

# **Forschungsmethoden und Statistik**

**für Psychologen und Sozialwissenschaftler**

**3., aktualisierte und erweiterte Auflage**

**Peter Sedlmeier  
Frank Renkewitz**

# Inhaltsverzeichnis

Kapitel 1	Methoden der Psychologie	1
Kapitel 2	Wahrnehmung und Denken	1
Kapitel 3	Emotionen und Motivation	1

## **Teil I Grundlagen und Konzepte** **1**

### Kapitel 1 Methoden der Psychologie 1

1.1	Die Fallstricke der Alltagspsychologie	5
1.1.1	Fehler beim Wahrnehmen	5
1.1.2	Fehler beim Erinnern	8
1.1.3	Fehler beim logischen Denken	10
1.1.4	Fehler beim Umgang mit Wahrscheinlichkeiten	11
1.2	Sprachgebrauch in Alltag und Wissenschaft	12
1.2.1	Missverständnisse beim Verstehen von Sprache im Alltag	12
1.2.2	Präzisierung der Sprache in der Wissenschaft	13
1.3	Die wissenschaftliche Methode	15
1.3.1	Theorien, Hypothesen und ihre Präzisierung	16
1.3.2	Design	17
1.3.3	Durchführung von Studien	17
1.3.4	Datenanalyse und -interpretation	17
1.4	Was gewinnen wir durch die wissenschaftliche Vorgehensweise?	18

### Kapitel 2 Wissenschaftstheorie, Theorien und Hypothesen 21

2.1	Was ist die Wirklichkeit und wie können wir sie erkennen?	23
2.1.1	Das Leib-Seele-Problem	24
2.1.2	Induktion vs. Deduktion	25
2.2	Wissenschaftstheoretische Ansätze im Überblick	26
2.2.1	Logischer Empirismus	27
2.2.2	Kritischer Rationalismus	29
2.2.3	Historisch-soziologische Analyse (Kuhn)	41
2.2.4	Methodologie wissenschaftlicher Forschungsprogramme (Lakatos)	43
2.2.5	Wirklichkeit als Konstruktion	43
2.3	Spezialprobleme der Psychologie	46
2.3.1	Latente Variablen	47
2.3.2	Verhältnis zwischen Forscher und „Erforschten“	47

2.4	Woher kommen Theorien? .....
2.4.1	Bed, Bathroom and Bicycle .....
2.4.2	Die systematische Suche nach Theorien .....
2.5	Von Theorien zu Hypothesen .....
2.5.1	Wie sehen Theorien in der Psychologie aus? .....
2.5.2	Von der Theorie zur Hypothesenprüfung: Grundlegende Vorgehensweise .....
2.5.3	Von der Theorie zur Hypothesenprüfung: Beispiele .....
2.5.4	Hypothesenprüfung und Wissenschaftstheorie .....

### Kapitel 3 Messen und Testen

3.1	Was ist Messen? .....
3.2	Messtheorie .....
3.2.1	Messtheoretische Probleme .....
3.3	Skalenniveaus .....
3.3.1	Nominalskala .....
3.3.2	Ordinalskala .....
3.3.3	Intervallskala .....
3.3.4	Verhältnisskala .....
3.3.5	Absolutskala .....
3.4	Tests .....
3.5	Gütekriterien beim Testen und Messen .....
3.5.1	Objektivität .....
3.5.2	Reliabilität .....
3.5.3	Validität .....

### Kapitel 4 Datenerhebung: Befragung und Beobachtung

4.1	Befragung: Unterschiedliche Perspektiven .....
4.1.1	Mündlich oder schriftlich? .....
4.1.2	Freie oder festgelegte Antwortmöglichkeiten? .....
4.1.3	Einzel- oder Gruppenbefragung? .....
4.1.4	Wie sehr standardisieren? .....
4.2	Befragung: Fehlermöglichkeiten und Gegenmaßnahmen .....
4.2.1	Potenzielle Probleme bei der Gestaltung und Anordnung von Items .....
4.2.2	Potenzielle Probleme bei der Durchführung der Befragung .....
4.3	Befragung: Ein kurzes Resümee .....
4.3.1	Wann welche Art von Befragung? .....
4.3.2	Einige abschließende Hinweise .....
4.4	Beobachtung: Unterschiedliche Perspektiven .....
4.5	Beobachtung: Fehlermöglichkeiten und Gegenmaßnahmen .....
4.6	Beobachtung: Ein kurzes Resümee .....
4.6.1	Wann welche Form von Beobachtung? .....
4.6.2	Einige abschließende Hinweise .....

