

Contents

Preface to the second edition	xv
Preface to the first edition	xvii
1 How to use this book	1
Introduction	1
The text of the chapters	1
What should you do if you run into trouble?	2
Elephants	3
The numerical examples in the text	3
Boxes	4
Spare-time activities	4
Executive summaries	5
Why go to all that bother?	5
The bibliography	7
2 Introduction	9
What are statistics?	9
Notation	10
Notation for calculating the mean	12
3 Summarising variation	13
Introduction	13
Different summaries of variation	14
Range	14
Total deviation	14
Mean deviation	15
Variance	16
Why $n - 1$?	17
Why are the deviations squared?	18

The standard deviation	19
The next chapter	21
Spare-time activities	21
4 When are sums of squares NOT sums of squares?	23
Introduction	23
Calculating machines offer a quicker method of calculating the sum of squares	24
Added squares	24
The correction factor	24
Avoid being confused by the term <i>sum of squares</i>	24
Summary of the calculator method for calculations as far as the standard deviation	25
Spare-time activities	26
5 The normal distribution	27
Introduction	27
Frequency distributions	27
The normal distribution	28
What percentage is a standard deviation worth?	30
Are the percentages always the same as these?	30
Other similar scales in everyday life	33
The standard deviation as an estimate of the frequency of a number occurring in a sample	33
From percentage to probability	34
EXECUTIVE SUMMARY 1 – The standard deviation	36
6 The relevance of the normal distribution to biological data	39
To recap	39
Is our observed distribution normal?	41
Checking for normality	42
What can we do about a distribution that clearly is not normal?	42
Transformation	42
Grouping samples	47
Doing nothing!	47
How many samples are needed?	47
Type 1 and Type 2 errors	48
Calculating how many samples are needed	49

7 Further calculations from the normal distribution	51
Introduction	51
Is A bigger than B?	52
The yardstick for deciding	52
The standard error of a difference between two means of three eggs	53
Derivation of the standard error of a difference between two means	53
Step 1: from variance of single data to variance of means	55
Step 2: From variance of single data to <i>variance of differences</i>	57
Step 3: The combination of Steps 1 and 2: the standard error of difference between means (s.e.d.m.)	58
Recap of the calculation of s.e.d.m. from the variance calculated from the individual values	61
The importance of the standard error of differences between means	61
Summary of this chapter	62
EXECUTIVE SUMMARY 2 – Standard error of a difference between two means	66
Spare-time activities	67
8 The <i>t</i>-test	69
Introduction	69
The principle of the <i>t</i> -test	70
The <i>t</i> -test in statistical terms	71
Why <i>t</i> ?	71
Tables of the <i>t</i> -distribution	72
The standard <i>t</i> -test	75
The procedure	76
The actual <i>t</i> -test	81
<i>t</i> -test for means associated with unequal variances	81
The s.e.d.m. when variances are unequal	82
A worked example of the <i>t</i> -test for means associated with unequal variances	85
The paired <i>t</i> -test	87
Pair when possible	90
EXECUTIVE SUMMARY 3 – The <i>t</i> -test	92
Spare-time activities	94

9 One tail or two?	95
Introduction	95
Why is the analysis of variance <i>F</i> -test one-tailed?	95
The two-tailed <i>F</i> -test	96
How many tails has the <i>t</i> -test?	98
The final conclusion on number of tails	99
10 Analysis of variance (ANOVA): what is it? How does it work?	101
Introduction	101
Sums of squares in ANOVA	102
Some 'made-up' variation to analyse by ANOVA	102
The sum of squares table	104
Using ANOVA to sort out the variation in Table C	104
Phase 1	104
Phase 2	105
SqADS: an important acronym	107
Back to the sum of squares table	108
How well does the analysis reflect the input?	109
End phase	109
Degrees of freedom in ANOVA	110
The completion of the end phase	112
The variance ratio	113
The relationship between <i>t</i> and <i>F</i>	114
Constraints on ANOVA	115
Adequate size of experiment	115
Equality of variance between treatments	117
Testing the homogeneity of variance	117
The element of chance: randomisation	118
Comparison between treatment means in ANOVA	119
The least significant difference	121
A caveat about using the LSD	123
EXECUTIVE SUMMARY 4 – The principle of ANOVA	124
11 Experimental designs for analysis of variance (ANOVA)	129
Introduction	129
Fully randomised	130
Data for analysis of a fully randomised experiment	131
Prelims	132
Phase 1	132

Phase 2	133
End phase	133
Randomised blocks	135
Data for analysis of a randomised block experiment	137
Prelims	138
Phase 1	139
Phase 2	140
End phase	141
Incomplete blocks	142
Latin square	145
Data for the analysis of a Latin square	145
Prelims	146
Phase 1	150
Phase 2	150
End phase	151
Further comments on the Latin square design	152
Split plot	154
Types of analysis of variance	154
One- and two-way analysis of variance	155
Fixed-, random-, and mixed-effects analysis of variance	156
EXECUTIVE SUMMARY 5 – Analysis of a one-way randomised block experiment	158
Spare-time activities	159
12 Introduction to factorial experiments	163
What is a factorial experiment?	163
Interaction: what does it mean biologically?	165
If there is no interaction	167
What if there IS interaction?	167
How about a biological example?	168
Measuring any interaction between factors is often the main/only purpose of an experiment	170
How does a factorial experiment change the form of the analysis of variance?	171
Degrees of freedom for interactions	171
The similarity between the <i>residual</i> in Phase 2 and the <i>interaction</i> in Phase 3	172
Sums of squares for interactions	172

