Contents

Preface

Part 1 An introduction to statistical analysis

1 Variables and their measurement 3
The conceptualization and operationalization of variables 4
Scales of measurement 7
Levels of measurement 8
Univariate, bivariate, and multivariate analysis 11
Descriptive statistics 14
Exercises 15

2 Setting up an SPSS data file 17
Obtaining a copy of SPSS 17
Alternatives to SPSS 17
Options for data entry in SPSS 18
The SPSS Data Editor 19
Assigning a variable name 21
Setting the data type 22
Setting the data width and decimal places 23
Defining variable labels 23
Defining value labels 24
Setting missing values 25
Setting the column format and alignment 26
Specifying the level of measurement 27
Generating variable definitions in SPSS 28
The SPSS Viewer window 32
Saving a data file 32
Data entry 33
Checking for incorrect values: Data cleaning 35
Summary 35
Exercises 35

Part 2 Descriptive statistics: Graphs and tables

3 The graphical description of data 39
Some general principles 39
Pie graphs 40
Bar graphs 42
Histograms and polygons 44
Interpreting a univariate distribution 46
Graphing two variables 47
Common problems and misuses of graphs 50
Exercises 53

4 The tabular description of data 55
Listed data tables 55
Simple frequency tables 55
Relative frequency tables: percentages and proportions 57
Cumulative frequency tables 60
Class intervals 61
Percentiles 64
Frequency tables using SPSS 65
Valid cases and missing values 67
Improving the look of tables 67
Exercises 68

5 Using tables to investigate the relationship between variables: Crosstabulations 70
Crosstabulations as descriptive statistics 70
Types of data suitable for crosstabulations 72
Crosstabulations with relative frequencies 73
Crosstabulations using SPSS 74
Interpreting a crosstabulation: The pattern and strength of a relationship 75
Interpreting a crosstabulation when both variables are at least ordinal 76
Summary 78
Exercises 78

6 Measures of association for crosstabulations: Nominal data 81
Measures of association as descriptive statistics 81
Measures of association for nominal scales 83
Properties of lambda 86
Lambda using SPSS 87
Limitations on the use of lambda 90
Standardizing table frequencies 92
Exercises 93

7 Measures of association for crosstabulations: Ranked data 95
Data considerations 95
Concordant pairs 96
Discordant pairs 97
Measures of association for ranked data 98
Gamma 99
Somers' *d* 101
Kendall's tau-*b* 102
Kendall's tau-*c* 102
Measures of association using SPSS 102
Summary 107
Exercises 107

8 Multivariate analysis of crosstabs: Elaboration 110
Direct relationship 110
Elaboration of crosstabs using SPSS 112
Partial gamma 113
Spurious or intervening relationship? 114
Conditional relationship 115
Summary 117
Exercises 118

Part 3 Descriptive statistics: Numerical measures

9 Measures of central tendency 123
Measures of central tendency 123
The mode 124
The median 125
The mean 126
Choosing a measure of central tendency 128
Measures of central tendency using SPSS: Univariate analysis 129
Measures of central tendency using SPSS: Bivariate and multivariate analysis 132
Summary 133
Exercises 134

10 Measures of dispersion 136
The range 136
The interquartile range 137
The standard deviation 138
Coefficient of relative variation 140
Index of qualitative variation 141
Measures of dispersion using SPSS 145
Summary 145
Exercises 146

11 The normal curve 147
The normal distribution 147
Using normal curves to describe a distribution 150
z-scores 151
Normal curves on SPSS 157
Exercises 159

12 Correlation and regression 161
Scatter plots 161
Linear regression 162
Pearson's product moment correlation coefficient 169
Explaining variance: The coefficient of determination 170
Plots, correlation, and regression using SPSS 172
The assumptions behind regression analysis 177
Spearman's rank-order correlation coefficient 179
Spearman's rho using SPSS 180
Correlation where the independent variable is categorical: Eta 182
Summary 183
Exercises 183

13 Multiple regression 187
Introduction to multiple regression 188
Multiple regression with SPSS 190
Testing for the significance of the multivariate model 193
Alternative methods for selecting variables in the regression model 193
Stepwise regression 194
Extending the basic regression analysis: Adding categorical independent variables 197
Further extensions to the basic regression analysis: Hierarchical regression 198
The assumptions behind multiple regression 198
Exercises 199

