

Wolf-Michael Kahler

Statistische Datenanalyse

**Verfahren verstehen und
mit SPSS gekonnt einsetzen**

5., verbesserte und erweiterte Auflage

Mit 292 Abbildungen



Inhaltsverzeichnis

1	Einleitung	1
2	Verteilungen	9
2.1	Datenaufbereitung und empirische Häufigkeitsverteilung	9
2.2	Absolute, relative und prozentuale Häufigkeiten.	11
2.3	Präsentation von empirischen Verteilungen.	14
2.4	Gliederung einer Verteilung.	17
2.5	Klassierung von Daten.	22
2.6	Verteilungsverläufe.	26
2.7	Normalverteilungen.	29
3	Das Skalenniveau von Merkmalen	33
4	Kennzeichnung des Zentrums	37
4.1	Zentrale Tendenz bei intervallskalierten Merkmalen.	37
4.2	Zentrale Tendenz bei ordinalskalierten Merkmalen.	44
4.3	Zentrale Tendenz bei nominalskalierten Merkmalen.	46
5	Kennzeichnung der Variabilität	49
5.1	Variabilität intervallskalierter Merkmale.	50
5.2	Variabilität ordinalskalierter Merkmale.	55
5.3	Schiefe und Wölbung.	56
5.4	Zusammenfassung.	57
6	Einsatz des Datenanalyseystems SPSS	59
6.1	Datenerfassung und Analyseanforderung.	59
6.2	Anzeige von Analyseergebnissen.	62
6.3	Auswahl, Klassierung und Sicherung	68

7	Standardisierung und Normalverteilung	73
7.1	Prozentränge und Profil-Diagramme	73
7.2	Die z-Transformation	77
7.3	Inverse z-Transformation und Flächengleichheit	82
7.4	Prüfung auf Normalverteilung	85
7.5	Bildung von Gesamt-Indikatoren	88
8	Statistische Beziehungen	91
8.1	Statistische Abhängigkeit und statistische Unabhängigkeit	91
8.2	Kontingenz-Tabellen	94
8.3	Partial-Tabellen	99
8.4	Boxplots	102
8.5	Stärke und Richtung von statistischen Zusammenhängen	103
9	Die Stärke des statistischen Zusammenhangs bei nominalskalierten Merkmalen	105
9.1	Der Chi-Quadrat-Koeffizient	105
9.2	Der Phi-Koeffizient für 2x2-Tabellen	108
9.3	Der Koeffizient "Cramer's V" für rxc-Tabellen	111
9.4	Der Kontingenz-Koeffizient "C"	111
9.5	PRE-Maße	112
9.6	Das PRE-Maß "Lambda"	114
9.7	Das PRE-Maß "Tau"	117
10	Die Stärke des statistischen Zusammenhangs bei ordinalskalierten Merkmalen	119
10.1	Konkordante und diskordante Paare	119
10.2	Die Statistik "Gamma"	124
10.3	Die Statistik "Somers' d"	129
10.4	Die Kendall'schen Statistiken	131
11	Die Stärke des statistischen Zusammenhangs bei intervallskalierten Merkmalen	133
11.1	Streudiagramme und gemeinsame Variation	133
11.2	Die Regressionsgerade	138
11.3	Das PRE-Maß "Determinationskoeffizient"	143
11.4	Der Produktmoment-Korrelationskoeffizient "r"	147
11.5	Trennschärfe und Regression zum Mittel	150

12 Weitere Statistiken zur Beschreibung von statistischen Beziehungen	153
12.1 Der Rangkorrelationskoeffizient von Spearman	153
12.2 Statistiken zur Beschreibung der Ähnlichkeit von Merkmalsträgern	156
12.2.1 Der Konkordanzkoeffizient von Kendall	156
12.2.2 Der Kappa-Koeffizient von Cohen	160
12.3 Der Korrelationskoeffizient "Eta" und der punkt-biseriale Korrelationskoeffizient	162
12.3.1 Nichtlineare Abhängigkeiten	162
12.3.2 Die Statistik "Eta-Quadrat".	163
12.3.3 PRE-Modell-Erklärung von "Eta-Quadrat".	165
12.3.4 Der punkt-biseriale Korrelationskoeffizient.	167
12.4 Mittelwertunterschiede und Korrelation.	168
13 Kontrolle von statistischen Beziehungen	179
13.1 Scheinkorrelationen und multivariate Zusammenhänge.	179
13.2 Die partielle Korrelation.	184
14 Multivariate Datenanalyse	189
14.1 Lineare Einfachregression und lineare Mehrfachregression.	189
14.1.1 Modell der "Linearen Einfachregression".	189
14.1.2 Modell der "Linearen Mehrfachregression".	190
14.1.3 Vektoren und Matrizen.	193
14.1.4 Bestimmung der Regressionskoeffizienten.	200
14.1.5 Probleme bei der Berechnung von Regressionskoeffizienten	205
14.2 Faktorenanalyse.	209
14.2.1 Das Hauptachsen-Modell und das Hauptkomponenten-Modell.	209
14.2.2 Matrix-Darstellung und Fundamentaltheorem.	213
14.2.3 Bestimmung der Komponenten-Matrix durch die Hauptachsen-Methode.	216
14.2.4 Die Extraktion von Faktoren.	221
14.2.5 Rotation zur Einfachstruktur.	224
14.2.6 Schiefwinklge Rotation.	228
15 Zufallsstichproben	231

