

Inhalt

Vorwort	11
1 Datenanalyse? Daten? Statistik?	13
1.1 Was für Daten gibt es?	17
1.2 Wo kommen Daten her?	20
Zusammenfassung	23
Aufgaben	23
2 Univariate, exploratorische Analyse	25
2.1 Häufigkeitstabellen und deren Darstellung	26
2.2 Auswertung der Urliste: Lage und Variabilität	38
2.2.1 Zur Lage eines Datensatzes	39
2.2.2 Zur Variabilität eines Datensatzes	46
2.3 Die empirische Verteilungsfunktion	52
2.4 Besondere Strukturen einer Verteilung	55
2.5 Konzentrationsmessung – LORENZ und GINI	61
2.6 Fallstudie – das 6 aus 49 Lotto	65
Zusammenfassung	73
Aufgaben	74
3 Bivariate, exploratorische Analyse	77
3.1 Korrelation von Merkmalen	78
3.2 Der Vergleich zweier Merkmale	87
Zusammenfassung	89
Aufgaben	89
4 Auf zur Modellierung	91
4.1 Konzepte am Beispiel der Binomialverteilung	92
4.1.1 Bernoulli-Experimente und Zufallsvariablen	92
4.1.2 Wahrscheinlichkeitsfunktion	96
4.1.3 Binomialverteilung	100
4.1.4 Verteilungsfunktion	102
4.1.5 Erwartungswerte	105
4.1.6 Erwartungswert der Binomialverteilung	108
4.1.7 Additivitätseigenschaft des Erwartungswertes	109
4.1.8 Binomialverteilung und Variabilität	111

4.1.9	Verteilung von Mittelwerten	114
4.2	Verschiedene diskrete Verteilungen	117
4.2.1	Die hypergeometrische Verteilung	117
4.2.2	Von der Binomial- zur Poisson-Verteilung	123
4.3	Stetige Modellwelt	127
4.3.1	Stetige Gleichverteilung	128
4.3.2	Über Summen zur Normalverteilung	131
4.3.3	Wartezeitverteilungen	141
4.3.4	Von geometrisch zu exponential	143
4.3.5	Von Poisson zu Exponential	145
4.3.6	Summe exponentialverteilter Zufallsvariablen	147
4.3.7	GLIVENKO und CANTELLI	148
	Zusammenfassung	149
	Aufgaben	150
5	Casino-Statistik	155
5.1	Würfelfragen	156
5.2	Wahrscheinlichkeit — was ist das?	157
5.3	Rechnen mit Wahrscheinlichkeiten	159
5.4	Axiome der Wahrscheinlichkeitsrechnung	162
5.5	Zusammengesetzte Ereignisse	163
5.6	Kombinatorik für das Gleichmöglichkeitsmodell	167
5.7	Wahrscheinlichkeiten und Bedingungen	168
5.8	Abhängigkeit und Unabhängigkeit	172
5.9	Totale Wahrscheinlichkeit	175
5.10	Lernen aus Zusatzinformationen	178
5.11	Zusammengesetzte Zufallsexperimente	180
	Zusammenfassung	184
	Aufgaben	185
6	Parameterschätzungen	189
6.1	Datengrundlage	190
6.2	Zur Identifikation des Modelltyps	194
6.3	Stichproben- und Schätzfunktionen	199
6.3.1	Eigenschaften von Schätzfunktionen	199
6.3.2	Die Stichprobenfunktionen \bar{X} und S^2	201
6.3.3	Experimente zur Untersuchung von Stichproben- funktionen	206
6.4	Zur Konstruktion von Schätzfunktionen	209
6.4.1	Parameterschätzung nach der Methode der Momente	209
6.4.2	Parameterschätzung nach der ML-Methode	213
6.4.3	Fragen an Schätzfunktionen	220
6.5	Check des gefundenen Modells	224

