



An Interactive e-Text

APPLIED STATISTICS AND PROBABILITY FOR ENGINEERS

THIRD EDITION

DOUGLAS C. MONTGOMERY
GEORGE C. RUNGER

Applied Statistics and Probability for Engineers

Third Edition

Douglas C. Montgomery
Arizona State University

George C. Runger
Arizona State University



John Wiley & Sons, Inc.

ACQUISITIONS EDITOR *Wayne Anderson*
ASSISTANT EDITOR *Jenny Welter*
MARKETING MANAGER *Katherine Hepburn*
SENIOR PRODUCTION EDITOR *Norine M. Pigliucci*
DESIGN DIRECTOR *Maddy Lesure*
ILLUSTRATION EDITOR *Gene Aiello*
PRODUCTION MANAGEMENT SERVICES *TechBooks*

This book was set in Times Roman by TechBooks and printed and bound by Donnelley/Willard.
The cover was printed by Phoenix Color Corp.

This book is printed on acid-free paper. (∞)

Copyright 2003 © John Wiley & Sons, Inc. All rights reserved.

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 605 Third Avenue, New York, NY 10158-0012, (212) 850-6011, fax (212) 850-6008, E-Mail: PERMREQ@WILEY.COM. To order books please call 1(800)-225-5945.

Library of Congress Cataloging-in-Publication Data

Montgomery, Douglas C.

Applied statistics and probability for engineers / Douglas C. Montgomery, George C. Runger.—3rd ed.

p. cm.

Includes bibliographical references and index.

ISBN 0-471-20454-4 (acid-free paper)

1. Statistics. 2. Probabilities. I. Runger, George C. II. Title.

QA276.12.M645 2002

519.5—dc21

2002016765

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

To:

Meredith, Neil, Colin, and Cheryl

Rebecca, Elisa, George, and Taylor

Preface

This is an introductory textbook for a first course in applied statistics and probability for undergraduate students in engineering and the physical or chemical sciences. These individuals play a significant role in designing and developing new products and manufacturing systems and processes, and they also improve existing systems. Statistical methods are an important tool in these activities because they provide the engineer with both descriptive and analytical methods for dealing with the variability in observed data. Although many of the methods we present are fundamental to statistical analysis in other disciplines, such as business and management, the life sciences, and the social sciences, we have elected to focus on an engineering-oriented audience. We believe that this approach will best serve students in engineering and the chemical/physical sciences and will allow them to concentrate on the many applications of statistics in these disciplines. We have worked hard to ensure that our examples and exercises are engineering- and science-based, and in almost all cases we have used examples of real data—either taken from a published source or based on our consulting experiences.

We believe that engineers in all disciplines should take at least one course in statistics. Unfortunately, because of other requirements, most engineers will only take one statistics course. This book can be used for a single course, although we have provided enough material for two courses in the hope that more students will see the important applications of statistics in their everyday work and elect a second course. We believe that this book will also serve as a useful reference.

ORGANIZATION OF THE BOOK

We have retained the relatively modest mathematical level of the first two editions. We have found that engineering students who have completed one or two semesters of calculus should have no difficulty reading almost all of the text. It is our intent to give the reader an understanding of the methodology and how to apply it, not the mathematical theory. We have made many enhancements in this edition, including reorganizing and rewriting major portions of the book.

Perhaps the most common criticism of engineering statistics texts is that they are too long. Both instructors and students complain that it is impossible to cover all of the topics in the book in one or even two terms. For authors, this is a serious issue because there is great variety in both the content and level of these courses, and the decisions about what material to delete without limiting the value of the text are not easy. After struggling with these issues, we decided to divide the text into two components; a set of core topics, many of which are most

likely to be covered in an engineering statistics course, and a set of supplementary topics, or topics that will be useful for some but not all courses. The core topics are in the printed book, and the complete text (both core and supplementary topics) is available on the CD that is included with the printed book. Decisions about topics to include in print and which to include only on the CD were made based on the results of a recent survey of instructors.

The *Interactive e-Text* consists of the complete text and a wealth of additional material and features. The text and links on the CD are navigated using Adobe Acrobat™. The links within the *Interactive e-Text* include the following: (1) from the Table of Contents to the selected *eText* sections, (2) from the Index to the selected topic within the *e-Text*, (3) from reference to a figure, table, or equation in one section to the actual figure, table, or equation in another section (all figures can be enlarged and printed), (4) from end-of-chapter Important Terms and Concepts to their definitions within the chapter, (5) from in-text **boldfaced terms** to their corresponding Glossary definitions and explanations, (6) from in-text references to the corresponding Appendix tables and charts, (7) from boxed-number end-of-chapter exercises (essentially most odd-numbered exercises) to their answers, (8) from some answers to the complete problem solution, and (9) from the opening splash screen to the textbook Web site.

Chapter 1 is an introduction to the field of statistics and how engineers use statistical methodology as part of the engineering problem-solving process. This chapter also introduces the reader to some engineering applications of statistics, including building empirical models, designing engineering experiments, and monitoring manufacturing processes. These topics are discussed in more depth in subsequent chapters.

Chapters 2, 3, 4, and 5 cover the basic concepts of probability, discrete and continuous random variables, probability distributions, expected values, joint probability distributions, and independence. We have given a reasonably complete treatment of these topics but have avoided many of the mathematical or more theoretical details.

Chapter 6 begins the treatment of statistical methods with random sampling; data summary and description techniques, including stem-and-leaf plots, histograms, box plots, and probability plotting; and several types of time series plots. Chapter 7 discusses point estimation of parameters. This chapter also introduces some of the important properties of estimators, the method of maximum likelihood, the method of moments, sampling distributions, and the central limit theorem.

Chapter 8 discusses interval estimation for a single sample. Topics included are confidence intervals for means, variances or standard deviations, and proportions and prediction and tolerance intervals. Chapter 9 discusses hypothesis tests for a single sample. Chapter 10 presents tests and confidence intervals for two samples. This material has been extensively rewritten and reorganized. There is detailed information and examples of methods for determining appropriate sample sizes. We want the student to become familiar with how these techniques are used to solve real-world engineering problems and to get some understanding of the concepts behind them. We give a logical, heuristic development of the procedures, rather than a formal mathematical one.

Chapters 11 and 12 present simple and multiple linear regression. We use matrix algebra throughout the multiple regression material (Chapter 12) because it is the only easy way to understand the concepts presented. Scalar arithmetic presentations of multiple regression are awkward at best, and we have found that undergraduate engineers are exposed to enough matrix algebra to understand the presentation of this material.

Chapters 13 and 14 deal with single- and multifactor experiments, respectively. The notions of randomization, blocking, factorial designs, interactions, graphical data analysis, and fractional factorials are emphasized. Chapter 15 gives a brief introduction to the methods and applications of nonparametric statistics, and Chapter 16 introduces statistical quality control, emphasizing the control chart and the fundamentals of statistical process control.

Each chapter has an extensive collection of exercises, including end-of-section exercises that emphasize the material in that section, supplemental exercises at the end of the chapter that cover the scope of chapter topics, and mind-expanding exercises that often require the student to extend the text material somewhat or to apply it in a novel situation. As noted above, answers are provided to most odd-numbered exercises and the *e-Text* contains complete solutions to selected exercises.

USING THE BOOK

This is a very flexible textbook because instructors' ideas about what should be in a first course on statistics for engineers vary widely, as do the abilities of different groups of students. Therefore, we hesitate to give too much advice but will explain how we use the book.

We believe that a first course in statistics for engineers should be primarily an applied statistics course, not a probability course. In our one-semester course we cover all of Chapter 1 (in one or two lectures); overview the material on probability, putting most of the emphasis on the normal distribution (six to eight lectures); discuss most of Chapters 6 through 10 on confidence intervals and tests (twelve to fourteen lectures); introduce regression models in Chapter 11 (four lectures); give an introduction to the design of experiments from Chapters 13 and 14 (six lectures); and present the basic concepts of statistical process control, including the Shewhart control chart from Chapter 16 (four lectures). This leaves about three to four periods for exams and review. Let us emphasize that the purpose of this course is to introduce engineers to how statistics can be used to solve real-world engineering problems, not to weed out the less mathematically gifted students. This course is not the "baby math-stat" course that is all too often given to engineers.

If a second semester is available, it is possible to cover the entire book, including much of the *e-Text* material, if appropriate for the audience. It would also be possible to assign and work many of the homework problems in class to reinforce the understanding of the concepts. Obviously, multiple regression and more design of experiments would be major topics in a second course.

USING THE COMPUTER

In practice, engineers use computers to apply statistical methods to solve problems. Therefore, we strongly recommend that the computer be integrated into the class. Throughout the book we have presented output from Minitab as typical examples of what can be done with modern statistical software. In teaching, we have used other software packages, including Statgraphics, JMP, and Statistica. We did not clutter up the book with examples from many different packages because how the instructor integrates the software into the class is ultimately more important than which package is used. All text data is available in electronic form on the *e-Text* CD. In some chapters, there are problems that we feel should be worked using computer software. We have marked these problems with a special icon in the margin.



In our own classrooms, we use the computer in almost every lecture and demonstrate how the technique is implemented in software as soon as it is discussed in the lecture. Student versions of many statistical software packages are available at low cost, and students can either purchase their own copy or use the products available on the PC local area networks. We have found that this greatly improves the pace of the course and student understanding of the material.

USING THE WEB

Additional resources for students and instructors can be found at www.wiley.com/college/montgomery/.