16 Prüfung der statistischen Beziehung und der Anpassung (χ^2-Test)	243
16.1 Nullhypothesen und Alternativhypothesen	243
16.2 Prüfung der statistischen Beziehung mit einem χ^2 -Test	245
16.3 Die Testverteilung " $\chi^2(4f)$ " • • •!	255
16.4 Durchführung des χ^2 -Tests zur Prüfung der statistischen Beziehung	260
16.5 Fehlerarten bei der Test-Entscheidung.	264
16.6 Die Prüfung von Verteilungseigenschaften mit einem χ^2 -Test (χ^2 -Anpassungstest).	271
16.7 Signifikanz-Tests und Kreuzvalidierung	276
17 Prüfung von Zentren (z-Test, t-Test)	279
17.1 Nullhypothesen über Parameter.	279
17.1.1 Parameter der Grundgesamtheit	279
17.1.2 Beispiele für Nullhypothesen	280
17.1.3 Parametrische und nichtparametrische Signifikanz-Tests	282
17.2 Der einseitige z-Test zur Prüfung einer Mitte.	283
17.2.1 Die Normalverteilung als Testverteilung	283
17.2.2 Null- und Alternativhypothese.	284
17.2.3 Durchführung des z-Tests (als einseitiger z-Test).	289
17.2.4 Der Fehler 2. Art (beim einseitigen z-Test).	292
17.2.5 Die Operationscharakteristik-und die Power-Kurve	295
17.3 Der zweiseitige z-Test zur Prüfung einer Mitte.	299
17.3.1 Durchführung des z-Tests (als zweiseitiger Test).	299
17.3.2 Der Fehler 2. Art (beim zweiseitigen z-Test).	302
17.4 Der z-Test zur Prüfung einer Mitte (für einen beliebigen Stichprobenumfang).	303
17.4.1 Die Teststatistik X	303
17.4.2 Strategie der Testdurchführung	305
17.4.3 Durchführung eines ein-und eines zweiseitigen z-Tests	306
17.4.4 Vergleich der Teststärken von ein- und zweiseitigen z-Tests	307
17.4.5 Durchführung eines z-Tests bei unbekannter Streuung	309
17.4.6 Verletzung der Test-Voraussetzungen beim z-Test	311
17.5 Der t-Test zur Prüfung einer Mitte.	312
17.5.1 Test-Voraussetzungen und Teststatistik	312
17.5.2 Durchführung eines ein-und eines zweiseitigen t-Tests	315
17.6 Zusammenfassung	317

18 Optimaler Stichprobenumfang und Effektgröße	319
18.1 Probleme bei zu großem bzw. zu geringem Stichprobenumfang . . .	319
18.2 Kontrolle des Fehlers 2. Art	323
18.3 Indifferenzbereich und optimaler Stichprobenumfang' . . . !	326
18.4 Effektgrößen und a-priori-Poweranalysen . . . • ' . . ' • •	328
18.5 Bestimmung optimaler Stichprobenumfänge beim t-Test	335
18.6 Durchführung von Post-hoc-Analysen.	339
18.7 Effektgröße und optimaler Stichprobenumfang beim χ^2 -Test	346
18.8 Statistische Testtheorien	351
19 Schätzung von Parametern und Ermittlung von Konfidenzintervallen	355
19.1 Schätzung von Parametern.	355
19.1.1 Schätzung der Mitte.	355
19.1.2 Schätzung der Populations-Varianz.	356
19.1.3 Eigenschaften der Schätzstatistik \bar{X}	357
19.2 Ermittlung von Konfidenzintervallen	358
19.2.1 Signifikanz-Test und Akzeptanzbereich	359
19.2.2 Konstruktion von Konfidenzintervallen	360
19.2.3 Berechnung von Konfidenzintervallen für die Mitte	363
19.2.4 Berechnung von Mindest-Stichprobenumfängen	366
19.2.5 Eigenschaften von Konfidenzintervallen.	367
19.3 Prüfung von Nullhypothesen durch die Berechnung von Konfidenz- intervallen.	369
19.3.1 Konfidenzintervall für den Korrelationskoeffizienten " ρ " . . .	369
19.3.2 Signifikanztest zur Prüfung des Korrelationskoeffizienten " ρ " und Bestimmung des optimalen Stichprobenumfangs . . .	371
19.3.3 Konfidenzintervall für den Prozentsatz 'V.	373
19.3.4 Signifikanztest zur Prüfung des Prozentsatzes "TT" und Be- stimmung des optimalen Stichprobenumfangs.	375
20 Parametrische Prüfung auf Unterschiede	377
20.1 Treatment-Effekte und Untersuchungspläne.	377
20.1.1 Kontrollgruppenplan	378
20.1.2 Unabhängige Stichproben	378
20.1.3 Nullhypothese.	379
20.1.4 Mittelwertdifferenz und Variation.	379

