



Auf einen Blick

Einleitung	17
Teil I: Datenanalyse und Modellbildung – Grundlagen	23
Kapitel 1: Datenanalyse als Kunst und Wissenschaft.....	25
Kapitel 2: Orientierung innerhalb der statistischen Techniken.....	47
Kapitel 3: Vertrauen aufbauen und Modelle testen.....	65
Teil II: Mit Hilfe der Regression Vorhersagen treffen	83
Kapitel 4: Einfache lineare Regression verstehen.....	85
Kapitel 5: Zwei Variablen besser als eine: Multiple Regression.....	103
Kapitel 6: Ein Schritt vor, zwei zurück: Auswahl des Regressionsmodells.....	125
Kapitel 7: Mit Daten in die Kurve gehen: Nichtlineare Regression.....	149
Kapitel 8: Ja, Nein, Vielleicht: Vorhersagen mit logistischer Regression.....	171
Teil III: Viele Mittelwerte vergleichen – mit der Varianzanalyse	181
Kapitel 9: Einfache Varianzanalyse.....	183
Kapitel 10: Mit multiplen Vergleichen Paare finden.....	199
Kapitel 11: Weiter mit der zweifachen ANOVA!.....	209
Kapitel 12: Grenzenlos: Wie Regression und ANOVA zusammenhängen.....	221
Teil IV: Starke Verbindungen mit Chi-Quadrat-Tests	233
Kapitel 13: Mit Kreuztabellen Assoziationen bilden.....	235
Kapitel 14: Ausreichend unabhängig für den Chi-Quadrat-Test.....	255
Kapitel 15: Der Chi-Quadrat-Test auf Güte der Anpassung.....	275
Teil V: Rebellen ohne Verteilung	285
Kapitel 16: Es wird nichtparametrisch!.....	287
Kapitel 17: Der Vorzeichentest und der Vorzeichen-Rangtest.....	299
Kapitel 18: Der Rangsummentest.....	313
Kapitel 19: Kruskal-Wallis und Wilcoxon.....	325
Kapitel 20: Korrelationen mit dem Spearman'schen Rang bestimmen.....	339
Teil VI: Der Top-Ten-Teil	347
Kapitel 21: Zehn Fehler in statistischen Schlüssen.....	349
Kapitel 22: Zehn Probleme aus der Praxis.....	359
Anhang A: Tabellen zum Nachschlagen	371
Stichwortverzeichnis	383







Inhaltsverzeichnis

Einleitung	17
Zu diesem Buch.....	17
Konventionen in diesem Buch.....	18
Was Sie nicht lesen müssen.....	19
Törichte Annahmen über den Leser.....	19
Wie dieses Buch aufgebaut ist.....	20
Die Symbole in diesem Buch.....	21
Wie es weitergeht.....	22
TEIL I	
DATENANALYSE UND MODELLBILDUNG – GRUNDLAGEN	23
Kapitel 1	
Datenanalyse als Kunst und Wissenschaft	25
Datenanalyse: Nicht mehr nur für Statistiker.....	26
Die gute alte Zeit.....	26
Der Nachteil der heutigen Statistik-Software.....	27
Regel Nr.1: Informieren Sie sich VOR der Verarbeitung!.....	28
Nichts ist ewig (nicht einmal eine Gerade).....	28
Datenschnüffeln ist nicht cool!.....	29
(Daten-)Fischen verboten!.....	31
Das große Ganze: Ein Überblick über weiterführende Statistik.....	35
Populationsparameter.....	36
Stichprobenkenngröße.....	36
Vertrauensintervall.....	37
Hypothesentest.....	37
Varianzanalyse (ANOVA, Analysis of Variance).....	38
Multiple Vergleiche.....	39
Interaktionseffekte.....	40
Korrelation.....	42
Lineare Regression.....	43
Chi-Quadrat-Tests.....	44
Nichtparametrische Statistik.....	46
Kapitel 2	
Orientierung innerhalb der statistischen Techniken	47
Qualitative und quantitative Variablen in der statistischen Analyse.....	48
Statistiken für qualitative Variablen.....	49
Anteile vergleichen.....	49
Einen Anteil abschätzen.....	50
Nach Zusammenhängen zwischen qualitativen Variablen suchen.....	50
Modelle für Vorhersagen erstellen.....	52





