## Contents

Preface	v
1 Statistical methodology	1
Defining the problem	1
The population	1
Levels of measurement	2 3 3
Factors and treatments	3
Collecting data	3
The law of large numbers	4
Replication	4
Sampling methodology	5
Primary and secondary data	5
Analysing data	5
Deterministic models	6
Probability	6 7
Expectation	8
Stochastic models	9
Scientific method	10
Statistical inference	10
Method 1: Estimation	10
Method 2: Hypothesis testing	12
Summary Everthor modeling	12
Further reading Exercises	13
Exercises	13
2 Describing data	15
Bar charts, pictograms and pie charts	15
Numerical measures for raw data	19
Frequency tables	21
Grouped data, histograms and ogives	23
Time series, frequency polygons and indices	28
Graphs and scatter diagrams	29
Summary	31
Further reading	31
Exercises	32
3 One sample with one variable	34
The normal distribution	34
Classical sampling theory	40
Sampling with known population standard deviation	41
Sampling with unknown population standard deviation	43

## viii Contents

Testing probability models         47           Yates' correction         49           The binomial distribution         50           The Poisson distribution         52           The Poisson approximation to the binomial distribution         54           Normal approximations to the Poisson and binomial distributions         54           Normal approximations to the Poisson and binomial distributions         54           Nonparametric methods         57           Summary         58           Further reading         59           Exercises         60           4 One sample with two variables         62           Contingency tables         62           Correlation and regression         67           Curvilinear regression         74           Paired data         78           The mitched-pair t test         79           The Milcoxon matched-pair signed-ranks test         79           Rank correlation         81           The McNemar test         82           Summary         83           Further reading         83           Exercises         86           Cluster analysis         87           Fyrincipal components         91	The $t$ and $\chi^2$ distributions	43 47
The binomial distribution         50           The Poisson distribution         52           The Poisson approximation to the binomial distribution         54           Normal approximations to the Poisson and binomial distributions         54           Population proportions         55           Konparametric methods         57           The sign test         57           Summary         58           Further reading         59           Exercises         60           4 One sample with two variables         62           Correlation         66           Curvilinear regression         67           Curvilinear regression         74           Paired data         78           The matched-pair t test         79           The Wilcoxon matched-pair signed-ranks test         79           Rank correlation         81           The McNemar test         82           Summary         83           Exercises         86           Cluster analysis         87           Further reading         83           Exercises         86           Cluster analysis         86           Cluster analysis         87           Princ		
The Poisson distribution         52           The Poisson approximation to the binomial distribution         54           Normal approximations to the Poisson and binomial distributions         54           Population proportions         55           Nonparametric methods         57           The sign test         57           Summary         58           Further reading         59           Exercises         60           4 One sample with two variables         62           Contingency tables         62           Correlation and regression         67           Curvilinear regression         67           Paired data         78           The matched-pair test         79           Rank correlation         81           The Milcoxon matched-pair signed-ranks test         79           Rank correlation         81           The Michemar test         82           Summary         83           Further reading         83           Exercises         86           Cluster analysis         87           Principal components         91           Multivariate regression         95           Selection of variables in a multivariate regression		
The Poisson approximation to the binomial distribution         54           Normal approximations to the Poisson and binomial distributions         55           Population proportions         55           Nonparametric methods         57           The sign test         57           Summary         58           Further reading         59           Exercises         60           4 One sample with two variables         62           Correlation and regression         67           Curvilinear regression         74           Paired data         78           The matched-pair t test         79           The Wilcoxon matched-pair signed-ranks test         79           Rank correlation         81           The McNemar test         82           Summary         83           Evercises         84           5 One sample with several variables         86           Cluster analysis         87           Principal components         96           Multivariate regression         98           Summary         90           Further reading         101           Exercises         102           6 Two samples         103		
