

CONTENTS

Introduction	1
CHAPTER 1	
Data Collection and its Graphical Presentation	5
1.1 Introduction to statistics	6
1.2 Data collection: Samples, surveys and experiments	8
1.3 Some sources of statistical data	16
1.4 Sorting and classifying data	18
1.5 Charts: Bar charts, pie charts and plotting in Excel	23
1.6 Graphs: Histograms and Ogives. Graphs in Excel	42
1.7 Line and Lorenz graphs	52
1.8 Misleading graphs	59
CHAPTER 2	
Descriptive Statistics	63
2.1 Summary statistics for raw data: Mean, quartiles and mode	64
2.2 Summary statistics for grouped data: Mean, quartiles and mode	72
2.3 Measures of dispersion for raw data. Variance, QD, IQR	85
2.4 Measures of dispersion for grouped data. Variance, QD, IQR	92
2.5 Use of calculator for descriptive statistics	94
2.6 Other descriptive statistics. CV, skewness and box plots	96
2.7 Descriptive statistics in Excel	101
CHAPTER 3	
Regression and Correlation: Introduction	113
3.1 Introduction to regression: Scatter plots and lines	114
3.2 The least-squares line. Criteria and equation for the best fit line	118
3.3 Excel: XY (scatter) plots, the least-squares lines and formulae	127
3.4 Coefficient of determination and correlation	133
3.5 Rank correlation. Calculate and interpret rank correlation	144
3.6 Use the calculator for linear regression and correlation	146
3.7 Why bother with formulae?	154

CHAPTER 4	
Probability	159
4.1 Introduction to probability	159
4.2 Multiplication and addition rules for probability	170
4.3 Joint, marginal and conditional probability	184
4.4 Bayes' Rule	192
CHAPTER 5	
Introduction to Probability Distributions	205
5.1 Introduction to probability distributions and random variables	206
5.2 The Binomial probability distribution	212
5.3 The Poisson probability distribution	228
5.4 The Normal probability distribution	236
5.5 Expected values (mathematical expectation)	253
CHAPTER 6	
Sampling Distributions for Means and Proportions	265
6.1 Statistical inference and the sampling distribution of the mean	265
6.2 Sampling distribution of proportions	275
6.3 Some desirable properties of estimators	282
CHAPTER 7	
Confidence Intervals for Means and Proportions	287
7.1 Confidence intervals for the mean	287
7.2 Confidence intervals for proportions	297
7.3 The precision of confidence intervals for the mean	300
7.4 Confidence intervals for differences between means and proportions	304
CHAPTER 8	
Tests of Hypothesis for Means and Proportions	319
8.1 Hypothesis tests for means	319
8.2 Hypothesis tests for proportions	334
8.3 Hypothesis tests for the difference between means and proportions	337
8.4 Minitab and Excel for confidence intervals and tests of hypothesis	349
CHAPTER 9	
Inference from Small Samples. Confidence Intervals and Tests of Hypothesis	359
9.1 Inference from small samples: Normal populations, σ known	360
9.2 The Student's t distribution	362
9.3 Inference from small samples: Normal populations, σ NOT known	365
9.4 Difference between means. Small independent samples	370
9.5 F -test for equality of two variances	378
9.6 Difference between means, paired samples	382

CONTENTS

CHAPTER 10	
Analysis of Variance	393
10.1 The rationale behind one-way analysis of variance	393
10.2 One-way analysis of variance	397
10.3 Two-way ANOVA	419
10.4 ANOVA and design of experiment	430
10.5 Excel and Minitab for ANOVA	435
CHAPTER 11	
Chi-squared Tests	453
11.1 Introduction	454
11.2 The χ^2 probability distribution	454
11.3 Contingency tables	455
11.4 χ^2 tests for independence (no association)	456
11.5 χ^2 test for homogeneous populations	464
11.6 χ^2 tests for the equality of several proportions	466
11.7 Goodness of fit tests	478
11.8 χ^2 tests in Minitab and Excel	490
CHAPTER 12	
Regression Analysis	499
12.1 The simple linear regression model	499
12.2 Inference about the population slope (rate of change)	504
12.3 Confidence intervals and prediction intervals at $x = x_0$	513
12.4 Checks on the model assumptions based on residuals plots	518
12.5 Regression analysis in Minitab and Excel	521
12.6 Multiple regression	534
APPENDIX A: Technicalities and Conventions for Defining Class Intervals, Mid-Interval, Widths, Limits, Boundaries	545
APPENDIX B: Formulae for Calculating the Quartiles for Grouped Data	547
APPENDIX C: Outline of Derivation of Formulae (3.2) and (3.3) for the Slope and Intercept of the Least-Squares Line, $Y = a + bX$	549
APPENDIX D: Brief Review of the Mathematics for the Binomial	551
APPENDIX E: The Number e	555
APPENDIX F: Calculation of Mean and Variance of Proportions by Expected Values	557
APPENDIX G: Confidence Intervals for Means and Proportions	559
APPENDIX H: Degrees of Freedom	561
APPENDIX I: Notes on Summations and Double Summations	563

APPENDIX J: Expressing the Estimates of Variance as Sums of Squares Divided by Degrees of Freedom in Their Simplest Form	567
APPENDIX K: Sums of Squares for the One-Way ANOVA	571
APPENDIX L: Fitted Values and Residuals	573
APPENDIX M: Sums of Squares Identity for Two-Way ANOVA	575
Chapter Solutions	577
Tables	637
Table 1 Cumulative Binomial Probability tables	637
Table 2 Cumulative Poisson Probabilities $\sum_{x=0}^{x=r} P(x \leq r) = \sum_{x=0}^{x=r} \frac{\lambda^x e^{-\lambda}}{x!}$	639
Table 3 Normal probability Distribution	641
Table 4 Percentage point for the Normal probability distribution	642
Table 5 Percentage points of the Student's <i>t</i> -distribution	643
Table 6 0.5% points of the <i>F</i> -distribution	644
Table 7 Percentage points for the Chi-squared distribution	648
Table 8 Table of random numbers	649
Glossary	651
Index	653