

Ludwig Fahrmeir • Rita Künstler  
Iris Pigeot • Gerhard Tutz

# Statistik

Der Weg zur Datenanalyse

**Vierte, verbesserte Auflage**

Mit 162 Abbildungen  
und 25 Tabellen

(pjj 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	29
<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	39
Unimodale und multimodale Verteilungen	47

Symmetrie und Schiefe	48
2.1.3 Kumulierte Häufigkeitsverteilung und empirische Verteilungsfunktion .	49
2.2 Beschreibung von Verteilungen	52
Lagemaße	53
Arithmetisches Mittel	53
Median	55
Modus	56
Berechnung der Lagemaße bei gruppierten Daten	58
Lageregeln	60
Das geometrische Mittel	61
Das harmonische Mittel	63
Das getrimmte Mittel	63
2.2.2 Quantile und Box-Plot	64
(2^3) Standardabweichung, Varianz und Variationskoeffizient	68
2.2.4 Maßzahlen für Schiefe und Wölbung	74
2.3 Konzentrationsmaße	76
2.3.1 Relative Konzentration: Lorenzkurve und Gini-Koeffizient	77
Lorenzkurve aus den geordneten Daten	77
Lorenzkurve bei gruppierten Daten	80
Gini-Koeffizient	82
2.3.2 Alternative Konzentrationsmaße	84
Konzentrationsrate $CR_g$	84
Herfindahl-Index	85
2.4 Dichtekurven und Normalverteilung	86
2.4.1 Dichtekurven	86
2.4.2 Normalverteilungen	89
*Normal-Quantil-Plots	94
*2.4.3 Approximation von Dichtekurven	97
2.5 Zusammenfassung und Bemerkungen	101
2.6 Aufgaben	103
<b>3 Multivariate Deskription und Exploration</b>	<b>107</b>
3.1 Diskrete und gruppierte Merkmale	107
3.1.1 Zweidimensionale Daten: Die Kontingenztafel	107
3.1.2 Bedingte Häufigkeiten	113
3.2 Zusammenhangsanalyse in Kontingenztafeln	117
3.2.1 Chancen und relative Chancen	117
3.2.2 Kontingenz- und $\chi^2$ -Koeffizient	120
3.3 Graphische Darstellungen quantitativer Merkmale	125

3.3.1	Streudiagramm	126
3.3.2	Zweidimensionale Histogramme und Dichten	128
3.3.3	Mehrdimensionale Darstellungen	131
3.4	Zusammenhangsmaße bei metrischen Merkmalen	132
3.4.1	Empirischer Korrelationskoeffizient nach Bravais-Pearson	132
3.4.2	Spearman's Korrelationskoeffizient	139
3.4.3	Invarianzeigenschaften	143
3.5	Korrelation und Kausalität	145
3.6	Regression	150
3.6.1	Das lineare Regressionsmodell	150
3.6.2	Die Berechnung der Ausgleichsgeraden	151
3.6.3	Bestimmtheitsmaß und Residualanalyse	156
*3.6.4	Nichtlineare Regression	163
3.7	Zusammenfassung und Bemerkungen	165
3.8	Aufgaben	167
<b>4</b>	<b>Wahrscheinlichkeitsrechnung</b>	<b>171</b>
( $\mathbb{N}$ )	Definition und Begriff der Wahrscheinlichkeit	172
4.1.1	Mengen und Mengenoperationen	173
4.1.2	Zufallsereignisse	177
4.1.3	Wahrscheinlichkeiten	179
( $\mathbb{R}$ )	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	192
	Zufallsstichproben und Kombinatorik	193
4.3.1	Modell mit Zurücklegen	194
4.3.2	Modell ohne Zurücklegen	195
4.3.3	Permutationen	196
4.3.4	Modell ohne Zurücklegen und ohne Berücksichtigung der Reihenfolge	197
( $\mathbb{P}$ )	Bedingte Wahrscheinlichkeiten	200
(4.5)	Unabhängigkeit von zwei Ereignissen	204
(4.6)	Totale Wahrscheinlichkeit	207
4.7	Der Satz von Bayes	209
4.8	Unendliche Grundgesamtheiten	214
4.9	Zusammenfassung und Bemerkungen	217
4.10	Aufgaben	218

