

Ludwig Fahrmeir • Rita Künstler  
Iris Pigeot • Gerhard Tutz

# Statistik

Der Weg zur Datenanalyse

Dritte, verbesserte Auflage

Mit 165 Abbildungen  
und 34 Tabellen



Springer

# Inhaltsverzeichnis

<b>Vorwort</b>	<b>v</b>
<b>1. Einführung</b>	<b>1</b>
1.1 Wo braucht man Statistik? . . . . .	1
1.2 Was macht man mit Statistik? . . . . .	11
1.3 Was steht am Anfang? . . . . .	13
1.3.1 Statistische Einheiten, Merkmale und Gesamtheiten . . . . .	13
1.3.2 Merkmalstypen . . . . .	15
1.4 Wie gewinnt man Daten? . . . . .	19
1.4.1 Elemente der Versuchsplanung . . . . .	20
1.4.2 Datengewinnung und Erhebungsarten . . . . .	22
Einfache Zufallsstichproben . . . . .	24
Geschichtete Zufallsstichproben . . . . .	24
Klumpenstichprobe . . . . .	25
Mehrstufige Auswahlverfahren . . . . .	25
Bewußte Auswahlverfahren . . . . .	26
1.5 Zusammenfassung und Bemerkungen . . . . .	27
1.6 Aufgaben . . . . .	28
<b>2 Univariate Deskription und Exploration von Daten</b>	<b>29</b>
2.1 Verteilungen und ihre Darstellungen . . . . .	29
2.1.1 Häufigkeiten . . . . .	30
2.1.2 Graphische Darstellungen . . . . .	32
Stab- und Kreisdiagramme . . . . .	32
Stamm-Blatt-Diagramme . . . . .	35
Histogramme . . . . .	38
Unimodale und multimodale Verteilungen . . . . .	45
Symmetrie und Schiefe . . . . .	46
2.1.3 Kumulierte Häufigkeitsverteilung und empirische Verteilungsfunktion . . . . .	46

2.2	Beschreibung von Verteilungen	51
2.2.1	Lagemaße	51
	Arithmetisches Mittel	51
	Median	53
	Modus	55
	Berechnung der Lagemaße bei gruppierten Daten	56
	Lageregeln	58
	Das geometrische Mittel	59
	Das harmonische Mittel	61
	Das getrimmte Mittel	62
2.2.2	Quantile und Box-Plot	62
2.2.3	Standardabweichung, Varianz und Variationskoeffizient	67
2.2.4	Maßzahlen für Schiefe und Wölbung	72
2.3	Konzentrationsmaße	74
2.3.1	Relative Konzentration: Lorenzkurve und Gini-Koeffizient	75
	Lorenzkurve aus den geordneten Daten	75
	Lorenzkurve bei gruppierten Daten	78
	Gini-Koeffizient	80
2.3.2	Alternative Konzentrationsmaße	82
	Konzentrationsrate $CR_g$	82
	Herfindahl-Index	83
2.4	Dichtekurven und Normalverteilung	84
2.4.1	Dichtekurven	84
2.4.2	Normalverteilungen	89
	*Normal-Quantil-Plots	93
*2.4.3	Approximation von Dichtekurven	98
2.5	Zusammenfassung und Bemerkungen	103
2.6	Aufgaben	104
<b>3</b>	<b>Multivariate Deskription und Exploration</b>	<b>109</b>
3.1	Diskrete und gruppierte Merkmale	109
3.1.1	Zweidimensionale Daten: Die Kontingenztafel	109
3.1.2	Bedingte Häufigkeiten	115
3.2	Zusammenhangsanalyse in Kontingenztafeln	119
3.2.1	Chancen und relative Chancen	119
3.2.2	Kontingenz- und $\chi^2$ -Koeffizient	122
3.3	Graphische Darstellungen quantitativer Merkmale	127
3.3.1	Streudiagramm	128
3.3.2	Zweidimensionale Histogramme und Dichten	130
3.3.3	Mehrdimensionale Darstellungen	132

