

Contents

Part I Basic Knowledge and Tools to Apply Statistics

1	Statistics is Fun	3
1.1	Why Statistics?	3
1.2	Checkpoints	4
1.3	Data	4
1.4	Checkpoints	7
1.5	Scales—Lifelong Important in Data Analysis	7
1.6	Checkpoints	9
1.7	Software: Excel, SPSS, or “R”	10
1.8	Case Studies—The Best Way to Learn	11
1.9	Case Study: Company Growth.	11
1.10	Application.	11
2	R and R Commander: A Brief Introduction	15
2.1	Interesting Facts about R and the R Commander	15
2.2	The Installation of R and the R Commander	16
2.3	The R Commander and its Functions.	20
2.4	Reading Data with the R Commander	23

Part II Describe, Nothing but Describe

3	Averages: How People and Objects Behave in General	29
3.1	Averages—Why We Need Them	29
3.2	The (Arithmetic) Mean	29
3.3	The Median	32
3.4	The Mode	33
3.5	The Geometric Mean and Growth Rates	33
3.6	Which Average Value Should We Use and What else Do We Need to Know?	35
3.7	Calculation of Averages with the R Commander.	37

3.8	Checkpoints	41
3.9	Application.	42
4	Variation: The Deviation from Average Behavior	43
4.1	Variation—The Flip Side of Average Behavior	43
4.2	The Range	44
4.3	The Standard Deviation	44
4.4	The Variance.	47
4.5	The Coefficient of Variation.	47
4.6	The Interquartile Range	48
4.7	The Boxplot	49
4.8	Calculation of Measures of Variation with the R Commander	51
4.9	Creating the Boxplot with the R Commander	53
4.10	Checkpoints	57
4.11	Application.	58
5	Charts: The Possibility to Display Data Visually	59
5.1	Charts: Why do We Need Them?	59
5.2	The Frequency Table	59
5.3	The Frequency Charts	61
5.4	Absolute Frequency Chart, Relative Frequency Chart, or Histogram?	63
5.5	More Ways to Display Data	63
5.6	Creating the Frequency Charts with the R Commander	66
5.7	Checkpoints	69
5.8	Application.	70
6	Correlation: From Relationships.	71
6.1	Correlation—The Joint Movement of Two Variables	71
6.2	The Correlation Coefficient of Bravais-Pearson for Metric Variables	72
6.3	The Scatterplot.	74
6.4	The Correlation Coefficient of Spearman for Ordinal Variables	78
6.5	The Phi Coefficient for Nominal Variables with Two Characteristics	79
6.6	The Contingency Coefficient for Nominal Variables.	81
6.7	Correlation, Causality, Third Variables, and Other Correlation Coefficients	83
6.8	Calculation of Correlation Coefficients with the R Commander.	85
6.9	Checkpoints	93
6.10	Application.	93

7 Ratios and Indices: The Opportunity to Generate New Knowledge from Old Ones 97

7.1 The Reference Number—The Quotient of Two Different Sizes 98

7.2 The Quota—The Quotient of a Partial Number to a Total Number 99

7.3 The Dynamic Index 100

7.4 The Price and Quantity Index of Laspeyres and Paasche 101

7.5 Generating New Variables with the R Commander 103

7.6 Checkpoints 105

7.7 Application. 105

Part III From Few to All

8 About Data and Truth 109

8.1 How do we Get Our Data: Or Primary or Secondary Data? 109

8.2 The Random Sample—The Best Estimator for Our Population 111

8.3 On Truth—Validity and Reliability 119

8.4 Checkpoints 120

8.5 Application. 121

9 Hypotheses: Just a Refinement of the Question 123

9.1 The Small, Big Thing of the (Research) Hypothesis 123

9.2 The Null Hypothesis H_0 and the Alternative Hypothesis H_A 124

9.3 Hypotheses, Non-Directional or Directional? 126

9.4 What Makes a Good Hypothesis? 127

9.5 Checkpoints 128

9.6 Application. 128

10 Normal Distribution and Other Test Distributions 131

10.1 The Normal Distribution 131

10.2 The Z-Value and the Standard Normal Distribution 135

10.3 Normal Distribution, t-Distribution, χ^2 -Distribution and (or rather) F-Distribution. 138

10.4 Drawing the Distributions with the R Commander 141

10.5 Checkpoints 144

10.6 Application. 145

11 Hypothesis Testing: What Applies? 147

11.1 What does Statistical Significance Mean? 147

11.2 The Significance Level α 149

11.3 Statistically Significant, but also Practically Relevant? 152

11.4	Steps in Performing the Hypothesis Test	153
11.5	How Do I Choose My Test Procedure?	154
11.6	Checkpoints	155
11.7	Application.	155