<b>5 ^ Diskrete Zufallsvariablen</b>	<b>221</b>
5.1 Zufallsvariablen	222
5.2 Verteilungen und Parameter von diskreten Zufallsvariablen	225
5.2.1 Definition und Verteilung	225
Diskrete Gleichverteilung	232
Geometrische Verteilung	232
5.2.2 Unabhängigkeit von diskreten Zufallsvariablen	236
5.2.3 Lageparameter, Quantile und Streuungsparameter einer diskreten Verteilung	240
Erwartungswert	240
5.2.4 Weitere Lageparameter	245
Varianz und Standardabweichung	246
5.3 Spezielle diskrete Verteilungsmodelle	250
5.3.1 Die Binomialverteilung	251
5.3.2 Die hypergeometrische Verteilung	256
5.3.3 Die Poisson-Verteilung	258
5.4 Zusammenfassung und Bemerkungen	263
5.5 Aufgaben	265
<b>Stetige Zufallsvariablen</b>	<b>267</b>
6.1 Definition und Verteilung	267
Unabhängigkeit von stetigen Zufallsvariablen	273
Exponentialverteilung	277
6.2 Lageparameter, Quantile und Varianz von stetigen Zufallsvariablen	280
Erwartungswert	280
Modus	284
Median und Quantile	284
Varianz	286
Standardisierung von Zufallsvariablen	289
Symmetrie und Schiefe	289
6.3 Spezielle stetige Verteilungsmodelle	291
6.3.1 Die Normalverteilung	291
Quantile	294
6.3.2 Die logarithmische Normalverteilung	299
6.3.3 Chi-Quadrat-, Student- und Fisher-Verteilung	299
Die Chi-Quadrat-Verteilung	300
Die Student-Verteilung	301
Die Fisher-Verteilung	303
6.4 Zusammenfassung und Bemerkungen	305
6.5 Aufgaben	306

<b>7 Mehr über Zufallsvariablen und Verteilungen</b>	<b>309</b>
7.1 Gesetz der großen Zahlen und Grenzwertsätze	309
7.1.1 Das Gesetz der großen Zahlen und der Hauptsatz der Statistik - - - -	311
7.1.2 Der zentrale Grenzwertsatz	313
7.2 Approximation von Verteilungen	316
*7.3 Zufallszahlen und Simulation	319
*7.4 Einige Ergänzungen	322
7.4.1 Zufallsvariablen als Abbildungen	322
7.4.2 Verteilungsfunktion und ihre Eigenschaften	324
7.4.3 Ungleichung von Tschebyscheff	326
7.4.4 Maßzahlen für Schiefe und Wölbung	328
7.5 Zusammenfassung und Bemerkungen	329
7.6 Aufgaben	330
<b>8 Mehrdimensionale Zufallsvariablen</b>	<b>333</b>
8.1 Begriff mehrdimensionaler Zufallsvariablen	333
8.2 Zweidimensionale diskrete Zufallsvariablen	336
8.3 Zweidimensionale stetige Zufallsvariablen	342
8.4 Unabhängigkeit von Zufallsvariablen	344
8.5 Kovarianz und Korrelation	347
8.6 Die zweidimensionale Normalverteilung	355
8.7 Zusammenfassung und Bemerkungen	359
8.8 Aufgaben	359
<b>9 Parameterschätzung</b>	<b>361</b>
9.1 Punktschätzung	362
9.2 Eigenschaften von Schätzstatistiken	364
9.2.1 Erwartungstreue	365
9.2.2 Erwartete mittlere quadratische Abweichung und Konsistenz	368
9.2.3 Wirksamste Schätzstatistiken	372
9.3 Konstruktion von Schätzfunktionen	374
9.3.1 Maximum Likelihood-Schätzung	374
9.3.2 Kleinste-Quadrate-Schätzung	377
9.3.3 Bayes-Schätzung	378
9.4 Intervallschätzung	383
9.4.1 Konfidenzintervalle für Erwartungswert und Varianz	385
9.4.2 Konfidenzintervalle für den Anteilswert	390
9.5 Zusammenfassung und Bemerkungen	391
9.6 Aufgaben	392

