

Inhaltsverzeichnis

Teil I Statistik in der Theorie

1 Explorative Werkzeuge	3
1.1 Scatterplots	3
1.1.1 Zwei metrische Merkmale	3
1.1.2 Ein metrisches und ein kategoriales Merkmal	7
1.2 Verdichtung von Daten	11
1.2.1 Lokations- und Streuungsmaße	12
1.2.2 Histogramme	19
1.2.3 Boxplots	24
1.3 Aufgaben	27
Literatur	30
2 Schätzen in Binomialmodellen	31
2.1 Motivation und Einführung	31
2.2 Schätzen eines Anteilswertes	35
2.3 Punktschätzer	38
2.4 Intervallschätzer	46
2.4.1 Konstruktion mittels Tschebyschow-Ungleichung	47
2.4.2 Konstruktion mittels Quantilen	50
2.4.3 Einseitige Konfidenzintervalle	55
2.4.4 Berechnung der Intervallgrenzen	56
2.5 Approximative Konfidenzintervalle	58
2.6 Planung des Stichprobenumfangs	64
2.7 Test oder Konfidenzintervall?	65
2.7.1 Der Binomialtest	67
2.7.2 Zusammenhang zwischen Konfidenzintervall und Test	72
2.8 Aufgaben	74
Literatur	77

3	Schätzen in Normalverteilungsmodellen	79
3.1	Motivation und Einführung	79
3.2	Punktschätzer	83
3.3	Intervallschätzer für den Mittelwert	85
3.3.1	Bekannte Standardabweichung	85
3.3.2	Unbekannte Standardabweichung	90
3.4	Konfidenzintervall für die Standardabweichung	97
3.4.1	Bekannter Mittelwert	98
3.4.2	Unbekannter Mittelwert	100
3.5	Überprüfen der Normalverteilungsannahme	101
3.5.1	Der Normal Probability Plot als grafisches Hilfsmittel	102
3.5.2	Der Shapiro-Wilk-Test auf Normalverteilung	105
3.6	Aufgaben	107
	Literatur	111
4	Schätzen in linearen Regressionsmodellen	113
4.1	Motivation	113
4.2	Die lineare Regression in der deskriptiven Statistik	115
4.2.1	Die Methode der kleinsten Quadrate	115
4.2.2	Die zweite Regressionsgerade	121
4.3	Das Standardmodell der einfachen linearen Regression	122
4.3.1	Kleinste-Quadrate-Schätzer	124
4.3.2	Schätzung der Standardabweichung	126
4.3.3	Konfidenzintervalle im Regressionsmodell	128
4.4	Lineares Regressionsmodell mit zufälligem Design	129
4.5	Regressionsanalyse mit dem R-Commander	133
4.6	Aufgaben	136
	Literatur	139
5	Zweistichproben-Tests	141
5.1	Zweistichproben-Problem	141
5.2	Zweistichproben-Gauß-Test	144
5.3	Zweistichproben- <i>t</i> -Test	146
5.4	Wilcoxon-Rangsummentest, Mann-Whitney-U-Test	150
5.5	Mehrstichproben – ein Ausblick	159
5.6	Aufgaben	162
	Literatur	164
6	Der χ^2-Anpassungstest	167
6.1	Einführung	167
6.2	Der χ^2 -Test auf spezielle Verteilungsparameter	170
6.2.1	Die Testgröße χ^2	171
6.2.2	Ein Grenzwertsatz	172
6.2.3	Durchführung des χ^2 -Tests	173
6.3	Der χ^2 -Test auf eine Verteilungsfamilie	174

