statistics*, Vol. 2, 3rd ed. Griffin, London, UK.

- Kenward, M. G. and Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, 53, 983–997.
- Kirk, R. E. (1968). *Experimental Design: Procedures for the Behavioral Sciences*. Brooks/Cole, Belmont, CA.
- Kramer, C. Y. (1956). Extension of multiple range tests to group means with unequal numbers of replications. *Biometrics*, 12, 307–310.
- Laundsby, R. G. and Weese, D. L. (1993). *Straight Talk on Designing Experiments: An Introductory Design of Experiments Reference Handbook*. Laundsby Consulting, Colorado Springs, CO.
- Lehmann, E. L. (1986). *Testing Statistical Hypotheses*. John Wiley and Sons, New York, NY.
- Lentner, M. and Bishop, T. (1986). *Experimental Design and Analysis*. Valley Book Company, Blacksburg, VA.
- Levene, H. (1960). Robust tests for the equality of variances. In *Contributions to Probability and Statistics*, pp. 278–292. I. Olkin, S. G. Ghurye, W. Hoeffding, W. G. Madow, and H. B. Mann (eds.). Stanford University Press, Palo Alto, CA.
- Littell, R., Milliken, G. A., Stroup, W., Wolfinger, R., and Schabenberger, O. (2006). *SAS for Mixed Models*, 2nd edn. SAS Institute, Inc., Cary, NC.
- Lu, T.-F. C., Graybill, F. A., and Burdick, R. K. (1989). Confidence intervals on the ratio of expected $(\theta_1 - d\theta_2)/\theta_3$. *J. Stat. Plan. Infer.*, 21, 179–190.
- McGaughey, K. B. (2003). Tests of scale parameters using data depth. Ph.D. dissertation. Department of Statistics, Kansas State University, Manhattan, KS.
- McLean, R. A. and Sanders, W. L. (1988). Approximating degrees of freedom for standard errors in mixed linear models. *ASA Proceedings of the Statistical Computing Section*, 50–59. American Statistical Association, Alexandria, VA.
- Meed, R. (1988). *The Design of Experiments: Statistical Principles for Practical Application*. Cambridge University Press, Cambridge, UK.
- Miller, R. G. (1967). *Simultaneous Statistical Inference*. McGraw-Hill Inc., New York, NY.
- Miller, R. G., Jr. (1981). *Simultaneous Statistical Inference*. Springer, New York, NY.
- Milliken, G. A. (2003). Mixed models and repeated measures: Some illustrative industrial examples. *Handbook of Statistics 22: Statistics in Industry*, pp. 171–207. Elsevier Science Publishing, New York, NY; North-Holland Publishing Co., Amsterdam.
- Milliken, G. A. (2003a). Mixed models and repeated measures: some illustrative industrial examples. In *Handbook of Statistics Vol. 22: Statistics in Industry*, Khattree, R. and Rao, C. R., eds. Elsevier Science Publishing, Amsterdam/NY.
- Milliken, G. A. (2003b). *Multi-Level Models and Their Analyses*. Proceeding of the 28th SUGI Conference, Seattle, WA. SAS Institute Inc., Cary, NC.
- Milliken, G. A. and Graybill, F. A. (1970). Extensions of the general linear hypothesis model. *J. Am. Stat. Assoc.*, 65, 797–807.
- Milliken, G. A. and Johnson, D. E. (1989). *Analysis of Messy Data, Vol. 2: Nonreplicated Experiments*. Chapman & Hall/CRC Press, Boca Raton, FL.
- Milliken, G. A. and Johnson, D. E. (1992). *Analysis of Messy Data, Volume 1: Designed Experiments*. Chapman & Hall/CRC Press, Boca Raton, FL.
- Milliken, G. A. and Johnson, D. E. (2002). *Analysis of Messy Data, Vol 3: Analysis of Covariance*. Chapman and Hall/CRC, Boca Raton, FL.
- Milliken, G. A., Shi, X., Mendicino, M., and Vasudev, P. K. (1998). Strip-plot design for two-step processes. *Qual. Reliab. Enging. Int.*, 14, 1–15.
- Montgomery, D. C. (1991). *Design and Analysis of Experiments*, 3rd edn. John Wiley and Sons, New York, NY.
- Montgomery, D. C. and Runger, G. C. (1993–1994). Gauge capability and designed experiments. Part II: Experimental design models and variance component estimation. *Quality Engineering*, 6, 289–305.
- Morrison, D. F. (1976). *Multivariate Statistical Methods*. McGraw-Hill, New York, NY.
- Njuho, P. M. and Milliken, G. A. (2005). Analysis of linear models with one factor having both fixed and random levels. *Commun. Stat. Theory Meth.*, 34, 1979–1999.

- Olejnik, S. F. and Algina, J. (1987). Type I error rates and power estimates of selected parametric and nonparametric tests of scale. *J. Educ. Stat.*, 12, 45–61.
- O'Brien, R. G. (1979). A general ANOVA method for robust tests of additive models for variances. *J. Am. Stat. Assoc.*, 74, 877–880.
- Ott, L. (1988). *An Introduction to Statistical Methods and Data Analysis*, 3rd ed. PWS-Kent Publishing Co., Boston, MA.
- Rao, C. R. (1971). Minimum variance quadratic unbiased estimation of variance components. *J. Multivariate Anal.*, 1, 445–456.
- SAS Institute, Inc. (1999). *SAS/STAT User's Guide, Version 8*. SAS Institute, Inc., Cary, NC.
- SAS Institute, Inc. (2005) *JMP Introductory Guide*, Release 6. SAS Institute, Inc., Cary, NC.
- Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin* 2, 110–114.
- Schwarz, Gideon. (1978). Estimating the dimension of a model. *Ann. Stat.*, 6, 461–464.
- Searle, S. R. (1971). *Linear Models*. John Wiley and Sons, New York, NY.
- Searle, S. R. (1987). *Linear Models for Unbalanced Data*. John Wiley and Sons, New York, NY.
- Searle, S. R., Cassella, G. and McCulloch, C. E. (1992). *Variance Components*. John Wiley and Sons, New York, NY.
- Searle, S. R., Speed, F. M. and Milliken, G. A. (1980). Population marginal means in the linear model: An alternative to least squares means. *Am. Stat.*, 34, 216–221.
- Šidák, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. *J. Am. Stat. Assoc.*, 62, 626–633.
- St John, R. C. and Draper, N. R. (1975). D-Optimality for regression designs: A review. *Technometrics*, 17, 15.
- Swallow, W. H. and Monahan, J. F. (1984). Monte Carlo Comparison of ANOVA, MIVQUE, REML, and ML estimators of variance components. *Technometrics*, 26, 47–57.
- Swallow, W. H. and Searles, S. R. (1978). Minimum variance quadratic unbiased estimation (MIVQUE) of variance components. *Technometrics*, 20, 265–272.
- Ting, N., Burdick, R. K., Graybill, F. A., Jeyaratnam, S. and Lu, T.-F. C. (1990). Confidence intervals on linear combinations of variance components that are unrestricted in sign. *J. Stat. Comput. Simul.*, 35, 135–143.
- Tukey, J. W. (1952). Allowances for various types of error rates. Unpublished IMS address, Chicago, IL.
- Tukey, J. W. (1953). The problem of multiple comparisons. Unpublished manuscript.
- Westfall, P. H. (2002). *From Farms to Pharmaceuticals: Multiple Comparisons Enter the 21st Century*. Proceedings of the 13th Annual Kansas State University Conference on Applied Statistics in Agriculture, G. A. Milliken, ed. Department of Statistics, Kansas State University, Manhattan, KS.
- Westfall, P. H., Tobias, R. D., Rom, D., Wolfinger, R. D., and Hochberg, Y. (1999). *Multiple Comparisons and Multiple Tests Using the SAS System*. SAS Institute, Inc., Cary, NC.
- Williams, E. J. (1949). Experimental designs balanced for the estimation of residual effects of treatments. *Aust. J. Sci. Res., Ser. A* 2, 149–168.
- Williams, J. S. (1962). A confidence interval for variance components. *Biometrika*, 49, 278–281.
- Winer, B. J. (1971). *Statistical Principles in Experimental Design*, 2nd ed. McGraw-Hill, New York, NY.
- Welch, B. L. (1951). On the comparison of several mean values: An alternative approach. *Biometrika*, 38, 330–336.

