

# Table of contents

## STATISTICAL MODELLING

<b>I.</b>	<b>Statistical inference .....</b>	<b>I-1</b>
I.A	Expected values and variances .....	I-1
I.B	The linear regression model .....	I-3
I.C	Model selection.....	I-10
	a) Obtaining parameter estimates .....	I-10
	b) Regression analysis of variance .....	I-16
I.D	Summary .....	I-22
I.E	Exercises .....	I-23
<b>II.</b>	<b>Designing experiments — some general aspects .....</b>	<b>II-1</b>
II.A	General considerations.....	II-1
	a) Basic purposes of experimentation .....	II-1
	b) Basic requirements .....	II-1
	c) The experimental process.....	II-2
	d) Definition of some key terms.....	II-4
II.B	Three key principles in designing experiments .....	II-4
	a) Replication .....	II-6
	b) Randomization .....	II-9
	c) Blocking .....	II-10
II.C	Choosing the factors and their levels.....	II-11
	a) Factorial experiments.....	II-12
	b) Choosing the levels of a factor.....	II-12
	c) Sequences of experiments .....	II-15
II.D	Experimental validity.....	II-15
	a) Internal validity .....	II-16
	b) External validity .....	II-16
	c) Key methods for overcoming internal and external validity problems .....	II-17
II.E	An example experiment.....	II-17
II.F	Summary .....	II-18
<b>III.</b>	<b>Completely Randomized Design.....</b>	<b>III-1</b>
III.A	Design of a CRD.....	III-1
III.B	Models and estimation for a CRD .....	III-2
	a) Maximal model.....	III-2
	b) Alternative indicator-variable, expectation models .....	III-6
III.C	Hypothesis testing using the ANOVA method .....	III-8
	a) Analysis of the rat example .....	III-8
	b) Sums of squares for the analysis of variance .....	III-8
	c) Expected mean squares .....	III-12
	d) Summary of the hypothesis test.....	III-15
	e) Comparison with traditional one-way ANOVA.....	III-16
	f) Computation of the ANOVA in R.....	III-16
III.D	Diagnostic checking.....	III-18
III.E	Treatment differences.....	III-25
	a) Multiple comparisons procedures for comparing all treatments.....	III-25
	b) Fitting submodels.....	III-27
	c) Comparison of treatment parametrizations .....	III-34

III.F	Summary .....	III-35
III.G	Exercises .....	III-36
<b>IV.</b>	<b>Randomized Complete Block Design (RCBD) .....</b>	<b>IV-1</b>
IV.A	Design of an RCBD .....	IV-1
a)	Obtaining a layout for an RCBD in R .....	IV-2
IV.B	Indicator-variable models and estimation for an RCBD .....	IV-4
a)	Maximal model.....	IV-4
b)	Alternative expectation models .....	IV-9
IV.C	Hypothesis testing using the ANOVA method for an RCBD .....	IV-10
a)	Analysis of the penicillin example .....	IV-10
b)	Sums of squares for the analysis of variance .....	IV-12
c)	Expected mean squares .....	IV-14
d)	Summary of the hypothesis test.....	IV-16
e)	Comparison with traditional two-way ANOVA .....	IV-17
f)	Computation of the ANOVA in R.....	IV-17
IV.D	Diagnostic checking.....	IV-19
IV.E	Treatment differences.....	IV-22
IV.F	Fixed versus random effects.....	IV-22
a)	Another maximal model for the RCBD .....	IV-22
b)	Estimation and analysis of variance for Blocks random .....	IV-25
IV.G	Generalized randomized complete block design .....	IV-26
IV.H	Summary .....	IV-28
IV.I	Exercises .....	IV-29
<b>V.</b>	<b>Latin squares designs (LS) .....</b>	<b>V-1</b>
V.A	Design of Latin squares.....	V-1
a)	Obtaining a layout for a Latin square in R.....	V-2
V.B	Indicator-variable models and estimation for a Latin square.....	V-4
a)	Maximal model.....	V-4
b)	Alternative expectation models .....	V-6
V.C	Hypothesis testing using the ANOVA method for a Latin square.....	V-7
a)	Analysis of an example .....	V-7
b)	Sums of squares for the analysis of variance .....	V-9
c)	Expected mean squares .....	V-11
d)	Summary of the hypothesis test.....	V-13
e)	Comparison with traditional Latin-square ANOVA table.....	V-14
f)	Computation of ANOVA and diagnostic checking in R.....	V-15
V.D	Diagnostic checking.....	V-18
V.E	Treatment differences.....	V-19
V.F	Design of sets of Latin squares .....	V-20
V.G	Hypothesis tests for sets of Latin squares .....	V-22
a)	Case 1 — same Drivers and Cars .....	V-22
b)	Case 2 — same Cars different Drivers .....	V-27
c)	Case 3 — different Drivers and Cars .....	V-29
d)	Comparison of Latin square analyses.....	V-32
e)	Computation of ANOVA in R.....	V-32
V.H	Summary .....	V-33
V.I	Exercises .....	V-33