6.3.1	Simulation einer Poisson-verteilten Stichprobe	175
6.3.2	Der Satz von Fisher-Cramér	177
6.3.3	Anwendung des Satzes von Fisher-Cramér	178
6.3.4	χ^2 -Test und Normalverteilung	180
6.4	Aufgaben	182
	Literatur	185
7	Unabhängigkeitstests und Homogenitätstests	187
7.1	Einführendes Beispiel und Darstellungsmöglichkeiten	187
7.1.1	2×2 -Kontingenztafeln	188
7.1.2	Säulendiagramm und Mosaikdiagramm	190
7.2	Der χ^2 -Unabhängigkeitstest	192
7.3	Fishers exakter Test	196
7.3.1	Das Chancenverhältnis bei 2×2 -Tafeln	196
7.3.2	Test auf Unabhängigkeit bei 2×2 -Kontingenztafeln	203
7.3.3	Der erweiterte Fisher-Test als Trendtest	205
7.3.4	Allgemeine Kontingenztafeln	207
7.4	Homogenitätstests	209
7.5	Aufgaben	215
	Literatur	217

Teil II Statistik in der Praxis

8	Erfolgsmessung im Kampagnenmanagement	221
8.1	Einführung ins Kampagnenmanagement	221
8.2	Messung von Kampagnen	223
8.3	Repräsentativität von Kontrollgruppen	227
8.4	Signifikanz des Hebeleffektes	233
8.5	Kontrollgruppengröße	238
8.6	Aufgaben	240
	Literatur	241
9	Präventives Kreditausfallmanagement	243
9.1	Ziele des präventiven Kreditausfallmanagements	243
9.2	Identifikation geeigneter Kunden	244
9.2.1	Datengrundlage	245
9.2.2	Scoring	248
9.2.3	Entscheidungsbäume	249
9.3	Interaktion mit den Kunden	260
	Literatur	261
10	Untersuchung der Zahngesundheit der Schüler in Bayern	263
10.1	Problemstellung der Studie	263
10.2	Stichprobenbestimmung	264
10.2.1	Zielgröße DMFT-Wert	264
10.2.2	Datenbasis	265

10.2.3 Mathematische Methoden	267
10.2.4 Ergebnisse der Größenbestimmung	275
10.2.5 Stichprobeneinschätzung	277
10.3 Stichprobenauswertung	278
10.3.1 Drop-out-Werte	278
10.3.2 Historischer Vergleich der DMFT-Werte	279
10.3.3 Vergleich der DMFT-Werte zwischen den Schularten	286
10.4 Weitere statistische Untersuchungen	290
10.5 Aufgaben	290
Literatur	292

Teil III Statistik mittels Simulationen

11 Computerintensive Statistik	295
11.1 Monte-Carlo-Verfahren	296
11.1.1 Berechnung der Kreiszahl π	297
11.1.2 Berechnung von bestimmten Integralen	301
11.1.3 Monte-Carlo in der Finanzwelt	305
11.2 Bootstrap – Der Münchhausen-Trick in der Statistik	310
11.2.1 Der klassische Lösungsansatz	311
11.2.2 Der Bootstrap-Ansatz	312
11.3 Aufgaben	318
Literatur	320

12 Zuckerbrot oder Peitsche? Drosophila Larven und Bootstrap	321
12.1 Biologische Hintergründe	321
12.2 Statistische Fragestellung	325
12.3 Anwendung des Bootstrap	326
12.4 Simulationsergebnisse	331
12.5 Biologische Resultate	333
12.6 Aufgaben	336
Literatur	336

Teil IV Statistik als Projekt im Unterricht

13 Kann man Münzen fälschen?	339
13.1 Themenstellung	340
13.2 Erwartungshorizont und Vorkenntnisse	340
13.3 Projektablauf	341
13.4 Varianten	349
13.5 Materialien	350
Literatur	350