Index

A

Absolute deviations, values, 27
Agricultural study, 148–150
AIC. *See* Akaike Information Criteria (AIC)
AICC, 557, 558, 567, 588, 593
 covariance structure, 565, 576, 577
Aircraft engines, 145–146
Akaike Information Criteria (AIC), 37, 557,
 558, 567, 577, 588
AR(1), 562, 593
 covariance structure, 565, 576
Analysis of variance (ANOVA), 35, 103, 240,
 241, 521, 607
 balanced incomplete block design
 structure, 95
 bread baking study, 124, 125
 bread recipes and baking temperature
 example, 423
 cheese-making experiment, 133, 134, 135
 column models, 127
 comfort study, 141, 148, 512
 companies and insecticide example, 635
 complex split-plot design, 130
 cooking beans example, 117, 118,
 119, 120
 crossover design, 86, 103, 106, 109,
 146, 602, 604
 display case study, 125, 130
 drug heart rate study, 503
 expected mean squares, 343
 family attitude study, 517
 flour milling experiment, 121
 IBD design structures, 96, 125

Latin square, 88, 89, 92, 93
loaf volume, 425–426
nested designs structure, 148
nitrogen and irrigation, 481
nitrogen by irrigation example, 486
one-way model, 236, 296
one-way treatment structure, 84, 88,
 106, 107, 108
paint-paving experiment, 190, 191
RCBD, 84, 85, 94, 104, 107, 108, 110, 144
repeated Latin square treatment
 structure, 90
row and column blocks example, 98
soybean varieties nested within
 maturity groups, 144
split-plot designs, 109, 110, 145
strip-plot designs, 111, 112, 473, 474, 486
strip-split-plot design, 484, 485, 489
three-way factorial arrangement, 87
two-way nested treatment structure, 146
two-way random effects model, 339
two-way treatment structures, 89, 90,
 92, 94, 103, 104, 109, 110, 112,
 250, 251
type I sums of squares, 342
type III, 605
unbalanced two-way experiment,
 201, 204
unweighted means, 233, 236, 237
variance components inferences
 methods, 338
wheat yield data, 427
whole-plot design structure, 109

- Animal comparison ratios, 453–456, 458
asymptotic covariance matrix of estimates, 458
covariance parameters, 454
fixed effects type III tests, 454
pairwise comparisons, 455
SAS-mixed code, 454
temperature force means, 455
Tukey adjustment, 455
- Animal genetics, 143
nested treatment structure, 143
strip-plot designs, 143
- ANOVA. *See* Analysis of variance (ANOVA)
- AR(1) structure, 556
AIC, 562, 593
drug heart rate study, 564
estimated R matrix, 561, 562
heterogeneous variances, 562
- Assembly line workers, 365–381
covariance parameter, 365, 367
data, 366
efficiency, 365
estimates of variance components, 371
expected mean square, 367, 371
JMP fit model screen, 380
mean square, 367
ML, 367
MVQ, 367
MVQNB, 367
reduced model, 371, 375, 381
REML, 367, 381
REML estimates, 374, 375
Satterthwaite approximation, 373, 375
Satterthwaite-type confidence intervals, 374
type I analysis, 367, 371, 377
type I sums of squares, 375
type I tests of hypotheses, 369
type II, 367
type III, 367
type III analysis of variance, 368, 371
type III sums of squares, 373
type III tests, 370, 373
variance components, 369
Wald-type confidence intervals, 375
- Asymptotic covariance matrix of estimates, 458
Asymptotic sampling distribution, 26
- B**
- Baking bread example. *See* Bread baking study
- Balanced incomplete block design structure, 94
analysis of variance, 95
- Balanced model
covariance matrix, 406
data set, 408
one-way, 319–321
- Balanced one-way random effects treatment structure, 356
- Balanced two-way experiments, 187–198
case study, 187–198
interaction effects contrasts, 188
main effect means contrasts, 187
multiple comparisons, 195
paint-paving example, 189–192
quantitative treatment factors analysis, 193–194
- Balanced two-way treatment structures, 181–184, 207, 243–244
computer analyses, 184
effects model, 182, 209–230
fat-surfactant example, 239–242
interactions, 182, 202–203
main effects, 183, 202–203
means model, 181
model definitions, 181–182, 199
multiple comparisons, 206
parameter estimation, 182, 199–200
population marginal means, 204–205
simultaneous interferences, 206
testing whether all means are equal, 201
unequal subclass numbers case study, 239–244
unequal subclass numbers means model analysis, 199–207
- Bartlett's test, 25, 26
one-way treatment structure, 24
- Batch error terms, 121
- Benjamini and Hochberg method, 45
adjusted p-values, 60
control FDR, 44, 58–59
LDC, 59
simultaneous inference procedures, 58
- Best estimates, 504
- Best linear unbiased estimator (BLUE), 395, 398
estimable functions, 397
machine person example, 406
- Best linear unbiased predictor (BLUP), 397, 398
- BF test GLM code, 35, 36
- Bias prevention, 73
- BIC. *See* Schwarz's Bayesian Criterion (BIC)
- Big block analysis of variance
nitrogen by irrigation example, 486
- Block design structure, 76

- Block effects, 74
Block interaction, 474
Blocking, 71, 97
Blocks and replications example, 94–97
BLUE. *See* Best linear unbiased estimator (BLUE)
BLUP. *See* Best linear unbiased predictor (BLUP)
Bonferroni–Holm method, 58
 simultaneous inference procedures, 58
Bonferroni's method, 45, 48, 49, 51, 55, 195
 contrast, 60
 differences, 53
 simultaneous inference procedures, 48
 widths of confidence intervals, 57
Bootstrap methodology, 51
Box's correction
 conservative estimate, 548–549
 sorghum varieties and fertilizer levels, 549
Bread baking study, 122–125
 baking bread, 122–124
 combinations multilevel designs, 122–124
 IBD design structures, 123, 124, 125
 split-plot design structures, 122–124
Bread recipes and baking temperature
 example, 422–426, 433–434
 analysis of variance, 423, 425–426
 means comparing, 434
 randomization, 423, 424
 SAS-mixed code, 426
 split-plot analysis, 422–425
Brown and Forsythe's test, 25, 27
 equality of variances, 35
 homogeneity of errors variance, 35
 one-way treatment structure, 24
- C**
- Cell means, 201
Change-over designs, 141–142. *See also*
 Crossover designs
Cheese making example, 131–135
 analysis of variance, 133, 135
 analysis of variance for batch
 experimental unit design, 134
batch design, 133
chamber design, 132
chamber experimental unit, 133
combinations multilevel designs, 131–134
randomization scheme, 131, 132, 133
repeated measures multilevel designs,
 131–134
split-plots designs, 131–134
strip-plot designs, 131–134
- Chi-square
 sphericity tests, 552
Coffee price example, 289–290, 629–630,
 632–634, 638–639
 analysis of variance, 633
 estimates of mean, 634
 estimates of variance components, 634
JMP data screen, 638
JMP fit model screen, 639
nested designs analysis, 629
random effects model, 632
REML estimates, 639
three-way, two-level, nested system, 289
type I sums of squares, mean squares, 633
- Column blocks, 80
Column marginal mean, 201
Column models
 analysis of variance, 127
Combinations multilevel designs, 101–154
 bread baking study, 122–124
 cheese-making experiment, 131–134
Comfort study, 139–141, 147–148, 510–516,
 524–527, 570–583, 631–632
 analysis of variance, 141, 148, 512, 632
 comparisons between clothing, 580
 comparisons between environments, 582
 comparisons between sexes, 581
 comparisons between times, 582
 complex, 510–516, 524–527, 570–583
 compound symmetry covariance
 structure, 576
covariance parameter estimates, 524, 577
data, 511, 574, 632
data arrangement, 140
environment × time means, 513
environment × time means with
 comparisons, 515
Fisher's LSD method, 512
fit statistics, 576
JMP fit model screen, 637
JMP screen, 636
least squares from JMP, 638
least squares means, 526
linear and quadratic contrasts, 578, 579
linear and quadratic contrasts for
 time, 525
means with LSD, 513
measures of linear and quadratic
 trends, 514
nested designs analysis, 629
pairwise comparisons, 526
persons and temperatures assignments
 to chambers, 147