Part IV Method for Testing Hypotheses

12	The Test for a Group Mean or One Sample t-Test.	159
12.1	Introduction to the Test for a Group Mean	159
12.2	The Research Question and Hypotheses in the Test for a Group Mean: Are Company Founders on Average 40 Years Old?	159
12.3	The Test Distribution and Test Statistic in the Test for a Group Mean	160
12.4	The Critical Value in the Test for a Group Mean	160
12.5	The z-Value	161
12.6	The Decision	162
12.7	The Test for a Group Mean With Unknown Standard Deviation in the Population or With a Small Sample $n \leq 30$	162
12.8	The Effect Size in the Test for a Group Mean	164
12.9	Calculation with the R Commander.	165
12.10	Checkpoints	168
12.11	Application.	169
13	The Test for a Difference Between Means with Independent Samples or Independent Samples t-Test	171
13.1	Introduction to the Test for a Difference Between Means with Independent Samples	171
13.2	The Research Question and Hypotheses in the Test: Are Women and Men the Same Age at the Time of Founding?	171
13.3	The Test Distribution and the Test Statistic	172
13.4	The Critical t-value	173
13.5	The t-value and the Decision	174
13.6	The Effect Size.	174
13.7	Equal or Unequal Variances	175
13.8	Calculation with the R Commander.	176
13.9	Checkpoints	182
13.10	Application.	183

14	The Test for a Difference Between Means with Dependent Samples or Dependent Samples t-Test	185
14.1	Introduction to the Test for a Difference Between Means with Dependent Samples	185
14.2	The Example: Training of Company Founders in the Pre-Startup Phase	185
14.3	The Research Question and the Hypotheses in the Test: Does the Training Have an Influence on the Assessment of Market Potential?	187
14.4	The Test Statistic	187
14.5	The Critical t-value	188
14.6	The t-value and the Decision	188
14.7	The Effect Size	189
14.8	Calculation with the R Commander	190
14.9	Checkpoints	192
14.10	Application	193
15	The Analysis of Variance for Testing Group Differences When There Are More Than Two Groups	195
15.1	Introduction to the Analysis of Variance	195
15.2	The Example: Do Company Founders Differ in Their Workload According to the Founding Motive?	196
15.3	The Research Question and the Hypotheses in the Analysis of Variance	198
15.4	The Basic Idea of the Analysis of Variance	198
15.5	The Test Statistic	200
15.6	The Critical F-value	201
15.7	The F-value and the Decision	202
15.8	The Analysis of Variance, an Omnibus Test, and the Post-hoc Tests	202
15.9	The Effect Size in the Analysis of Variance	203
15.10	Calculating the Analysis of Variance With the R Commander	204
15.11	Checkpoints	207
15.12	Application	208
16	The Test for Correlation with Metric, Ordinal, and Nominal Data	209
16.1	The Correlation Test with Metric Data	209
16.2	The Test for a Correlation with Ordinal Data	213
16.3	The Test for a Correlation with Nominal Data	215
16.4	Calculation of the Correlation Tests with the R Commander	221
16.5	Checkpoints	226
16.6	Application	226

17 More Tests for Nominal Variables 229

17.1 The χ^2 -Test with One Sample: Does the Proportion of Female Founders Correspond to the Gender Ratio in Society? 229

17.2 The χ^2 -Test with Two Independent Samples: Are the Founding Motives the Same for Sservice and Industrial Companies?. 231

17.3 The χ^2 Test with Two Dependent Samples: Is my Advertising Campaign Effective? 233

17.4 Calculation with the R Commander 235

17.5 Checkpoints 238

17.6 Application. 238

18 Summary Part IV—Overview of the Test Procedures. 239

Part V Regression Analysis

19 The Simple Linear Regression. 245

19.1 Objectives of Regression Analysis. 245

19.2 The Linear Regression Line and the Method of Ordinary Least Squares. 246

19.3 How Well and How Much Can We Explain, the R^2 251

19.4 Calculation with the R Commander. 253

19.5 Is One Independent Variable Enough, Out-of-Sample Predictions, and More Warnings. 258

19.6 Checkpoints 260

19.7 Application. 261

20 The Multiple Regression Analysis 263

20.1 The Multiple Regression Analysis - More Than One Independent Variable 263

20.2 F-Test, t-Test and Adjusted- R^2 265

20.3 Calculation with the R Commander. 267

20.4 When is the Ordinary Least Squares Estimate BLUE? 270

20.5 Checkpoints 277

20.6 Application. 277

Part VI What Happens Next?

21 How to Present Results 281

21.1 Contents of an Empirical Paper 281

21.2 Example I for a Report: Is There a Difference in Founding Age Between Male and Female Founders (Fictional). 282

21.3 Example II for a Report: Professional Experience and Company Performance (Fictional). 283

21.4 Application. 284

22

Advanced Statistical Methods

285

23

Interesting and Advanced Statistical Textbooks

287

24

Another Dataset for Practice

291

Appendix

295