Normal approximations to the Poisson and binomial distributions Population proportions S5 Nonparametric methods S7 The sign test S7 Summary S8 Further reading S9 Exercises 60  4 One sample with two variables Contingency tables Contingency tables Correlation and regression Curvilinear regression The matched-pair t test The matched-pair t test The Milcoxon matched-pair signed-ranks test The McNemar test Summary S8 Further reading Exercises  5 One sample with several variables Cluster analysis Frincipal components Multivariate regression Selection of variables in a multivariate regression Summary Selection of variables in a multivariate regression Summary S1 S2 S2 S2 S4 S4 S6 CN S6 S6 CN S7 S7 S8		
Population proportions   55     Nonparametric methods   57     The sign test   57     Summary   58     Further reading   59     Exercises   60     4 One sample with two variables   62     Contingency tables   62     Correlation and regression   67     Curvilinear regression   74     Paired data   78     The matched-pair t test   79     Rank correlation matched-pair signed-ranks test   79     Rank correlation   81     The McNemar test   82     Summary   83     Further reading   83     Exercises   84    5 One sample with several variables   86     Cluster analysis   87     Principal components   97     Multivariate regression   98     Sultivariate regression   98     Sultivariate regression   98     Summary   100     Further reading   101     Exercises   102     6 Two samples   103     Testing proportions   103     Testing proportions   103     Testing proportions   105     The two-sample t test   107     The analysis of covariance   108     The two-sample t test   107     The analysis of covariance   108     Population size estimation   113     Method of moments   113     Method of moments   113     Method of maximum likelihood   114     Summary   115     Further reading   115     Further readi		
Nonparametric methods         57           The sign test         57           Summary         58           Further reading         59           Exercises         60           4 One sample with two variables         62           Correlation and regression         67           Curvilinear regression         74           Paired data         78           The matched-pair t test         79           The Wilcoxon matched-pair signed-ranks test         79           Rank correlation         81           The McNemar test         82           Summary         83           Further reading         83           Exercises         84           5 One sample with several variables         86           Cluster analysis         87           Principal components         91           Multivariate regression         95           Selection of variables in a multivariate regression         95           Selection of variables in a multivariate regression         95           Summary         100           Further reading         101           Exercises         103           Testing proportions         103           The two		
The sign test         57           Summary         58           Further reading         59           Exercises         60           4 One sample with two variables         62           Contingency tables         62           Correlation and regression         67           Curvilinear regression         74           Paired data         78           The matched-pair t test         79           The Wilcoxon matched-pair signed-ranks test         79           Rank correlation         81           The McNemar test         82           Summary         83           Further reading         83           Exercises         84           5 One sample with several variables         86           Cluster analysis         87           Principal components         91           Multivariate regression         98           Summary         100           Further reading         101           Exercises         102           6 Two samples in a multivariate regression         98           Summary         103           Testing proportions         103           The two-sample t test         107		
Further reading         59           Exercises         60           4 One sample with two variables         62           Contingency tables         62           Correlation and regression         74           Paired data         78           The matched-pair t test         79           The Wilcoxon matched-pair signed-ranks test         79           Rank correlation         81           The McNemar test         82           Summary         83           Further reading         83           Exercises         84           5 One sample with several variables         86           Cluster analysis         87           Principal components         91           Multivariate regression         95           Selection of variables in a multivariate regression         98           Summary         100           Further reading         101           Exercises         102           6 Two samples         103           Testing proportions         103           Testing proportions         103           The Mann-Whitney test         105           The analysis of covariance         106           Population size esti	The sign test	57
