

Ludwig Fahrmeir • Rita Künstler  
Iris Pigeot • Gerhard Tutz

# Statistik

Der Weg zur Datenanalyse

Siebte, neu bearbeitete Auflage

*kl* 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?	14
1.3.1 Statistische Einheiten, Merkmale und Gesamtheiten	14
1.3.2 Merkmalstypen	16
Stetige und diskrete Merkmale	16
Skalen	17
Quantitative und qualitative Merkmale	19
1.4 Wie gewinnt man Daten?	20
1.4.1 Elemente der Versuchsplanung	21
1.4.2 Datengewinnung und Erhebungsarten	23
Einfache Zufallsstichproben	25
Geschichtete Zufallsstichproben	26
Klumpenstichprobe	26
Mehrstufige Auswahlverfahren	27
Bewußte Auswahlverfahren	27
Studiendesigns	28
1.5 Zusammenfassung und Bemerkungen	28
1.6 Aufgaben	30
<b>2 Univariate Deskription und Exploration von Daten</b>	<b>31</b>
2.1 Verteilungen und ihre Darstellungen	31
2.1.1 Häufigkeiten	32
2.1.2 Graphische Darstellungen	35
Stab- und Kreisdiagramme	35
Stamm-Blatt-Diagramme	37
Histogramme	40
Unimodale und multinodale Verteilungen	47

	Symmetrie und Schiefe . . . . .	48
2.1.3	Kumulierte Häufigkeitsverteilung und empirische Verteilungsfunktion . . . . .	49
2.2	Beschreibung von Verteilungen . . . . .	53
2.2.1	Lagemaße . . . . .	53
	Arithmetisches Mittel . . . . .	53
	Mediän . . . . .	55
	Modus . . . . .	57
	Berechnung der Lagemaße bei gruppierten Daten . . . . .	58
	Lageregeln . . . . .	60
	Das geometrische Mittel . . . . .	61
	Das harmonische Mittel . . . . .	63
	Das getrimmte Mittel . . . . .	64
2.2.2	Quartile und Box-Plot . . . . .	64
2.2.3	Standardabweichung, Varianz und Variationskoeffizient . . . . .	69
2.2.4	Maßzahlen für Schiefe und Wölbung . . . . .	74
2.3	Konzentrationsmaße . . . . .	77
2.3.1	Relative Konzentration: Lorenzkurve und Gini-Koeffizient . . . . .	78
	Lorenzkurve aus den geordneten Daten . . . . .	78
	Lorenzkurve bei gruppierten Daten . . . . .	80
	Gini-Koeffizient . . . . .	82
2.3.2	Alternative Konzentrationsmaße . . . . .	85
	Konzentrationsrate $CR_g$ . . . . .	85
	Herfindalil-Index . . . . .	86
2.4	Dichtekurven und Normal Verteilung . . . . .	87
2.4.1	Dichtekurven . . . . .	87
2.4.2	Normalverteilungen . . . . .	90
	*Normal-Quantil-Plots . . . . .	95
*2.4.3	Approximation von Dichtekurven . . . . .	98
2.5	Zusammenfassung und Bemerkungen . . . . .	102
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 Kontingenztabelle . . . . .	109
3.1.2	Bedingte Häufigkeiten . . . . .	115
3.2	Zusammenhangsanalyse in Kontingenztabellen . . . . .	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	Zusammenhänge bei metrischen Merkmalen	135
3.4.1	Empirischer Korrelationskoeffizient nach Bravais-Pearson	135
3.4.2	Spearman's Korrelationskoeffizient	142
3.4.3	Invarianzeigenschaften	146
3.5	Korrelation und Kausalität	148
3.6	Regression	153
3.6.1	Das lineare Regressionsmodell	153
3.6.2	Die Berechnung der Ausgleichsgeraden	154
3.6.3	Bestimmtheitsmaß und Residualanalyse	159
*3.6.4	Nichtlineare Regression	166
3.7	Zusammenfassung und Bemerkungen	167
3.8	Aufgaben	169
<b>4</b>	<b>Wahrscheinlichkeitsrechnung</b>	<b>173</b>
4.1	Definition und Begriff der Wahrscheinlichkeit	174
4.1.1	Mengen und Mengenoperationen	175
4.1.2	Zufallsereignisse	179
4.1.3	Wahrscheinlichkeiten	181
4.2	Zur empirischen Interpretation von Wahrscheinlichkeiten	187
4.2.1	Die Laplace-Wahrscheinlichkeit	188
4.2.2	Objektive Wahrscheinlichkeiten als Grenzwert relativer Häufigkeiten	191
4.2.3	Subjektive Wahrscheinlichkeiten	194
4.3	Zufallsstichproben und Kombinatorik	195
4.3.1	Modell mit Zurücklegen	196
4.3.2	Modell ohne Zurücklegen	197
4.3.3	Permutationen	198
4.3.4	Modell ohne Zurücklegen und ohne Berücksichtigung der Reihenfolge	199
4.4	Bedingte Wahrscheinlichkeiten	202
4.5	Unabhängigkeit von zwei Ereignissen	206
4.6	Totale Wahrscheinlichkeit	209
4.7	Der Satz von Bayes	211
4.8	Unendliche Grundgesamtheiten	216
4.9	Zusammenfassung und Bemerkungen	219
4.10	Aufgaben	220