3.4	Zusammenhangsmaße bei metrischen Merkmalen . . . . .	135
3.4.1	Empirischer Korrelationskoeffizient nach Bravais-Pearson . . . . .	135
3.4.2	Spearman's Korrelationskoeffizient . . . . .	141
3.4.3	Invarianzeigenschaften . . . . .	146
3.5	Korrelation und Kausalität . . . . .	147
3.6	Regression . . . . .	152
3.6.1	Das lineare Regressionsmodell . . . . .	152
3.6.2	Die Berechnung der Ausgleichsgeraden . . . . .	153
3.6.3	Bestimmtheitsmaß und Residualanalyse . . . . .	158
*3.6.4	Nichtlineare Regression . . . . .	165
3.7	Zusammenfassung und Bemerkungen . . . . .	167
3.8	Aufgaben . . . . .	169
<b>4</b>	<b>Wahrscheinlichkeitsrechnung</b> . . . . .	<b>171</b>
4.1	Definition und Begriff der Wahrscheinlichkeit . . . . .	172
4.1.1	Mengen und Mengenoperationen . . . . .	173
4.1.2	Zufallsereignisse . . . . .	177
4.1.3	Wahrscheinlichkeiten . . . . .	179
4.2	Zur empirischen Interpretation von Wahrscheinlichkeiten . . . . .	185
4.2.1	Die Laplace-Wahrscheinlichkeit . . . . .	186
4.2.2	Objektive Wahrscheinlichkeiten als Grenzwert relativer Häufigkeiten . . . . .	189
4.2.3	Subjektive Wahrscheinlichkeiten . . . . .	191
4.3	Zufallsstichproben und Kombinatorik . . . . .	192
4.3.1	Modell mit Zurücklegen . . . . .	193
4.3.2	Modell ohne Zurücklegen . . . . .	194
4.3.3	Permutationen . . . . .	195
4.3.4	Modell ohne Zurücklegen und ohne Berücksichtigung der Reihenfolge . . . . .	196
4.4	Bedingte Wahrscheinlichkeiten . . . . .	199
4.5	Unabhängigkeit von zwei Ereignissen . . . . .	203
4.6	Totale Wahrscheinlichkeit . . . . .	206
4.7	Der Satz von Bayes . . . . .	209
4.8	Unendliche Grundgesamtheiten . . . . .	213
4.9	Zusammenfassung und Bemerkungen . . . . .	215
4.10	Aufgaben . . . . .	217
<b>5</b>	<b>Diskrete Zufallsvariablen</b> . . . . .	<b>221</b>
5.1	Zufallsvariablen . . . . .	221
5.2	Verteilungen und Parameter von diskreten Zufallsvariablen . . . . .	225
5.2.1	Definition und Verteilung . . . . .	225
5.2.2	Unabhängigkeit von diskreten Zufallsvariablen . . . . .	236

5.2.3	Lageparameter, Quantile und Streuungsparameter einer diskreten Verteilung	239
5.3	Spezielle diskrete Verteilungsmodelle	250
5.3.1	Die Binomialverteilung	250
5.3.2	Die hypergeometrische Verteilung	256
5.3.3	Die Poisson-Verteilung	258
5.4	Zusammenfassung und Bemerkungen	262
5.5	Aufgaben	264
<b>Stetige Zufallsvariablen</b>		<b>267</b>
6.1	Definition und Verteilung	267
6.2	Lageparameter, Quantile und Varianz von stetigen Zufallsvariablen	279
6.3	Spezielle stetige Verteilungsmodelle	289
6.3.1	Die Normalverteilung	289
6.3.2	Die logarithmische Normalverteilung	297
6.3.3	Chi-Quadrat-, Student- und Fisher-Verteilung	298
6.4	Zusammenfassung und Bemerkungen	302
6.5	Aufgaben	304
<b>Mehr über Zufallsvariablen und Verteilungen</b>		<b>307</b>
7.1	Gesetz der großen Zahlen und Grenzwertsätze	307
7.1.1	Das Gesetz der großen Zahlen und der Hauptsatz der Statistik	309
7.1.2	Der zentrale Grenzwertsatz	312
7.2	Approximation von Verteilungen	315
*7.3	Zufallszahlen und Simulation	318
*7.4	Einige Ergänzungen	321
7.4.1	Zufallsvariablen als Abbildungen	321
7.4.2	Verteilungsfunktion und ihre Eigenschaften	323
7.4.3	Ungleichung von Tschebyscheff	325
7.4.4	Maßzahlen für Schiefe und Wölbung	327
7.5	Zusammenfassung und Bemerkungen	328
7.6	Aufgaben	329
<b>Mehrdimensionale Zufallsvariablen</b>		<b>331</b>
8.1	Begriff mehrdimensionaler Zufallsvariablen	331
8.2	Zweidimensionale diskrete Zufallsvariablen	334
8.3	Zweidimensionale stetige Zufallsvariablen	339
8.4	Unabhängigkeit von Zufallsvariablen	342
8.5	Kovarianz und Korrelation	345
8.6	Die zweidimensionale Normalverteilung	353