- Comfort study (*Continued*)
 repeated measures complex examples, 573–582
 repeated measures factor, 575
 response over time, 582
 SAS commands, 524
 SAS-mixed statements, 579
 Satterthwaite estimated degrees of freedom, 514
 sex × clothing means with LSD, 513
 split-plot-in-time analysis, 575
 standard error of comparison, 513
 strip-plot designs, 139–140, 147
 three-way least squares means, 580
 Tukey multiple comparisons, 638
 two-way least squares means, 580
 type III tests of fixed effects, 578
 type II tests of fixed effects, 525
- Companies and insecticide example, 627–628, 630–631, 634–635
 analysis of variance, 635
 data, 628
 means estimates, 631
 model, 634
 nested designs analysis, 627–628
 treatment structure, 628
- Comparisons
 area means, 519
 comfort study, 515
 error rate, 44
 family attitude study, 592
 mean to mean, 54
 percentage points, 54
- Complete block design structures, 79
- Completely randomized design (CRD), 71
 analysis of variance, 86, 103, 106, 109, 146
 cupcakes, 102
 design structure, 79, 84
 error term and variance source, 113
 means and effects models, 86
 one-way treatment structure, 1–17, 21–42
 two-way treatment structures, 159
 whole-plot structure, 109
- Complex split-plot design, 130
- Complex three-way random effect, 339–343
- Compound symmetry covariance
 structure, 556
 comfort study, 576
 estimated R matrix, 562
 heterogeneous variances, 562
- Computational algorithms, 321
- Computer analyses, 184
- Conditional error, 13–14
- Confidence intervals, 247–249, 358
 graph, 354
 region, 378
- Construct sequence treatments, 141
- CONTRAST statement, 216
- Cooking beans experiment example, 116–120
 analysis of variance, 117, 118, 119, 120
 batch analysis, 119
 cooking methods, 118
 experimental unit size, 120
 randomization scheme for assigning varieties, 116
 randomization scheme of assigning cooking methods, 117
 randomly assign to batches, 118
 row analysis, 117
 strip-plot designs, 116–119
 whole plot or row design structure, 119
- Covariance matrix, 590
- Covariance parameter estimates, 37
 comfort study, 577
 crossover designs, 615
 crossover designs six sequences, 620
 family attitude study, 528, 589
- Covariance structures comparison, 587
- CRD. *See* Completely randomized design (CRD)
- Crossover designs, 71, 599–618, 623–625
 analysis, 599–626
 ANOVA, 602, 604
 contrasts, 615
 covariance parameter estimates, 615
 definitions, assumptions, and models, 599
 ESTIMATE options, 607
 expected main squares, 605
 fixed effects, 609
 F-statistic for testing, 602
 GRIZ, 604
 least squares means, 608, 616
 milk yields, 622
 more than two periods, 609–615
 more than two treatments, 616–625
 pain score, 622
 period cell means sequencing, 613
 SAS code, 613
 SAS-GLM code, 604, 605
 SAS-mixed code, 606
 sequences set, 600
 strip-plot designs, 141
 subject error mean square, 604
 subjects contrast, 612
 three period/two treatment, 610, 617–618
 treatment, 609, 616
 treatment main effect means, 605

- treatment/4 period, 623
two period/two treatment designs, 600–608
two-way means, 609
type III ANOVA, 605
type III tests of fixed effects, 608, 615
wash-out period, 607
Crossover designs six sequences, 617–618, 619–621
covariance parameter estimates, 620
estimable least squares means for treatments, 621
least squares means, 620
pairwise comparisons, 621
type III tests of fixed effects, 620
Cupcake experiment, 102–115
design structure, 102, 108
IBD design structures, 110
RCBD, 103
split-plot design, 104, 106
split-plot design structure, 109
strip-plot designs, 110, 111
volume evaluation, 102
- D**
- Data-driven comparisons, 45
Data snooping, 45
Designing experiments, 71–100
blocks and replications, 94–96
design structure, 77
diets, 83–87
examples, 83–97
house paint, 88–89
nitrogen and potassium levels example, 92–93
row and column blocks, 97
steel plates, 90–91
structures, 66–79
treatment structure, 77
treatment structure types, 80–82
types, 78–79
Design matrix, 155
Design structure, 71
blocks and replications, 94
cupcake treatment structure, 102, 108
nested or hierarchical structure, 105
nitrogen and potassium levels example, 92
oven treatment structure, 108
row and column blocks, 97
split-plot and repeated measures, 83
Diets example, 83–87
designing experiments, 83–87
treatment structure, 83, 87
Display case study, 125–130
analysis of variance, 125, 130
intensities assignments, 125
levels of lighting at temperature, 128
levels of packaging at temperature, 129
strip-plot designs, 125–130
temperature intensity levels, 127
temperatures assignments, 125
Drug heart rate study, 500–510, 519–523, 535–537, 558–570, 592–596
analysis of variance, 506
AR(1) conditions, 564
comparisons of time means, 522
compound symmetry, 563
covariance parameter estimates, 521
differences of least squares means, 523
drug and time means, 521
drug effects, 500, 506
estimated covariance matrices, 563
estimated standard error, 507
estimate options, 523
graph of means, 508
H-F conditions, 536, 560, 563
linear and quadratic contrasts in time, 523
linear and quadratic trends estimates, 508
LSD value, 507
pairwise comparisons, 522
SAS-mixed code, 520
Satterthwaite estimated degrees of freedom, 509
Satterthwaite's method, 504
statistical tests, 521
time means comparisons, 507, 509
unstructured conditions, 564
vector of responses, 535
Drugs and errors examples, 25–38
one-way treatment structure, 25–28, 31–32
Duncan's method
multiple comparison procedure, 67
multiple range tests, 61, 64
proc GLM code, 66
simultaneous inference procedures, 64
Dunnett's method, 45, 55
simultaneous inference procedures, 53
- E**
- EER. *See* Experimentwise error rate (EER)
EERC. *See* Experimentwise error rate under the complete null hypothesis (EERC)

- Effects model analysis, 164, 209–230
 balanced two-way treatment structures, 209–230
 completely randomized design structure, 86
 computer analyses, 226–227
 defined, 182
 estimable functions in SAS, 213–216
 exercises, 229
 hypotheses, 219, 220
 model definition, 209
 parameters estimates and type I analysis, 209–212
 population marginal means and least squares means, 226
 type III hypotheses, 220
 types I–IV estimable functions in SAS-GLM, 222–225
 types I–IV hypotheses, 217–221
 unequal subclass numbers, 209–230
- Engines on aircraft example. *See Aircraft engines*
- Environmental conditions, 139
- Equality of means
 hypothesis testing, 11
 one-way model, 14
 task and pulse rate data, 15
- Equal variances, 22
- Error rates
 degrees of freedom, 113, 117
 nitrogen and potassium levels example, 92
 ratio, 66
 simultaneous inference procedures, 44
 structure, 92
- Error terms
 designs, 113
 split-plot designs, 113
 strip-plot designs, 113
 variability sources, 113
- Error variances, 22
- Estimable functions
 basis set, 215
 general forms, 214
- ESTIMATE options, 261
- ESTIMATE statement, 216
- Estimated population marginal means, 173
- Estimated standard errors, 504
 cell means, 176
 difference, 408
 GPA data, 176
 machine person example, 408
- Estimating variance components method, 309–336
 algebraic method, 293–294
 assembly line workers, 371
 balanced one-way model, 318–321, 319–321
 balanced one-way model REML solution, 323–324
 computing expected mean squares, 292–306
 description, 325–326
 exercises, 308, 334–336
 general random effects model in matrix notation, 290–291
- Hartley's method of synthesis, 295–306
- JMP variance components estimation, 329–333
- maximum likelihood estimators, 318–321
- MIVQUE method, 325–328
- moments method, 309–317
- one-way random effects model, 290–291, 313–314
- random effects nested treatment structure, 289
- restricted or residual maximum likelihood estimation, 322–324
- two-way design, 315–317
- unbalanced one-way model, 309–312
- unbalanced one-way design, 327–328
- Examples. *See specific name*
- Expected mean square
 analysis of variance, 343
 assembly line workers, 367, 371
 and F-statistic, 343
- Experimental unit, 71, 72
 cooking beans example, 120
 flour milling experiment, 121
 horse feet, 136
 selection, 74
 size, 120, 121
- Experiments. *See also specific name*
 applications, 72
 combinations and generalizations, 80
 design, 73
 designing, 83–87
 objectives, 72
 randomization scheme, 78
 statistical objective, 73
 times, 72
 treatments, 72
- Experimentwise error rate (EER), 44, 45
- Experimentwise error rate under the complete null hypothesis (EERC), 44, 61