<b>VI.</b>	<b>Determining the analysis of variance table.....</b>	<b>VI-1</b>
VI.A	The procedure .....	VI-1
a)	Description of pertinent features of the study.....	VI-2
b)	The experimental structure .....	VI-7
c)	Sources derived from the structure formulae .....	VI-11
d)	Degrees of freedom and sums of squares .....	VI-11
e)	The analysis of variance table .....	VI-17
f)	Maximal expectation and variation models .....	VI-18
g)	The expected mean squares .....	VI-22
VI.B	The Latin square example .....	VI-23
VI.C	Usage of the procedure .....	VI-26
VI.D	Rules for determining the analysis of variance table — summary .....	VI-27
a)	Description of pertinent features of the study.....	VI-27
b)	The experimental structure .....	VI-27
c)	Sources derived from the structure formulae .....	VI-28
d)	Degrees of freedom and sums of squares .....	VI-28
e)	The analysis of variance table .....	VI-29
f)	Maximal expectation and variation models .....	VI-29
g)	The expected mean squares .....	VI-30
VI.E	Determining the analysis of variance table – further examples.....	VI-31
VI.F	Exercises .....	VI-62
<b>VII.</b>	<b>Factorial experiments .....</b>	<b>VII-1</b>
VII.A	Design of factorial experiments .....	VII-2
a)	Obtaining a layout for a factorial experiment in R .....	VII-2
VII.B	Advantages of factorial experiments.....	VII-3
a)	Interaction in factorial experiments .....	VII-3
b)	Advantages over one-factor-at-a-time experiments .....	VII-7
VII.C	An example two-factor CRD experiment.....	VII-8
a)	Determining the ANOVA table for a two-Factor CRD.....	VII-8
b)	Analysis of an example .....	VII-11
VII.D	Indicator-variable models and estimation for factorial experiments ..	VII-13
a)	Maximal model for two-factor CRD experiments.....	VII-13
b)	Alternative expectation models .....	VII-15
VII.E	Hypothesis testing using the ANOVA method for factorial experiments .....	VII-18
a)	Sums of squares for the analysis of variance .....	VII-19
b)	Expected mean squares .....	VII-20
c)	Summary of the hypothesis test.....	VII-25
d)	Computation of ANOVA and diagnostic checking in R.....	VII-27
e)	Box-Cox transformations for correcting transformable non-additivity .....	VII-30
VII.F	Treatment differences.....	VII-33
a)	Multiple comparisons procedures .....	VII-33
b)	Polynomial submodels .....	VII-36
VII.G	Nested factorial structures .....	VII-48
VII.H	Models and hypothesis testing for three-factor experiments.....	VII-55
a)	Using the rules to determine the ANOVA table for a 3-factor CRD experiment .....	VII-56
b)	Indicator-variable models and estimation for the three-factor CRD .....	VII-58

