Contents

Preface vii

1 Representation and summary of data 1

Discrete data 1
Continuous data 2
Stem and leaf diagrams (stemplots) 4
Ways of grouping data 9
Histograms 11
Frequency polygons 17
Frequency curves 19
Circular diagrams or pie charts 24
The mean 28
Variability of data 38
The standard deviation, s, and the variance, s^2 38
Combining sets of data 47
Scaling sets of data 51
Using a method of coding to find the mean and standard deviation 56
Cumulative frequency 58
Cumulative percentage frequency diagrams 65
Median, quartiles and percentiles 68
Skewness 84
The normal distribution 89
Box and whisker diagrams (box plots) 92
Summary 102

2 Regression and correlation 118

Scatter diagrams 118
Regression function 119
Linear correlation and regression lines 119
The product-moment correlation coefficient, r 139
Spearman's coefficient of rank correlation, r_s 146
Summary 154

3 Probability 168

Experimental probability 169
Probability when outcomes are equally likely 171
Subjective probabilities 171
Probability notation and probability laws 171
Illustrating two or more events using Venn diagrams 175
Probability rule for combined events 175
Exclusive (or mutually exclusive) events 179

Exhaustive events 180
Conditional probability 182
Independent events 185
Probability trees 193
Bayes' Theorem 197
Some useful methods 204
Arrangements 206
Permutations of r objects from n objects 214
Combinations of r objects from n objects 214
Summary 221

4 Probability distributions I – discrete variables 233

Probability distributions 233
Expectation of X, $E(X)$ 237
Expectation of any function of X, $E(g(X))$ 245
Variance, $\mathrm{Var}(X)$ or $V(X)$ 248
The Cumulative distribution function, $F(x)$ 253
Two independent random variables 256
Distribution of $X_1 + X_2 + \cdots + X_n$ 258
Comparing the distributions of $X_1 + X_2$ and $2X$ 259
Summary 262

5 Special discrete probability distributions 270

The uniform distribution 270
The geometric distribution 271
Expectation and variance of the geometric distribution 275
The binomial distribution 278
Expectation and variance of the binomial distribution 286
The Poisson distribution 291
Using the Poisson distribution as an approximation to the binomial distribution 299
The sum of independent Poisson variables 301
Summary 304

6 Probability distributions II – continuous variables 314

Continuous random variables 314
Probability density function (p.d.f.) 314
Expectation of X, $E(X)$ 320
Expectation of any function of X 324
Variance of X, $\mathrm{Var}(X)$ 327
The mode 329
Cumulative distribution function $F(x)$ 334
Obtaining the p.d.f., $f(x)$, from the cumulative distribution function 341
The continous uniform (or rectangular) distribution 345
Expectation and variance of the uniform distribution 347
The cumulative distribution function, $F(x)$, for a uniform distribution 348
Summary 351

7	**The normal distribution**	**360**
	Finding probabilities	361
	The standard normal variable, Z	361
	Using standard normal tables	362
	Using standard normal tables for any normal variable, X	368
	Using the standard normal tables in reverse to find z when $\Phi(z)$ is known	371
	Using the tables in reverse for *any* normal variable, X	374
	Value of μ or σ or both	378
	The normal approximation to the binomial distribution	382
	Continuity corrections	383
	Deciding when to use a normal approximation and when to use a Poisson approximation for a binomial distribution	387
	The normal approximation to the Poisson distribution	390
	Summary	392
8	**Linear combinations of normal variables**	**403**
	The sum of independent normal variables	403
	The difference of independent normal variables	407
	Multiples of independent normal variables	410
	Summary	414
9	**Sampling and estimation**	**421**
	Sampling	421
	Surveys	422
	Sampling methods	424
	Simulating random samples from given distributions	431
	Sample statistics	436
	The distribution of the sample mean	436
	Central limit theorem	441
	The distribution of the sample proportion, p	444
	Unbiased estimates of population parameters	447
	Point estimates	447
	Interval estimates	449
	The t-distribution	462
	Confidence intervals for the population proportion, p	469
	Summary	472
10	**Hypothesis tests: discrete distributions**	**483**
	Hypothesis test for a binomial proportion, p (small sample size)	483
	Procedure for carrying out a hypothesis test	486
	One-tailed and two-tailed tests	489
	Summary of stages of a hypothesis (significance) test	492
	Type I and Type II errors	493
	Significance test for a Poisson mean λ	496
	Summary of stages of a significance test	500
	Summary of Type I and Type II errors	501

11 Hypothesis testing (z-tests and t-tests) 507

Hypothesis testing 507
One-tailed and two-tailed tests 511
Critical z-values 512
Summary of critical values and rejection criteria 513
Stages in the hypothesis test 513
Hypothesis test 1: testing μ (the mean of a population) 514
Type I and Type II errors 520
Hypothesis test 2: testing a binomial proportion p when n is large 528
Hypothesis test 3: testing $\mu_1 - \mu_2$, the difference between means of two normal populations 534
Summary 547

12 The χ^2 significance test 560

The χ^2 significance test 560
Performing a χ^2 goodness-of-fit test 563
Summary of the procedure for performing a χ^2 goodness-of-fit test 566
Test 1 – goodness-of-fit test for a uniform distribution 567
Test 2 – goodness-of-fit test for a distribution in a given ratio 568
Test 3 – goodness-of-fit test for a binomial distribution 571
Test 4 – goodness-of-fit test for a Poisson distribution 573
Test 5 – goodness-of-fit test for a normal distribution 576
Summary of the number of degrees of freedom for a goodness-of-fit test 579
The χ^2 significance test for independence 582
Summary 590

13 Significance tests for correlation coefficients 600

Significance tests for correlation coefficients 600
Test for the product-moment correlation coefficient, r 600
Spearman's coefficient of rank correlation, r_s 605
Summary 608

ICT statistics supplement 617

Appendix 645

Cumulative binomial probabilities 645
Cumulative Poisson probabilities 648
The standard normal distribution function 649
Critical values for the normal distribution 649
Critical values for the t-distribution 650
Critical values for the χ^2 distribution 651
Critical values for correlation coefficients 652
Random numbers 653

Answers 655