F

- Factorial arrangement, 82
 treatment structure, 81
- False discovery rate (FDR), 44, 45
 Benjamini and Hochberg method, 58
- Family attitude study, 501–502, 516–519,
 525, 527–529, 583–592
 analysis of variance, 517
 comparison of person means, 518
 comparison of time means, 518
 comparisons, 592
 covariance parameter estimates,
 528, 589
 covariance structures comparison, 587
 data, 517, 583
 estimated covariance matrix, 590
 fit statistics, 588
 fixed effects, 589
 least squares means, 591
 pairwise comparison, 529, 591
 repeated measures complex examples,
 583–591
 SAS commands, 527–529, 587
 two-way least squares means, 591
 two-way means, 528
 type III tests of fixed effects, 528, 590
- Familywise error rate (FWER), 44, 45, 47,
 49–51, 54, 58, 60, 64
- Fat–surfactant example, 239–242, 265–267
 balanced two-way treatment structures,
 239–242
 least squares means, 241, 268
 pairwise comparisons, 242
 SAS analyses, 240
 specific volumes, 240
 treatment combination, 266
 two-way combinations, 265
 volumes, 266
- FDR. *See* False discovery rate (FDR)
- Fertilizer. *See* Moisture and fertilizer
 example; Sorghum varieties
 and fertilizer levels; Wheat
 varieties example
- Final regression model, 448
- First-order Taylor's series, 349
- Fisher's least significant difference
 method, 64, 65
 comfort study, 512
 simultaneous inference procedures, 47
- Fit statistics
 comfort study, 576

- family attitude study, 588
 four covariance structures, 576
- Fitting-constants method, 306
- Fixed effects, 287
 determination, 288
 drug heart rate study, 568
 family attitude study, 589
 SAS-mixed commands, 568
 type III tests, 568
- Fractional factorial arrangement
 treatment structure, 81
- F*-statistic, 9, 47, 66–68, 202–203, 216, 248–249
 analysis of variance, 343
 crossover designs, 602
 multiple comparisons, 195
- Full model, 13
 covariance parameter estimates, 446
 evaluate likelihood function, 346
 fixed effects solution, 447
 proc mixed code and results for fitting, 346
 response surface, 447
 SAS-mixed code, 446
- FWER. *See* Familywise error rate (FWER)
- G**
- General random effects model
 components, 291
 maximum likelihood estimators, 318

G–G. *See* Greenhouse and Geisser (G–G)
 estimate

Grade point average data, 175–178
 estimated standard errors, 176
 least squares estimator, 176
 SAS-GLM code, 178
 students to classes population
 distribution, 178
 two-way treatment structures, 175, 176

Greenhouse and Geisser (G–G) estimate
 adjusted degrees of freedom, 552
 sorghum varieties and fertilizer levels,
 549, 552
- Grinding wheat example, 120–121
 analysis of variance, 121
 batch error terms, 121
 error terms, 121
 experimental unit size, 121
 randomization process, 121
 strip-plot designs, 120–121
 whole-plot analysis, 121
 whole-plot method, 120

Grizzle's data using crossover designs,
 602, 603, 604

H

- Hartley's method of synthesis
 computing expectations, 299–300
 estimating variance components, 295–306
 expected values, 305
 random effects models, 295–306
 sums of squares expectation, 300
- Hartley's test, 25
 F-Max, 23
- Henderson's method I
 sums of squares, 315
- Henderson's method III, 306
 sums of squares, 316–317
- Henderson's methods I, II, III, and IV, 305
- Herbicide. *See* Nitrogen by irrigation
 example; Soybeans and herbicides
- Heterogeneous AR(1) structure, 556
- Heterogeneous compound symmetry structure, 556
- Heterogeneous errors
 one-way treatment structure, 21–42
 variances, 22
- H-F. *See* Huynh–Feldt (H-F) structure
- Hierarchical design structure
 analysis of variance table, 115
 split-plot designs, 627
 teaching method study, 115
- Higher-order treatment structures
 analysis, 271–276
 analysis, 284
 cross classified, 284
- Holm modification, 58
- Homogeneity of errors variance, 35
- Homogeneity of variances tests, 23–25
- Homogeneous errors
 linear combinations simultaneous tests, 7–8
 one-way treatment structure, 1–17
- Horse feet experiment, 136–139
 analysis of variance, 139
 experimental units sizes, 136
 IBD design structures, 138
 strip-plot designs, 136–138
 three-way treatment structures, 137
 treatment combinations assignment, 136, 138
- House paint example, 88–90
 designing experiments, 88–89
 nested within squares design, 143
 two-way treatment structure, 89
- Huynh–Feldt (H-F) structure, 555, 556
 correction, 549
 drug heart rate study, 536, 560, 563

- estimated R matrix, 560
 sorghum varieties and fertilizer levels, 549

- Hypothesis testing, 247–249
 mixed models, 396
 nested designs analysis, 633–635
 two-way treatment structures, 247–248
 variance components inferences methods, 337–346

I

- IBD. *See* Incomplete block design (IBD) structures
- Incomplete block design (IBD) structures, 71, 79, 80
 analysis of variance, 96, 124, 125
 bread baking study, 123, 124, 125
 cupcake experiment, 110
 horse foot experiment, 138
 temperature assignments, 123
 whole-plot, 123–125
- Insect damage in wheat, 349–350
 computations, 314
 REML, 349
- Interaction hypotheses, 190

J

- JMP
 coffee prices example, 638
 comfort study, 636, 638
 data, 331, 638
 data set, 329
 estimating variance components methods, 329–333
 fit model screen, 330
 REML, 331, 332, 333
 semiconductor industry, 490–491
 strip-plot designs, 486–490
 variance components estimation, 329–333
- JMP computations
 fertilizer, 459–465
 random effects model case study, 379
 split-plot designs, 459–465
- JMP fit model, 330, 332
 assembly line workers, 380
 coffee prices example, 639
 comfort study, 637
 unbounded variance components, 333

K

Kackar–Harville type of adjustment, 395
Kenward–Roger's method, 555
Kronecker or direct product, 319

L

Large balanced two-way experiments
 having subclass numbers,
 231–238
computer analyses, 236
exercises, 238
feasibility problems, 231
method of unweighted means, 235
simultaneous inference and multiple
 comparisons, 233–234
unweighted means, 232

Latin square, 71
 analysis of variance, 89, 92, 93
 analysis of variance table, 88
 arrangement, 90, 92, 93
 design, 80, 90
 design structure, 88, 90, 97
 house paint, 88
 one-way treatment structure, 88
 row and columns, 97
 steel plates, 90
 temperature level, 91
 treatment structure, 88, 89, 91, 93

Least significant difference (LSD)
 procedure, 45, 53, 64, 65, 195
comfort study, 512
comparing, 68
drug heart rate study, 507
simulated error rates, 47
simultaneous inference procedures, 46

Least squares estimator, 164
 GPA data, 176
Least squares means
 comfort study, 526
 crossover designs, 608, 616
 crossover designs six sequences, 620
 family attitude study, 591
 fat–surfactant example, 241, 267, 268
 machine person example, 408
 multicenter drug experiment, 595
 pairwise comparisons, 267
 plot, 242
 three way interactions, 465
 Tukey HSD option, 465

