

Auf einen Blick

Über den Autor	9
Einführung	21
Teil I: Erste Schritte bei der statistischen Analyse mit R.....	25
Kapitel 1: Daten, Statistiken und Entscheidungen	27
Kapitel 2: R: Was R kann und wie R das macht.....	37
Teil II: Daten beschreiben	69
Kapitel 3: Daten grafisch darstellen	71
Kapitel 4: Suchen Sie Ihre Mitte	109
Kapitel 5: Abweichungen vom Durchschnitt	119
Kapitel 6: Standards und Wertungen kennenlernen	127
Kapitel 7: Alles zusammenfassen	137
Kapitel 8: Was ist normal?	157
Teil III: Rückschlüsse aus Daten ziehen	175
Kapitel 9: Die Sache mit dem Vertrauen: Schätzung	177
Kapitel 10: Ein-Stichproben-Hypothesentest	191
Kapitel 11: Zwei-Stichproben-Hypothesentest	217
Kapitel 12: Mehr als zwei Stichproben testen	243
Kapitel 13: Komplexere Tests	267
Kapitel 14: Lineare, multiple und allgemeine lineare Regression	287
Kapitel 15: Korrelation: Aufstieg und Fall von Zusammenhängen	321
Kapitel 16: Kurvenförmige Regression: Wenn Beziehungen kompliziert werden	341
Teil IV: Umgang mit der Wahrscheinlichkeit	365
Kapitel 17: Einführung in die Wahrscheinlichkeit	367
Kapitel 18: Einführung in die statistische Modellierung	387
Teil V: Der Top-Ten-Teil	407
Kapitel 19: (Fast) zehn nützliche R-Onlinere Ressourcen	409
Abbildungsverzeichnis	413
Stichwortverzeichnis	421

Inhaltsverzeichnis

Über den Autor	9
Einführung	21
Über dieses Buch	21
Ähnlichkeiten mit diesem anderen »Für Dummies«-Buch	22
Was Sie nicht lesen müssen	22
Törichte Annahmen über den Leser	22
Wie dieses Buch aufgebaut ist	22
Teil I: Erste Schritte bei der statistischen Analyse mit R	23
Teil II: Daten beschreiben	23
Teil III: Rückschlüsse aus Daten ziehen	23
Teil IV: Umgang mit der Wahrscheinlichkeit	23
Teil V: Der Top-Ten-Teil	23
Symbole, die in diesem Buch verwendet werden	24
Wie es weitergeht	24
 TEIL I	
ERSTE SCHRITTE BEI DER STATISTISCHEN ANALYSE MIT R	25
Kapitel 1	
Daten, Statistiken und Entscheidungen	27
Die statistischen (und damit verwandten) Begriffe, die Sie einfach kennen müssen	28
Stichproben und Grundgesamtheiten	28
Abhängige und unabhängige Variablen	29
Arten von Daten	30
Ein bisschen Wahrscheinlichkeit	31
Inferenzstatistik: Testen von Hypothesen	32
Nullhypothese und Alternativhypothese	33
Zwei Arten von Fehlern	34
 Kapitel 2	
R: Was R kann und wie R das macht	37
R und RStudio herunterladen	37
Eine Session mit R	41
Das Arbeitsverzeichnis	41
Jetzt geht es richtig los	42
Fehlende Daten	45
R-Funktionen	46
Benutzerdefinierte Funktionen	48
Kommentare	48
R-Strukturen	49
Vektoren	49
Numerische Vektoren	50