10 Inhaltsverzeichnis

Statistik für quantitative Variablen.....	53
Vergleiche anstellen	53
Zusammenhänge erkennen	54
Vorhersagen treffen	55
Verzerrung vermeiden.....	56
Betrachtung der Verzerrung durch die Brille der Statistik.....	57
Der Varianzstreitfall: der Kampf zwischen $n - 1$ und n	58
Höchste Genauigkeit erzielen	59
Genauigkeit aus statistischer Perspektive verstehen.....	60
Genauigkeit mit dem Fehlerspielraum messen	60
Schlussfolgerungen treffen und Grenzen erkennen	63

Kapitel 3

Vertrauen aufbauen und Modelle testen..... 65

Parameter anhand von Vertrauensintervallen schätzen	66
Die Grundlagen: Die allgemeine Form eines Vertrauensintervalls.....	66
Das Vertrauensintervall für einen Populationsmittelwert finden.....	67
Was ändert den Fehlerspielraum?.....	68
Ein Vertrauensintervall interpretieren	72
Modelle aufstellen und testen.....	72
Was sind H_0 und H_a wirklich?.....	73
Ihre Hinweise in einer Teststatistik zusammenfassen.....	74
Die Stärke des Beweises mit einem p-Wert bestimmen	74
Typ-I- und Typ-II-Fehler ausschließen.....	75
Die Güte eines Hypothesentests	77

TEIL II

MIT HILFE DER REGRESSION VORHERSAGEN TREFFEN..... 83

Kapitel 4

Einfache lineare Regression verstehen..... 85

Mit Streudiagrammen und Korrelationen Beziehungen untersuchen	86
Mit Hilfe von Streudiagrammen Beziehungen untersuchen	87
Informationen mit Hilfe des Korrelationskoeffizienten zuordnen.....	88
Ein einfaches lineares Regressionsmodell erstellen	89
Die beste Gerade für die Modellierung Ihrer Daten	89
Der y-Schnittpunkt der Regressionsgeraden	90
Die Steigung der Regressionsgeraden	92
Schätzungen anhand der Regressionsgeraden.....	93
Prüfen, ob das Modell passt.....	93
Definition der Bedingungen.....	93
Die Residuen finden und untersuchen	95
Mit r^2 messen, ob das Modell geeignet ist	98
Ausreißer	99





Inhaltsverzeichnis 11

Korrekte Schlüsse ziehen.....	100
Vermeiden Sie, Ursache- und Wirkungsdiskussionen.....	100
Extrapolation: No-No!.....	101
Die Grenzen eines einfachen linearen Regressionsmodells kennen.....	102

Kapitel 5

Zwei Variablen besser als eine: Multiple Regression..... 103

Das multiple Regressionsmodell.....	103
Die Verwendungszwecke der multiplen Regression erkennen.....	104
Die allgemeine Form des multiplen Regressionsmodells.....	104
Die Analyseschritte.....	105
Alle x und y betrachten.....	105
Daten sammeln.....	106
Mögliche Beziehungen erkennen.....	107
Streudiagramme erstellen.....	107
Korrelationen: Untersuchung der Verbindung.....	109
Auf Multikollinearität prüfen.....	112
Das am besten angepasste Modell finden.....	113
Die Koeffizienten des multiplen Regressionsmodells bestimmen.....	113
Die Koeffizienten interpretieren.....	114
Koeffizienten testen.....	116
Vorhersage von y anhand der x-Variablen.....	117
Prüfen, wie gut das Modell angepasst ist.....	118
Die Voraussetzungen.....	118
Ein Plan für die Prüfung der Voraussetzungen.....	119

