

Contents

- 1 Basics..... 1**
 - 1.1 Objectives of Statistics 1
 - A Definition of Statistics..... 1
 - Explained: Descriptive and Inductive Statistics 3
 - 1.2 Statistical Investigation 4
 - Conducting a Statistical Investigation 4
 - Sources of Economic Data 4
 - Explained: Public Sources of Data 6
 - More Information: Statistical Processes 6
 - 1.3 Statistical Element and Population 8
 - Statistical Elements 8
 - Population 8
 - Explained: Statistical Elements and Population 8
 - 1.4 Statistical Variable 10
 - 1.5 Measurement Scales 11
 - 1.6 Qualitative Variables..... 11
 - Nominal Scale 11
 - Ordinal Scale 12
 - 1.7 Quantitative Variables 13
 - Interval Scale 13
 - Ratio Scale..... 13
 - Absolute Scale..... 13
 - Discrete Variable 14
 - Continuous Variable..... 14
 - 1.8 Grouping Continuous Data 14
 - Explained: Grouping of Data 16
 - 1.9 Statistical Sequences and Frequencies..... 16
 - Statistical Sequence 16
 - Frequency 17
 - Explained: Absolute and Relative Frequency 18

2	One-Dimensional Frequency Distributions	21
2.1	One-Dimensional Distribution	21
2.1.1	Frequency Distributions for Discrete Data	21
	Frequency Table	21
2.1.2	Graphical Presentation	22
	Explained: Job Proportions in Germany	25
	Enhanced: Evolution of Household Sizes	25
2.2	Frequency Distribution for Continuous Data	26
	Frequency Table	27
	Graphical Presentation	27
	Explained: Petrol Consumption of Cars	30
	Explained: Net Income of German Nationals	31
2.3	Empirical Distribution Function	34
2.3.1	Empirical Distribution Function for Discrete Data	35
2.3.2	Empirical Distribution Function for Grouped	
	Continuous Data	36
	Explained: Petrol Consumption of Cars	37
	Explained: Grades in Statistics Examination	38
2.4	Numerical Description of One-Dimensional Frequency	
	Distributions	40
	Measures of Location	40
	Explained: Average Prices of Cars	47
	Interactive: Dotplot with Location Parameters	49
	Interactive: Simple Histogram	49
2.5	Location Parameters: Mean Values—Harmonic Mean,	
	Geometric Mean	50
	Harmonic Average	50
	Geometric Average	52
2.6	Measures of Scale or Variation	55
	Range	56
	Interquartile Range	57
	Mean Absolute Deviation	57
	The Variance and the Standard Deviation	58
	Explained: Variations of Pizza Prices	60
	Enhanced: Parameters of Scale for Cars	61
	Interactive: Dotplot with Scale Parameters	62
2.7	Graphical Display of the Location and Scale Parameters	64
	Boxplot (Box-Whisker-Plot)	64
	Explained: Boxplot of Car Prices	66
	Interactive: Visualization of One-Dimensional	
	Distributions	67
3	Probability Theory	69
3.1	The Sample Space, Events, and Probabilities	69
	Venn Diagram	70

3.2	Event Relations and Operations	70
	Subsets and Complements	70
	Union of Sets	71
	Intersection of Sets	71
	Logical Difference of Sets or Events	73
	Disjoint Decomposition of the Sample Space	74
	Some Set Theoretic Laws	75
3.3	Probability Concepts	75
	Classical Probability	76
	Statistical Probability	76
	Axiomatic Foundation of Probability	78
	Addition Rule of Probability	78
	More Information: Derivation of the Addition Rule	79
	More Information: Implications of the Probability Axioms	80
	Explained: A Deck of Cards	81
3.4	Conditional Probability and Independent Events	82
	Conditional Probability	82
	Multiplication Rule	83
	Independent Events	83
	Two-Way Cross-Tabulation	84
	More Information: Derivation of Rules for Independent Events	85
	Explained: Two-Way Cross-Tabulation	85
	Explained: Screws	86
3.5	Theorem of Total Probabilities and Bayes' Rule	87
	Theorem of Total Probabilities	87
	Bayes' Rule	88
	Explained: The Wine Cellar	88
	Enhanced: Virus Test	90
	Interactive: Monty Hall Problem	91
	Interactive: Die Rolling Sisters	94
4	Combinatorics	97
4.1	Introduction	97
	Different Ways of Grouping and Ordering	97
	Use of Combinatorial Theory	98
4.2	Permutation	98
	Permutations Without Repetition	98
	Permutations with Repetition	99
	Permutations with More Groups of Identical Elements	99
	Explained: Beauty Competition	100
4.3	Variations	100
	Variations with Repetition	100
	Variations Without Repetition	101
	Explained: Lock Picking	101