20.1.5	Abhängige Stichproben	381
20.1.6	Paarbildung und Randomisierung	382
20.2	t-Test für abhängige Stichproben	383
20.2.1	Nullhypothese und Teststatistik	384
20.2.2	Testdurchführung (mit SPSS)	385
20.2.3	Der "Vortest-Nachtest-Plan"	387
20.2.4	Poweranalyse beim abhängigen t-Test	389
20.3	t-Test für unabhängige Stichproben	392
20.3.1	Teststatistiken	392
20.3.2	Testdurchführung (mit SPSS)	394
20.3.3	Poweranalyse beim unabhängigen t-Test	398
20.4	Prüfung der Varianzhomogenität bei unabhängigen Stichproben durch den Levene-Test	401
20.5	Prüfung der Varianzhomogenität bei unabhängigen Stichproben durch einen F-Test	403
20.5.1	Nullhypothese und Teststatistik	403
20.5.2	Bestimmung des Akzeptanzbereichs	404
20.5.3	Inferenzschluss beim F-Test	405
20.5.4	Testdurchführung	406
21	Nichtparametrische Prüfung auf Unterschiede	409
21.1	Nichtparametrische und parametrische Tests	409
21.2	Test für zwei unabhängige Stichproben (U-Test von Mann-Whitney)	411
21.2.1	Nullhypothese, Teststatistik und kritische Werte	411
21.2.2	Testdurchführung	415
21.3	Test für zwei abhängige Stichproben (Wilcoxon-Test)	419
21.3.1	Nullhypothese, Teststatistik und kritische Werte	419
21.3.2	Testdurchführung	422
22	Varianzanalyse	425
22.1	Statistische Beziehungen	425
22.2	Voraussetzungen und Nullhypothese der Varianzanalyse	427
22.3	Zerlegung der Gesamt-Stichprobenvariation	428
22.4	Der F-Test der Varianzanalyse	430
22.5	Durchführung des F-Tests	433

22.6	Vergleiche einzelner Faktorstufen	435
22.7	Überprüfung der Voraussetzungen der Varianzanalyse.	439
22.8	Poweranalyse bei der 1-faktoriellen Varianzanalyse.	440
22.9	Weitere Mehrstichprobenvergleiche.	447
22.9.1	Der H-Test von Kruskal-Wallis für unabhängige Stichproben	448
22.9.2	Varianzanalyse für abhängige Stichproben.	450
22.9.3	Poweranalyse bei der Varianzanalyse für abhängige Stichproben.	453
22.9.4	Friedman'sche Rang varianzanalyse für abhängige Stichproben.	458
Anhang		461
A.1	Kodierung des Fragebogens.	461
A.2	Flächenanteile der Standardnormalverteilung.	462
A.3	Das empirische und das numerische Relativ.	464
A.4	Wahrscheinlichkeiten.	467
A.5	Zentrum und Dispersion von theoretischen Verteilungen.	477
A.6	Zufallszahlen-Tafel	480
A.7	Kritische Werte bei χ^2 -Verteilungen.	482
A.8	Kritische Werte bei t-Verteilungen.	482
A.9	Kritische Werte bei F-Verteilungen.	483
A. 10	Optimale Stichprobenumfänge.	487
A.11	Kritische Werte für den U-Test	489
A. 12	Kritische Werte für den Wilcoxon-Test.	490
A.13	Datenbasis.	490
A. 14	Werte der inversen Fisher'schen z-Transformation.	493
Literaturverzeichnis		494
Index		495