ACKNOWLEDGMENTS

We would like to express our grateful appreciation to the many organizations and individuals who have contributed to this book. Many instructors who used the first two editions provided excellent suggestions that we have tried to incorporate in this revision. We also thank Professors Manuel D. Rossetti (University of Arkansas), Bruce Schmeiser (Purdue University), Michael G. Akritas (Penn State University), and Arunkumar Pennathur (University of Texas at El Paso) for their insightful reviews of the manuscript of the third edition. We are also indebted to Dr. Smiley Cheng for permission to adapt many of the statistical tables from his excellent book (with Dr. James Fu), *Statistical Tables for Classroom and Exam Room*. John Wiley and Sons, Prentice Hall, the Institute of Mathematical Statistics, and the editors of Biometrics allowed us to use copyrighted material, for which we are grateful. Thanks are also due to Dr. Lora Zimmer, Dr. Connie Borrer, and Dr. Alejandro Heredia-Langner for their outstanding work on the solutions to exercises.

Douglas C. Montgomery

George C. Runger

Contents

CHAPTER 1 The Role of Statistics in Engineering 1

- 1-1 The Engineering Method and Statistical Thinking 2
- 1-2 Collecting Engineering Data 5
 - 1-2.1 Basic Principles 5
 - 1-2.2 Retrospective Study 5
 - 1-2.3 Observational Study 6
 - 1-2.4 Designed Experiments 6
 - 1-2.5 A Factorial Experiment for the Connector Pull-Off Force Problem (CD Only) 8**
 - 1-2.6 Observing Processes Over Time 8
- 1-3 Mechanistic and Empirical Models 11
- 1-4 Probability and Probability Models 14

CHAPTER 2 Probability 16

- 2-1 Sample Spaces and Events 17
 - 2-1.1 Random Experiments 17
 - 2-1.2 Sample Spaces 18
 - 2-1.3 Events 22
 - 2-1.4 Counting Techniques (CD Only) 25**
- 2-2 Interpretations of Probability 27
 - 2-2.1 Introduction 27
 - 2-2.2 Axioms of Probability 30
- 2-3 Addition Rules 33
- 2-4 Conditional Probability 37
- 2-5 Multiplication and Total Probability Rules 42
 - 2-5.1 Multiplication Rule 42
 - 2-5.2 Total Probability Rule 43
- 2-6 Independence 46
- 2-7 Bayes' Theorem 51
- 2-8 Random Variables 53

CHAPTER 3 Discrete Random Variables and Probability Distributions 59

- 3-1 Discrete Random Variables 60
- 3-2 Probability Distributions and Probability Mass Functions 61

- 3-3 Cumulative Distribution Functions 63
- 3-4 Mean and Variance of a Discrete Random Variable 66
- 3-5 Discrete Uniform Distribution 70
- 3-6 Binomial Distribution 72
- 3-7 Geometric and Negative Binomial Distributions 78
 - 3-7.1 Geometric Distribution 78
 - 3-7.2 Negative Binomial Distribution 80
- 3-8 Hypergeometric Distribution 84
- 3-9 Poisson Distribution 89

CHAPTER 4 Continuous Random Variables and Probability Distributions 97

- 4-1 Continuous Random Variables 98
- 4-2 Probability Distributions and Probability Density Functions 98
- 4-3 Cumulative Distribution Functions 102
- 4-4 Mean and Variance of a Continuous Random Variable 105
- 4-5 Continuous Uniform Distribution 107
- 4-6 Normal Distribution 109
- 4-7 Normal Approximation to the Binomial and Poisson Distributions 118
- 4-8 Continuity Corrections to Improve the Approximation (CD Only) 122**
- 4-9 Exponential Distribution 122
- 4-10 Erlang and Gamma Distribution 128
 - 4-10.1 Erlang Distribution 128
 - 4-10.2 Gamma Distribution 130
- 4-11 Weibull Distribution 133
- 4-12 Lognormal Distribution 135

CHAPTER 5 Joint Probability Distributions 141

-
- 5-1 Two Discrete Random Variables 142
 - 5-1.1 Joint Probability Distributions 142
 - 5-1.2 Marginal Probability Distributions 144
 - 5-1.3 Conditional Probability Distributions 146
 - 5-1.4 Independence 148
 - 5-2 Multiple Discrete Random Variables 151
 - 5-2.1 Joint Probability Distributions 151
 - 5-2.2 Multinomial Probability Distribution 154
 - 5-3 Two Continuous Random Variables 157
 - 5-3.1 Joint Probability Distributions 157
 - 5-3.2 Marginal Probability Distributions 159
 - 5-3.3 Conditional Probability Distributions 162
 - 5-3.4 Independence 164
 - 5-4 Multiple Continuous Random Variables 167
 - 5-5 Covariance and Correlation 171
 - 5-6 Bivariate Normal Distribution 177
 - 5-7 Linear Combinations of Random Variables 180
 - 5-8 Functions of Random Variables (CD Only) 185**
 - 5-9 Moment Generating Functions (CD Only) 185**
 - 5-10 Chebyshev's Inequality (CD Only) 185**

CHAPTER 6 Random Sampling and Data Description 189

-
- 6-1 Data Summary and Display 190
 - 6-2 Random Sampling 195
 - 6-3 Stem-and-Leaf Diagrams 197
 - 6-4 Frequency Distributions and Histograms 203
 - 6-5 Box Plots 207
 - 6-6 Time Sequence Plots 209
 - 6-7 Probability Plots 212
 - 6-8 More About Probability Plotting (CD Only) 216**

CHAPTER 7 Point Estimation of Parameters 220

-
- 7-1 Introduction 221
 - 7-2 General Concepts of Point Estimation 222
 - 7-2.1 Unbiased Estimators 222
 - 7-2.2 Proof that \bar{S} is a Biased Estimator of σ (CD Only) 224**
 - 7-2.3 Variance of a Point Estimator 224
 - 7-2.4 Standard Error: Reporting a Point Estimator 225
 - 7-2.5 Bootstrap Estimate of the Standard Error (CD Only) 226**
 - 7-2.6 Mean Square Error of an Estimator 226
 - 7-3 Methods of Point Estimation 229
 - 7-3.1 Method of Moments 229
 - 7-3.2 Method of Maximum Likelihood 230
 - 7-3.3 Bayesian Estimation of Parameters (CD Only) 237**
 - 7-4 Sampling Distributions 238
 - 7-5 Sampling Distribution of Means 239

CHAPTER 8 Statistical Intervals for a Single Sample 247

-
- 8-1 Introduction 248
 - 8-2 Confidence Interval on the Mean of a Normal Distribution, Variance Known 249
 - 8-2.1 Development of the Confidence Interval and Its Basic Properties 249
 - 8-2.2 Choice of Sample Size 252
 - 8-2.3 One-sided Confidence Bounds 253
 - 8-2.4 General method to Derive a Confidence Interval 253
 - 8-2.5 A Large-Sample Confidence Interval for μ 254
 - 8-2.6 Bootstrap Confidence Intervals (CD Only) 256**
 - 8-3 Confidence Interval on the Mean of a Normal Distribution, Variance Unknown 257
 - 8-3.1 The t Distribution 258
 - 8-3.2 Development of the t Distribution (CD Only) 259**
 - 8-3.3 The t Confidence Interval on μ 259

- 8-4 Confidence Interval on the Variance and Standard Deviation of a Normal Distribution 261
- 8-5 A Large-Sample Confidence Interval for a Population Proportion 265
- 8-6 A Prediction Interval for a Future Observation 268
- 8-7 Tolerance Intervals for a Normal Distribution 270

CHAPTER 9 Tests of Hypotheses for a Single Sample 277

-
- 9-1 Hypothesis Testing 278
 - 9-1.1 Statistical Hypotheses 278
 - 9-1.2 Tests of Statistical Hypotheses 280
 - 9-1.3 One-Sided and Two-Sided Hypotheses 286
 - 9-1.4 General Procedure for Hypothesis Testing 287
 - 9-2 Tests on the Mean of a Normal Distribution, Variance Known 289
 - 9-2.1 Hypothesis Tests on the Mean 289
 - 9-2.2 *P*-Values in Hypothesis Tests 292
 - 9-2.3 Connection Between Hypothesis Tests and Confidence Intervals 293
 - 9-2.4 Type II Error and Choice of Sample Size 293
 - 9-2.5 Large Sample Test 297
 - 9-2.6 Some Practical Comments on Hypothesis Tests 298
 - 9-3 Tests on the Mean of a Normal Distribution, Variance Unknown 300
 - 9-3.1 Hypothesis Tests on the Mean 300
 - 9-3.2 *P*-Value for a *t*-Test 303
 - 9-3.3 Choice of Sample Size 304
 - 9-3.4 Likelihood Ratio Approach to Development of Test Procedures (CD Only) 305**
 - 9-4 Tests on the Variance and Standard Deviation of a Normal Distribution 307
 - 9-4.1 The Hypothesis Testing Procedures 307
 - 9-4.2 β -Error and Choice of Sample Size 309

- 9-5 Tests on a Population Proportion 310
 - 9-5.1 Large-Sample Tests on a Proportion 310
 - 9-5.2 Small-Sample Tests on a Proportion (CD Only) 312**
 - 9-5.3 Type II Error and Choice of Sample Size 312
- 9-6 Summary of Inference Procedures for a Single Sample 315
- 9-7 Testing for Goodness of Fit 315
- 9-8 Contingency Table Tests 320