Kapitel 6

Ein Schritt vor, zwei zurück: Auswahl des Regressionsmodells..... 125

Abstoßdistanzen schätzen – Der ultimative Kick.....	126
Variablen-Brainstorming und -Datensammlung.....	126
Streudiagramme und Korrelationen untersuchen.....	127
Beziehungen mit Hilfe von Streudiagrammen erkennen.....	128
Verbindungen unter Verwendung von Korrelationen erkennen.....	129
Die Vorwärtsselektion.....	131
Variablen hinzufügen – eine nach der anderen.....	131
Wie gut passt das Modell?.....	132
Die Vorwärtsselektion für die Abstoßdistanzen.....	134
Und jetzt das Ganze von hinten: Die Rückwärtsselektion.....	136
Variablen nacheinander eliminieren.....	138
Passt das Modell?.....	140
Variablen ausschließen, um die Abstoßdistanz zu bestimmen.....	140
Das Verfahren der besten Teilmengen.....	143
Alle Modelle erzeugen und dann das beste davon auswählen.....	144
Die besten Teilmengen für das Beispiel der Abstoßdistanz.....	145



12 Inhaltsverzeichnis

Vergleich der Modellauswahlverfahren	147
Warum erzeugen nicht alle Verfahren dieselben Ergebnisse?	147
Welche Vor- und Nachteile haben die einzelnen Verfahrensweisen?	147

Kapitel 7

Mit Daten in die Kurve gehen: Nichtlineare Regression 149

Am Anfang war das Streudiagramm	150
Polynome für Kurven nutzen	151
Zurück zu den Polynomen	151
Suche nach dem besten Polynomialmodell	152
Die Prüfung mit einem Polynom zweiten Grades bestehen	154
Bewertung der Anpassung eines Polynomialmodells	158
Vorhersagen treffen	162
Nach oben? Nach unten? Exponentiell!	163
Zurück zu den exponentiellen Modellen	163
Die Suche nach dem besten exponentiellen Modell	164
Geheimnisse in exponentieller Geschwindigkeit verbreiten	166

Kapitel 8

Ja, Nein, Vielleicht: Vorhersagen mit logistischer

Regression 171

Aufstellung des logistischen Regressionsmodells	171
Definition eines logistischen Regressionsmodells	172
Mit einer S-Kurve Wahrscheinlichkeiten abschätzen	172
Interpretation der Koeffizienten des logistischen Regressionsmodells	173
Abschätzen der Wahrscheinlichkeit, dass ein Film ein Kassenschlager wird – mit der logistischen Regression	173
Allgemeine Schritte für die logistische Regression	174
Die Analyse in SPSS	175
Die Koeffizienten finden und das Modell erstellen	176
Schätzung von p	177
Die Anpassung des Modells überprüfen	178

TEIL III

VIELE MITTELWERTE VERGLEICHEN – MIT DER VARIANZANALYSE

181

Kapitel 9

Einfache Varianzanalyse..... 183

Zwei Mittelwerte mit einem t-Test vergleichen	183
Mehr Mittelwerte mit einer ANOVA vergleichen	185
Kerne spucken: Eine Situation, die nach einer ANOVA schreit	185
Die ANOVA-Schritte	186



Inhaltsverzeichnis 13

Die Bedingungen prüfen	187
Unabhängigkeit prüfen.....	187
Was ist schon normal?	187
Die Streuung.....	189
Die Hypothesen aufstellen.....	189
Der F-Test	190
Die ANOVA in SPSS.....	191
Die Varianz in Quadratsummen zerlegen	191
Die mittleren Quadratsummen suchen	192
Die F-Statistik bestimmen	193
Schlüsse aus der ANOVA ziehen.....	194
Was kommt als Nächstes?.....	195
Die Anpassung des ANOVA-Modells überprüfen	196

Kapitel 10 **Mit multiplen Vergleichen Paare finden 199**

Nach der ANOVA.....	199
Mit Fisher und Tukey unterschiedliche Mittelwerte dingfest machen.....	203
Mit dem Fisher-LSD-Test Unterschiede erkennen.....	203
Mit dem Tukey-Test vergleichen.....	205