13 2-Factor factorial experiments	175
Introduction	175
An example of a 2-factor experiment	175
Analysis of the 2-factor experiment	176
Prelims	176
Phase 1	177
Phase 2	177
End phase (of Phase 2)	178
Phase 3	179
End phase (of Phase 3)	183
Two important things to remember about factorials before tackling the next chapter	185
Analysis of factorial experiments with unequal replication	185
EXECUTIVE SUMMARY 6 – Analysis of a 2-factor randomised block experiment	188
Spare-time activity	190
 14 Factorial experiments with more than two factors – leave this out if you wish!	 191
Introduction	191
Different 'orders' of interaction	191
Example of a 4-factor experiment	192
Prelims	194
Phase 1	196
Phase 2	196
Phase 3	197
To the end phase	205
Spare-time activity	214
 15 Factorial experiments with split plots	 217
Introduction	217
Deriving the split plot design from the randomised block design	218
Degrees of freedom in a split plot analysis	221
Main plots	221
Sub-plots	222
Numerical example of a split plot experiment and its analysis	224
Calculating the sums of squares	225
End phase	229
Comparison of split plot and randomised block experiments	229

Uses of split plot designs	233
Spare-time activity	235
16 The <i>t</i>-test in the analysis of variance	237
Introduction	237
Brief recap of relevant earlier sections of this book	238
Least significant difference test	239
Multiple range tests	240
Operating the multiple range test	242
Testing differences between means	246
My rules for testing differences between means	246
Presentation of the results of tests of differences between means	247
The results of the experiments analysed by analysis of variance in Chapters 11–15	249
Fully randomised design (p. 131)	250
Randomised block experiment (p. 137)	251
Latin square design (p. 146)	253
2-Factor experiment (p. 176)	255
4-Factor experiment (p. 195)	257
Split plot experiment (p. 224)	259
Some final advice	261
Spare-time activities	261
17 Linear regression and correlation	263
Introduction	263
Cause and effect	264
Other traps waiting for you to fall into	264
Extrapolating beyond the range of your data	264
Is a straight line appropriate?	265
The distribution of variability	268
Regression	268
Independent and dependent variables	272
The regression coefficient (<i>b</i>)	272
Calculating the regression coefficient (<i>b</i>)	275
The regression equation	281
A worked example on some real data	282
The data	282
Calculating the regression coefficient (<i>b</i>), i.e. the slope of the regression line	282
Calculating the intercept (<i>a</i>)	284

Drawing the regression line	285
Testing the significance of the slope (b) of the regression	286
How well do the points fit the line? The coefficient of determination (r^2)	290
Correlation	291
Derivation of the correlation coefficient (r)	291
An example of correlation	292
Is there a correlation line?	293
Extensions of regression analysis	296
Nonlinear regression	297
Multiple linear regression	298
Multiple nonlinear regression	300
EXECUTIVE SUMMARY – Linear regression	301
Spare time activities	303
18 Analysis of covariance (ANCOVA)	305
Introduction	305
A worked example of ANCOVA	307
Data: cholesterol levels of subjects given different diets	307
Data: ages of subjects in experiment	308
Regression of cholesterol level on age	309
The structure of the ANCOVA table	312
Total sum of squares	313
Residual sum of squares	314
Corrected means	316
Test for significant difference between means	316
Executive Summary 8 – Analysis of covariance (ANCOVA)	319
Spare-time activity	320
19 Chi-square tests	323
Introduction	323
When not and where not to use χ^2	324
The problem of low frequencies	325
Yates' correction for continuity	325
The χ^2 test for <i>goodness of fit</i>	326
The case of more than two classes	328
χ^2 with heterogeneity	331
Heterogeneity χ^2 Analysis with 'Covariance'	333
Association (or contingency) χ^2	335
2 $\times$ 2 contingency table	336

Fisher's exact test for a 2×2 table	338
Larger contingency tables	340
Interpretation of contingency tables	341
Spare-time activities	343
20 Nonparametric methods (what are they?)	345
Disclaimer	345
Introduction	346
Advantages and disadvantages of parametric and nonparametric methods	347
Where nonparametric methods score	347
Where parametric methods score	349
Some ways data are organised for nonparametric tests	349
The sign test	350
The Kruskal-Wallis analysis of ranks	350
Kendall's rank correlation coefficient	352
The main nonparametric methods that are available	353
Analysis of two replicated treatments as in the <i>t</i> -test (Chapter 8)	353
Analysis of more than two replicated treatments as in the analysis of variance (Chapter 11)	354
Correlation of two variables (Chapter 17)	354
Appendix A How many replicates?	355
Appendix B Statistical tables	365
Appendix C Solutions to spare-time activities	373
Appendix D Bibliography	393
Index	397