
Contents

Preface	v
1 Statistical methodology	1
Defining the problem	1
The population	1
Levels of measurement	2
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
Expectation	7
Stochastic models	8
Scientific method	9
Statistical inference	10
Method 1: Estimation	10
Method 2: Hypothesis testing	10
Summary	12
Further reading	12
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

The t and χ^2 distributions	43
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
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
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
Exercises	115

7 Many samples	117
One-way analysis of variance	117
The Kruskal–Wallis test	122
Two-way analysis of variance without replication	123
Two-way analysis of variance with replication	126
Summary	130
Further reading	131
Exercises	131
8 Experimental design	133
Blocking	133
Factorial designs	136
Latin square designs	140
Nested or hierarchical designs	142
Summary	144
Further reading	145
Exercises	145
9 Sampling and sample surveys	148
Running a survey	148
Selecting a sample	150
Random sampling	150
Systematic sampling	150
Stratified sampling	150
Cluster sampling	151
Other methods	151
Sample size	151
Designing a questionnaire	152
Question structure	152
Sequencing the questions	153
Question wording	154
Presentation	155
Data-processing aspects	155
Confidentiality	157
Cosmetic responses	157
Analysing the data	157
Summary	159
Further reading	160
Exercises	160
10 Quality control	162
Fitting a normal distribution	162
Life testing	165
Control charts	168
Acceptance sampling	172
Summary	175
Further reading	175
Exercises	176

11 Computing aspects of statistics	177
Multivariate analysis with matrices	178
Regression	178
Measures of distance	180
Principal components	181
Computer programs for matrix manipulation	181
Probabilities for test statistics	182
The normal distribution	182
The chi-squared distribution	186
The F distribution	189
The t distribution	193
The Wilcoxon matched-pair signed-ranks statistic	193
Spearman's rank correlation coefficient	195
Summary	196
Further reading	197
Exercises	197
Bibliography	198
Answers to exercises	201
Index	205