Exercises       60         4 One sample with two variables       62         Contingency tables       62         Correlation and regression       74         Curvilinear regression       74         Paired data       78         The matched-pair t test       79         The Wilcoxon matched-pair signed-ranks test       79         Rank correlation       81         The McNemar test       82         Summary       83         Further reading       83         Exercises       84         5 One sample with several variables       86         Cluster analysis       87         Principal components       91         Multivariate regression       95         Selection of variables in a multivariate regression       98         Summary       100         Further reading       101         Exercises       102         6 Two samples       103         Testing proportions       103         The two-sample t test       105         The two-sample t test       105         The malysis of covariance       108         Population size estimation       113         Method of maximum likelihood	Summary	58
4 One sample with two variables       62         Contingency tables       62         Correlation and regression       67         Curvilinear regression       74         Paired data       78         The matched-pair t test       79         The Wilcoxon matched-pair signed-ranks test       79         Rank correlation       81         The McNemar test       82         Summary       83         Further reading       83         Exercises       86         Cluster analysis       87         Principal components       91         Multivariate regression       95         Selection of variables in a multivariate regression       98         Summary       100         Further reading       101         Exercises       102         6 Two samples       103         Testing proportions       103         The Mann-Whitney test       105         The two-sample t test       105         The analysis of covariance       108         Population size estimation       113         Method of moments       113         Method of maximum likelihood       114         Summary       1		59
Contingency tables         62           Correlation and regression         67           Curvilinear regression         74           Paired data         78           The matched-pair t test         79           The Wilcoxon matched-pair signed-ranks test         79           Rank correlation         81           The McNemar test         82           Summary         83           Further reading         83           Exercises         86           Cluster analysis         87           Principal components         91           Multivariate regression         95           Selection of variables in a multivariate regression         98           Summary         100           Further reading         101           Exercises         102           6 Two samples         103           Testing proportions         103           The Mann-Whitney test         105           The two-sample t test         107           The analysis of covariance         108           Population size estimation         113           Method of moments         113           Method of maximum likelihood         114           Summary	Exercises	60
Contingency tables         62           Correlation and regression         67           Curvilinear regression         74           Paired data         78           The matched-pair t test         79           The Wilcoxon matched-pair signed-ranks test         79           Rank correlation         81           The McNemar test         82           Summary         83           Further reading         83           Exercises         84           5 One sample with several variables         86           Cluster analysis         87           Principal components         91           Multivariate regression         95           Selection of variables in a multivariate regression         98           Summary         100           Further reading         101           Exercises         102           6 Two samples         103           Testing proportions         103           The Mann-Whitney test         105           The wo-sample t test         107           The analysis of covariance         108           Population size estimation         113           Method of maximum likelihood         114	4 One sample with two variables	62
Curvilinear regression         74           Paired data         78           The matched-pair t test         79           The Wilcoxon matched-pair signed-ranks test         79           Rank correlation         81           The McNemar test         82           Summary         83           Further reading         83           Exercises         84           5 One sample with several variables         86           Cluster analysis         87           Principal components         91           Multivariate regression         95           Selection of variables in a multivariate regression         98           Summary         100           Further reading         101           Exercises         102           6 Two samples         103           Testing proportions         103           The two-sample t test         105           The analysis of covariance         108           Population size estimation         113           Method of moments         113           Method of maximum likelihood         114           Summary         114           Further reading         115	Contingency tables	62
Paired data The matched-pair t test The Wilcoxon matched-pair signed-ranks test The Wilcoxon matched-pair signed-ranks test Rank correlation The McNemar test Summary Sultivariate regression Summary		67