CHAPTER 10 Statistical Inference for Two Samples 327

-
- 10-1 Introduction 328
 - 10-2 Inference For a Difference in Means of Two Normal Distributions, Variances Known 328
 - 10-2.1 Hypothesis Tests for a Difference in Means, Variances Known 329
 - 10-2.2 Choice of Sample Size 331
 - 10-2.3 Identifying Cause and Effect 333
 - 10-2.4 Confidence Interval on a Difference in Means, Variances Known 334
 - 10-3 Inference For a Difference in Means of Two Normal Distributions, Variances Unknown 337
 - 10-3.1 Hypothesis Tests for a Difference in Means, Variances Unknown 337
 - 10-3.2 More About the Equal Variance Assumption (CD Only) 344**
 - 10-3.3 Choice of Sample Size 344
 - 10-3.4 Confidence Interval on a Difference in Means, Variances Unknown 345
 - 10-4 Paired *t*-Test 349
 - 10-5 Inference on the Variances of Two Normal Distributions 355
 - 10-5.1 The *F* Distribution 355
 - 10-5.2 Development of the *F* Distribution (CD Only) 357**
 - 10-5.3 Hypothesis Tests on the Ratio of Two Variances 357
 - 10-5.4 β -Error and Choice of Sample Size 359
 - 10-5.5 Confidence Interval on the Ratio of Two Variances 359

10-6	Inference on Two Population Proportions	361
10-6.1	Large-Sample Test for $H_0: p_1 = p_2$	361
	10-6.2 Small Sample Test for $H_0: p_1 = p_2$ (CD Only)	364
10-6.3	β -Error and Choice of Sample Size	364
10-6.4	Confidence Interval for $P_1 - P_2$	365
10-7	Summary Table for Inference Procedures for Two Samples	367

CHAPTER 11 Simple Linear Regression and Correlation 372

11-1	Empirical Models	373
11-2	Simple Linear Regression	375
11-3	Properties of the Least Squares Estimators	383
	11-4 Some Comments on Uses of Regression (CD Only)	384
11-5	Hypothesis Tests in Simple Linear Regression	384
11-5.1	Use of t -Tests	384
11-5.2	Analysis of Variance Approach to Test Significance of Regression	387
11-6	Confidence Intervals	389
11-6.1	Confidence Intervals on the Slope and Intercept	389
11-6.2	Confidence Interval on the Mean Response	390
11-7	Prediction of New Observations	392
11-8	Adequacy of the Regression Model	395
11-8.1	Residual Analysis	395
11-8.2	Coefficient of Determination (R^2)	397
	11-8.3 Lack-of-Fit Test (CD Only)	398
11-9	Transformations to a Straight Line	400
	11-10 More About Transformations (CD Only)	400
11-11	Correlation	400

CHAPTER 12 Multiple Linear Regression 410

12-1	Multiple Linear Regression Model	411
------	----------------------------------	-----

12-1.1	Introduction	411
12-1.2	Least Squares Estimation of the Parameters	414
12-1.3	Matrix Approach to Multiple Linear Regression	417
12-1.4	Properties of the Least Squares Estimators	421
12-2	Hypothesis Tests in Multiple Linear Regression	428
12-2.1	Test for Significance of Regression	428
12-2.2	Tests on Individual Regression Coefficients and Subsets of Coefficients	432
	12-2.3 More About the Extra Sum of Squares Method (CD Only)	435
12-3	Confidence Intervals in Multiple Linear Regression	437
12-3.1	Confidence Intervals on Individual Regression Coefficients	437
12-3.2	Confidence Interval on the Mean Response	438
12-4	Prediction of New Observations	439
12-5	Model Adequacy Checking	441
12-5.1	Residual Analysis	441
12-5.2	Influential Observations	444
12-6	Aspects of Multiple Regression Modeling	447
12-6.1	Polynomial Regression Models	447
12-6.2	Categorical Regressors and Indicator Variables	450
12-6.3	Selection of Variables and Model Building	452
12-6.4	Multicollinearity	460
	12-6.5 Ridge Regression (CD Only)	461
	12-6.6 Nonlinear Regression Models (CD Only)	461

CHAPTER 13 Design and Analysis of Single-Factor Experiments: The Analysis of Variance 468

13-1	Designing Engineering Experiments	469
13-2	The Completely Randomized Single-Factor Experiment	470
13-2.1	An Example	470
13-2.2	The Analysis of Variance	472
13-2.3	Multiple Comparisons Following the ANOVA	479

13-2.4	More About Multiple Comparisons (CD Only)	481
13-2.5	Residual Analysis and Model Checking	481
13-2.6	Determining Sample Size	482
13-2.7	Technical Details about the Analysis of Variance (CD Only)	485
13-3	The Random Effects Model	487
13-3.1	Fixed Versus Random Factors	487
13-3.2	ANOVA and Variance Components	487
13-3.3	Determining Sample Size in the Random Model (CD Only)	490
13-4	Randomized Complete Block Design	491
13-4.1	Design and Statistical Analysis	491
13-4.2	Multiple Comparisons	497
13-4.3	Residual Analysis and Model Checking	498
13-4.4	Randomized Complete Block Design with Random Factors (CD Only)	498
 CHAPTER 14 Design of Experiments with Several Factors 505		
14-1	Introduction	506
14-2	Some Applications of Designed Experiments (CD Only)	506
14-3	Factorial Experiments	506
14-4	Two-Factor Factorial Experiments	510
14-4.1	Statistical Analysis of the Fixed-Effects Model	511
14-4.2	Model Adequacy Checking	517
14-4.3	One Observation Per Cell	517
14-4.4	Factorial Experiments with Random Factors: Overview	518
14-5	General Factorial Experiments	520
14-6	Factorial Experiments with Random Factors (CD Only)	523
14-7	2^k Factorial Designs	523
14-7.1	2^2 Design	524
14-7.2	2^k Design for $k \geq 3$ Factors	529
14-7.3	Single Replicate of the 2^k Design	537
14-7.4	Addition of Center Points to a 2^k Design (CD Only)	541

14-8	Blocking and Confounding in the 2^k Design	543
14-9	Fractional Replication of the 2^k Design	549
14-9.1	One Half Fraction of the 2^k Design	549
14-9.2	Smaller Fractions: The 2^{k-p} Fractional Factorial	555
 14-10 Response Surface Methods and Designs (CD Only) 564		
 CHAPTER 15 Nonparametric Statistics 571		
15-1	Introduction	572
15-2	Sign Test	572
15-2.1	Description of the Test	572
15-2.2	Sign Test for Paired Samples	576
15-2.3	Type II Error for the Sign Test	578
15-2.4	Comparison to the t -Test	579
15-3	Wilcoxon Signed-Rank Test	581
15-3.1	Description of the Test	581
15-3.2	Large-Sample Approximation	583
15-3.3	Paired Observations	583
15-3.4	Comparison to the t -Test	584
15-4	Wilcoxon Rank-Sum Test	585
15-4.1	Description of the Test	585
15-4.2	Large-Sample Approximation	587
15-4.3	Comparison to the t -Test	588
15-5	Nonparametric Methods in the Analysis of Variance	589
15-5.1	Kruskal-Wallis Test	589
15-5.2	Rank Transformation	591
 CHAPTER 16 Statistical Quality Control 595		
16-1	Quality Improvement and Statistics	596
16-2	Statistical Quality Control	597
16-3	Statistical Process Control	597
16-4	Introduction to Control Charts	598
16-4.1	Basic Principles	598
16-4.2	Design of a Control Chart	602
16-4.3	Rational Subgroups	603
16-4.4	Analysis of Patterns on Control Charts	604
16-5	\bar{X} and R or S Control Chart	607
16-6	Control Charts for Individual Measurements	615

16-7	Process Capability	619
16-8	Attribute Control Charts	625
16-8.1	<i>P</i> Chart (Control Chart for Proportion)	625
16-8.2	<i>U</i> Chart (Control Chart for Defects per Unit)	627
16-9	Control Chart Performance	630
16-10	Cumulative Sum Control Chart	632
16-11	Other SPC Problem-Solving Tools	639
16-12	Implementing SPC	641

APPENDICES 649

APPENDIX A: Statistical Tables and Charts 651

Table I	Summary of Common Probability Distributions	652
Table II	Cumulative Standard Normal Distribution	653
Table III	Percentage Points $\chi^2_{\alpha, \nu}$ of the Chi-Squared Distribution	655
Table IV	Percentage Points $t_{\alpha, \nu}$ of the <i>t</i> -distribution	656

Table V	Percentage Points f_{α, ν_1, ν_2} of the <i>F</i> -distribution	657
---------	----------------------------------------------------------------------------	-----

Chart VI	Operating Characteristic Curves	662
----------	---------------------------------	-----

Table VII	Critical Values for the Sign Test	671
-----------	-----------------------------------	-----

Table VIII	Critical Values for the Wilcoxon Signed-Rank Test	671
------------	---------------------------------------------------	-----

Table IX	Critical Values for the Wilcoxon Rank-Sum Test	672
----------	------------------------------------------------	-----

Table X	Factors for Constructing Variables Control Charts	673
---------	---------------------------------------------------	-----

Table XI	Factors for Tolerance Intervals	674
----------	---------------------------------	-----

APPENDIX B: Bibliography 677

APPENDIX C: Answers to Selected Exercises 679

GLOSSARY 689

INDEX 703

PROBLEM SOLUTIONS

1

The Role of Statistics in Engineering

CHAPTER OUTLINE

- | | |
|-----------------------------------------------------|-----------------------------------------------------------------------|
| 1-1 THE ENGINEERING METHOD AND STATISTICAL THINKING | 1-2.5 A Factorial Experiment for the Pull-off Force Problem (CD Only) |
| 1-2 COLLECTING ENGINEERING DATA | 1-2.6 Observing Processes Over Time |
| 1-2.1 Basic Principles | 1-3 MECHANISTIC AND EMPIRICAL MODELS |
| 1-2.2 Retrospective Study | 1-4 PROBABILITY AND PROBABILITY MODELS |
| 1-2.3 Observational Study | |
| 1-2.4 Designed Experiments | |
-

LEARNING OBJECTIVES

After careful study of this chapter you should be able to do the following:

1. Identify the role that statistics can play in the engineering problem-solving process
2. Discuss how variability affects the data collected and used for making engineering decisions
3. Explain the difference between enumerative and analytical studies
4. Discuss the different methods that engineers use to collect data
5. Identify the advantages that designed experiments have in comparison to other methods of collecting engineering data
6. Explain the differences between mechanistic models and empirical models
7. Discuss how probability and probability models are used in engineering and science

CD MATERIAL

8. Explain the factorial experimental design.
 9. Explain how factors can Interact.
-

Answers for most odd numbered exercises are at the end of the book. Answers to exercises whose numbers are surrounded by a box can be accessed in the e-Text by clicking on the box. Complete worked solutions to certain exercises are also available in the e-Text. These are indicated in the Answers to Selected Exercises section by a box around the exercise number. Exercises are also

available for some of the text sections that appear on CD only. These exercises may be found within the e-Text immediately following the section they accompany.