2.4	Woher kommen Theorien? .....
2.4.1	Bed, Bathroom and Bicycle .....
2.4.2	Die systematische Suche nach Theorien .....
2.5	Von Theorien zu Hypothesen .....
2.5.1	Wie sehen Theorien in der Psychologie aus? .....
2.5.2	Von der Theorie zur Hypothesenprüfung: Grundlegende Vorgehensweise .....
2.5.3	Von der Theorie zur Hypothesenprüfung: Beispiele .....
2.5.4	Hypothesenprüfung und Wissenschaftstheorie .....

### Kapitel 3 Messen und Testen

3.1	Was ist Messen? .....
3.2	Messtheorie .....
3.2.1	Messtheoretische Probleme .....
3.3	Skalenniveaus .....
3.3.1	Nominalskala .....
3.3.2	Ordinalskala .....
3.3.3	Intervallskala .....
3.3.4	Verhältnisskala .....
3.3.5	Absolutskala .....
3.4	Tests .....
3.5	Gütekriterien beim Testen und Messen .....
3.5.1	Objektivität .....
3.5.2	Reliabilität .....
3.5.3	Validität .....

### Kapitel 4 Datenerhebung: Befragung und Beobachtung

4.1	Befragung: Unterschiedliche Perspektiven .....
4.1.1	Mündlich oder schriftlich? .....
4.1.2	Freie oder festgelegte Antwortmöglichkeiten? .....
4.1.3	Einzel- oder Gruppenbefragung? .....
4.1.4	Wie sehr standardisieren? .....
4.2	Befragung: Fehlermöglichkeiten und Gegenmaßnahmen .....
4.2.1	Potenzielle Probleme bei der Gestaltung und Anordnung von Items .....
4.2.2	Potenzielle Probleme bei der Durchführung der Befragung. ....
4.3	Befragung: Ein kurzes Resümee .....
4.3.1	Wann welche Art von Befragung? .....
4.3.2	Einige abschließende Hinweise .....
4.4	Beobachtung: Unterschiedliche Perspektiven .....
4.5	Beobachtung: Fehlermöglichkeiten und Gegenmaßnahmen .....
4.6	Beobachtung: Ein kurzes Resümee .....
4.6.1	Wann welche Form von Beobachtung? .....
4.6.2	Einige abschließende Hinweise .....

