

CONTENTS

| | |
|---|-----------|
| <i>About the author</i> | xiii |
| <i>About the online resources</i> | xv |
| <i>Acknowledgements</i> | xvii |
| <i>Introduction</i> | xix |
| | |
| Part One Understanding quantitative data and R | 1 |
| | |
| 1 Introduction to information, knowledge and quantitative data | 3 |
| 1.1 Quantitative data, information and knowledge in education | 4 |
| 1.1.1 What is quantitative data? | 5 |
| 1.2 Measurement and scales of measurement | 6 |
| 1.2.1 What is a measurement? What is a scale? | 6 |
| 1.3 From concepts to constructs and variables | 8 |
| 1.4 Types of variables | 9 |
| 1.5 Quantitative data and R | 10 |
| 1.5.1 Why use R? | 10 |
| Further reading | 11 |
| | |
| 2 An introduction to R and RStudio | 13 |
| 2.1 Installing R | 14 |
| 2.2 Upgrading R | 14 |
| 2.3 Installing and using RStudio | 15 |

| | | |
|--|---|-----------|
| 2.4 | Functions, packages and libraries | 18 |
| 2.5 | Working with data in R | 22 |
| 2.5.1 | Data structures and value types in R | 24 |
| 2.6 | Importing and exporting data sets | 28 |
| 2.6.1 | Importing data sets in R | 29 |
| 2.6.2 | Exporting data sets from R | 32 |
| 2.7 | Getting help on R and RStudio | 33 |
| | Further reading | 34 |
| Part Two Data visualisation | | 37 |
| 3 | Graphical representation of data | 39 |
| 3.1 | Using tables | 42 |
| 3.2 | Using graphs | 44 |
| 3.2.1 | Bar graph | 45 |
| 3.2.2 | Pareto graph | 47 |
| 3.2.3 | Pie graph | 49 |
| 3.2.4 | Histogram | 51 |
| 3.2.5 | Box and whisker graph | 52 |
| 3.2.6 | Line graph | 53 |
| 3.2.7 | Scatter graph | 55 |
| | Further reading | 58 |
| Part Three Providing information about data | | 61 |
| 4 | Descriptive statistics | 63 |
| 4.1 | From raw data to frequency distributions | 64 |
| 4.2 | Measures of location or central tendency | 70 |
| 4.2.1 | Mode | 71 |
| 4.2.2 | Median | 72 |
| 4.2.3 | Mean | 73 |
| 4.3 | Advantages and disadvantages of using the mode, median and mean | 76 |
| 4.4 | Measures of location and graphical display | 77 |
| | Further reading | 82 |
| 5 | Measures of dispersion and distributions | 83 |
| 5.1 | Range | 84 |
| 5.2 | Percentiles, deciles and quartiles | 87 |
| 5.3 | Interquartile range | 94 |
| 5.4 | Mean deviation | 96 |
| 5.5 | Standard deviation | 99 |
| 5.6 | Coefficient of variation | 104 |

| | | |
|------------------|--|------------|
| 5.7 | Shape of distributions and skewness | 106 |
| 5.8 | Advantages and disadvantages of using the range, mean deviation and standard deviation | 110 |
| 5.9 | Measures of dispersion and graphical display | 111 |
| | Further reading | 116 |
| 6 | Normal distribution and standardised scores | 117 |
| 6.1 | From histogram to normal distribution curve | 118 |
| 6.2 | Other visual methods for assessing normality of data | 121 |
| 6.2.1 | The boxplot and the normal distribution | 122 |
| 6.2.2 | The QQ plot and the normal distribution | 123 |
| 6.3 | Normal distribution and standard deviation | 123 |
| 6.4 | Statistical tests for normality | 125 |
| 6.5 | Standard normal distribution and z-scores | 129 |
| 6.6 | Transforming data values into z-scores | 131 |
| | Further reading | 138 |
| Part Four | Making estimations and predictions from data | 139 |
| 7 | Fundamentals of inferential statistics | 141 |
| 7.1 | What is inferential statistics and how does it work? | 142 |
| 7.2 | From sample to population | 143 |
| 7.3 | Sampling strategies | 144 |
| 7.3.1 | Random sampling methods | 145 |
| 7.3.2 | Non-random sampling methods | 151 |
| 7.4 | Making decisions about the population based on the information about the sample | 152 |
| 7.5 | Standard distributions | 154 |
| 7.5.1 | Standard normal distribution | 154 |
| 7.5.2 | t-distributions | 154 |
| 7.5.3 | Chi-squared distributions | 157 |
| | Further reading | 160 |
| 8 | Estimation and hypothesis testing | 161 |
| 8.1 | Making estimations | 162 |
| 8.1.1 | Standard error | 162 |
| 8.1.2 | Sample size | 164 |
| 8.1.3 | Confidence interval and confidence level | 166 |
| 8.2 | Statistical hypothesis testing process | 169 |
| 8.2.1 | Null and alternative hypotheses | 169 |
| 8.2.2 | Directional and non-directional hypotheses | 172 |
| 8.2.3 | Decisions about the null hypothesis: statistical levels, types of error and power | 173 |