14 Inhaltsverzeichnis

Matrizen	51
Faktoren	52
Listen	53
Listen und Statistik	54
Datensätze (Data Frames).....	55
Daten aus einem Datensatz extrahieren.....	57
Packages.....	58
Weitere Packages	60
Die R-Formelschnittstelle	62
Lesen und Schreiben	63
Tabellenkalkulation	63
CSV-Dateien	65
Textdateien	66
TEIL II	
DATEN BESCHREIBEN	69
Kapitel 3	
Daten grafisch darstellen	71
Muster erkennen.....	71
Verteilung grafisch darstellen.....	72
Säulensprünge	72
Die Torte schneiden.....	74
Das verstreute Diagramm.....	75
Kastengrafik: Kästchen und Antennen.....	76
Diagramme mit dem R-Basispaket erstellen	77
Histogramme	77
Diagrammfeatures hinzufügen.....	79
Säulendiagramme.....	80
Kreisdiagramme	82
Punktdiagramme	82
Noch einmal Säulendiagramme.....	83
Streudiagramme.....	86
Boxplots	90
Zu ggplot2 aufsteigen	90
Histogramme	91
Säulendiagramme.....	94
Punktdiagramme	95
Noch einmal Säulendiagramme, die Zweite	98
Streudiagramme.....	101
Matrix von Streudiagrammen.....	103
Boxplots	105
Zusammenfassung und Ausblick.....	108
Kapitel 4	
Suchen Sie Ihre Mitte	109
Mittelwert: Die Lehre vom Durchschnitt.....	109
Der Mittelwert in R: mean().....	111

Wie lauten Ihre Bedingungen?.....	111
Mit with() die Dollarzeichen weglassen.....	112
Die Daten erforschen.....	112
Ausreißer: Schönheitsfehler der Mittelwerte.....	114
Und schließlich noch ein paar andere Mittel.....	114
Mediane: Auf halber Strecke erwisch.....	116
Der Median in R: median().....	117

Kapitel 5

Abweichungen vom Durchschnitt 119

Die Streuung berechnen.....	119
Mittelwert von quadratischen Abweichungen: Varianz, und wie sie berechnet wird.....	120
Varianz einer Stichprobe.....	122
Varianz in R.....	123
Zurück zu den Wurzeln: Standardabweichung.....	124
Standardabweichung einer Grundgesamtheit.....	124
Standardabweichung einer Stichprobe.....	124
Standardabweichung in R.....	125
Bedingungen, Bedingungen, Bedingungen.....	125

Kapitel 6

Standards und Wertungen kennenlernen..... 127

z-Werte einfangen.....	127
Eigenschaften von z-Werten.....	128
Bonds und Ruth.....	128
Prüfungsergebnisse.....	129
Standardwerte in R.....	130
Wo stehen Sie?.....	132
Rangermittlung in R.....	132
Gleiche Werte.....	132
k-kleinsten und k-größten Wert.....	133
Quantile.....	133
Prozentrang.....	134
Zusammenfassen.....	136

Kapitel 7

Alles zusammenfassen 137

Wie viele?.....	137
Groß und klein.....	139
Im Moment leben.....	139
Ein lehrreicher Moment.....	140
Zurück zu den Beschreibungen.....	140
Schiefe.....	141
Kurtosis – Wölbung.....	144
Nun kommt die Häufigkeit ins Spiel.....	145
Nominalskalierte Variablen: table() et al.....	145

Numerische Variable: hist()	146
Kumulierte Häufigkeit	147
Schritt für Schritt: Die empirische kumulative Verteilungsfunktion	147
Numerische Variable: stem()	151
Einen Datensatz zusammenfassen	153

Kapitel 8
Was ist normal? 157

So kratzen Sie die Kurve	157
Tiefer graben	158
Parameter einer Normalverteilung	159
Mit Normalverteilungen arbeiten	161
Verteilungen in R	161
Dichtefunktion	161
Normalverteilung als Kurve darstellen	162
Kumulierte Dichtefunktion	165
Die Verteilungsfunktion zeichnen	167
Quantile der Normalverteilungen	168
Die Verteilungsfunktion mit Quartilen zeichnen	170
Zufällige Stichproben	171
Eine ganz besondere Verteilung	171
Die Standardnormalverteilung in R	173
Die Standardnormalverteilung als Graphen darstellen	174

TEIL III
RÜCKSCHLÜSSE AUS DATEN ZIEHEN 175

Kapitel 9
Die Sache mit dem Vertrauen: Schätzung 177

Stichprobenverteilungen verstehen	177
Ein BESONDERS wichtiges Konzept: der zentrale Grenzwertsatz	179
(Näherungsweise) den zentralen Grenzwertsatz simulieren	180
Vorhersagen des zentralen Grenzwertsatzes	185
Vertrauen: Es gibt Grenzen	186
So ermitteln Sie die Vertrauensgrenzen für einen Mittelwert	186
Passend für ein t	188