1-1 THE ENGINEERING METHOD AND STATISTICAL THINKING

An engineer is someone who solves problems of interest to society by the efficient application of scientific principles. Engineers accomplish this by either refining an existing product or process or by designing a new product or process that meets customers' needs. The **engineering, or scientific, method** is the approach to formulating and solving these problems. The steps in the engineering method are as follows:

1. Develop a clear and concise description of the problem.
2. Identify, at least tentatively, the important factors that affect this problem or that may play a role in its solution.
3. Propose a model for the problem, using scientific or engineering knowledge of the phenomenon being studied. State any limitations or assumptions of the model.
4. Conduct appropriate experiments and collect data to test or validate the tentative model or conclusions made in steps 2 and 3.
5. Refine the model on the basis of the observed data.
6. Manipulate the model to assist in developing a solution to the problem.
7. Conduct an appropriate experiment to confirm that the proposed solution to the problem is both effective and efficient.
8. Draw conclusions or make recommendations based on the problem solution.

The steps in the engineering method are shown in Fig. 1-1. Notice that the engineering method features a strong interplay between the problem, the factors that may influence its solution, a model of the phenomenon, and experimentation to verify the adequacy of the model and the proposed solution to the problem. Steps 2–4 in Fig. 1-1 are enclosed in a box, indicating that several cycles or iterations of these steps may be required to obtain the final solution. Consequently, engineers must know how to efficiently plan experiments, collect data, analyze and interpret the data, and understand how the observed data are related to the model they have proposed for the problem under study.

The field of **statistics** deals with the collection, presentation, analysis, and use of data to make decisions, solve problems, and design products and processes. Because many aspects of engineering practice involve working with data, obviously some knowledge of statistics is important to any engineer. Specifically, statistical techniques can be a powerful aid in designing new products and systems, improving existing designs, and designing, developing, and improving production processes.

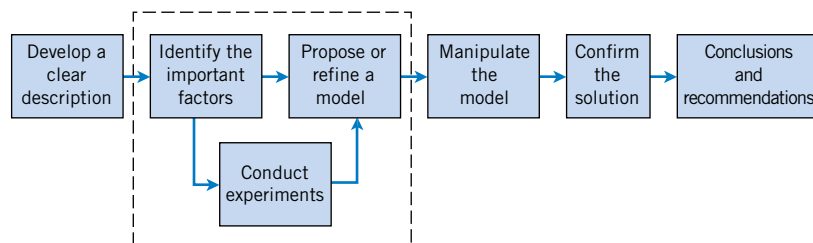


Figure 1-1 The engineering method.

Statistical methods are used to help us describe and understand **variability**. By variability, we mean that successive observations of a system or phenomenon do not produce exactly the same result. We all encounter variability in our everyday lives, and **statistical thinking** can give us a useful way to incorporate this variability into our decision-making processes. For example, consider the gasoline mileage performance of your car. Do you always get exactly the same mileage performance on every tank of fuel? Of course not—in fact, sometimes the mileage performance varies considerably. This observed variability in gasoline mileage depends on many factors, such as the type of driving that has occurred most recently (city versus highway), the changes in condition of the vehicle over time (which could include factors such as tire inflation, engine compression, or valve wear), the brand and/or octane number of the gasoline used, or possibly even the weather conditions that have been recently experienced. These factors represent potential **sources of variability** in the system. Statistics gives us a framework for describing this variability and for learning about which potential sources of variability are the most important or which have the greatest impact on the gasoline mileage performance.

We also encounter variability in dealing with engineering problems. For example, suppose that an engineer is designing a nylon connector to be used in an automotive engine application. The engineer is considering establishing the design specification on wall thickness at 3/32 inch but is somewhat uncertain about the effect of this decision on the connector pull-off force. If the pull-off force is too low, the connector may fail when it is installed in an engine. Eight prototype units are produced and their pull-off forces measured, resulting in the following data (in pounds): 12.6, 12.9, 13.4, 12.3, 13.6, 13.5, 12.6, 13.1. As we anticipated, not all of the prototypes have the same pull-off force. We say that there is variability in the pull-off force measurements. Because the pull-off force measurements exhibit variability, we consider the pull-off force to be a **random variable**. A convenient way to think of a random variable, say X , that represents a measurement, is by using the model

$$X = \mu + \epsilon \quad (1-1)$$

where μ is a constant and ϵ is a random disturbance. The constant remains the same with every measurement, but small changes in the environment, test equipment, differences in the individual parts themselves, and so forth change the value of ϵ . If there were no disturbances, ϵ would always equal zero and X would always be equal to the constant μ . However, this never happens in the real world, so the actual measurements X exhibit variability. We often need to describe, quantify and ultimately reduce variability.

Figure 1-2 presents a **dot diagram** of these data. The dot diagram is a very useful plot for displaying a small body of data—say, up to about 20 observations. This plot allows us to see easily two features of the data; the **location**, or the middle, and the **scatter** or **variability**. When the number of observations is small, it is usually difficult to identify any specific patterns in the variability, although the dot diagram is a convenient way to see any unusual data features.

The need for statistical thinking arises often in the solution of engineering problems. Consider the engineer designing the connector. From testing the prototypes, he knows that the average pull-off force is 13.0 pounds. However, he thinks that this may be too low for the



Figure 1-2 Dot diagram of the pull-off force data when wall thickness is 3/32 inch.



Figure 1-3 Dot diagram of pull-off force for two wall thicknesses.

intended application, so he decides to consider an alternative design with a greater wall thickness, 1/8 inch. Eight prototypes of this design are built, and the observed pull-off force measurements are 12.9, 13.7, 12.8, 13.9, 14.2, 13.2, 13.5, and 13.1. The average is 13.4. Results for both samples are plotted as dot diagrams in Fig. 1-3, page 3. This display gives the impression that increasing the wall thickness has led to an increase in pull-off force. However, there are some obvious questions to ask. For instance, how do we know that another sample of prototypes will not give different results? Is a sample of eight prototypes adequate to give reliable results? If we use the test results obtained so far to conclude that increasing the wall thickness increases the strength, what risks are associated with this decision? For example, is it possible that the apparent increase in pull-off force observed in the thicker prototypes is only due to the inherent variability in the system and that increasing the thickness of the part (and its cost) really has no effect on the pull-off force?

Often, physical laws (such as Ohm's law and the ideal gas law) are applied to help design products and processes. We are familiar with this reasoning from general laws to specific cases. But it is also important to reason from a specific set of measurements to more general cases to answer the previous questions. This reasoning is from a sample (such as the eight connectors) to a population (such as the connectors that will be sold to customers). The reasoning is referred to as **statistical inference**. See Fig. 1-4. Historically, measurements were obtained from a sample of people and generalized to a population, and the terminology has remained. Clearly, reasoning based on measurements from some objects to measurements on all objects can result in errors (called sampling errors). However, if the sample is selected properly, these risks can be quantified and an appropriate sample size can be determined.

In some cases, the sample is actually selected from a well-defined population. The sample is a subset of the population. For example, in a study of resistivity a sample of three wafers might be selected from a production lot of wafers in semiconductor manufacturing. Based on the resistivity data collected on the three wafers in the sample, we want to draw a conclusion about the resistivity of all of the wafers in the lot.

In other cases, the population is conceptual (such as with the connectors), but it might be thought of as future replicates of the objects in the sample. In this situation, the eight prototype connectors must be representative, in some sense, of the ones that will be manufactured in the future. Clearly, this analysis requires some notion of **stability** as an additional assumption. For example, it might be assumed that the sources of variability in the manufacture of the prototypes (such as temperature, pressure, and curing time) are the same as those for the connectors that will be manufactured in the future and ultimately sold to customers.