<b>5</b>	<b>Diskrete Zufallsvariablen</b>	<b>223</b>
5.1	Zufallsvariablen	224
5.2	Verteilungen und Parameter von diskreten Zufallsvariablen	227
5.2.1	Definition und Verteilung	227
	Diskrete Gleichverteilung	234
	Geometrische Verteilung	234
5.2.2	Unabhängigkeit, von diskreten Zufallsvariablen	238
5.2.3	Lageparameter. Quantile und Streuungsparameter einer diskreten Verteilung	242
	Erwartungswert	242
5.2.4	Weitere Lageparameter	247
	Varianz und Standardabweichung	248
5.3	Spezielle diskrete Verteilungsmodelle	252
5.3.1	Die Binomialverteilung	253
5.3.2	Die hypergeometrische Verteilung	258
5.3.3	Die Poisson-Verteilung	260
5.4	Zusammenfassung und Bemerkungen	265
5.5	Aufgaben	267
<b>6</b>	<b>Stetige Zufallsvariablen</b>	<b>269</b>
6.1	Definition und Verteilung	269
	Unabhängigkeit von stetigen Zufallsvariablen	275
	Exponentialverteilung	279
6.2	Lageparameter. Quantile und Varianz von stetigen Zufallsvariablen	282
	Erwartungswert	282
	Modus	286
	Mediän und Quantile	286
	Varianz	288
	Standardisierung von Zufallsvariablen	291
	Symmetrie und Schiefe	291
6.3	Spezielle stetige Verteilungsmodelle	293
6.3.1	Die Normalverteilung	293
	Quantile	296
6.3.2	Die logarithmische Normalverteilung	301
6.3.3	Chi-Quadrat-, Student- und Fisher-Verteilung	301
	Die Chi-Quadrat-Verteilung	302
	Die Student-Verteilung	303
	Die Fisher-Verteilung	305
6.4	Zusammenfassung und Bemerkungen	307
6.5	Aufgaben	308

<b>7 Mehr über Zufallsvariablen und Verteilungen</b>	<b>311</b>
7.1 Gesetz der großen Zahlen und Grenzwertsätze	311
7.1.1 Das Gesetz der großen Zahlen und der Hauptsatz der Statistik	313
7.1.2 Der zentrale Grenzwertsatz	315
7.2 Approximation von Verteilungen	318
*7.3 Zufallszahlen und Simulation	321
*7.4 Einige Ergänzungen	324
7.4.1 Zufallsvariablen als Abbildungen	324
7.4.2 Verteilungsfunktion und ihre Eigenschaften	326
7.4.3 Ungleichung von Tschebyscheff	328
7.4.4 Maßzahlen für Schiefe und Wölbung	330
7.5 Zusammenfassung und Bemerkungen	331
7.6 Aufgaben	332
<b>8 Mehrdimensionale Zufallsvariablen</b>	<b>335</b>
8.1 Begriff mehrdimensionaler Zufallsvariablen	335
8.2 Zweidimensionale diskrete Zufallsvariablen	338
8.3 Zweidimensionale stetige Zufallsvariablen	344
8.4 Unabhängigkeit von Zufallsvariablen	346
8.5 Kovarianz und Korrelation	349
8.6 Die zweidimensionale Normalverteilung	357
8.7 Zusammenfassung und Bemerkungen	361
8.8 Aufgaben	361
<b>9 Parameterschätzung</b>	<b>363</b>
9.1 Punktschätzung	364
9.2 Eigenschaften von Schätzstatistiken	366
9.2.1 Erwartungstreue	367
9.2.2 Erwartete mittlere quadratische Abweichung und Konsistenz	370
9.2.3 Wirksamste Schätzstatistiken	374
9.3 Konstruktion von Schätzfunktionen	376
9.3.1 Maximum Likelihood-Schätzung	376
9.3.2 Kleinste-Quadrate-Schätzung	379
9.3.3 Bayes-Schätzung	380
9.4 Intervallschätzung	385
9.4.1 Konfidenzintervalle für Erwartungswert und Varianz	387
9.4.2 Konfidenzintervalle für den Anteilswert	392
9.5 Zusammenfassung und Bemerkungen	393
9.6 Aufgaben	394