4.4	Combinations	102
	Combinations Without Repetition	102
	Combinations with Repetition	103
	Explained: German Lotto	103
4.5	Properties of Euler's Numbers (Combination Numbers)	104
	Symmetry	104
	Specific Cases	104
	Sum of Two Euler's Numbers	104
	Euler's Numbers and Binomial Coefficients	105
5	Random Variables	107
5.1	The Definition	107
	More Information	107
	Explained: The Experiment	108
	Enhanced: Household Size I	108
5.2	One-Dimensional Discrete Random Variables	109
	Discrete Random Variable	109
	Explained: One-Dimensional Discrete Random Variable ...	110
	Enhanced: Household Size II	111
5.3	One-Dimensional Continuous Random Variables	113
	Density Function	113
	Distribution Function	113
	More Information: Continuous Random Variable, Density, and Distribution Function	114
	Explained: Continuous Random Variable	116
	Enhanced: Waiting Times of Supermarket Costumers	116
5.4	Parameters	119
	Expected Value	120
	Variance	121
	Standard Deviation	121
	Standardization	122
	Chebyshev's Inequality	122
	Explained: Continuous Random Variable	123
	Explained: Traffic Accidents	124
5.5	Two-Dimensional Random Variables	124
	Marginal Distribution	125
	The Conditional Marginal Distribution Function	126
	Explained: Two-Dimensional Random Variable	127
	Enhanced: Link Between Circulatory Diseases and Patient Age	129
5.6	Independence	131
	Conditional Distribution	132
	More Information	133
	Explained: Stochastic Independence	134
	Enhanced: Economic Conditions in Germany	136

5.7	Parameters of Two-Dimensional Distributions	139
	Covariance	140
	Correlation Coefficient	140
	More Information	141
	Explained: Parameters of Two-Dimensional Random Variables	144
	Enhanced: Investment Funds	146
6	Probability Distributions	149
6.1	Important Distribution Models	149
6.2	Uniform Distribution	149
	Discrete Uniform Distribution	149
	Continuous Uniform Distribution	150
	More Information	151
	Explained: Uniform Distribution	152
6.3	Binomial Distribution	154
	More Information	155
	Explained: Drawing Balls from an Urn	157
	Enhanced: Better Chances for Fried Hamburgers	158
	Enhanced: Student Jobs	160
	Interactive: Binomial Distribution	162
6.4	Hypergeometric Distribution	163
	More Information	164
	Explained: Choosing Test Questions	166
	Enhanced: Selling Life Insurances	167
	Enhanced: Insurance Contract Renewal	168
	Interactive: Hypergeometric Distribution	169
6.5	Poisson Distribution	170
	More Information	171
	Explained: Risk of Vaccination Damage	172
	Enhanced: Number of Customers in Service Department	173
	Interactive: Poisson Distribution	175
6.6	Exponential Distribution	176
	More Information	177
	Explained: Number of Defects	178
	Enhanced: Equipment Failures	180
	Interactive: Exponential Distribution	181
6.7	Normal Distribution	181
	Standardized Random Variable	183
	Standard Normal Distribution	183
	Confidence Interval	184
	More Information	186
	Other Properties of the Normal Distribution	187
	Standard Normal Distribution	188

	Explained: Normal Distributed Random Variable	188
	Interactive: Normal Distribution.....	195
6.8	Central Limit Theorem	196
	Central Limit Theorem.....	197
	More Information	197
	Explained: Application to a Uniform Random Variable	197
6.9	Approximation of Distributions	199
	Normal Distribution as Limit of Other Distributions	199
	Explained: Wrong Tax Returns	201
	Enhanced: Storm Damage	203
6.10	Chi-Square Distribution	204
	More Information	205
6.11	t-Distribution (Student t-Distribution)	206
	More Information	207
6.12	F-Distribution	207
	More Information	208
7	Sampling Theory	209
7.1	Basic Ideas	209
	Population	209
	Sample	210
	Statistic	211
	More Information	213
	Explained: Illustrating the basic Principles of Sampling Theory	213
7.2	Sampling Distribution of the Mean	218
	Distribution of the Sample Mean	218
	More Information	221
	Explained: Sampling Distribution	225
	Enhanced: Gross Hourly Earnings of a Worker.....	228
7.3	Distribution of the Sample Proportion	233
	Explained: Distribution of the Sample Proportion.....	237
	Enhanced: Drawing Balls from a Urn	239
7.4	Distribution of the Sample Variance	242
	Distribution of the Sample Variance S^2	243
	Probability Statements About S^2	243
	More Information	244
	Explained: Distribution of the Sample Variance	247
8	Estimation	251
8.1	Estimation Theory	251
	Point Estimation.....	251
	The Estimator or Estimating Function	251
	Explained: Basic Examples of Estimation Procedures	252
8.2	Properties of Estimators	253
	Mean Squared Error.....	255