Kapitel 10
Ein-Stichproben-Hypothesentest 191

Hypothesen, Tests und Fehler	191
Hypothesentests und Stichprobenverteilungen	193
Noch einmal z-Werte	195
z-Test in R	197
t-Test für eine Stichprobe	199
t-Tests in R	200
Mit t-Verteilungen arbeiten	200

t-Verteilungen visualisieren	201
t mit den R-Basisfunktionen darstellen.....	202
Diagramm mit ggplot2 erstellen.....	204
Eine Sache noch zu ggplot2.....	208
Testen einer Varianz	209
In R testen	210
Mit Chi-Quadrat-Verteilungen arbeiten.....	212
Chi-Quadrat-Verteilungen visualisieren.....	212
Chi-Quadrat mit den R-Basisfunktionen darstellen.....	212
Chi-Quadrat mit ggplot2 darstellen.....	214

Kapitel 11

Zwei-Stichproben-Hypothesentest..... 217

Hypothesen für zwei.....	217
Noch einmal Stichprobenverteilungen.....	218
Den zentralen Grenzwertsatz anwenden.....	219
Noch einmal z-Werte.....	220
Z-Test für zwei Stichproben in R.....	222
t-Test für zwei Stichproben	223
Wie ein Ei dem anderen: gleiche Varianzen	224
t-Test in R.....	225
Zwei Vektoren verwenden	226
Einen Datensatz und eine Formel verwenden	226
Die Ergebnisse visualisieren.....	227
Wie Äpfel und Birnen: unterschiedliche Varianzen.....	230
Ein passendes Paar: Hypothesentest für abhängige Stichproben.....	231
t-Test für abhängige Stichproben in R.....	233
Zwei Varianzen testen.....	233
F-Test in R.....	235
F zusammen mit t.....	236
Mit F-Verteilungen arbeiten	236
F-Verteilungen visualisieren.....	237

Kapitel 12

Mehr als zwei Stichproben testen..... 243

Mehr als zwei Stichproben testen.....	243
Eine harte Nuss.....	244
Eine Lösung.....	245
Wichtige Zusammenhänge	249
ANOVA in R.....	249
Die Ergebnisse visualisieren.....	250
Nach der ANOVA.....	251
Kontraste in R.....	254
Nicht geplante Vergleiche.....	255
Eine andere Art Hypothese, eine andere Art Test	256
Mit wiederholten Messungen bei der Varianzanalyse arbeiten	256

Varianzanalyse für wiederholte Messungen in R.....	258
Die Ergebnisse visualisieren.....	260
Jetzt wird es trendy.....	261
Trendanalyse in R.....	264

Kapitel 13

Komplexere Tests 267

Die Kombinationen knacken.....	267
Interaktionen.....	269
Die Analyse.....	269
Zweifaktorielle Varianzanalyse in R.....	271
Ergebnisse der zweifaktoriellen Varianzanalyse visualisieren.....	272
Zwei Arten von Variablen, und zwar gleichzeitig.....	275
Gemischte ANOVA in R.....	277
Ergebnisse der gemischten ANOVA visualisieren.....	279
Nach der Analyse.....	280
Multivariate Varianzanalyse.....	280
MANOVA in R.....	282
MANOVA-Ergebnisse visualisieren.....	283
Nach der Analyse.....	285

Kapitel 14

Lineare, multiple und allgemeine lineare Regression 287

Das Streudiagramm.....	287
Geraden zeichnen.....	289
Regression: Was für eine Gerade!.....	291
Die Regression für Schätzungen verwenden.....	293
Streuung um die Regressionsgerade.....	293
Hypothesen über die Regression testen.....	295
Lineare Regression in R.....	300
Features des linearen Modells.....	301
Vorhersagen treffen.....	301
Das Streudiagramm und die Regressionsgerade visualisieren.....	302
Residuendiagramm erstellen.....	303
Irrsinnig viele Zusammenhänge auf einmal: multiple Regression.....	304
Multiple Regression in R.....	305
Vorhersagen treffen.....	306
Das 3-D-Streudiagramm und die Regressionsebene visualisieren.....	307
ANOVA: Eine andere Perspektive.....	310
Kovarianzanalyse: Die letzte Komponente des allgemeinen linearen Modells ...	313
Moment bitte – da gibt's noch mehr.....	319