Least squares solution, 168

Le^ve_ne's test, 25

 computations, 27
 one-way treatment structure, 24

Likelihood ratio test, 344–345

 statistic, 538
 variance components inferences
 methods, 344

Wilks', 538

Linear and quadratic trends

 comfort study, 514, 578
 drug heart rate study, 508
 SAS-mixed code with estimate
 statements, 442

Linear comparisons, 59

 drug heart rate study, 505
 one-way treatment structure, 4
 proc mixed code, 57
 SAS-Mixed code, 56

Linear model

 matrix form, 171–172
 mixed matrix notation, 386
 testing hypotheses, 171–172

LSD. *See* Least significant difference (LSD)
 procedure

M

Machine person example, 401–417

 BLUE, 406
 estimate of standard error of
 difference, 408
 JMP analysis, 415–417
 nested designs analysis, 636–639
 pairwise difference, 408
 productivity of machines, 401
 productivity scores, 402
 Satterthwaite confidence intervals, 412
 two-way mixed model, 401
 unbalanced two-way mixed model, 409

Main effect contrasts, 190

Main effects means, 283

Making cheese example. *See* Cheese
 making example

MANOVA. *See* Multivariate analysis of
 variance (MANOVA) methods

Marginal means

 best estimates, 227
 unbalanced two-way experiment, 201

Matrix form of model, 155–180
 basic notation, 155–160
 connectedness, 171

- Matrix form of model (*Continued*)
 estimability and connected designs, 169–171
 estimable functions, 170
 least squares equations, 161–168
 linear model testing hypotheses, 171–172
 one-way treatment structure, 156–157, 167–168
 population marginal means, 172–178
 set-to-zero restrictions, 166
 simple linear regression model, 156
 sum-to-zero restrictions, 165–166
 two-way treatment structure means model, 158–160
- Mauchly's criterion sphericity tests, 552
- Maximum likelihood estimates (ML), 556
 assembly line workers covariance parameter, 367
 difference between machines, 414
 estimates of mean for machine, 413
 estimates of variance components, 403, 411
 estimating variance components methods, 318–321
 general random effects model, 318
 machine person example, 402
 models variance components, 347
 proc mixed code to compute, 321
 proc mixed code to obtains, 322
 solutions, 348
 two-way mixed model, 390
- Means model, 158
 balanced two-way treatment structures, 181
 character, 21
 completely randomized design structure, 86
 defined, 181
 matrix form, 160
 raw, least squares, and weighted GPA data, 176, 177
 type III hypotheses, 220
- Mean square, 367
- Means/Walker option, 66
- Meat display case study. *See* Display case study
- Method of moment
 estimators, 311
 machine person example, 402
 technique, 316
- Method of unweighted means, 235
- Milk yields, 622
- Minimum variance quadratic unbiased estimators (MIVQUE), 309, 325
 assembly line workers covariance parameter, 367
 difference between machines, 414
 estimates of mean for machine, 413
 estimates of variance components, 403, 411
 estimating variance components methods, 325–328
 machine person example, 402
 no bound, 367
 proc mixed code to compute, 328
 proc mixed code to obtain, 328
 variance components, 329
- MINQUE method, 387, 394
- Missing treatment combinations, 245–247
 two-way experiment, 246
- MIVQUE. *See* Minimum variance quadratic unbiased estimators (MIVQUE)
- MIVQUE0. *See* Minimum variance quadratic unbiased estimators (MIVQUE)
- Mixed models
 analysis, 385–400
 balanced and unbalanced, 401
 best linear unbiased prediction, 397
 case studies, 401–420
 confidence intervals, 395
 estimation, 394
 exercises, 399–400, 418–420
 fixed effects, 394–396
 maximum likelihood, 390–391
 MINQUE method, 394
 mixed model equations, 397–398
 moments method, 387–389
 parts, 385
 procedures, 555
 random effects, 387–393
 residual maximum likelihood, 392–393
 testing hypotheses, 396
 two-way mixed model, 401–408
 unbalanced two-way data set JMP analysis, 415–416
 unbalanced two-way mixed model, 409–414
- Mixed-up split-plot design, 448–450
- ML. *See* Maximum likelihood estimates (ML)
- Model
 comparison procedure, 15
 fixed or fixed effects, 288
 mixed or mixed effects, 288
 random or random effects, 288
- MODEL II, 309

- MODEL statement SAS-GLM, 223
Moisture and fertilizer example,
 440–448, 449
 dry matter graphic, 445
 dry matter measurements, 440–442
 linear and quadratic trends, 443, 444
 prediction surface for dry matter, 449
Satterthwaite approximation, 444
split-plot designs, 440–443
Moments method
 estimating variance components
 methods, 309–317
 mixed models, 387–389
Multicenter drug experiment, 592–596
 covariance parameter estimates, 595
 hypothesis tests of fixed effects, 595
 least squares means, 595
 main effect means, 596
 pairwise comparison, 596
 repeated measures complex examples,
 592–596
 SAS commands, 594
 time least squares means, 596
Multilevel designs
 bread baking study, 122–124
 characteristics, 114
 cupcake experiment, 114
Multilocation study
 alfalfa varieties assignment, 149
 analysis of variance, 150
 experimental units display, 149
 repeated measures, 150
 repeated measures complex examples,
 592–596
 strip-plot designs, 148–149
Multiple comparisons, 43–70, 52, 66
 critical differences, 55
 SAS system code, 54
 and simultaneous inference
 procedures, 43–70
 types, 45
Multiple range tests
 pairwise comparisons example, 66–67
 simultaneous inference
 procedures, 61–63
Multivariate analysis of variance
 (MANOVA) methods
 cross-products matrix, 546
 error sums of squares, 546
 interaction tests, 547
 repeated measures experiments, 537–546
 SAS-GLM code, 545
sorghum varieties and fertilizer levels,
 545, 547, 553
test for time, 552
Multivariate *t* method, 55–56
 simultaneous inference procedures, 55
 widths of confidence intervals, 57
MVQ. *See* Minimum variance quadratic
 unbiased estimators (MIVQUE)
MVQNB. *See* Minimum variance quadratic
 unbiased estimators (MIVQUE)
- N
- Nested designs analysis, 627–650
 coffee price example revisited, 629
 comfort experiment revisited, 629
 companies and insecticides, 627–628
 confidence interval construction, 633–635
 definitions, assumptions, models, 627–629
 machine person example, 636–639
 parameter estimation, 629–632
 testing hypotheses, 633–635
Nesting
 analysis of variance, 148
 animal genetics, 143
 coffee prices example, 289
 observable engine-aircraft
 configurations, 146
 strip-plot designs, 142–150
 treatment structure, 143
Nitrogen and potassium levels
 example, 92–94
 designing experiments, 92–94
 design structure, 92
 treatment combinations, 93
Nitrogen by irrigation example, 477–490
 analysis of variance, 481
 big block analysis of variance, 486
 covariance parameter estimates, 479
 data comparisons, 479
 estimate statements, 480
 field layout, 481
 fixed effects tests, 479
 REML analysis, 479
 SAS-mixed code, 478, 479
 seeding rates, 484, 486
 split-plot design, 481, 482
 strip-plot as subplots, 486
 strip-plot data, 478
 strip-plot design, 481, 482
 strip-split-plot design, 484
 type III sums of squares analysis, 478