	Unbiasedness	255
	Asymptotic Unbiasedness	256
	Efficiency	256
	consistency	257
	More Information	257
	Explained: Properties of Estimators	262
	Enhanced: Properties of Estimation Functions	263
8.3	Construction of Estimators	264
	Maximum Likelihood	264
	Least Squares Estimation	266
	More Information	266
	Applications of ML	266
	Application of Least Squares	270
	Explained: ML Estimation of an Exponential Distribution	271
	Explained: ML Estimation of a Poisson Distribution	272
8.4	Interval Estimation	273
8.5	Confidence Interval for the Mean	275
	Confidence Interval for the Mean with Known Variance....	276
	Confidence Interval for the Mean with Unknown Variance.....	278
	Explained: Confidence Intervals for the Average Household Net Income.....	280
	Enhanced: Confidence Intervals for the Lifetime of a Bulb	285
	Interactive: Confidence Intervals for the Mean	287
8.6	Confidence Interval for Proportion	288
	Properties of Confidence Intervals	290
	Explained: Confidence Intervals for the Percentage of Votes	291
	Interactive: Confidence Intervals for the Proportion.....	291
8.7	Confidence Interval for the Variance	292
	Properties of the Confidence Interval	293
	Explained: Confidence Intervals for the Variance of Household Net Income.....	294
	Interactive: Confidence Intervals for the Variance	295
8.8	Confidence Interval for the Difference of Two Means	295
	1. Case: The Variances σ_1^2 and σ_2^2 of the Two Populations Are Known.....	297
	Properties of the Confidence Interval	297
	2. Case: The Variances σ_1^2 and σ_2^2 of the Two Populations Are Unknown	298
	Properties of Confidence Intervals When Variances Are Unknown	299

	Explained: Confidence Interval for the Difference of Car Gas Consumptions	300
	Enhanced: Confidence Intervals of the Difference of Two Mean Stock Prices	301
	Interactive: Confidence Intervals for the Difference of Two Means	304
8.9	Confidence Interval Length	305
	(a) Confidence Interval for μ	306
	(b) Confidence Interval for π	306
	Explained: Finding a Required Sample Size	307
	Enhanced: Finding the Sample Size for an Election Threshold	308
	Interactive: Confidence Interval Length for the Mean	309
9	Statistical Tests	311
9.1	Key Concepts	311
	Formulating the Hypothesis	313
	Test Statistic	314
	Decision Regions and Significance Level	314
	Non-rejection Region of Null Hypothesis	315
	Rejection Region of Null Hypothesis	315
	Power of a Test	323
	OC-Curve	324
	A Decision-Theoretical View on Statistical Hypothesis Testing	324
	More Information: Examples	325
	More Information: Hypothesis Testing Using Statistical Software	327
9.2	Testing Normal Means	330
	Hypotheses	331
	Test Statistic, Its Distribution, and Derived Decision Regions	332
	Calculating the Test Statistic from an Observed Sample	336
	Test Decision and Interpretation	337
	Power	338
	More Information: Conducting a Statistical Test	342
	Explained: Testing the Population Mean	348
	Enhanced: Average Life Time of Car Tires	352
	Hypothesis	353
	1st Alternative	354
	2nd Alternative	355
	3rd Alternative	357
	Interactive: Testing the Population Mean	358
	Interactive: Testing the Population Mean with Type I and II Error	359

9.3	Testing the Proportion in a Binary Population	360
	Hypotheses	361
	Test Statistic and Its Distribution: Decision Regions	361
	Sampling and Computing the Test Statistic	363
	Test Decision and Interpretation	363
	Power Curve $P(\pi)$	364
	Explained: Testing a Population Proportion	364
	Enhanced: Proportion of Credits with Repayment Problems	369
	Interactive: Testing a Proportion in a Binary Population	376
9.4	Testing the Difference of Two Population Means	377
	Hypotheses	377
	Test Statistic and Its Distribution: Decision Regions	378
	Sampling and Computing the Test Statistic	380
	Test Decision and Interpretation	380
	Explained: Testing the Difference of Two Population Means	381
	Enhanced: Average Age Difference of Female and Male Bank Employees	383
	1st Dispute	384
	2nd Dispute	386
	3rd Dispute	387
	Interactive: Testing the Difference of Two Population Means	388
9.5	Chi-Square Goodness-of-Fit Test	389
	Hypothesis	390
	How Is p_j Computed?	391
	Test Statistic and Its Distribution: Decision Regions	391
	Approximation Conditions	392
	Sampling and Computing the Test Statistic	393
	Test Decision and Interpretation	394
	More Information	394
	Explained: Conducting a Chi-Square Goodness-of-Fit Test	397
	Enhanced: Goodness-of-Fit Test for Product Demand	399
	1st Version	400
	2nd Version	401
9.6	Chi-Square Test of Independence	404
	Hypothesis	405
	Test Statistic and Its Distribution: Decision Regions	406
	Sampling and Computing the Test Statistic	407
	Test Decision and Interpretation	408
	More Information	408