<b>10 Testen von Hypothesen</b>	<b>395</b>
^10.1 Der Binomial- und der Gauß-Test	395
10.1.1 Der exakte Binomialtest	399
10.1.2 Der approximative Binomialtest	402
10.1.3 Der Gauß-Test	406
10.2 Prinzipien des Testens	409
10.2.1 Fehlentscheidungen	413
10.2.2 Statistische Tests und Konfidenzintervalle	416
10.2.3 Überschreitungswahrscheinlichkeit	417
10.2.4 Gütefunktion	418
*Multiple Testprobleme	425
10.3 Zusammenfassung und Bemerkungen	426
10.4 Aufgaben	428
<b>11 Spezielle Testprobleme</b>	<b>431</b>
11.1 Ein-Stichproben-Fall	433
11.1.1 Tests zu Lagealternativen	433
11.1.2 Anpassungstests	443
11.2 Vergleiche aus unabhängigen Stichproben	452
11.2.1 Tests zu Lagealternativen	453
11.2.2 $\chi^2$ -Homogenitätstest	460
11.3 Vergleiche aus verbundenen Stichproben	463
11.4 Zusammenhangsanalyse	464
11.4.1 $\chi^2$ -Unabhängigkeitstest	465
11.4.2 Korrelation bei metrischen Merkmalen	467
11.5 Zusammenfassung und Bemerkungen	469
11.6 Aufgaben	470
<b>12 Regressionsanalyse</b>	<b>473</b>
12.1 Lineare Einfachregression	474
12.1.1 Das Modell der linearen Einfachregression	474
12.1.2 Schätzen, Testen und Prognose	478
Schätzen	478
Testen	483
Prognose	486
12.1.3 Residualanalyse	488
12.2 Multiple lineare Regression	491
12.2.1 Das multiple lineare Regressionsmodell	492
12.2.2 Schätzen. Testen und Prognose	494
Schätzen	494

Testen	497
Prognose	500
*12.2.3 Multiple lineare Regression in Matrixnotation	501
12.3 Binäre Regression	504
*12.4 Nichtlineare und nichtparametrische Regression	506
12.5 Zusammenfassung und Bemerkungen	510
12.6 Aufgaben	511
<b>13 Varianzanalyse</b>	<b>515</b>
13.1 Einfaktorielle Varianzanalyse	517
Modellformulierung (I)	518
Modellformulierung (II)	519
13.2 Zweifaktorielle Varianzanalyse mit festen Effekten	526
Modellformulierung (I)	528
Modellformulierung (II)	529
13.3 Zusammenfassung und Bemerkungen	540
13.4 Aufgaben	541
<b>14 Zeitreihen</b>	<b>545</b>
14.1 Indizes	548
14.2 Komponentenmodelle	551
14.3 Globale Regressionsansätze	553
14.3.1 Trendbestimmung	554
14.3.2 Bestimmung der Saisonkomponente	555
14.4 Lokale Ansätze	557
14.4.1 Trendbestimmung	558
Gleitende Durchschnitte	558
Lokale Regression	559
*Spline-Glättung	561
14.4.2 Bestimmung der Saisonkomponente	564
Gleitende Durchschnitte und lokale Regression	564
*Spline-Glättung	566
14.5 Zusammenfassung und Bemerkungen	567
14.6 Aufgaben	567



<b>Tabellen</b>	<b>571</b>
A Standardnormalverteilung	571
B Binomialverteilung	572
C $\chi^2$ -Verteilung	581
D Students $t$ -Verteilung	582
E $F$ -Verteilung	583
F Wilcoxon-Vorzeichen-Rang-Test	588
G Wilcoxon-Rangsummen-Test	589
<b>Literatur</b>	<b>591</b>
<b>Verzeichnis der Beispiele</b>	<b>595</b>
<b>Sachregister</b>	<b>601</b>