2.4	Woher kommen Theorien? . . . . .	49
2.4.1	Bed, Bathroom and Bicycle . . . . .	49
2.4.2	Die systematische Suche nach Theorien. . . . .	51
2.5	Von Theorien zu Hypothesen . . . . .	52
2.5.1	Wie sehen Theorien in der Psychologie aus? . . . . .	52
2.5.2	Von der Theorie zur Hypothesenprüfung: Grundlegende Vorgehensweise . . . . .	53
2.5.3	Von der Theorie zur Hypothesenprüfung: Beispiele . . . . .	55
2.5.4	Hypothesenprüfung und Wissenschaftstheorie . . . . .	58
Kapitel 7 Messen und Testen . . . . .		61
3.1	Was ist Messen? . . . . .	63
3.2	Messtheorie . . . . .	66
3.2.1	Messtheoretische Probleme . . . . .	68
3.3	Skalenniveaus . . . . .	71
3.3.1	Nominalskala . . . . .	71
3.3.2	Ordinalskala . . . . .	72
3.3.3	Intervallskala . . . . .	74
3.3.4	Verhältnisskala . . . . .	75
3.3.5	Absolutskala . . . . .	77
3.4	Tests . . . . .	77
3.5	Gütekriterien beim Testen und Messen . . . . .	79
3.5.1	Objektivität . . . . .	80
3.5.2	Reliabilität . . . . .	81
3.5.3	Validität . . . . .	85
Kapitel 4 Datenerhebung: Befragung und Beobachtung . . . . .		91
4.1	Befragung: Unterschiedliche Perspektiven . . . . .	93
4.1.1	Mündlich oder schriftlich? . . . . .	93
4.1.2	Freie oder festgelegte Antwortmöglichkeiten? . . . . .	96
4.1.3	Einzel- oder Gruppenbefragung? . . . . .	97
4.1.4	Wie sehr standardisieren? . . . . .	98
4.2	Befragung: Fehlermöglichkeiten und Gegenmaßnahmen . . . . .	102
4.2.1	Potenzielle Probleme bei der Gestaltung und Anordnung von Items . . . . .	103
4.2.2	Potenzielle Probleme bei der Durchführung der Befragung. . . . .	109
4.3	Befragung: Ein kurzes Resümee . . . . .	110
4.3.1	Wann welche Art von Befragung? . . . . .	111
4.3.2	Einige abschließende Hinweise . . . . .	112
4.4	Beobachtung: Unterschiedliche Perspektiven . . . . .	113
4.5	Beobachtung: Fehlermöglichkeiten und Gegenmaßnahmen . . . . .	120
4.6	Beobachtung: Ein kurzes Resümee . . . . .	125
4.6.1	Wann welche Form von Beobachtung? . . . . .	125
4.6.2	Einige abschließende Hinweise . . . . .	126

4.7	Generalisierbarkeit von Befragungs- und Beobachtungsergebnissen . . . .	126
4.7.1	Auswahl der Situation . . . . .	126
4.7.2	Auswahl der Studienteilnehmer . . . . .	127
<b>Kapitel 5 Experimentelle Design</b> . . . . . 131		
5.1	Warum werden Experimente durchgeführt? . . . . .	131
5.2	Die Logik des Experiments . . . . .	132
5.2.1	Grundlage für Kausalschlüsse . . . . .	133
5.2.2	Interne Validität . . . . .	138
5.3	Kontrolltechniken . . . . .	139
5.3.1	Kontrolle personengebundener Störvariablen . . . . .	140
5.3.2	Kontrolle von Störvariablen in der Versuchssituation . . . . .	144
5.4	Externe Validität . . . . .	150
5.4.1	Wie wichtig ist die externe Validität? . . . . .	151
5.4.2	Wie kann die externe Validität erhöht werden? . . . . .	153
5.5	Within-Subjects-Designs . . . . .	154
5.5.1	Warum werden Within-Subjects-Designs eingesetzt? . . . . .	157
5.5.2	Positionseffekte und ihre Kontrolle . . . . .	162
5.5.3	Carry-Over-Effekte . . . . .	168
5.6	Mehrfaktorielle Designs . . . . .	169
5.6.1	Haupteffekte und Interaktionen in $2 \times 2$ -Designs . . . . .	171
5.6.2	Komplexere Designs . . . . .	176
5.6.3	Interaktionen und externe Validität . . . . .	178
5.7	Quasi-Experimente . . . . .	179
<b>Teil II Deskriptive und explorative Datenanalyse</b>		<b>185</b>
<b>Kapitel 6 Lage- und Streuungsmaße</b> . . . . . 187		
6.1	Warum brauchen wir Streuungsmaße? . . . . .	189
6.2	Lage und Streuung auf einen Blick . . . . .	190
6.2.1	Stamm-Blatt-Diagramme . . . . .	190
6.2.2	Box-Plots . . . . .	194
6.3	Lagemaße im Detail . . . . .	197
6.3.1	Arithmetisches Mittel . . . . .	197
6.3.2	Median und Quantile . . . . .	198
6.3.3	Modalwert . . . . .	199
6.3.4	Weitere Lagemaße . . . . .	199
6.4	Streuungsmaße im Detail . . . . .	200
6.4.1	Standardabweichung und Varianz . . . . .	200
6.4.2	Interquartilsabstand und andere Quantilsabstände . . . . .	201
6.4.3	Weitere Streuungsmaße . . . . .	201