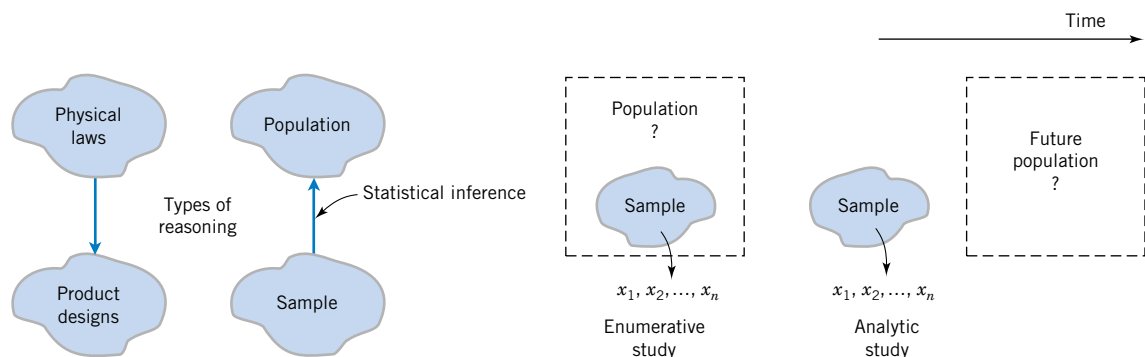


Figure 1-4 Statistical inference is one type of reasoning.

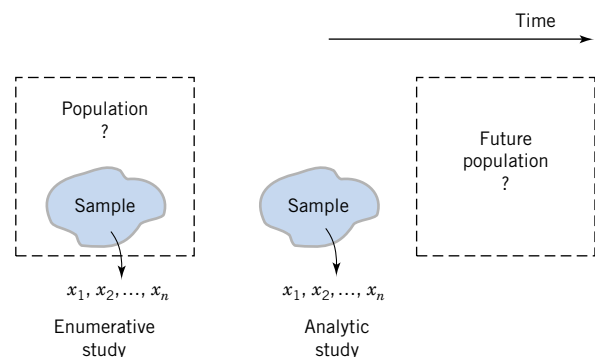


Figure 1-5 Enumerative versus analytic study.

The wafers-from-lots example is called an **enumerative** study. A sample is used to make an inference to the population from which the sample is selected. The connector example is called an **analytic** study. A sample is used to make an inference to a conceptual (future) population. The statistical analyses are usually the same in both cases, but an analytic study clearly requires an assumption of stability. See Fig. 1-5, on page 4.

1-2 COLLECTING ENGINEERING DATA

1-2.1 Basic Principles

In the previous section, we illustrated some simple methods for summarizing data. In the engineering environment, the data is almost always a sample that has been selected from some population. Three basic methods of collecting data are

- A **retrospective** study using historical data
- An **observational** study
- A **designed experiment**

An effective data collection procedure can greatly simplify the analysis and lead to improved understanding of the population or process that is being studied. We now consider some examples of these data collection methods.

1-2.2 Retrospective Study

Montgomery, Peck, and Vining (2001) describe an acetone-butyl alcohol distillation column for which concentration of acetone in the distillate or output product stream is an important variable. Factors that may affect the distillate are the reboil temperature, the condensate temperature, and the reflux rate. Production personnel obtain and archive the following records:

- The concentration of acetone in an hourly test sample of output product
- The reboil temperature log, which is a plot of the reboil temperature over time
- The condenser temperature controller log
- The nominal reflux rate each hour

The reflux rate should be held constant for this process. Consequently, production personnel change this very infrequently.

A retrospective study would use either all or a sample of the historical process data archived over some period of time. The study objective might be to discover the relationships among the two temperatures and the reflux rate on the acetone concentration in the output product stream. However, this type of study presents some problems:

1. We may not be able to see the relationship between the reflux rate and acetone concentration, because the reflux rate didn't change much over the historical period.
2. The archived data on the two temperatures (which are recorded almost continuously) do not correspond perfectly to the acetone concentration measurements (which are made hourly). It may not be obvious how to construct an approximate correspondence.

3. Production maintains the two temperatures as closely as possible to desired targets or set points. Because the temperatures change so little, it may be difficult to assess their real impact on acetone concentration.
4. Within the narrow ranges that they do vary, the condensate temperature tends to increase with the reboil temperature. Consequently, the effects of these two process variables on acetone concentration may be difficult to separate.

As you can see, a retrospective study may involve a lot of **data**, but that data may contain relatively little useful **information** about the problem. Furthermore, some of the relevant data may be missing, there may be transcription or recording errors resulting in **outliers** (or unusual values), or data on other important factors may not have been collected and archived. In the distillation column, for example, the specific concentrations of butyl alcohol and acetone in the input feed stream are a very important factor, but they are not archived because the concentrations are too hard to obtain on a routine basis. As a result of these types of issues, statistical analysis of historical data sometimes identify interesting phenomena, but solid and reliable explanations of these phenomena are often difficult to obtain.

1-2.3 Observational Study

In an observational study, the engineer observes the process or population, disturbing it as little as possible, and records the quantities of interest. Because these studies are usually conducted for a relatively short time period, sometimes variables that are not routinely measured can be included. In the distillation column, the engineer would design a form to record the two temperatures and the reflux rate when acetone concentration measurements are made. It may even be possible to measure the input feed stream concentrations so that the impact of this factor could be studied. Generally, an observational study tends to solve problems 1 and 2 above and goes a long way toward obtaining accurate and reliable data. However, observational studies may not help resolve problems 3 and 4.

1-2.4 Designed Experiments

In a designed experiment the engineer makes *deliberate or purposeful changes* in the controllable variables of the system or process, observes the resulting system output data, and then makes an inference or decision about which variables are responsible for the observed changes in output performance. The nylon connector example in Section 1-1 illustrates a designed experiment; that is, a deliberate change was made in the wall thickness of the connector with the objective of discovering whether or not a greater pull-off force could be obtained. Designed experiments play a very important role in engineering design and development and in the improvement of manufacturing processes. Generally, when products and processes are designed and developed with designed experiments, they enjoy better performance, higher reliability, and lower overall costs. Designed experiments also play a crucial role in reducing the lead time for engineering design and development activities.

For example, consider the problem involving the choice of wall thickness for the nylon connector. This is a simple illustration of a designed experiment. The engineer chose two wall thicknesses for the connector and performed a series of tests to obtain pull-off force measurements at each wall thickness. In this simple **comparative experiment**, the

engineer is interested in determining if there is any difference between the 3/32- and 1/8-inch designs. An approach that could be used in analyzing the data from this experiment is to compare the mean pull-off force for the 3/32-inch design to the mean pull-off force for the 1/8-inch design using statistical **hypothesis testing**, which is discussed in detail in Chapters 9 and 10. Generally, a **hypothesis** is a statement about some aspect of the system in which we are interested. For example, the engineer might want to know if the mean pull-off force of a 3/32-inch design exceeds the typical maximum load expected to be encountered in this application, say 12.75 pounds. Thus, we would be interested in testing the hypothesis that the mean strength exceeds 12.75 pounds. This is called a **single-sample hypothesis testing problem**. It is also an example of an **analytic study**. Chapter 9 presents techniques for this type of problem. Alternatively, the engineer might be interested in testing the hypothesis that increasing the wall thickness from 3/32- to 1/8-inch results in an increase in mean pull-off force. Clearly, this is an **analytic study**; it is also an example of a **two-sample hypothesis testing problem**. Two-sample hypothesis testing problems are discussed in Chapter 10.

Designed experiments are a very powerful approach to studying complex systems, such as the distillation column. This process has three factors, the two temperatures and the reflux rate, and we want to investigate the effect of these three factors on output acetone concentration. A good experimental design for this problem must ensure that we can separate the effects of all three factors on the acetone concentration. The specified values of the three factors used in the experiment are called **factor levels**. Typically, we use a small number of levels for each factor, such as two or three. For the distillation column problem, suppose we use a “high,” and “low,” level (denoted +1 and −1, respectively) for each of the factors. We thus would use two levels for each of the three factors. A very reasonable experiment design strategy uses every possible combination of the factor levels to form a basic experiment with eight different settings for the process. This type of experiment is called a **factorial experiment**. Table 1-1 presents this experimental design.

Figure 1-6, on page 8, illustrates that this design forms a cube in terms of these high and low levels. With each setting of the process conditions, we allow the column to reach equilibrium, take a sample of the product stream, and determine the acetone concentration. We then can draw specific inferences about the effect of these factors. Such an approach allows us to proactively study a population or process. Designed experiments play a very important role in engineering and science. Chapters 13 and 14 discuss many of the important principles and techniques of experimental design.

Table 1-1 The Designed Experiment (Factorial Design) for the Distillation Column

Reboil Temp.	Condensate Temp.	Reflux Rate
−1	−1	−1
+1	−1	−1
−1	+1	−1
+1	+1	−1
−1	−1	+1
+1	−1	+1
−1	+1	+1
+1	+1	+1

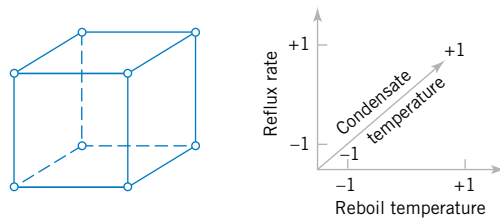


Figure 1-6 The factorial design for the distillation column.

1-2.5 A Factorial Experiment for the Connector Pull-off Force Problem (CD Only)

1-2.6 Observing Processes Over Time

Often data are collected over time. In this case, it is usually very helpful to plot the data versus time in a **time series plot**. Phenomena that might affect the system or process often become more visible in a time-oriented plot and the concept of stability can be better judged.

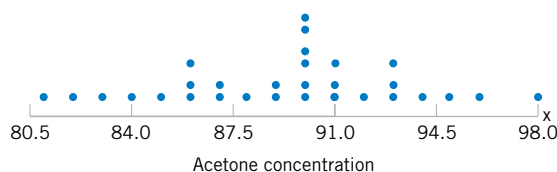
Figure 1-7 is a dot diagram of acetone concentration readings taken hourly from the distillation column described in Section 1-2.2. The large variation displayed on the dot diagram indicates a lot of variability in the concentration, but the chart does not help explain the reason for the variation. The time series plot is shown in Figure 1-8, on page 9. A shift in the process mean level is visible in the plot and an estimate of the time of the shift can be obtained.