Explained: The Chi-Square Test of Independence in Action	411
Enhanced: Chi-Square Test of Independence for Economic Situation and Outlook	413
10 Two-Dimensional Frequency Distribution	419
10.1 Introduction	419
10.2 Two-Dimensional Frequency Tables	419
Realizations $m \cdot r$	420
Absolute Frequency	420
Relative Frequency	420
Properties	420
Explained: Two-Dimensional Frequency Distribution	421
Enhanced: Department Store	422
Interactive: Example for Two-Dimensional Frequency Distribution	423
10.3 Graphical Representation of Multidimensional Data	423
Frequency Distributions	423
Scatterplots	424
Explained: Graphical Representation of a Two- or Higher Dimensional Frequency Distribution	426
Interactive: Example for the Graphical Representation of a Two- or Higher Dimensional Frequency Distribution	429
10.4 Marginal and Conditional Distributions	429
Marginal Distribution	429
Conditional Distribution	430
Explained: Conditional Distributions	432
Enhanced: Smokers and Lung Cancer	433
Enhanced: Educational Level and Age	434
10.5 Characteristics of Two-Dimensional Distributions	435
Covariance	435
More Information	437
Explained: How the Covariance Is Calculated	437
10.6 Relation Between Continuous Variables (Correlation, Correlation Coefficients)	438
Properties of the Correlation Coefficient	439
Relation of Correlation and the Scatterplot of X and Y Observations	440
Explained: Relationship of Two Metrically Scaled Variables	443
Interactive: Correlation Coefficients	444
10.7 Relation Between Discrete Variables (Rank Correlation)	445
Spearman's Rank Correlation Coefficient	445

Kendall's Rank Correlation Coefficient	447
Explained: Relationship Between Two Ordinally Scaled Variables	448
Interactive: Example for the Relationship Between Two Ordinally Scaled Variables.....	450
10.8 Relationship Between Nominal Variables (Contingency)	450
Explained: Relationship Between Two Nominally Scaled Variables	452
Interactive: Example for the Relationship Between Two Nominally Scaled Variables	454
11 Regression	455
11.1 Regression Analysis	455
The Objectives of Regression Analysis	455
11.2 One-Dimensional Regression Analysis	457
One-Dimensional Linear Regression Function	457
Quality (Fit) of the Regression Line.....	463
One-Dimensional Nonlinear Regression Function	466
Explained: One-Dimensional Linear Regression	468
Enhanced: Crime Rates in the US	471
Enhanced: Linear Regression for the Car Data	472
Interactive: Simple Linear Regression	473
11.3 Multi-Dimensional Regression Analysis	474
Multi-Dimensional Regression Analysis.....	474
12 Time Series Analysis	477
12.1 Time Series Analysis	477
Definition	477
Graphical Representation	477
The Objectives of Time Series Analysis	477
Components of Time Series	479
12.2 Trend of Time Series	479
Method of Moving Average	479
Least-Squares Method	481
More Information: Simple Moving Average.....	483
Explained: Calculation of Moving Averages	485
Interactive: Test of Different Filters for Trend Calculation	486
12.3 Periodic Fluctuations	487
Explained: Decomposition of a Seasonal Series.....	489
Interactive: Decomposition of Time Series	491
12.4 Quality of the Time Series Model	492
Mean Squared Dispersion (Estimated Standard Deviation)	493
Interactive: Comparison of Time Series Models.....	494

A Data Sets in the Interactive Examples	495
A.1 ALLBUS Data	495
A.1.1 ALLBUS1992, ALLBUS2002, and ALLBUS2012: Economics	495
A.1.2 ALLBUS1994, ALLBUS2002, and ALLBUS2012: Trust	496
A.1.3 ALLBUS2002, ALLBUS2004, and ALLBUS2012: General	497
A.2 Boston Housing Data	499
A.3 Car Data	499
A.4 Credit Data	500
A.5 Decathlon Data	501
A.6 Hair and Eye Color of Statistics Students	502
A.7 Index of Basic Rent	502
A.8 Normally Distributed Data	503
A.9 Telephone Data	503
A.10 Titanic Data	504
A.11 US Crime Data	504
Glossary	507

Introduction to Statistics

Using Interactive MM*Stat Elements

Härdle, W.K.; Klinke, S.; Rönz, B.

2015, XX, 516 p. 205 illus., 173 illus. in color.,

Hardcover

ISBN: 978-3-319-17703-8