- Noint option, 249
 Null hypothesis deviations, 172
 Null model likelihood ratio test, 37
 Nutrition scores example, 277–284
 main-effects means, 283
 SAS-GLM, 283
 three-way treatment structures, 277
 type IV hypotheses, 282
- O**
- O'Brien test, 25
 scores, 28
 statistic, 27
 One-fold nested design, 358
 One-way model
 analysis of variance, 296
 analyzing treatment, 22
 effects, 162
 equality of means, 14
 means, 162
 normal equations, 162
 random effects, 296, 362
 simulation code and results, 362
 One-way random effects model
 estimating variance components, 290–291
 estimating variance components methods, 313–314
 random effects models, 290–291
 One-way treatment structure, 81
 analysis of variance, 84, 88, 106, 107, 108
 arrangement, 90
 Bartlett's test, 24
 Brown and Forsythe's test, 24
 classification, 83
 combined with two-way treatment structure, 82
 comfort experiment, 147
 comparing all means, 34–37
 completely randomized design structure, 1–17, 21–42
 computer analyses, 15
 diets, 83
 drugs and errors examples, 25–28, 31–32
 Hartley's F-Max test, 23
 heterogeneous errors, 21–42
 homogeneity of variances tests, 23–25
 homogeneous errors, 1–17
 homogeneous errors linear combinations simultaneous tests, 7–8
 Latin square design structure, 88
 Levene's test, 24
- linear combinations, 4
 linear combinations inferences, 29–30
 loaf design, 424
 matrix form of model, 156–157, 167–168
 means and effects, 167
 model definitions, 22
 model definitions and assumptions, 1
 O'Brien's test, 25
 parameter estimation, 2–3, 22
 principle of conditional error, 13–14
 Satterthwaite approximation for degrees of freedom, 33
 tasks and pulse rate example, 5–6, 15
 testing and equality of all means, 11
 tests and confidence intervals, 4
 Optimal design treatment structures, 82
 Orthogonal components, 551
 Orthogonal polynomials expected values, 194
 Oven level analysis. *See* Bread baking study
 Overspecified or singular models, 164
- P**
- Pain score, 622
 Paint–paving example, 189–192
 analysis of variance, 190, 191
 balanced two-way experiments, 189–192
 cell means, 189
 interaction hypotheses, 190
 main effect contrasts, 190
 single-degree-of-freedom tests, 191
 Paired-association learning task experiment, 26
 Pairwise comparisons, 21, 28, 45
 all least squares means, 281
 animal comparisons, 455
 comfort study, 526
 crossover designs six sequences, 621
 drug heart rate study, 522
 equality test, 53
 example, 51–53
 family attitude study, 529, 591
 fat–surfactant example, 242
 least squares means, 242
 least squares means for effect fat × surfactant, 267
 multicenter drug experiment, 596
 multiple range, 66–68
 percentage points, 52
 ration by packaging force, 456
 SAS system code, 52
 simultaneous inference procedures, 51–52
 unequal variance model, 38

- Pairwise difference, 408
Parameters
 estimable functions, 170
 functions, 504
Pooled estimate, 3
 variance, 10
Population marginal means, 204–205
Principle of conditional error, 13, 15
PRIORTRT, 614
- Q**
- Quadratic contrasts
 comfort study, 514, 578
- R**
- Random effects, 287, 386
 batch design, 133
 bread recipes and baking temperature example, 423, 424
 cheese-making experiment, 131, 132, 133
 cooking beans example, 117
 determination, 288
 flour milling experiment, 121
 one-way model, 296, 362
 SAS-mixed commands, 567
Random effects model
 algebraic method, 293–294
 case study, 365–381
 computing expected means squares, 292–306
 confidence intervals, 374–378
 data set, 365–366
 estimating variance components, 290–291
 estimation, 367
 general random effects model, 290–291
 Hartley's method of synthesis, 295–306
 inference procedures, 337
 JMP computations, 379
 in matrix notation, 290–291
 model building, 368–369
 one-way random effects model, 290–291
 random effects nested treatment structure, 289
 reduced model, 370–373
 and variance components, 287–308
Randomized complete block design (RCBD), 71, 77
 analysis of variance, 84, 85, 94, 104, 107, 108, 110, 144
 cupcake experiment, 103
- cupcake experimental unit, 107
error term, 113
structure, 79, 84, 85, 94, 103, 104, 107, 108, 113, 144
variance source, 113
whole-plot design, 105, 110
- Raw, least squares, and weighted means, 176, 177
- RCBD. *See* Randomized complete block design (RCBD)
- Reduced model, 13
 assembly line workers, 371, 375, 381
 evaluate likelihood function, 346
 proc mixed code and results
 for fitting, 346
- Reduced regression model
 covariance parameter estimates, 447
 SAS-mixed code, 447
 solution for fixed effects, 450
- REML. *See* Restricted maximum likelihood estimates (REML)
- Repeated Latin square treatment structure, 90
- Repeated measures designs, 71, 101–154
 cheese-making experiment, 131–134
 comfort experiment, 575
 complex comfort experiment, 573–582
 complex examples, 573–598
 family attitudes experiment, 583–591
 fit statistics, 567
 multilevel, 131–134
 multilocation experiment, 592–596
 multilocation study analysis of variance, 150
 vs. split-plot design, 499
 strip-plot designs, 135–141
- Repeated measures experiments
 analysis, 499–534
 complex comfort experiment, 510–515
 covariance structures, 556
 drugs effect on heart rate, 502–509
 family attitudes, 516–518
 ideal conditions not satisfied, 535–572
 layout, 500
 MANOVA methods, 537–546
 maximum likelihood method, 555
 mixed model methods, 553–567
 model specifications and ideal conditions, 500–501

restricted maximum likelihood method, 555–567

SAS-mixed procedure, 519–529

split-plot in time analyses, 502–518

- Repeated option
 covariance structure, 559
 drug heart rate study, 559
 SAS-GLM code, 552
 SAS-mixed code, 559
 sorghum varieties and fertilizer levels, 552
- Replication, 71
vs. subsample, 77
- Residual maximum likelihood. *See* Restricted maximum likelihood estimates (REML)
- Residuals, 28
- Response over time, 515
- Restricted maximum likelihood estimates (REML), 309, 326, 556
 analysis, 405, 411
 assembly line workers, 374, 375, 381
 assembly line workers covariance parameter, 367
 coffee prices example, 639
 confidence intervals, 378
 covariance, 405
 difference between machines, 414
 estimates, 392–393, 403, 411, 413
 insect damage in wheat, 349
 JMP, 331, 333
 machine person example, 402
 mean for machine, 413
 mixed models, 392–393
 nitrogen by irrigation example, 479
 proc mixed code to compute, 324
 proc mixed code to obtain, 324
 repeated measures experiments, 555–567
 SAS-mixed code, 349, 350, 403, 411
 Satterthwaite type confidence intervals, 360
 solutions, 348
 variance components, 403, 411
 wheat yield data, 428
 worker data, 361
- Restricted model, 13
- Row and column blocks example, 97–98
 analysis of variance, 98
 designing experiments, 97
- Rows
 error computed, 474
 marginal mean, 201
- S**
- Sampling distributions, 23
- SAS
 comfort study, 524
 crossover designs, 613
 family attitude study, 527–529, 587
- multicenter drug experiment, 594
 multiple comparison problems, 54
 pairwise comparisons, 52
 type I sums of squares, 303
 type II estimable functions, 224
 type II sums of squares, 303
 types I–III sums of squares, 305, 316–317
 unweighted means, 237
- SAS-GLM
 analysis, 249
 code, 35
 commands, 250
 contrast statements, 283
 crossover designs, 604, 605
 estimable functions identification, 213
 grade point average study, 178
 MANOVA, 545
 model statement, 215
 nutrition scores example, 283
 procedure, 213, 228, 249
 repeated option, 552
 SOLUTION option, 215, 216
 three-way treatment structures, 277–281
 type I analysis and type III analysis, 228
 type I sums of squares, 303
 type I estimable functions, 223
- SAS mixed code, 51
 analysis of variance, 487
 animal comparisons, 454
 comfort study, 579
 crossover designs, 606
 estimate statements, 442
 final regression model, 448
 fixed effects, 568
 full regression model, 446
 HRT_RATE, 520
 linear and quadratic trends, 442
 linearly independent comparisons, 56
 loaf volume, 426
 nitrogen by irrigation example, 478, 479
 procedure, 36
 random subject effect, 567
 reduced regression model, 447
 REML, 349, 350, 403, 411
 repeated measures experiments, 519–529
 shear force data, 454
 split-plot designs, 575
 strip-split-plot design, 482, 487
 three-way interaction term, 132
 type III sums of squares, 343
 unequal variance model, 37
 using options, 348
 worker data, 357, 361