W. Edwards **Deming**, a very influential industrial statistician, stressed that it is important to understand the nature of variability in processes and systems over time. He conducted an experiment in which he attempted to drop marbles as close as possible to a target on a table. He used a funnel mounted on a ring stand and the marbles were dropped into the funnel. See Fig. 1-9. The funnel was aligned as closely as possible with the center of the target. He then used two different strategies to operate the process. (1) He never moved the funnel. He just dropped one marble after another and recorded the distance from the target. (2) He dropped the first marble and recorded its location relative to the target. He then moved the funnel an equal and opposite distance in an attempt to compensate for the error. He continued to make this type of adjustment after each marble was dropped.

After both strategies were completed, he noticed that the variability of the distance from the target for strategy 2 was approximately 2 times larger than for strategy 1. The adjustments to the funnel increased the deviations from the target. The explanation is that the error (the deviation of the marble's position from the target) for one marble provides no information about the error that will occur for the next marble. Consequently, adjustments to the funnel do not decrease future errors. Instead, they tend to move the funnel farther from the target.

This interesting experiment points out that adjustments to a process based on random disturbances can actually *increase* the variation of the process. This is referred to as **overcontrol**

Figure 1-7 The dot diagram illustrates variation but does not identify the problem.



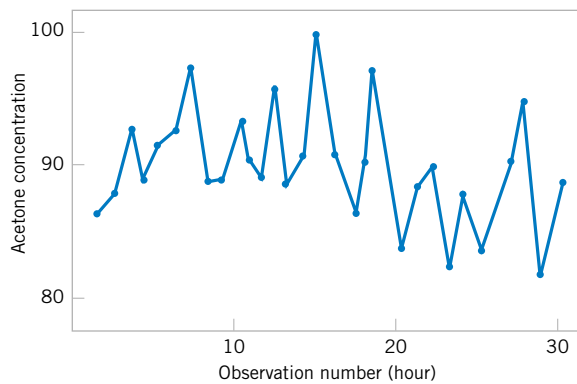


Figure 1-8 A time series plot of concentration provides more information than the dot diagram.

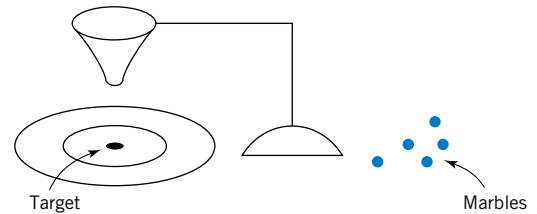


Figure 1-9 Deming's funnel experiment.

or **tampering**. Adjustments should be applied only to compensate for a nonrandom shift in the process—then they can help. A computer simulation can be used to demonstrate the lessons of the funnel experiment. Figure 1-10 displays a time plot of 100 measurements (denoted as y) from a process in which only random disturbances are present. The target value for the process is 10 units. The figure displays the data with and without adjustments that are applied to the process mean in an attempt to produce data closer to target. Each adjustment is equal and opposite to the deviation of the previous measurement from target. For example, when the measurement is 11 (one unit above target), the mean is reduced by one unit before the next measurement is generated. The overcontrol has increased the deviations from the target.

Figure 1-11 displays the data without adjustment from Fig. 1-10, except that the measurements after observation number 50 are increased by two units to simulate the effect of a shift in the mean of the process. When there is a true shift in the mean of a process, an adjustment can be useful. Figure 1-11 also displays the data obtained when one adjustment (a decrease of

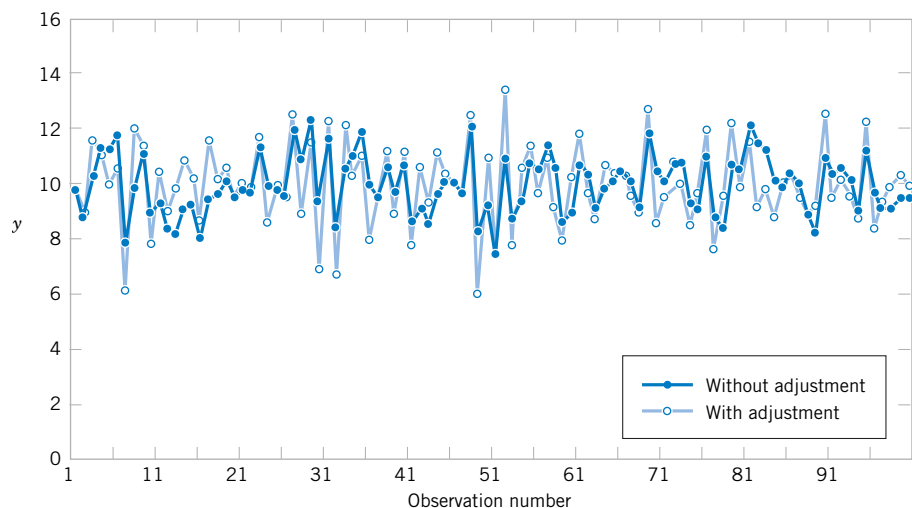


Figure 1-10 Adjustments applied to random disturbances overcontrol the process and increase the deviations from the target.

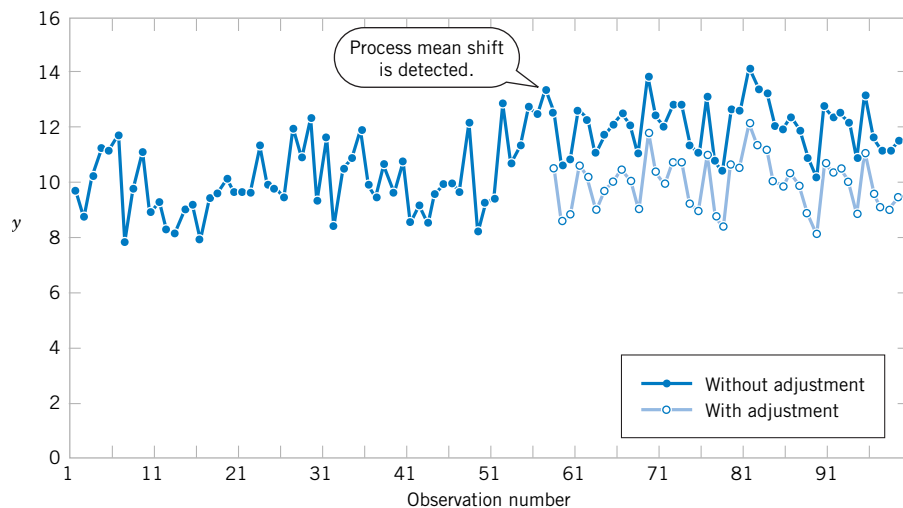


Figure 1-11 Process mean shift is detected at observation number 57, and one adjustment (a decrease of two units) reduces the deviations from target.

two units) is applied to the mean after the shift is detected (at observation number 57). Note that this adjustment decreases the deviations from target.

The question of when to apply adjustments (and by what amounts) begins with an understanding of the types of variation that affect a process. A **control chart** is an invaluable way to examine the variability in time-oriented data. Figure 1-12 presents a control chart for the concentration data from Fig. 1-8. The **center line** on the control chart is just the average of the concentration measurements for the first 20 samples ($\bar{x} = 91.5$ g/l) when the process is stable. The **upper control limit** and the **lower control limit** are a pair of statistically derived limits that reflect the inherent or natural variability in the process. These limits are located three standard deviations of the concentration values above and below the center line. If the process is operating as it should, without any external sources of variability present in the system, the concentration measurements should fluctuate randomly around the center line, and almost all of them should fall between the control limits.

In the control chart of Fig. 1-12, the visual frame of reference provided by the center line and the control limits indicates that some upset or disturbance has affected the process around sample 20 because all of the following observations are below the center line and two of them

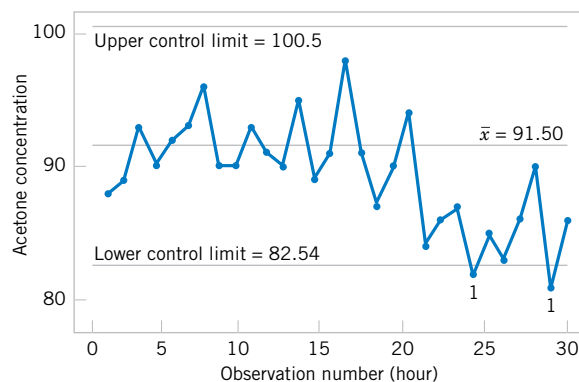


Figure 1-12 A control chart for the chemical process concentration data.

actually fall below the lower control limit. This is a very strong signal that corrective action is required in this process. If we can find and eliminate the underlying cause of this upset, we can improve process performance considerably.

Control charts are a very important application of statistics for monitoring, controlling, and improving a process. The branch of statistics that makes use of control charts is called **statistical process control**, or **SPC**. We will discuss SPC and control charts in Chapter 16.