14 David oder Goliath – Welche Ameisen sind bessere Erntehelfer?	353
14.1 Hintergründe und Themenstellung	353
14.2 Erwartungshorizont und Vorkenntnisse	355
14.3 Projektablauf	355
14.4 Varianten	365
14.5 Materialien	366
Literatur	367
15 Die Menge macht das Gift?	369
15.1 Themenstellung	369
15.2 Erwartungshorizont und Vorkenntnisse	370
15.3 Projektablauf	370
15.4 Varianten	384
Literatur	386
16 Topmanagergehälter im Vergleich – Wer „verdient“ sein Geld wirklich?	387
16.1 Hintergründe und Themenstellung	387
16.2 Erwartungshorizont und Vorkenntnisse	388
16.3 Projektablauf	389
16.4 Varianten	401
Literatur	402
17 Blätter und Blattformen	403
17.1 Sachzusammenhänge und Themenstellung	403
17.1.1 Biologische Basisinformationen zu Blättern	403
17.1.2 Blattfläche	404
17.1.3 Lineare Modelle ohne Interzeptparameter	406
17.1.4 Mögliche Fragestellungen und Vorkenntnisse	408
17.1.5 „Technische“ Voraussetzungen	409
17.2 Hinweise zu möglichen Projektaktivitäten	409
17.2.1 Zusammenhang zwischen Länge und Breite	409
17.2.2 Der Quotient Länge/Breite als typspezifischer Parameter?	411
17.2.3 Von der Flächendichte zur Untersuchung des Zusammenhangs zwischen Blattfläche und Länge · Breite	415
17.2.4 Weitere Problemstellungen	417
Literatur	418

Teil V Statistik für Einsteiger mit R und R-Commander

18 Grundlagen von R	421
18.1 Was ist R?	421
18.1.1 Kurze Entstehungsgeschichte	421
18.1.2 Warum R?	422
18.2 Installation von R	422
18.2.1 Installation auf Windows-Rechnern	422

18.2.2 Installation mit anderen Betriebssystemen	424
18.2.3 Update von R	425
18.3 Erste Schritte mit R	425
18.3.1 R als Taschenrechner	426
18.3.2 Zuweisungen	428
18.3.3 Funktionen in R	429
18.3.4 Eigene Funktionen schreiben	430
18.4 Zusatzpakete	431
18.4.1 Pakete installieren	432
18.4.2 Pakete laden	433
18.4.3 Pakete update	433
18.5 Das Hilfesystem von R	434
18.5.1 Allgemeine Hilfe	435
18.5.2 Hilfe zu Funktionen	435
18.5.3 Online-Hilfe und weitere Hilfen	436
18.6 Aufgaben	437
Literatur	438
19 Arbeiten mit R	439
19.1 Der Aufbau von R	439
19.1.1 Der Workspace	439
19.1.2 Das Skriptfenster	441
19.2 Objekte	442
19.2.1 Vektoren	442
19.2.2 Matrizen	444
19.2.3 Datensätze	445
19.3 Der R-Commander	446
19.3.1 Installation und Öffnen des R-Commanders	446
19.3.2 Aufbau und Funktionsweise	447
19.3.3 Laden und Speichern	450
19.3.4 Schließen des R-Commanders	450
19.4 Aufgaben	450
Literatur	452
20 Datenmanagement	453
20.1 Datenimport und Datenexport	453
20.1.1 Dateneingabe	454
20.1.2 Datenimport	455
20.1.3 Datenexport	457
20.2 Daten bearbeiten	459
20.2.1 Eine schnelle Übersicht	459
20.2.2 Berechnung neuer Variablen	460
20.2.3 Umbenennen und Löschen von Variablen	461
20.3 Variablen konvertieren und umkodieren	462
20.3.1 Vektoren in Faktoren konvertieren	462

20.3.2 Umkodierung von Variablen	463
20.4 Fälle auswählen	466
20.5 Daten zusammenfügen	468
20.5.1 Fälle hinzufügen	468
20.5.2 Variablen hinzufügen	469
20.6 Aufgaben	470
Literatur	471
21 Darstellung von Daten	473
21.1 Deskriptive Statistiken	473
21.1.1 Kategoriale Daten	474
21.1.2 Metrische Daten	476
21.1.3 Kategoriale und metrische Daten	478
21.2 Säulendiagramme	479
21.2.1 Einfache Säulendiagramme	480
21.2.2 Grafiken bearbeiten	481
21.2.3 Gruppierte Säulendiagramme	483
21.2.4 Grafiken speichern	485
21.3 Aufgaben	486
Literatur	487
Bezeichnungen und Begriffsklärungen	489
Sachverzeichnis	493
R-Befehle	504
Eigene Funktionen	506
Datensätze	507