<b>10 Testen von Hypothesen</b>	<b>397</b>
10.1 Der Binomial- und der Gauß-Test	397
10.1.1 Der exakte Binomialtest	401
10.1.2 Der approximative Binomialtest	404
10.1.3 Der Gauß-Test	408
10.2 Prinzipien des Testens	411
10.2.1 Fehlentscheidungen	415
10.2.2 Statistische Tests und Konfidenzintervalle	418
10.2.3 Überschreitungswahrscheinlichkeit	419
10.2.4 Gütefunktion	420
*Multiple Testprobleme	427
10.3 Zusammenfassung und Bemerkungen	428
10.4 Aufgaben	430
<b>11 Spezielle Testprobleme</b>	<b>433</b>
11.1 Ein-Stichproben-Fall	435
11.1.1 Tests zu Lagealternativen	435
11.1.2 Anpassungstests	445
11.2 Vergleiche aus unabhängigen Stichproben	454
11.2.1 Tests zu Lagealternativen	455
11.2.2 $\chi^2$ -Homogenitätstest	462
11.3 Vergleiche aus verbundenen Stichproben	465
11.4 Zusammenhangsanalyse	466
11.4.1 $\chi^2$ -Unabhängigkeitstest	467
11.4.2 Korrelation bei metrischen Merkmalen	469
11.5 Zusammenfassung und Bemerkungen	471
11.6 Aufgaben	472
<b>12 Regressionsanalyse</b>	<b>475</b>
12.1 Lineare Einfachregression	476
12.1.1 Das Modell der linearen Einfachregression	476
12.1.2 Schätzen, Testen und Prognose	480
Schätzen	480
Testen	485
Prognose	488
12.1.3 Residualanalyse	490
12.2 Multiple lineare Regression	492
12.2.1 Das multiple lineare Regressionsmodell	494
12.2.2 Schätzen, Testen und Prognose	495
Schätzen	496

Testen . . . . .	499
Prognose . . . . .	502
*12.2.3 Multiple, lineare Regression in Matrixnotation . . . . .	503
12.3 Binäre Regression . . . . .	506
*12.4 Nichtlineare und nichtparanietrische Regression . . . . .	508
12.5 Zusammenfassung und Bemerkungen . . . . .	513
12.6 Aufgaben . . . . .	514
<b>13 Varianzanalyse</b>	<b>517</b>
13.1 Einfaktorielle Varianzanalyse . . . . .	519
Modellformulierung (I) . . . . .	520
ModellformulieruHg (II) . . . . .	521
13.2 Zweifaktorielle Varianzanalyse mit festen Effekten . . . . .	528
Modellformulierung (I) . . ." . . . . .	530
Modellformulierung (II) . . . . .	531
13.3 Zusammenfassung und Bemerkungen . . . . .	542
13.4 Aufgaben . . . . .	543
<b>14 Zeitreihen</b>	<b>547</b>
14.1 Indizes . . . . .	550
14.2 Komponentenmodelle .! . . . .	553
14.3 Globale Regressionsansätze . . . . .	555
14.3.1 Trendbestimmung . . . . .	556
14.3.2 Bestimmung der Saisonkomponente . . . . .	557
14.4 Lokale Ansätze . . . . .	559
14.4.1 Trendbestimmung . . . . .	560
Gleitende Durchschnitte . . . . .	560
Lokale Regression . . . . .	561
*Spline-Glättung . . . . .	563
14.4.2 Bestimmung der Saisorikoinporieite . . . . .	566
Gleitende Durchschnitte und lokale Regression . . . . .	566
*Spline-Glättung . . . . .	568
14.5 Zusammenfassung und Bemerkungen . . . . .	569
14.6 Aufgaben . . . . .	569



<b>Tabellen</b>	<b>573</b>
A Standardnormal Verteilung . . . . .	573
B Binomialverteilung . . . . .	574
C $\chi^2$ -Verteilung . . . . .	583
D Students $t$ -Verteilung . . . . .	584
E $F$ -Verteilung . . . . .	585
F Wilcoxon-Vorzeichen-Rang-Test . . . . .	590
G Wilcoxon-Rangsummen-Test . . . . .	591
<b>Literatur</b>	<b>593</b>
<b>Verzeichnis der Beispiele</b>	<b>597</b>
<b>Sachregister</b>	<b>603</b>