- SAS-MULTTEST, 45, 51
Satterthwaite approximation, 30, 342, 343
assembly line workers, 373, 375
comfort study, 514
degrees of freedom, 33, 343, 509, 514, 519
drug heart rate study, 504, 509
machine person example, 404
moisture and fertilizer example, 444
one-way treatment structure, 33
sorghum varieties and fertilizer
levels, 555
strip-plot designs, 477
type I sums of squares, 377
variance components inferences
methods, 348
Satterthwaite confidence intervals, 362, 378
assembly line workers, 374
degrees of freedom, 376
machine person example, 412
REML solution, 360
variance components, 350
Scheffé's method, 45, 49, 51, 53, 55, 60, 195
simultaneous inference procedures, 48
widths of confidence intervals, 57
Schwarz's Bayesian Criterion (BIC), 557, 558,
567, 588, 593
covariance structure, 565, 566, 576, 577
Semiconductor industry, 486–493
Sequential analysis. *See* Type I analysis
Sequential rejective methods, 57–58
Shear force data, 453
SAS-mixed code, 454
split-split-plot model, 454
Šidák–Holm method, 45
simultaneous inference procedures, 58
Šidák method, 51, 55
contrast, 53, 60
simultaneous inference procedures, 51
Simple linear regression model, 156
Simultaneous inference procedures, 50
Benjamini and Hochberg method to
control FDR, 58
Bonferroni–Holm method, 58
Bonferroni's method, 48
caution, 68
comparing with control example, 54
Duncan's new multiple range method, 64
Dunnett's procedure, 53
error rates, 44
Fisher's LSD procedure, 47
least significant difference, 46
linearly independent comparisons
example, 59–60
multiple comparisons, 43–70
multiple range for pairwise comparisons
example, 66–67
multiple range tests, 61–63
multivariate t , 55
pairwise comparisons example, 51–52
proc MULTITEST code, 59
recommendations, 45
Scheffé's procedure, 48
sequential rejective methods, 57–58
Šidák–Holm method, 58
Šidák procedure, 51
Student–Newman–Keul's method, 61–63
Tukey–Kramer method, 49
Waller–Duncan procedure, 65
Single-degree-of-freedom tests, 191
Six-sequence design, 617
SNK. *See* Student–Newman–Keul's
method (SNK)
Soil study experiment, 596–597
SOLUTION option, 215, 216
Sorghum varieties and fertilizer levels,
540–555
G-G adjusted degrees of freedom, 552
leaf area index, 540
MANOVA, 545, 547, 553
matrix describing transformed variables,
546
split-plot designs, 544, 552
tests, 553
variety main effects tests, 551
Soybeans and herbicides, 448, 450–452, 487
cell means, 452
soybeans weights, 451
varieties combinations, 450
Soybeans in maturity groups example,
143–145
analysis of variance, 144
strip-plot designs, 143–144
treatment structures, 144
varieties nested maturity groups, 144
Sphericity tests, 552
Split-plot designs, 71
adjusted p -values, 553
analysis, 421–470
analysis of variance, 109, 110, 145
analysis of variance table, 115
bread baking study, 122–124
bread recipes and baking temperatures,
422–425
cheese-making experiment, 131–134
comfort experiment, 575

- Split-plot designs (*Continued*)**
- complex, 130
 - concepts, 422
 - CR whole-plot structure, 113
 - cupcake experiment, 104
 - cupcake treatment structure, 109
 - errors regressions, 444–447
 - error term and variance source, 113
 - exercises, 467–470
 - family attitude study, 584
 - general contrasts comparison, 436–439
 - hierarchical designs, 627
 - JMP computations, 459–465
 - means comparisons standard errors, 430–433
 - means standard errors of differences computing, 434–435
 - mixed-up split-plot design, 448–450
 - model definition and parameter estimation, 428–429
 - moisture and fertilizer, 440–443
 - oven design and treatment structures, 106
 - oven treatment structure, 109
 - randomization scheme, 105
 - randomized complete block whole-plot, 109
 - RCB whole-plot, 113
 - repeated measures, 627
 - vs. repeated measures design, 499
 - and repeated measures designs, 287
 - repeated measures experiments, 502–518
 - sample size and power, 455–458
 - SAS commands, 584
 - SAS-GLM procedure, 550
 - SAS-mixed code, 575
 - sorghum varieties and fertilizer levels, 544, 551, 552, 553
 - soybean varieties nested within maturity groups, 145
 - split-split-plot design, 452–454
 - and strip-plot designs, 115–130
 - teaching method study, 115
 - time analysis, 502–518, 544, 550–553, 575, 584
 - variation, 108
 - wheat varieties, 426–427
 - whole plot errors, 429
- Split-split-plot design**
- shear force data, 454
 - split-plot designs, 452–454
- SSCOLUMNS, 301**
- SSERROR, 301**
- SSINTERACTION, 301**
- SSROWS, 301**
- Standard error of comparison, 513**
- Statistical analysis objectives, 2**
- Steel plates example, 90–92**
- designing experiments, 90–91
- Stepdown Bonferroni contrast, 60**
- Stepdown Šidák contrast, 60**
- Strip-plot designs, 71, 76, 101–154**
- aircraft engines, 145–146
 - analysis, 471–498
 - analysis of variance, 111, 112, 473, 474, 486
 - animal genetics, 143
 - baking bread, 122–124
 - bread baking study, 122–124
 - change-over designs, 141
 - cheese making, 131–134
 - comfort study, 139–140
 - comparisons, estimators, and variance, 476
 - cooking beans experiment example, 116–119
 - cupcake experiment, 110, 111
 - description, 471–474
 - error term, 113
 - estimated standard errors, 477
 - experimental unit size identification, 101–113
 - grinding wheat, 120–121
 - hierarchical design, 114
 - horse feet, 136–138
 - inferences, 475–476
 - JMP, 486–490
 - meat in display case, 125–130
 - multilocation study with repeated measures, 148–149
 - nested factors, 142–150
 - nitrogen by irrigation, 477–479, 486
 - repeated measures designs, 135–141
 - Satterthwaite approximation, 477
 - simple comfort experiment, 147
 - soybeans in maturity groups, 143–144
 - split-plot 1, 480–481
 - split-plot 2, 482
 - split-plot 3, 483–484
 - split-plot 4, 485
 - split-plot design structures, 115–130
 - trial, 76
 - two-way treatment structure, 474
 - variance source, 113
- Strip-split-plot design**
- analysis of variance, 484, 485, 489
 - data, 482
 - nitrogen by irrigation example, 484, 487

- SAS-mixed code, 482
SAS-mixed code and analysis
 of variance, 487
semiconductor industry, 489
Studentized range critical point, 64
Student–Newman–Keul’s method (SNK),
 64, 66
 multiple comparison procedure, 67
 simultaneous inference procedures, 61–63
Subsample *vs.* replication, 77
Summation notation, 155
Sum of squares
 computing methods, 305
 deviations, 9
 generating, 222
 Hartley’s method of synthesis, 300
 null hypothesis deviations, 172
 testing, 9
Sum-to-zero restriction, 168
- T**
- Tasks and pulse rate example, 5–16
 analysis of variance table, 15
 equality of means, 15
one-way treatment structure, 5–6, 15
proc GLM code, 16
proc IML code, 16
pulsation data, 6
Taylor’s series approach, 351
 first-order approximation, 34
Testing hypothesis
 age, 283
 variance components techniques, 337
Three by four experiment
 orthogonal contrast coefficients, 193
 orthogonal polynomials expected
 values, 194
Three-factor experiments, 272
Three step process, 488
Three-way, two-level, nested system, 289
Three-way experiments, 274
Three-way factorial arrangement, 87
Three-way interaction term, 132
Three-way least squares means, 280
 comfort study, 580
Three-way random effect, 339–343
Three-way treatment structures, 386
 analysis, 271–276
 balanced and unbalanced
 experiments, 273
 complete analysis, 282–283
 horse foot experiment, 137
missing treatment combinations
 case study, 277–286
nutrition scores example, 277
SAS-GLM analysis, 277–281
strategy, 271–272
type I and II analyses, 273
Treatment combinations
 case study, 265–270
horse feet experiment, 138
nitrogen and potassium levels
 example, 93
two-way treatment structures,
 265–270
Treatment structure, 71
 absolute deviations values, 27
 comparison methods, 141
 diets, 83
 diets example, 87
 nitrogen and potassium levels, 92
 two-way factorial arrangement, 87
Tukey–Kramer method, 51, 65
 comparison, 454
 critical differences, 53
 HSD procedure, 66
 simultaneous inference procedures, 49
Tukey’s method, 45, 456
 animal comparisons, 455
 comfort study, 638
 HSD option, 465
 least squares means for three way
 interactions, 465
 multiple comparisons, 638
Two-level system, 289
Two-period two-treatment crossover
 design, 142
Two-way design
 cell mean parameters, 246
 estimating variance components
 methods, 315–317
 fat \times surfactant combinations, 265
 having subclass numbers, 231–238
 missing treatment combinations, 246
Two-way effects model
 estimable functions, 170
 matrix form, 159
 normal equations, 163, 166
 type I and type II sums of squares, 217
 type I hypotheses, 219
Two-way factorial arrangement, 86
 treatment structure, 87
Two-way least squares means, 252
 comfort study, 580
 family attitude study, 591