| | | |
|--|---|------------|
| 8.2.4 | Regions of rejection | 174 |
| 8.3 | Selection of statistical tests | 176 |
| | Further reading | 179 |
| Part Five From sample to population | | 181 |
| 9 | One-sample tests | 183 |
| 9.1 | Parameter hypothesis testing using sample statistics | 184 |
| 9.2 | One-sample statistical tests for interval/ratio data | 185 |
| 9.2.1 | z-test | 187 |
| 9.2.2 | t-test | 189 |
| 9.2.3 | Sign test | 195 |
| 9.3 | One-sample statistical tests for ordinal data | 198 |
| 9.3.1 | Wilcoxon signed-rank test | 198 |
| 9.4 | One-sample statistical tests for nominal data | 200 |
| 9.4.1 | Binomial test | 200 |
| 9.4.2 | Pearson chi-squared goodness-of-fit test | 203 |
| | Further reading | 208 |
| 10 | Differences between two independent or dependent samples | 209 |
| 10.1 | Differences between two independent samples | 210 |
| 10.1.1 | Mann–Whitney test (or Wilcoxon rank-sum test) | 211 |
| 10.1.2 | Independent samples t-test | 214 |
| 10.1.3 | Chi-squared test | 217 |
| 10.2 | Differences between two dependent samples | 220 |
| 10.2.1 | Wilcoxon signed-rank test | 220 |
| 10.2.2 | Paired samples t-test | 224 |
| 10.2.3 | McNemar’s test | 225 |
| | Further reading | 227 |
| 11 | Differences between more than two independent samples | 229 |
| 11.1 | The analysis of variance (ANOVA) | 230 |
| 11.2 | One-way ANOVA | 233 |
| 11.2.1 | Calculating one-way ANOVA by hand | 234 |
| 11.2.2 | Computing one-way ANOVA in R | 237 |
| 11.2.3 | Post-hoc tests for one-way ANOVA | 241 |
| 11.2.4 | Measuring the effect size in a one-way ANOVA | 243 |
| 11.3 | Two-way ANOVA | 246 |
| 11.3.1 | Computing a two-way ANOVA in R | 247 |
| 11.3.2 | Interaction plot for two-way ANOVA test results | 250 |
| 11.3.3 | Post-hoc analysis in two-way ANOVA | 251 |

| | | |
|-----------------|---|------------|
| 11.3.4 | Measuring the effect size in two-way ANOVA | 253 |
| 11.4 | Kruskal–Wallis ANOVA test | 254 |
| 11.4.1 | Post-hoc analysis for the Kruskal–Wallis test | 256 |
| | Further reading | 259 |
| 12 | Differences between more than two dependent samples | 261 |
| 12.1 | Repeated measures ANOVA | 262 |
| 12.2 | One-way repeated measures ANOVA | 264 |
| 12.2.1 | Checking assumptions for one-way repeated measures ANOVA | 265 |
| 12.2.2 | Computing one-way repeated measures ANOVA and Mauchly's test of sphericity | 269 |
| 12.2.3 | Post-hoc analysis for one-way repeated measures ANOVA | 271 |
| 12.2.4 | Measuring the effect size in a one-way repeated measures ANOVA | 272 |
| 12.3 | Two-way repeated measures ANOVA | 273 |
| 12.3.1 | Checking assumptions for two-way repeated measures ANOVA | 274 |
| 12.3.2 | Computing two-way repeated measures ANOVA in R | 277 |
| 12.4 | Friedman's test | 278 |
| 12.4.1 | Computing Friedman's test in R | 278 |
| 12.4.2 | Post-hoc analysis for Friedman's test | 280 |
| 12.4.3 | Measuring the effect size in Friedman's test | 281 |
| 12.5 | Cochran's Q-test | 282 |
| 12.5.1 | Manual calculation of Q-statistic | 284 |
| 12.5.2 | Computing Cochran's Q-test in R | 285 |
| 12.5.3 | Post-hoc analysis for Cochran's Q-test | 286 |
| 12.5.4 | Effect size for Cochran's Q-test | 288 |
| | Further reading | 290 |
| Part Six | Relationships and predictions | 291 |
| 13 | Relationships between variables | 293 |
| 13.1 | Covariance and correlation between two variables | 294 |
| 13.1.1 | Visual representation of the correlation | 297 |
| 13.1.2 | Coefficient of determination | 299 |
| 13.1.3 | Errors of the correlation coefficient | 299 |
| 13.2 | Correlations for more than two variables | 300 |
| 13.2.1 | Visual representation of the correlation matrix | 302 |
| 13.3 | Correlations and scales of measurement | 303 |
| 13.3.1 | Pearson's correlation coefficient | 303 |
| 13.3.2 | Spearman's correlation coefficient | 307 |
| 13.3.3 | Lambda, phi and Cramér's V correlation coefficients | 310 |
| | Further reading | 315 |

| | |
|--|------------|
| 14 Predictions for independent and dependent variables | 317 |
| 14.1 Linear regression models | 319 |
| 14.2 Ordinary least squares regression | 321 |
| 14.2.1 Creating the OLS regression model | 325 |
| 14.2.2 Checking for statistical significance | 327 |
| 14.2.3 Assessing the assumptions of the linear regression model | 330 |
| 14.3 Multiple linear regression | 333 |
| 14.3.1 Creating the multiple linear regression model | 334 |
| 14.3.2 Checking statistical significance | 335 |
| 14.3.3 Assessing the assumptions of the multiple linear regression model | 337 |
| Further reading | 343 |
| | |
| <i>Bibliography</i> | 345 |
| <i>Index</i> | 349 |