6.5.1	Modellcheck	224
6.5.2	Beispiel: Unfalldaten	227
	Zusammenfassung	230
	Aufgaben	231
7	Konfidenzintervalle	235
7.1	Konfidenzintervall für den Median	237
7.2	Was kostet der Wunsch?	238
7.2.1	Kann es nicht noch etwas vertrauenswürdiger sein?	239
7.2.2	Kann es nicht etwas kürzer sein?	239
7.2.3	Welches k zu vorgegebenem Konfidenzniveau γ ?	240
7.3	Konstruktionsprinzip für Konfidenzintervalle	241
7.4	Konfidenzintervall für einen Anteil p	242
7.5	Fragen an Konfidenzintervalle	245
7.6	Konfidenzintervalle für die Normalverteilung	248
7.7	Anwendung: Raucherrisiken	250
7.8	Caveat – Mahnung	251
	Zusammenfassung	252
	Aufgaben	253
8	Statistik und BAYES	255
8.1	Ein Problem in klassischer Sicht	255
8.1.1	Euro keine Zufallswährung?	255
8.1.2	Zutreffend oder nicht? 1. Versuch	256
8.1.3	Zutreffend oder nicht: 2. Versuch	259
8.1.4	Welches p ?	260
8.1.5	Ein Modell	260
8.1.6	Zwei Lösungsvorschläge für das Schätzproblem	261
8.1.7	Das p aus den p 's	264
8.2	BAYES und der Euro	265
8.2.1	Sammlung und Typisierung von Informationen	265
8.2.2	Beschreibung durch Wahrscheinlichkeitsverteilungen	266
8.2.3	Per aspera ad astra	268
8.2.4	Eine Rechtfertigung?	269
8.2.5	Ein Bayesscher Schätzer	270
8.2.6	Uniform prior	270
8.2.7	Glaubwürdiger	271
8.3	Prior – Sample – Posterior	271
8.3.1	Beta als Prior	271
8.3.2	Gut oder schlecht?	273
8.3.3	Parameterfortschreibung	273
8.3.4	Der Euro und seine Prior	273
8.3.5	$\Pi\alpha\nu\tau\alpha \rho\epsilon\iota$	274

8.3.6	Play it again, Sam!	275
8.4	Beta-Verteilung	276
8.4.1	Dichte	276
8.4.2	Porträt	276
8.4.3	Try yourself	276
8.4.4	Momente	277
8.5	Es hilft auch im Weltall	277
8.5.1	Ärger mit Ariane	277
8.5.2	Noch einmal Richtung BAYES	278
8.5.3	Truncated uniform prior	279
8.5.4	Anwendbarkeit des Bayesschen Ansatzes	280
	Zusammenfassung	281
	Aufgaben	281
9	Testen	283
9.1	Kochen und Testen	283
9.1.1	Das Problem: Zwiebelstatistik	283
9.1.2	Datenbeschaffung	284
9.1.3	Datensichtung und -reduktion	284
9.1.4	Vermutung	285
9.1.5	Idee eines Tests	286
9.2	Der Aufbau eines Tests	289
9.2.1	Hypothesen	289
9.2.2	Entscheidungen und Fehler	290
9.2.3	Teststatistik	291
9.2.4	Entscheidungsregel	291
9.2.5	Allgemeiner Fahrplan eines Tests mit Demonstration	293
9.2.6	Binomialtest mit R	294
9.2.7	Die Gütefunktion eines Tests	295
9.3	Der χ^2 -Test: Ein vielseitiger Geselle	297
9.3.1	Ist Lotto fair – passt die Gleichverteilung?	297
9.3.2	Der χ^2 -Anpassungstest	300
9.3.3	Opfer und Täter – der χ^2 -Unabhängigkeitstest	302
9.4	Eine kleine Testgalerie	303
9.4.1	Kolmogorov-Smirnov-Test	304
9.4.2	Normalverteilung: Test auf μ bei bekanntem σ	305
9.4.3	Normalverteilung: Test auf μ bei unbekanntem σ	306
9.4.4	Test auf Gleichheit der Mittelwerte	307
9.4.5	Vorzeichentest im Einstichprobenfall	309
9.4.6	Vorzeichentest im Zweistichprobenfall	310
9.4.7	Wilcoxon-Test für verbundene Stichproben	312
	Zusammenfassung	313
	Aufgaben	314

10 Regressionsanalyse	317
10.1 Eine Reise für den Überblick	318
10.2 Das lineare Regressionsmodell	322
10.3 Modell-Schätzung und -Check	323
10.3.1 Die Methode der kleinsten Quadrate	323
10.3.2 Ein Anwendungsbeispiel	325
10.3.3 Residualanalyse	327
10.4 Modell-Interpretation	331
10.4.1 Das Bestimmtheitsmaß	331
10.4.2 Konfidenzintervalle für Achsenabschnitt und Steigung	335
10.4.3 $E(Y x_0)$ und Prognose von $Y x_0$	337
10.4.4 Test und Modellvergleich	337
10.5 Ausblick	341
10.5.1 Mehrere erklärende Variablen	341
10.5.2 Nicht lineare Zusammenhänge	344
10.5.3 Variablentransformationen	346
10.5.4 Polynome und lokale Glätter	347
Zusammenfassung	349
Aufgaben	349
11 R-Einführung	353
11.1 Hintergrund, Installation und erste Schritte mit R	353
11.2 Daten einlesen und Statistiken berechnen	355
11.3 Graphiken erstellen	357
11.4 R als Rechenmaschine	359
11.5 Bequemes Arbeiten mit diesem Buch in R	363
11.6 Statistische R-Idioms	365
11.7 Weitere Infos	369
Glossar	371
Literatur	375
Index	377