- Two-way means model
 family attitude study, 528
 normal equations, 163
- Two-way mixed model, 390
- Two-way nested treatment structure, 146
- Two-way random effects model, 315
 analysis of variance, 339
 example, 301
 matrix form, 301
 TS in CR DS, 338
 variance components inferences
 methods, 338
- Two-way treatment structure model, 81, 86, 98
 additive levels, 91
 analysis of variance, 89, 90, 92, 94, 103, 104, 109, 110, 112
 ANOVA table, 251
 blocks, 95
 computer analyses, 249–252, 262
 connected and unconnected, 171
 contrast statements, 251
 CRD structure data, 159
 cupcakes, 102
 estimable functions, 251
 example, 247–248
 exercises, 253, 263
 GPA data, 176
 grade point average data, 175
 hypothesis testing and confidence intervals, 247–248
 matrix form of model, 158–160
 means, 158–160
 missing treatment combinations
 case study, 265–270
 missing treatment combinations effects
 model analysis, 255–263
 missing treatment combinations means
 model analysis, 245–253
 model statement solution options, 252
 parameter estimation, 245–246
 population marginal means and least squares means, 261
 randomization, 102, 103
 strip-plot designs, 474
 12 treatment combinations, 81
 type I and II hypothesis, 255–256
 type III hypothesis, 257–258
 type IV hypothesis, 259–260
- Type I analysis, 213, 409
 analysis of variance, 210, 213
 assembly line workers, 367, 371, 377
- assembly line workers covariance parameter, 367
- difference between machines, 414
- estimates of mean for machine, 413
- estimates of variance components, 411
- expected means squares and error terms, 344
- hypotheses, 219
- machine person example, 409
- test statistic for interaction, 213
- three-way treatment structures, 273
- Type I hypothesis
 assembly line workers, 369
 two-way treatment structures, 255–256
- Type I sums of squares
 analysis of variance, 342
 assembly line workers, 375
 coffee prices example, 633
 definitions, 217
 each column, 304
 proc mixed code, 342
 SAS code to evaluate, 303
 of SAS-GLM, 303
 SAS-mixed code, 343
 sums, 304
- Type II analysis, 218
 assembly line workers covariance parameter, 367
- difference between machines, 414
- estimable functions, 225
- estimates of mean for machine, 413
- estimates of variance components, 411
- hypotheses, 219, 220
- means model hypotheses, 218
- three-way treatment structures, 273
- Type II hypothesis, 255–256
- Type II sums of squares
 definitions, 217
 each column, 304
 SAS code to evaluate, 303
 SAS-GLM, 303
 sums, 304
- Type II tests, 525
- Type III analysis, 220, 221
 assembly line workers, 368, 371
 assembly line workers covariance parameter, 367
- covariance parameter estimates, 405
- crossover designs, 605
- difference between machines, 414
- estimable functions, 225
- estimates of mean for machine, 413

- estimates of variance components, 403, 411
machine person example, 405
means model, 258
Type III hypothesis
 effects models, 221, 258
 two-way treatment structures, 257–258
Type III sums of squares
 analysis of variance, 403
 assembly line workers, 373
 each column, 304
 machine person example, 403
 nitrogen by irrigation example, 478
 SAS code to evaluate, 303
 of SAS-GLM, 303
 SAS-mixed code, 343
 sums, 304
 wheat yield data, 428
Type III tests
 assembly line workers, 370, 373
 comfort study, 578
 crossover designs, 608, 615
 crossover designs six sequences, 620
 family attitude study, 528, 590
Type IV analysis of variance, 261
 nutrition scores example, 278
Type IV hypothesis
 fat-surfactant example, 266, 267
 nutrition scores example, 282
 SAS-GLM tests, 259, 279
 two-way treatment structures, 259–260
Type IV sums of squares, 303
- U**
- Unbalanced incomplete block design
 structures, 95
Unbalanced one-way design, 327–328
Unbalanced one-way model, 309–312
Unbalanced two-way, 347
 analysis of variance, 201, 204
 cell means and marginal means, 201
 data set JMP analysis, 415–416
 mixed models, 409–414, 415–416
Unbounded variance components, 333
Unconnected incomplete block design
 structure, 96
Unequal variance model, 22
 drug group means, 38
 fitting results, 37
 pair wise comparisons, 38
 SAS-mixed code, 37
- Unrestricted model, 13
Unstructured case, 556
 estimated R matrix, 561
Unweighted means
 analysis of variance, 233, 236, 237
 method, 231
 SAS code, 237
- V**
- Variance components, 287–308
 analysis, 14
 assembly line workers, 369
 procedures for testing equality, 22
Variance components inferences
 methods, 337–364
 analysis of variance table, 338
 approximate confidence interval, 349–352
 balanced one-way random effects
 treatment structure, 356
 character, 21
 complex three-way random effect TS,
 339–343
 confidence intervals construction,
 347–360
 exact confidence intervals, 353–355
 exercises, 363–364
 general Satterthwaite approximation, 348
 hypotheses testing, 337–346
 likelihood ratio test, 344
 one-way random effects model, 345–346
 residual variance, 348
 simulation study, 361
 two-way random effects TS in a
 CR DS, 338
 unbalanced two-way, 347
 Wald-type confidence intervals, 353
- W**
- Wald confidence intervals, 330
 assembly line workers, 375
 machine person example, 412
Waller–Duncan methods, 66
 multiple comparison procedure, 68
 simultaneous inference procedures, 65
Wash-out period, 607
Weighted means, 176, 177
Welch test, 34
 computing, 35
 equal means, 35
 statistic, 35

Wheat varieties example, 313–315, 345–347, 426–428, 459–465, 480–482
analysis of variance, 427
insect damage computations, 314
JMP computations, 459–465
type III sums of squares, 428
Whole-plot method
flour milling experiment, 121
grain-milling experiment, 120
IBD structures, 123–125
Whole-plot or row design structure, 119

Wilks' likelihood ratio criterion, 538
Williams' interval, 356
crossover designs, 617
Workers in assembly line. *See* Assembly line workers
Y
Yates' method, 221
weighted squares of means technique, 220

ANALYSIS OF MESSY DATA

VOLUME 1

DESIGNED EXPERIMENTS

SECOND EDITION

A bestseller for nearly 25 years, **Analysis of Messy Data, Volume 1: Designed Experiments** helps you analyze the kinds of data sets encountered in the real world. Written by two long-time researchers and professors, this second edition has been fully updated to reflect the many developments that have occurred since the original publication.

New to the Second Edition

- Several modern suggestions for multiple comparison procedures
- Additional examples of split-plot designs and repeated measures designs
- The use of SAS-GLM to analyze an effects model
- The use of SAS-MIXED and JMP to analyze data in random effects experiments, mixed model experiments, and repeated measures experiments

The book explores various techniques for multiple comparison procedures, random effects models, mixed models, split-plot experiments, and repeated measures designs. The authors implement the techniques using several statistical software packages and emphasize the distinction between design structure and the structure of treatments. They introduce each topic with examples, follow up with a theoretical discussion, and conclude with a case study. Bringing a classic work up to date, this edition will continue to show you how to effectively analyze real-world, nonstandard data sets.

C3340