Kapitel 11 **Weiter mit der zweifachen ANOVA! 209**

Das Modell für die zweifache ANOVA aufstellen.....	210
Die Behandlungen festlegen	210
Die Quadratsummen	211
Interaktionseffekte verstehen	212
Was ist eine Interaktion überhaupt?.....	213
Interaktion und Interaktionsdiagramme	213
Die Terme in der zweifachen ANOVA testen.....	216
Die zweifache ANOVA-Tafel erstellen	217
Ergebnisse interpretieren: Zahlen und Grafiken	217

Kapitel 12 **Grenzenlos: Wie Regression und ANOVA zusammenhängen 221**

Regression in den Augen der Variation	221
Variabilität in den y feststellen und nach x suchen, die sie erklären.....	222
Ergebnisse durch Regression.....	224
Bewertung der Anpassung des Regressionsmodells	225
Regression und ANOVA: Treffen der Modelle	226
Quadratsummen vergleichen.....	226
Die Freiheitsgrade aufteilen	228
Regression und die ANOVA-Tafel.....	229
F- und t-Statistik: Die letzte Hürde.....	230





14 Inhaltsverzeichnis

TEIL IV

STARKE VERBINDUNGEN MIT CHI-QUADRAT-TESTS 233

Kapitel 13

Mit Kreuztabellen Assoziationen bilden 235

Eine Kreuztabelle erstellen.....	236
Daten in einer Kreuztabelle anordnen.....	236
Die Zellenwerte eintragen.....	237
Randsummen erstellen.....	237
Die Wahrscheinlichkeiten zerlegen.....	238
Randwahrscheinlichkeiten.....	239
Gemeinsame Wahrscheinlichkeiten.....	240
Bedingte Wahrscheinlichkeiten.....	241
Unabhängigkeit anstreben.....	247
Auf Unabhängigkeit zwischen zwei Kategorien prüfen.....	247
Auf Unabhängigkeit zwischen zwei Variablen prüfen.....	248
Das Simpson-Paradoxon – keine Zauberei.....	249
Das Simpson-Paradoxon untersuchen.....	250
Die Frage nach dem Warum: Simpson-Paradoxon.....	252
Achten Sie immer auf das Simpson-Paradoxon!.....	253

Kapitel 14

Ausreichend unabhängig für den Chi-Quadrat-Test..... 255

Ein Hypothesentest auf Unabhängigkeit.....	256
Daten sammeln.....	257
Die Hypothesen festlegen.....	258
Bestimmung der erwarteten Zellenwerte.....	258
Die Bedingungen für den Test prüfen.....	261
Die Chi-Quadrat-Teststatistik berechnen.....	261
Ihre Ergebnisse in der Chi-Quadrat-Tabelle nachschlagen.....	264
Ihre Schlüsse.....	267
Zwei Tests für den Vergleich von zwei Anteilen vergleichen.....	269
Zurück zum Z-Test für zwei Populationsanteile.....	269
Chi-Quadrat-Tests und Z-Tests für eine 2 x 2-Tabelle gleichsetzen.....	270

Kapitel 15

Der Chi-Quadrat-Test auf Güte der Anpassung..... 275

Die Teststatistik für die Güte der Anpassung bestimmen.....	275
Beobachtetes mit Erwartetem vergleichen.....	276
Die Statistik für die Güte der Anpassung berechnen.....	278
Interpretation der Statistik für die Güte der Anpassung mit Hilfe von Chi-Quadrat.....	280
Als Erstes werden die Bedingungen geprüft.....	281
Die Schritte für den Chi-Quadrat-Test auf Güte der Anpassung.....	282