8.7	Zusammenfassung und Bemerkungen . . . . .	357
8.8	Aufgaben . . . . .	357
<b>9</b>	<b>Parameterschätzung</b> . . . . .	<b>359</b>
9.1	Punktschätzung . . . . .	360
9.2	Eigenschaften von Schätzstatistiken . . . . .	362
9.2.1	Erwartungstreue . . . . .	362
9.2.2	Erwartete mittlere quadratische Abweichung und Konsistenz . . . . .	366
9.2.3	Wirksamste Schätzstatistiken . . . . .	369
9.3	Konstruktion von Schätzfunktionen . . . . .	371
9.3.1	Maximum Likelihood-Schätzung . . . . .	372
9.3.2	Kleinste-Quadrate-Schätzung . . . . .	375
9.4	Intervallschätzung . . . . .	376
9.4.1	Konfidenzintervalle für Erwartungswert und Varianz . . . . .	378
9.4.2	Konfidenzintervalle für den Anteilswert . . . . .	382
9.5	Zusammenfassung und Bemerkungen . . . . .	383
9.6	Aufgaben . . . . .	384
<b>10</b>	<b>Testen von Hypothesen</b> . . . . .	<b>387</b>
10.1	Der Binomial- und der Gauß-Test . . . . .	387
10.1.1	Der exakte Binomialtest . . . . .	391
10.1.2	Der approximative Binomialtest . . . . .	394
10.1.3	Der Gauß-Test . . . . .	397
10.2	Prinzipien des Testens . . . . .	401
	Fehlentscheidungen . . . . .	404
	Statistische Tests und Konfidenzintervalle . . . . .	407
	Überschreitungswahrscheinlichkeit . . . . .	408
	Gütefunktion . . . . .	409
	*Multiple Testprobleme . . . . .	416
10.3	Zusammenfassung und Bemerkungen . . . . .	417
10.4	Aufgaben . . . . .	419
<b>11</b>	<b>Spezielle Testprobleme</b> . . . . .	<b>421</b>
11.1	Ein-Stichproben-Fall . . . . .	422
11.1.1	Tests zu Lagealternativen . . . . .	423
11.1.2	Anpassungstests . . . . .	432
11.2	Vergleiche aus unabhängigen Stichproben . . . . .	441
11.2.1	Tests zu Lagealternativen . . . . .	442
11.2.2	$\chi^2$ -Homogenitätstest . . . . .	448
11.3	Vergleiche aus verbundenen Stichproben . . . . .	450

11.4	Zusammenhangsanalyse. . . . .	452
11.4.1	$\chi^2$ -Unabhängigkeitstest.....	452
11.4.2	Korrelation bei metrischen Merkmalen. . . . .	454
11.5	Zusammenfassung und Bemerkungen. . . . .	456
11.6	Aufgaben. . . . .	457
<b>12</b>	<b>Regressionsanalyse</b>	<b>459</b>
12.1	Lineare Einfachregression. . . . .	460
12.1.1	Das Modell der linearen Einfachregression. . . . .	460
12.1.2	Schätzen, Testen und Prognose. . . . .	463
12.1.3	Residualanalyse. . . . .	473
12.2	Multiple lineare Regression. . . . .	476
12.2.1	Das multiple lineare Regressionsmodell. . . . .	478
12.2.2	Schätzen, Testen und Prognose. . . . .	479
* 12.2.3	Multiple lineare Regression in Matrixnotation. . . . .	488
* 12.3	Nichtlineare und nichtparametrische Regression. . . . .	490
12.4	Zusammenfassung und Bemerkungen. . . . .	494
12.5	Aufgaben. . . . .	495
<b>13</b>	<b>Varianzanalyse</b>	<b>497</b>
13.1	Einfaktorielle Varianzanalyse. . . . .	498
13.2	Zweifaktorielle Varianzanalyse mit festen Effekten. . . . .	507
13.3	Zusammenfassung und Bemerkungen. . . . .	521
13.4	Aufgaben. . . . .	522
<b>14</b>	<b>Zeitreihen</b>	<b>525</b>
14.1	Indizes. . . . .	528
14.2	Komponentenmodelle. . . . .	532
14.3	Globale Regressionsansätze. . . . .	534
14.3.1	Trendbestimmung. . . . .	534
14.3.2	Bestimmung der Saisonkomponente. . . . .	536
14.4	Lokale Ansätze. . . . .	537
14.4.1	Trendbestimmung. . . . .	538
	Gleitende Durchschnitte. . . . .	538
	Lokale Regression. . . . .	539
	*Spline-Glättung. . . . .	542
14.4.2	Bestimmung der Saisonkomponente. . . . .	544
	Gleitende Durchschnitte und lokale Regression. . . . .	544
	*Spline-Glättung. . . . .	546
14.5	Zusammenfassung und Bemerkungen. . . . .	547

---

14.6 Aufgaben . . . . .	547
<b>Tabellen</b>	<b>551</b>
A Standardnormalverteilung . . . . .	551
B Binomialverteilung . . . . .	552
C $\chi^2$ -Verteilung . . . . .	567
D Students $t$ -Verteilung . . . . .	568
E $F$ -Verteilung . . . . .	569
F Wilcoxon-Vorzeichen-Rang-Test . . . . .	578
G Wilcoxon-Rangsummen-Test . . . . .	578
<b>Literatur</b>	<b>581</b>
<b>Sachverzeichnis</b>	<b>585</b>