Kapitel 15

Korrelation: Aufstieg und Fall von Zusammenhängen 321

Noch einmal Streudiagramme.....	321
Grundlegendes zur Korrelation.....	322

Korrelation und Regression	324
Hypothesen über Korrelationen testen	327
Ist ein Korrelationskoeffizient größer als null?	327
Unterscheiden sich zwei Korrelationskoeffizienten voneinander?	328
Korrelation in R	329
Korrelationskoeffizient berechnen	329
Korrelationskoeffizient testen	330
Die Differenz zwischen zwei Korrelationskoeffizienten testen	330
Eine Korrelationsmatrix berechnen	331
Korrelationsmatrizen visualisieren	331
Multiple Korrelation	334
Multiple Korrelation in R	334
Das Bestimmtheitsmaß korrigieren	335
Partialkorrelation	336
Partialkorrelation in R	337
Semipartialkorrelation	338
Semipartialkorrelation in R	338

Kapitel 16

Kurvenförmige Regression: Wenn Beziehungen

kompliziert werden 341

Was ist ein Logarithmus?	342
Was ist e ?	344
Potenzregression	346
Exponentielle Regression	352
Logarithmische Regression	356
Polynomische Regression: Eine größere Potenz	359
Welches Modell sollten Sie verwenden?	363

TEIL IV

UMGANG MIT DER WAHRSCHEINLICHKEIT 365

Kapitel 17

Einführung in die Wahrscheinlichkeit 367

Was ist Wahrscheinlichkeit?	367
Experimente, Versuche, Ereignisse und Stichprobenräume	368
Wahrscheinlichkeitsräume und Wahrscheinlichkeit	368
Zusammengesetzte Ereignisse	369
Vereinigung und Schnitt	369
Noch mehr zum Schnitt	370
Bedingte Wahrscheinlichkeit	371
Mit Wahrscheinlichkeiten arbeiten	372
Die Grundlage des Testens von Hypothesen	372
Große Wahrscheinlichkeitsräume	372
Permutationen	373
Kombinationen	374
R-Funktionen für Zählregeln	375

20 Inhaltsverzeichnis

Zufallsvariablen: diskret und stetig.....	376
Wahrscheinlichkeitsverteilungen und Dichtefunktionen	377
Die Binomialverteilung.....	379
Binomial- und negative Binomialverteilung in R.....	380
Binomialverteilung.....	380
Negative Binomialverteilung.....	382
Hypothesen mit der Binomialverteilung testen.....	383
Weitere Informationen zum Testen von Hypothesen: R vs. Tradition	385

Kapitel 18

Einführung in die statistische Modellierung..... 387

Die Modellierung einer Verteilung.....	387
Näheres zur Poissonverteilung.....	388
Modellierung mit der Poissonverteilung.....	389
Prüfen, ob das Modell passt.....	393
Ein kurzer Hinweis zu <code>chisq.test()</code>	394
Modelle für Baseballstatistiken	396
Simulationen	399
Es darauf ankommen lassen: die Monte-Carlo-Methode	399
Den Würfel »zinken«	399
Simulation des zentralen Grenzwertsatzes	403

TEIL V

DER TOP-TEN-TEIL..... 407

Kapitel 19

(Fast) zehn nützliche R-Onlinere Ressourcen 409

Websites für R-Anwender.....	409
R-bloggers	409
Microsoft R Application Network.....	409
Quick-R.....	410
Stack Overflow	410
Online-Bücher und weitere Dokumentation	410
R-Handbuch.....	410
R-Dokumentation	411
RDocumentation	411
The R Journal.....	411

Abbildungsverzeichnis..... 413

Stichwortverzeichnis..... 421