Inhaltsverzeichnis 15

TEIL V	
REBELLEN OHNE VERTEILUNG	285
Kapitel 16	
Es wird nichtparametrisch!	287
Argumente für die nichtparametrische Statistik.....	287
Kein Problem, wenn Bedingungen nicht erfüllt sind	288
Der Median	288
Wo also liegt der Haken?.....	291
Die Grundlagen der nichtparametrischen Statistik	292
Vorzeichen (Signum).....	292
Rang.....	294
Vorzeichen-Rang	295
Rangsumme	296
Kapitel 17	
Der Vorzeichentest und der Vorzeichen-Rangtest	299
Die Vorzeichen erkennen: Der Vorzeichentest.....	299
Den Median testen	301
Den Median schätzen	303
Tests für verbundene Stichproben	305
Noch einen Schritt weiter: Mit dem Vorzeichen-Rangtest.....	307
Eine Einschränkung des Vorzeichentests.....	307
Die Schritte für einen Vorzeichen-Rangtest.....	308
Gewichtsverlust mit Vorzeichen-Rängen	309
Kapitel 18	
Der Rangsummentest	313
Den Rangsummentest durchführen.....	314
Die Bedingungen prüfen.....	314
Der schrittweise Test	314
Die Stichprobengröße heraufsetzen	316
Einen Rangsummentest durchführen: Wer verkauft die Häuser schneller?.....	317
Die Bedingungen für diesen Test überprüfen	318
Die Hypothesen testen.....	319
Kapitel 19	
Kruskal-Wallis und Wilcoxon	325
Mit dem Kruskal-Wallis-Test mehr als zwei Populationen vergleichen	325
Die Bedingungen prüfen.....	326
Aufstellung des Testproblems.....	329
Den Test schrittweise durchführen	330
Die Differenzen erkennen: Der Wilcoxon-Rangsummentest.....	332
Paarweise Vergleiche	333
Vergleichstest stellen fest, wer anders ist.....	334
Die Mediane untersuchen, um festzustellen, wie sie sich unterscheiden... ..	336



16 Inhaltsverzeichnis

Kapitel 20

Korrelationen mit dem Spearman'schen Rang bestimmen 339

Pearson und seine aufwändigen Bedingungen	340
Bewertungen mit Hilfe der Spearman-Rangkorrelation	341
Den Spearman-Rangkorrelationskoeffizienten bestimmen	342
Spearman in der Praxis: Talent und Leistung in eine Beziehung bringen ...	343

TEIL VI

DER TOP-TEN-TEIL.....347

Kapitel 21

Zehn Fehler in statistischen Schlüssen 349

Diese Statistiken beweisen	349
Es ist technisch nicht statistisch signifikant, aber	350
Das bedeutet, x verursacht y	350
Ich nahm an, die Daten seien normalverteilt	351
Ich berichte nur über »wichtige« Ergebnisse.....	352
Eine größere Stichprobe ist immer besser	353
Es ist technisch nicht zufällig, aber	353
1.000 Antworten sind 1.000 Antworten	354
Natürlich gelten diese Ergebnisse für die gesamte Population!	356
Ich habe beschlossen, das wegzulassen!	357

Kapitel 22

Zehn Probleme aus der Praxis 359

Mittelwerte mit der einfachen ANOVA vergleichen	359
Multiple Vergleiche	360
Mit der zweifachen ANOVA zwei Faktoren betrachten.....	361
Vorhersage einer quantitativen Variablen mit Hilfe der Regression.....	362
Eine Wahrscheinlichkeit mit der logistischen Regression vorhersagen	363
Nichtlineare Regression für gekrümmte Daten.....	365
Mit Chi-Quadrat auf Unabhängigkeit testen.....	366
Spezielle Modelle mit dem Test auf Güte der Anpassung testen	367
Den Median mit dem Rangsummentest schätzen	367
Die Modellanpassung mit R^2 überprüfen.....	368

Anhang A: Tabellen zum Nachschlagen 371

t-Tabelle	371
Binomialtabelle.....	373
Chi-Quadrat-Tabelle	376
Rangsummentabelle	378
F-Tabelle	379

Stichwortverzeichnis 383