The matched-pair t test         79           The Wilcoxon matched-pair signed-ranks test         79           Rank correlation         81           The McNemar test         82           Summary         83           Further reading         83           Exercises         84           5 One sample with several variables         86           Cluster analysis         87           Principal components         91           Multivariate regression         95           Selection of variables in a multivariate regression         98           Summary         100           Further reading         101           Exercises         102           6 Two samples         103           Testing proportions         103           The Mann-Whitney test         105           The two-sample t test         107           The analysis of covariance         108           Population size estimation         113           Method of moments         113           Method of maximum likelihood         114           Summary         114           Further reading         115		74
The Wilcoxon matched-pair signed-ranks test 79 Rank correlation 81 The McNemar test 82 Summary 83 Further reading 83 Exercises 84  5 One sample with several variables 85 Cluster analysis 87 Principal components 95 Selection of variables in a multivariate regression 95 Summary 100 Further reading 101 Exercises 102  6 Two samples 103 Testing proportions 103 The Mann-Whitney test 105 The two-sample t test 107 The analysis of covariance 108 Population size estimation 113 Method of moments 113 Method of maximum likelihood 114 Summary 114 Further reading 115		
Rank correlation The McNemar test Summary Further reading Exercises  5 One sample with several variables Cluster analysis Principal components Multivariate regression Selection of variables in a multivariate regression Summary Further reading Exercises  6 Two samples 102  6 Two samples 103 Testing proportions 104 The Mann-Whitney test 105 The two-sample t test 107 The analysis of covariance Population size estimation Method of moments Method of moments Method of maximum likelihood 114 Summary 115 Further reading 115		
The McNemar test       82         Summary       83         Further reading       83         Exercises       84         5 One sample with several variables       86         Cluster analysis       87         Principal components       91         Multivariate regression       95         Selection of variables in a multivariate regression       98         Summary       100         Further reading       101         Exercises       102         6 Two samples       103         Testing proportions       103         The Mann-Whitney test       105         The two-sample t test       107         The analysis of covariance       108         Population size estimation       113         Method of moments       113         Method of maximum likelihood       114         Summary       114         Further reading       115		
Summary         83           Further reading         83           Exercises         84           5 One sample with several variables         86           Cluster analysis         87           Principal components         91           Multivariate regression         95           Selection of variables in a multivariate regression         98           Summary         100           Further reading         101           Exercises         102           6 Two samples         103           Testing proportions         103           The Mann-Whitney test         105           The two-sample t test         107           The analysis of covariance         108           Population size estimation         113           Method of moments         113           Method of maximum likelihood         114           Summary         114           Further reading         115		
Further reading Exercises  5 One sample with several variables Cluster analysis Principal components Multivariate regression Selection of variables in a multivariate regression Summary Further reading Exercises  100  6 Two samples 101 Exercises 102  6 Two samples 103 Testing proportions 103 The Mann-Whitney test 105 The two-sample t test 107 The analysis of covariance Population size estimation 113 Method of moments Method of maximum likelihood 114 Summary 115 Further reading 115		
Exercises845 One sample with several variables86Cluster analysis87Principal components91Multivariate regression95Selection of variables in a multivariate regression98Summary100Further reading101Exercises1026 Two samples103Testing proportions103The Mann-Whitney test105The two-sample t test107The analysis of covariance108Population size estimation113Method of moments113Method of maximum likelihood114Summary114Further reading115		
5 One sample with several variables86Cluster analysis87Principal components91Multivariate regression95Selection of variables in a multivariate regression98Summary100Further reading101Exercises1026 Two samples103Testing proportions103The Mann-Whitney test105The two-sample t test107The analysis of covariance108Population size estimation113Method of moments113Method of maximum likelihood114Summary114Further reading115		
Cluster analysis 87 Principal components 91 Multivariate regression 95 Selection of variables in a multivariate regression 98 Summary 100 Further reading 101 Exercises 102  6 Two samples 103 Testing proportions 103 The Mann-Whitney test 105 The two-sample t test 107 The analysis of covariance 108 Population size estimation 113 Method of moments 113 Method of maximum likelihood 114 Summary 114 Further reading 115		0.