1-3 MECHANISTIC AND EMPIRICAL MODELS

Models play an important role in the analysis of nearly all engineering problems. Much of the formal education of engineers involves learning about the models relevant to specific fields and the techniques for applying these models in problem formulation and solution. As a simple example, suppose we are measuring the flow of current in a thin copper wire. Our model for this phenomenon might be Ohm's law:

$$\text{Current} = \text{voltage}/\text{resistance}$$

or

$$I = E/R \quad (1-2)$$

We call this type of model a **mechanistic model** because it is built from our underlying knowledge of the basic physical mechanism that relates these variables. However, if we performed this measurement process more than once, perhaps at different times, or even on different days, the observed current could differ slightly because of small changes or variations in factors that are not completely controlled, such as changes in ambient temperature, fluctuations in performance of the gauge, small impurities present at different locations in the wire, and drifts in the voltage source. Consequently, a more realistic model of the observed current might be

$$I = E/R + \epsilon \quad (1-3)$$

where ϵ is a term added to the model to account for the fact that the observed values of current flow do not perfectly conform to the mechanistic model. We can think of ϵ as a term that includes the effects of all of the unmodeled sources of variability that affect this system.

Sometimes engineers work with problems for which there is no simple or well-understood mechanistic model that explains the phenomenon. For instance, suppose we are interested in the number average molecular weight (M_n) of a polymer. Now we know that M_n is related to the viscosity of the material (V), and it also depends on the amount of catalyst (C) and the temperature (T) in the polymerization reactor when the material is manufactured. The relationship between M_n and these variables is

$$M_n = f(V, C, T) \quad (1-4)$$

say, where the *form* of the function f is unknown. Perhaps a working model could be developed from a first-order Taylor series expansion, which would produce a model of the form

$$M_n = \beta_0 + \beta_1 V + \beta_2 C + \beta_3 T \quad (1-5)$$

where the β 's are unknown parameters. Now just as in Ohm's law, this model will not exactly describe the phenomenon, so we should account for the other sources of variability that may affect the molecular weight by adding another term to the model; therefore

$$M_n = \beta_0 + \beta_1 V + \beta_2 C + \beta_3 T + \epsilon \quad (1-6)$$

is the model that we will use to relate molecular weight to the other three variables. This type of model is called an **empirical model**; that is, it uses our engineering and scientific knowledge of the phenomenon, but it is not directly developed from our theoretical or first-principles understanding of the underlying mechanism.

To illustrate these ideas with a specific example, consider the data in Table 1-2. This table contains data on three variables that were collected in an observational study in a semiconductor manufacturing plant. In this plant, the finished semiconductor is wire bonded to a frame. The variables reported are pull strength (a measure of the amount of force required to break the bond), the wire length, and the height of the die. We would like to find a model relating pull strength to wire length and die height. Unfortunately, there is no physical mechanism that we can easily apply here, so it doesn't seem likely that a mechanistic modeling approach will be successful.

Table 1-2 Wire Bond Pull Strength Data

Observation Number	Pull Strength y	Wire Length x_1	Die Height x_2
1	9.95	2	50
2	24.45	8	110
3	31.75	11	120
4	35.00	10	550
5	25.02	8	295
6	16.86	4	200
7	14.38	2	375
8	9.60	2	52
9	24.35	9	100
10	27.50	8	300
11	17.08	4	412
12	37.00	11	400
13	41.95	12	500
14	11.66	2	360
15	21.65	4	205
16	17.89	4	400
17	69.00	20	600
18	10.30	1	585
19	34.93	10	540
20	46.59	15	250
21	44.88	15	290
22	54.12	16	510
23	56.63	17	590
24	22.13	6	100
25	21.15	5	400

Figure 1-13 Three-dimensional plot of the wire and pull strength data.

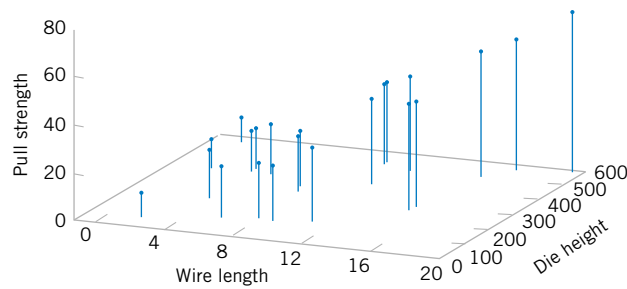


Figure 1-13 presents a three-dimensional plot of all 25 observations on pull strength, wire length, and die height. From examination of this plot, we see that pull strength increases as both wire length and die height increase. Furthermore, it seems reasonable to think that a model such as

$$\text{Pull strength} = \beta_0 + \beta_1(\text{wire length}) + \beta_2(\text{die height}) + \epsilon$$

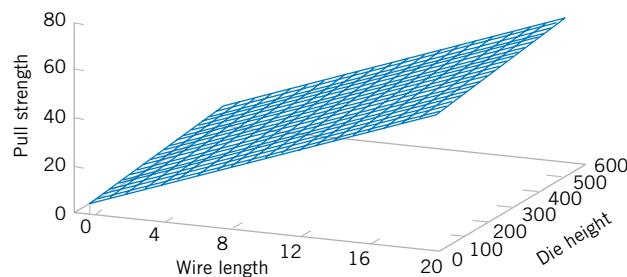
would be appropriate as an empirical model for this relationship. In general, this type of empirical model is called a **regression model**. In Chapters 11 and 12 we show how to build these models and test their adequacy as approximating functions. We will use a method for estimating the parameters in regression models, called the method of **least squares**, that traces its origins to work by Karl Gauss. Essentially, this method chooses the parameters in the empirical model (the β 's) to minimize the sum of the squared distances between each data point and the plane represented by the model equation. Applying this technique to the data in Table 1-2 results in

$$\widehat{\text{Pull strength}} = 2.26 + 2.74(\text{wire length}) + 0.0125(\text{die height}) \quad (1-7)$$

where the “hat,” or circumflex, over pull strength indicates that this is an estimated or predicted quantity.

Figure 1-14 is a plot of the predicted values of pull strength versus wire length and die height obtained from Equation 1-7. Notice that the predicted values lie on a plane above the wire length–die height space. From the plot of the data in Fig. 1-13, this model does not appear unreasonable. The empirical model in Equation 1-7 could be used to predict values of pull strength for various combinations of wire length and die height that are of interest. Essentially, the empirical model could be used by an engineer in exactly the same way that a mechanistic model can be used.

Figure 1-14 Plot of predicted values of pull strength from the empirical model.



1-4 PROBABILITY AND PROBABILITY MODELS

In Section 1-1, it was mentioned that decisions often need to be based on measurements from only a subset of objects selected in a sample. This process of reasoning from a sample of objects to conclusions for a population of objects was referred to as statistical inference. A sample of three wafers selected from a larger production lot of wafers in semiconductor manufacturing was an example mentioned. To make good decisions, an analysis of how well a sample represents a population is clearly necessary. If the lot contains defective wafers, how well will the sample detect this? How can we quantify the criterion to “detect well”? Basically, how can we quantify the risks of decisions based on samples? Furthermore, how should samples be selected to provide good decisions—ones with acceptable risks? **Probability** models help quantify the risks involved in statistical inference, that is, the risks involved in decisions made every day.

More details are useful to describe the role of probability models. Suppose a production lot contains 25 wafers. If all the wafers are defective or all are good, clearly any sample will generate all defective or all good wafers, respectively. However, suppose only one wafer in the lot is defective. Then a sample might or might not detect (include) the wafer. A probability model, along with a method to select the sample, can be used to quantify the risks that the defective wafer is or is not detected. Based on this analysis, the size of the sample might be increased (or decreased). The risk here can be interpreted as follows. Suppose a series of lots, each with exactly one defective wafer, are sampled. The details of the method used to select the sample are postponed until randomness is discussed in the next chapter. Nevertheless, assume that the same size sample (such as three wafers) is selected in the same manner from each lot. The proportion of the lots in which the defective wafer is included in the sample or, more specifically, the limit of this proportion as the number of lots in the series tends to infinity, is interpreted as the probability that the defective wafer is detected.

A probability model is used to calculate this proportion under reasonable assumptions for the manner in which the sample is selected. This is fortunate because we do not want to attempt to sample from an infinite series of lots. Problems of this type are worked in Chapters 2 and 3. More importantly, this probability provides valuable, quantitative information regarding any decision about lot quality based on the sample.

Recall from Section 1-1 that a population might be conceptual, as in an analytic study that applies statistical inference to future production based on the data from current production. When populations are extended in this manner, the role of statistical inference and the associated probability models becomes even more important.

In the previous example, each wafer in the sample was only classified as defective or not. Instead, a continuous measurement might be obtained from each wafer. In Section 1-2.6, concentration measurements were taken at periodic intervals from a production process. Figure 1-7 shows that variability is present in the measurements, and there might be concern that the process has moved from the target setting for concentration. Similar to the defective wafer, one might want to quantify our ability to detect a process change based on the sample data. Control limits were mentioned in Section 1-2.6 as decision rules for whether or not to adjust a process. The probability that a particular process change is detected can be calculated with a probability model for concentration measurements. Models for continuous measurements are developed based on plausible assumptions for the data and a result known as the central limit theorem, and the associated normal distribution is a particularly valuable probability model for statistical inference. Of course, a check of assumptions is important. These types of probability models are discussed in Chapter 4. The objective is still to quantify the risks inherent in the inference made from the sample data.

Throughout Chapters 6 through 15, decisions are based statistical inference from sample data. Continuous probability models, specifically the normal distribution, are used extensively to quantify the risks in these decisions and to evaluate ways to collect the data and how large a sample should be selected.

IMPORTANT TERMS AND CONCEPTS

In the E-book, click on any term or concept below to go to that subject.