6.5	Wann welches Maß? .....	202
6.5.1	Skalenniveau .....	202
6.5.2	Form der Verteilung.....	202
6.6	Standardisierung: z-Werte .....	204
6.7	Population vs. Stichprobe .....	205
Kapitel 7 Korrelation .....		207
7.1	Die grafische Darstellung von Korrelationen: Streudiagramme .....	209
7.2	Korrelationsmuster .....	212
7.2.1	Lineare und kurvilineare Zusammenhänge .....	212
7.2.2	Richtung und Stärke von Zusammenhängen .....	213
7.2.3	Die Bedeutung des Korrelationsmusters für die weitere Analyse .....	216
7.3	Der Produkt-Moment-Korrelationskoeffizient .....	217
7.3.1	z-Werte und der Produkt-Moment-Korrelationskoeffizient .....	224
7.4	Verzerrungen des Produkt-Moment-Korrelationskoeffizienten .....	226
7.4.1	Ausreißerwerte.....	227
7.4.2	Einschränkungen der Variabilität .....	228
7.4.3	Zusammenfassung von heterogenen Untergruppen.....	230
7.5	Korrelation und Kausalität .....	231
7.6	Partialkorrelation.....	234
7.7	Andere Zusammenhangsmaße .....	235
7.7.1	Korrelation zweier dichotomer Merkmale – der Phi-Koeffizient.....	236
7.7.2	Korrelation zweier ordinalskalierten Merkmale – Kendalls Tau .....	239
Kapitel 8 Lineare Regression .....		245
8.1	Grundbegriffe der Regressionsrechnung.....	247
8.1.1	Prädiktor und Kriterium .....	247
8.1.2	Deterministische Zusammenhänge und die Geradengleichung .....	248
8.1.3	Stochastische Zusammenhänge und die Regressionsgerade .....	250
8.1.4	Das Kriterium der kleinsten Quadrate .....	253
8.1.5	Bestimmung der Regressionsgeraden .....	254
8.1.6	Die Beziehung zwischen der Korrelation und dem Regressionsgewicht $b$ .....	256
8.1.7	Regression mit z-standardisierten Variablen .....	259
8.1.8	Der Regressionseffekt.....	261
8.1.9	Die Vorhersage von $X$ aus $Y$ .....	263
8.2	Die Güte der Vorhersage .....	265
8.2.1	Varianzzerlegung .....	267
8.2.2	Der Determinationskoeffizient $r^2$ .....	271
8.2.3	Der Standardschätzfehler.....	273
8.3	Probleme und Verzerrungen in der Regressionsrechnung.....	275
8.4	Ein Ausblick auf die multiple Regression .....	276
8.4.1	Multiple Regression mit z-standardisierten Variablen.....	277
8.4.2	Eine Illustration mit zwei Prädiktoren .....	278
8.4.3	Gütemaße in der multiplen Regression.....	282

<b>Kapitel 9</b>	<b>Effektgrößen</b>	<b>309</b>
9.1	Was sind Effektgrößen? . . . . .	291
9.2	Abstandsmaße . . . . .	291
9.3	Zusammenhangsmaße. . . . .	296
9.4	Effektgrößen aus Effektgrößen . . . . .	298
9.4.1	Abstandsmaße aus Abstandsmaßen . . . . .	299
9.4.2	Korrelationen aus Abstandsmaßen. . . . .	300
9.4.3	Abstandsmaße aus Korrelationen . . . . .	301
9.5	Wie bedeutsam ist eine Effektgröße? . . . . .	301
9.6	Weitere Effektgrößen-Maße. . . . .	304
9.6.1	Relatives Risiko . . . . .	304
9.6.2	Odds Ratio . . . . .	305
9.6.3	Mehr zu Effektgrößen in diesem Buch . . . . .	306