c)	Expected mean squares under alternative models .....	VII-59
d)	The hypothesis test.....	VII-60
VII.I	Summary .....	VII-63
VII.J	Exercises .....	VII-64
<b>VIII.</b>	<b>Factorial designs at two levels .....</b>	<b>VIII-1</b>
VIII.A	Replicated $2^k$ experiments .....	VIII-2
a)	Design of replicated $2^k$ experiments, including R expressions.....	VIII-2
b)	Analysis of variance .....	VIII-4
c)	Calculation of responses and Yates effects .....	VIII-8
d)	Yates algorithm .....	VIII-12
e)	Treatment differences .....	VIII-13
VIII.B	Economy in experimentation .....	VIII-16
a)	Design of unreplicated $2^k$ experiments, including R expressions.....	VIII-16
b)	Initial analysis of variance .....	VIII-17
c)	Analysis assuming no 3-factor or 4-factor interactions.....	VIII-18
d)	Probability plot of Yates effects.....	VIII-19
e)	Fitted values .....	VIII-20
f)	Diagnostic checking .....	VIII-21
g)	Treatment differences .....	VIII-23
VIII.C	Confounding in factorial experiments .....	VIII-24
a)	Total confounding of effects.....	VIII-24
b)	Partial confounding of effects.....	VIII-31
VIII.D	Fractional factorial designs at two levels .....	VIII-35
a)	Half-fractions of full factorial experiments .....	VIII-42
b)	More on construction and use of half-factions .....	VIII-49
c)	The concept of design resolution .....	VIII-50
d)	Resolution III designs (also called main effect plans) .....	VIII-52
e)	Table of fractional factorial designs.....	VIII-60
VIII.E	Summary .....	VIII-61
VIII.F	Exercises .....	VIII-62
<b>IX.</b>	<b>Split-plot experiments .....</b>	<b>IX-1</b>
IX.A	Design of split-plot experiments.....	IX-1
IX.B	The standard split-plot experiment .....	IX-3
a)	Designing a standard split-plot experiment .....	IX-3
b)	Determining the analysis of variance table .....	IX-4
c)	Analysis of the example .....	IX-7
d)	Treatment differences for the standard split-plot.....	IX-10
IX.C	Systematic or Unreplicated Main Plots .....	IX-12
IX.D	A Complex Split-Plot Experiment.....	IX-14
IX.E	Summary .....	IX-25
IX.F	Exercises .....	IX-26

<b>X.</b>	<b>Sample size and power.....</b>	<b>X-1</b>
X.A	How it is done .....	X-1
X.B	Power .....	X-3
	a) Type I and II errors.....	X-3
	b) Power of a hypothesis test about expectation model terms .....	X-4
X.C	Computing the required sample size for the CRD and RCBD with a single treatment factor .....	X-8
X.D	Sample size for the Latin square design.....	X-9
X.E	Sample size for factorial experiments.....	X-10
X.F	Sample size for the standard split-plot experiment .....	X-12
X.G	Summary .....	X-12
X.H	Exercises .....	X-13
<b>XI.</b>	<b>Estimation of linear model parameters .....</b>	<b>XI-1</b>
XI.A	Linear models for designed experiments.....	XI-1
XI.B	Least squares estimation of the expectation parameters in simple linear models .....	XI-6
	a) Ordinary least squares estimators for full-rank expectation models .....	XI-6
	b) Ordinary least squares estimators for less-than-full-rank expectation models .....	XI-12
	c) Estimable functions.....	XI-20
	d) Properties of estimable functions .....	XI-23
	e) Properties of the estimators in the full rank case .....	XI-28
	f) Estimation for the maximal model for the RCBD.....	XI-28
XI.C	Generalized least squares (GLS) estimation of the expectation parameters in general linear models .....	XI-31
XI.D	Maximum likelihood estimation of the expectation parameters.....	XI-37
XI.E	Estimating the variance .....	XI-40
	a) Estimating the variance for the simple linear model.....	XI-40
	b) Estimating the variance for the general linear model .....	XI-41
XI.F	Summary .....	XI-42
XI.G	Exercises .....	XI-43
<b>XII.</b>	<b>Justifying the ANOVA-based hypothesis test .....</b>	<b>XII-1</b>
XII.A	The sources for an ANOVA .....	XII-1
XII.B	The sums of squares for an ANOVA .....	XII-2
XII.C	Degrees of freedom of the sums of squares for an ANOVA .....	XII-3
XII.D	Expected mean squares for an ANOVA .....	XII-4
XII.E	The distribution of the F statistics for an ANOVA.....	XII-6
XII.F	Application of theory for ANOVA-based hypothesis test to an example.....	XII-8
	a) The sources for an ANOVA .....	XII-8
	b) The sums of squares for an ANOVA.....	XII-9
	c) Degrees of freedom of the sums of squares for an ANOVA .....	XII-10
	d) Expected mean squares for an ANOVA.....	XII-12
	e) The distribution of the F statistics for an ANOVA .....	XII-15
XII.G	Summary .....	XII-17
XII.H	Exercises .....	XII-18