Principal components Multivariate regression Selection of variables in a multivariate regression Summary Further reading Exercises  6 Two samples Testing proportions The Mann-Whitney test The two-sample t test The analysis of covariance Population size estimation Method of moments Method of maximum likelihood Summary Further reading  91 100 100 100 100 101 102 103 103 103 103 105 104 105 105 106 107 107 108 108 109 109 109 109 109 109 109 109 109 109		86
Multivariate regression       95         Selection of variables in a multivariate regression       98         Summary       100         Further reading       101         Exercises       102         6 Two samples       103         Testing proportions       103         The Mann-Whitney test       105         The two-sample t test       107         The analysis of covariance       108         Population size estimation       113         Method of moments       113         Method of maximum likelihood       114         Summary       114         Further reading       115		87
Selection of variables in a multivariate regression98Summary100Further reading101Exercises1026 Two samples103Testing proportions103The Mann-Whitney test105The two-sample $t$ test107The analysis of covariance108Population size estimation113Method of moments113Method of maximum likelihood114Summary114Further reading115		91
Summary       100         Further reading       101         Exercises       102         6 Two samples       103         Testing proportions       103         The Mann-Whitney test       105         The two-sample t test       107         The analysis of covariance       108         Population size estimation       113         Method of moments       113         Method of maximum likelihood       114         Summary       114         Further reading       115		
Further reading       101         Exercises       102         6 Two samples       103         Testing proportions       103         The Mann-Whitney test       105         The two-sample t test       107         The analysis of covariance       108         Population size estimation       113         Method of moments       113         Method of maximum likelihood       114         Summary       114         Further reading       115		
Exercises       102         6 Two samples       103         Testing proportions       103         The Mann-Whitney test       105         The two-sample t test       107         The analysis of covariance       108         Population size estimation       113         Method of moments       113         Method of maximum likelihood       114         Summary       114         Further reading       115		
6 Two samples       103         Testing proportions       103         The Mann-Whitney test       105         The two-sample t test       107         The analysis of covariance       108         Population size estimation       113         Method of moments       113         Method of maximum likelihood       114         Summary       114         Further reading       115		
Testing proportions103The Mann-Whitney test105The two-sample $t$ test107The analysis of covariance108Population size estimation113Method of moments113Method of maximum likelihood114Summary114Further reading115	Lacreises	102
The Mann-Whitney test105The two-sample $t$ test107The analysis of covariance108Population size estimation113Method of moments113Method of maximum likelihood114Summary114Further reading115		103
The two-sample $t$ test107The analysis of covariance108Population size estimation113Method of moments113Method of maximum likelihood114Summary114Further reading115		
The analysis of covariance108Population size estimation113Method of moments113Method of maximum likelihood114Summary114Further reading115		
Population size estimation 113 Method of moments 113 Method of maximum likelihood 114 Summary 114 Further reading 115		
Method of moments113Method of maximum likelihood114Summary114Further reading115		
Method of maximum likelihood 114 Summary 114 Further reading 115		
Summary 114 Further reading 115		
Further reading 115		
· · · · · · · · · · · · · · · · · · ·		

	Contents	ix
7 Maria annual a	1	17
7 Many samples One-way analysis of variance		17 17
The Kruskal-Wallis test		22
Two-way analysis of variance without replication		23
Two-way analysis of variance with replication	1	26
Summary	1	30
Further reading	_	31
Exercises	1	31
8 Experimental design	1	33
Blocking	_	33
Factorial designs		36
Latin square designs	_	40
Nested or hierarchical designs		42
Summary	_	44
Further reading	_	45
Exercises	1	143
9 Sampling and sample surveys		48
Running a survey		148
Selecting a sample	_	150 150
Random sampling Systematic sampling		150
Stratified sampling		150
Cluster sampling		151
Other methods		151
Sample size	1	151
Designing a questionnaire		152
Question structure		152
Sequencing the questions		153
Question wording		154
Presentation		155 155
Data-processing aspects		157
Confidentiality Cosmetic responses		157
Analysing the data		157
Summary		159
Further reading	1	160
Exercises	1	160
10 Quality control	1	62
Fitting a normal distribution		62
Life testing		165
Control charts		168
Acceptance sampling		172
Summary		175
Further reading		175
Exercises	1	176

## x Contents

11 Computing aspects of statistics	177		
Multivariate analysis with matrices Regression Measures of distance Principal components	178		
	178		
	180		
	181		
Computer programs for matrix manipulation	181		
Probabilities for test statistics  The normal distribution  The chi-squared distribution  The F distribution  The t distribution	182		
	182		
	186		
	189		
	193		
The Wilcoxon matched-pair signed-ranks statistic	193		
Spearman's rank correlation coefficient Summary Further reading	195 196 197		
		Exercises	197
		Bibliography	198
Answers to exercises	201		
Index	201		