Part 4 Inferential statistics: Tests for a mean

14 Sampling distributions 203
Random samples 204
The sampling distribution of a sample statistic 205
The central limit theorem 210
Generating random samples using SPSS 210
Summary 212
Exercises 212

15 Introduction to hypothesis testing and the one sample *z*-test for a mean 214
Step 1: State the null and alternative hypotheses 217
Step 2: Choose the test of significance 219
Step 3: Describe the sample and derive the *p*-score 220
Step 4: Decide at what alpha level, if any, the result is statistically significant 222
Step 5: Report results 224
What does it mean when we 'fail to reject the null hypothesis'? 226
What does it mean to 'reject the null hypothesis'? 226
A two-tail *z*-test for a single mean 227
The debate over one-tail and two-tail tests of significance 228
A one-tail *z*-test for a single mean 229
Summary 230
Appendix: Hypothesis testing using critical values of the test statistic 230
Exercises 231

16 The one sample *t*-test for a mean 233
The Student's *t*-distribution 233
The one sample *t*-test for a mean 234
The one sample *t*-test using SPSS 238
Summary 239
Exercises 240

17 Inference using estimation and confidence intervals 242
The sampling distribution of sample means 242
Estimation 243
Changing the confidence level 247
Changing the sample size 250
Estimation using SPSS 250
Confidence intervals and hypothesis testing 252
Exercises 253

18 The two samples *t*-test for the equality of means 255
Dependent and independent variables 256
The sampling distribution of the difference between two means 257
The two samples *t*-test for the equality of means 259
The two samples *t*-test using SPSS 261
Exercises 264

19 The *F*-test for the equality of more than two means: Analysis of variance 266
The one-way analysis of variance *F*-test 269
ANOVA using SPSS 272
Summary 277
Exercises 279

20 The two dependent samples *t*-test for the mean difference 280
Dependent and independent samples 280
The two dependent samples *t*-test for the mean difference 281
The two dependent samples *t*-test using SPSS 283
Exercises 286

Part 5 Inferential statistics: Tests for frequency distributions

21 One sample tests for a binomial distribution 291
Data considerations 291
The sampling distribution of sample percentages 292
The *z*-test for a binomial percentage 293
The *z*-test for a binomial percentage using SPSS 295
Estimating a population percentage 297
The runs test for randomness 299
The runs test using SPSS 302
Exercises 303

22 One sample tests for a multinomial distribution 305
The chi-square goodness-of-fit test 305
Chi-square goodness-of-fit test using SPSS 308
The chi-square goodness-of-fit test for normality 312
Summary 313
Exercises 314

23 The chi-square test for independence 316
The chi-square test and other tests of significance 316
Statistical independence 317
The chi-square test for independence 317
The distribution of chi-square 322
The chi-square test using SPSS 323
Problems with small samples 328
Problems with large samples 329
Appendix: hypothesis testing for two percentages 331
Exercises 333

24 Frequency tests for two dependent samples 335
The McNemar chi-square test for change 335
The McNemar test using SPSS 337
The sign test 338
Summary 340
Exercises 340

Part 6 Inferential statistics: Other tests of significance

25 Rank-order tests for two or more samples 343
Data considerations 343
The rank sum and mean rank as descriptive statistics 344
The *z*-test for the rank sum for two independent samples 348
Wilcoxon's rank sum *z*-test using SPSS 352
The Wilcoxon signed-ranks *z*-test for two dependent samples 353
The Wilcoxon signed-ranks test using SPSS 356

Other non-parametric tests for two or more samples 357
Appendix: the Mann–Whitney *U* test 358
Exercises 359

26 The *t*-test for a correlation coefficient 362
The *t*-test for Pearson's correlation coefficient 362
Testing the significance of Pearson's correlation coefficient using SPSS 364
The *t*-test for Spearman's rank-order correlation coefficient 365
Testing the significance of Spearman's correlation coefficient using SPSS 366
Testing for significance in multiple regression 367
Exercises 368

Appendix 369
Table A1 Area under the standard normal curve 369
Table A2 Critical values for *t*-distributions 370
Table A3 Critical values for *F*-distributions ($\alpha = 0.05$) 371
Table A4 Critical values for chi-square distributions 372
Table A5 Sampling errors for a binomial distribution (95% confidence level) 373
Table A6 Sampling errors for a binomial distribution (99% confidence level) 373

Key equations 374

Glossary 379

Answers 383

Index 397