<b>Appendix A Introduction to R.....</b>	<b>A-1</b>
A.1. Introduction to R .....	A-1
A.2. Some basic concepts .....	A-2
a) Performing tasks .....	A-2
b) Type of data objects.....	A-3
c) Naming conventions .....	A-3
A.3. An R session .....	A-4
a) Initializing Tinn-R and R.....	A-4
b) Data Entry .....	A-4
c) Initial graphs .....	A-5
d) Statistical analysis.....	A-5
e) Report generation .....	A-6
f) Finishing up .....	A-6
A.4. Managing R usage .....	A-6
a) Directories, the workspace and objects .....	A-6
b) Getting help in R .....	A-7
c) Functions in R .....	A-8
A.5. Entering data in R.....	A-9
a) Creating data .....	A-9
b) Opening previously stored data .....	A-10
c) Import data stored by other programs.....	A-10
A.6. Manipulating data .....	A-11
A.7. Examples.....	A-12
A.8. Numeric and Character Vectors versus Factors .....	A-16
A.9. Exercises .....	A-19
<b>Appendix B Randomized layouts and sample size computations in R .....</b>	<b>B-1</b>
B.1. Completely randomized design .....	B-2
B.2. Randomized complete block design .....	B-3
B.3. Generalized randomized complete block design .....	B-5
B.4. Latin square design .....	B-7
B.5. Sets of Latin Squares .....	B-8
B.6. Factorial experiments .....	B-10
B.7. Two-level factorial experiments .....	B-12
a) Replicated two-level factorial experiments.....	B-13
b) Unreplicated two-level factorial experiments.....	B-13
c) Confounded two-level factorial experiments .....	B-13
d) Fractional two-level factorial experiments .....	B-16
B.8. Split-plot experiments .....	B-18
B.9. Incomplete block designs .....	B-20
B.10. Balanced lattice square designs .....	B-21
B.11. Youden square designs.....	B-22
B.12. Power and sample size for designed experiments .....	B-23
a) Computing the power for given sample size .....	B-24
b) Computing the sample size to achieve specified power.....	B-25

<b>Appendix C Analysis of designed experiments in R.....</b>	<b>C-1</b>
C.1. Entering the results of an experiment into a data.frame .....	C-1
a) Adding the response variable to a randomized layout .....	C-1
b) Creating a data.frame from scratch with the factors in standard order.....	C-2
c) Creating a data.frame from scratch with the data recorded against the randomized layout .....	C-4
C.2. The elements of the analysis of experiments.....	C-4
C.3. Completely randomized design .....	C-5
C.4. Randomized complete block design .....	C-7
C.5. Latin square design .....	C-9
C.6. A set of Latin squares design .....	C-10
C.7. Factorial experiments .....	C-12
C.8. Two-level factorial experiments .....	C-18
a) Replicated two-level factorial experiments.....	C-18
b) Unreplicated two-level factorial experiments.....	C-19
c) Confounded two-level factorial experiments .....	C-20
d) Fractional two-level factorial experiments .....	C-20
C.9. Split-plot experiment.....	C-22