Analytic study
Designed experiment
Empirical model
Engineering method
Enumerative study

Mechanistic model
Observational study
Overcontrol
Population
Probability model
Problem-solving method
Retrospective study
Sample

Statistical inference
Statistical Process Control
Statistical thinking
Tampering
Variability

CD MATERIAL

Factorial Experiment
Fractional factorial experiment
Interaction

1-2.5 A Factorial Experiment for the Connector Pull-off Force Problem (CD only)

Much of what we know in the engineering and physical-chemical sciences is developed through testing or experimentation. Often engineers work in problem areas in which no scientific or engineering theory is directly or completely applicable, so experimentation and observation of the resulting data constitute the only way that the problem can be solved. Even when there is a good underlying scientific theory that we may rely on to explain the phenomena of interest, it is almost always necessary to conduct tests or experiments to confirm that the theory is indeed operative in the situation or environment in which it is being applied. We have observed that statistical thinking and statistical methods play an important role in planning, conducting, and analyzing the data from engineering experiments.

To further illustrate the factorial design concept introduced in Section 1-2.4, suppose that in the connector wall thickness example, there are two additional factors of interest, time and temperature. The cure times of interest are 1 and 24 hours and the temperature levels are 70°F and 100°F. Now since all three factors have two levels, a factorial experiment would consist of the eight test combinations shown at the corners of the cube in Fig. S1-1. Two trials, or **replicates**, would be performed at each corner, resulting in a 16-run factorial experiment. The observed values of pull-off force are shown in parentheses at the cube corners in Fig. S1-1. Notice that this experiment uses eight 3/32-inch prototypes and eight 1/8-inch prototypes, the same number used in the simple comparative study in Section 1-1, but we are now investigating *three* factors. Generally, factorial experiments are the most efficient way to study the joint effects of several factors.

Some very interesting tentative conclusions can be drawn from this experiment. First, compare the average pull-off force of the eight 3/32-inch prototypes with the average pull-off force of the eight 1/8-inch prototypes (these are the averages of the eight runs on the left face and right face of the cube in Fig. S1-1, respectively), or $14.1 - 13.45 = 0.65$. Thus, increasing the wall thickness from 3/32 to 1/8-inch increases the average pull-off force by 0.65 pounds. Next, to measure the effect of increasing the cure time, compare the average of the eight runs in the back face of the cube (where time = 24 hours) with the average of the eight runs in the front face (where time = 1 hour), or $14.275 - 13.275 = 1$. The effect of increasing the cure time from 1 to 24 hours is to increase the average pull-off force by 1 pound; that is, cure time apparently has an effect that is larger than the effect of increasing the wall

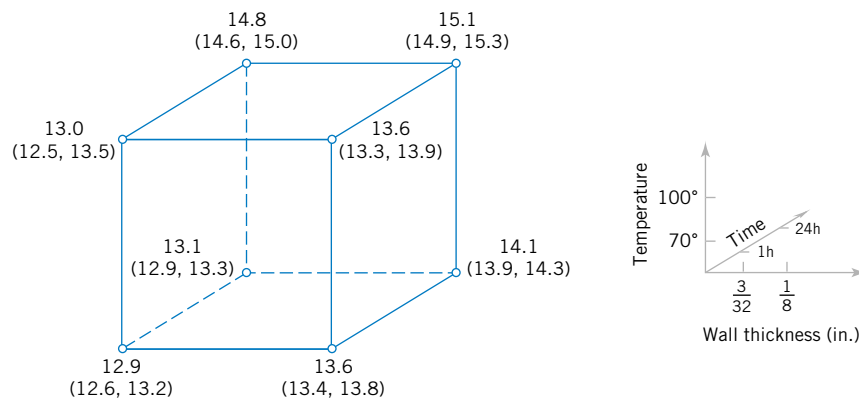


Figure S1-1 The factorial experiment for the connector wall thickness problem.

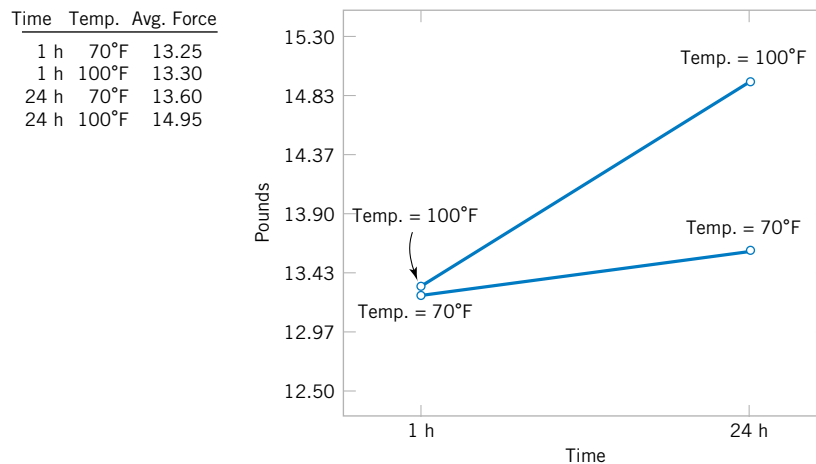


Figure S1-2 The two-factor interaction between cure time and cure temperature.

thickness. The cure temperature effect can be evaluated by comparing the average of the eight runs in the top of the cube (where temperature = 100°F) with the average of the eight runs in the bottom (where temperature = 70°F), or $14.125 - 13.425 = 0.7$. Thus, the effect of increasing the cure temperature is to increase the average pull-off force by 0.7 pounds. Thus, if the engineer's objective is to design a connector with high pull-off force, there are apparently several alternatives, such as increasing the wall thickness and using the "standard" curing conditions of 1 hour and 70°F or using the original 3/32-inch wall thickness but specifying a longer cure time and higher cure temperature.

There is an interesting relationship between cure time and cure temperature that can be seen by examination of the graph in Fig. S1-2. This graph was constructed by calculating the average pull-off force at the four different combinations of time and temperature, plotting these averages versus time and then connecting the points representing the two temperature levels with straight lines. The slope of each of these straight lines represents the effect of cure time on pull-off force. Notice that the slopes of these two lines do not appear to be the same, indicating that the cure time effect is *different* at the two values of cure temperature. This is an example of an **interaction** between two factors. The interpretation of this interaction is very straightforward; if the standard cure time (1 hour) is used, cure temperature has little effect, but if the longer cure time (24 hours) is used, increasing the cure temperature has a large effect on average pull-off force. Interactions occur often in physical and chemical systems, and factorial experiments are the only way to investigate their effects. In fact, if interactions are present and the factorial experimental strategy is not used, incorrect or misleading results may be obtained.

We can easily extend the factorial strategy to more factors. Suppose that the engineer wants to consider a fourth factor, type of adhesive. There are two types: the standard adhesive and a new competitor. Figure S1-3 illustrates how all four factors, wall thickness, cure time, cure temperature, and type of adhesive, could be investigated in a factorial design. Since all four factors are still at two levels, the experimental design can still be represented geometrically as a cube (actually, it's a *hypercube*). Notice that as in any factorial design, all possible combinations of the four factors are tested. The experiment requires 16 trials.

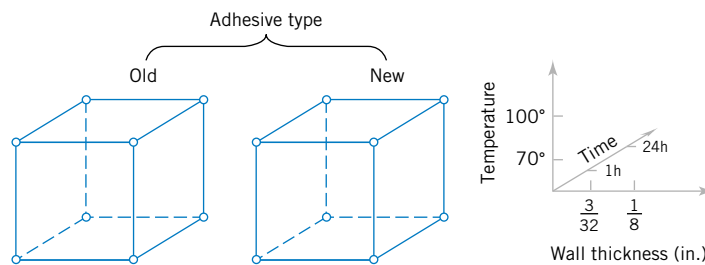


Figure S1-3 A four-factorial experiment for the connector wall thickness problem.

Generally, if there are k factors and they each have two levels, a factorial experimental design will require 2^k runs. For example, with $k = 4$, the 2^4 design in Fig. S1-3 requires 16 tests. Clearly, as the number of factors increases, the number of trials required in a factorial experiment increases rapidly; for instance, eight factors each at two levels would require 256 trials. This quickly becomes unfeasible from the viewpoint of time and other resources. Fortunately, when there are four to five or more factors, it is usually unnecessary to test all possible combinations of factor levels. A **fractional factorial experiment** is a variation of the basic factorial arrangement in which only a subset of the factor combinations are actually tested. Figure S1-4 shows a fractional factorial experimental design for the four-factor version of the connector experiment. The circled test combinations in this figure are the only test combinations that need to be run. This experimental design requires only 8 runs instead of the original 16; consequently it would be called a **one-half fraction**. This is an excellent experimental design in which to study all four factors. It will provide good information about the individual effects of the four factors and some information about how these factors interact.

Factorial and fractional factorial experiments are used extensively by engineers and scientists in industrial research and development, where new technology, products, and processes are designed and developed and where existing products and processes are improved. Since so much engineering work involves testing and experimentation, it is essential that all engineers understand the basic principles of planning efficient and effective experiments. We discuss these principles in Chapter 13. Chapter 14 concentrates on the factorial and fractional factorials that we have introduced here.

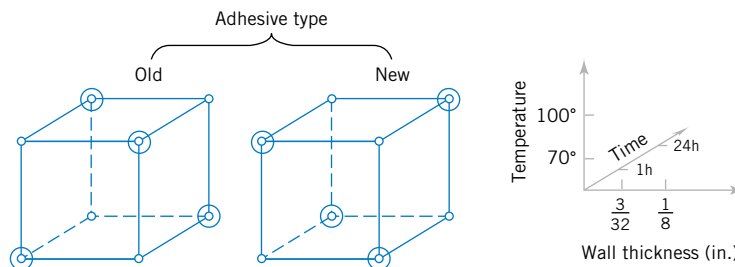


Figure S1-4 A fractional factorial experiment for the connector wall thickness problem.