**Teil III Inferenzstatistik 309**

<b>Kapitel 10</b>	<b>Wahrscheinlichkeiten, Simulationen, Stichprobenverteilung</b>	<b>313</b>
10.1	Wahrscheinlichkeiten, kurz gefasst . . . . .	313
10.1.1	Was ist Wahrscheinlichkeit? . . . . .	313
10.1.2	Wahrscheinlichkeit von Konjunktionen und bedingte Wahrscheinlichkeiten . . . . .	315
10.2	Von der Population über Stichproben zur Stichprobenverteilung . . . . .	318
10.2.1	Simulationsbeispiel für Anteile . . . . .	318
10.2.2	Simulationsbeispiel für Mittelwerte. . . . .	320
10.2.3	Die tatsächliche Vorgehensweise: Von der Stichprobe zur Population . . . . .	322
10.3	Stichprobenverteilung für Anteile . . . . .	323
10.3.1	Binomialverteilung „per Hand“ . . . . .	324
10.3.2	Binomialverteilung mit Binomialformel . . . . .	325
10.4	Lage- und Streuungsmaße von Stichprobenverteilungen . . . . .	326
10.4.1	Binomialverteilung . . . . .	327
10.4.2	Stichprobenverteilungen für Mittelwerte. . . . .	330
10.5	Der Einfluss der Stichprobengröße auf die Stichprobenverteilung . . . . .	335
10.5.1	Empirisches Gesetz der großen Zahlen . . . . .	335
10.5.2	Zentraler Grenzwertsatz . . . . .	337
10.6	Rekapitulation und Ausblick . . . . .	340
<b>Kapitel 11</b>	<b>Konfidenzintervalle</b>	<b>343</b>
11.1	Was ist ein Konfidenzintervall? . . . . .	345
11.1.1	Wahrscheinlichkeitsintervalle: Ein Gedankenexperiment . . . . .	345
11.1.2	Konfidenzintervalle für Anteile . . . . .	346
11.1.3	Auswirkungen der Höhe der Konfidenz und der Stichprobengröße. . . . .	348
11.1.4	Die Berechnung von Konfidenzintervallen . . . . .	350



11.2	Konfidenzintervalle für Mittelwerte . . . . .	353
11.3	Konfidenzintervalle für Mittelwertsunterschiede . . . . .	356
11.3.1	Unabhängige Messungen . . . . .	356
11.3.2	Abhängige (gepaarte) Messungen . . . . .	359
11.4	Die Interpretation von Konfidenzintervallen . . . . .	365
Kapitel 12 Signifikanztests . . . . .		
12.1	Wie funktioniert ein Signifikanztest? . . . . .	371
12.2	Vorgehensweise nach R. A. Fisher . . . . .	373
12.2.1	Beispiel 1: Vorzeichenstest . . . . .	374
12.2.2	Beispiel 2: <i>t</i> -Test für Mittelwert . . . . .	376
12.2.3	Probleme mit der Vorgehensweise nach Fisher . . . . .	377
12.3	Neymans & Pearsons Verbesserungsvorschläge . . . . .	378
12.3.1	Warum braucht man die Alternativhypothese und wie wird sie bestimmt? . . . . .	378
12.3.2	Fehler erster und zweiter Art ( $\alpha$ und $\beta$ ) . . . . .	380
12.3.3	Die „Verhaltensinterpretation“ des Signifikanztestergebnisses. . . . .	380
12.4	Welche Faktoren beeinflussen das Ergebnis eines Signifikanztests? . . . . .	381
12.4.1	Populations-Effektgröße. . . . .	381
12.4.2	Stichprobengröße. . . . .	382
12.4.3	Abwägung der Fehler erster und zweiter Art . . . . .	384
12.4.4	Minimierung des „experimentellen Fehlers“ . . . . .	385
12.4.5	Homogenität der Population(en) . . . . .	386
12.5	Poweranalyse . . . . .	386
12.5.1	Die Suche nach der Stichprobengröße: „A priori-Analyse“ . . . . .	387
12.5.2	Die Suche nach einem Kompromiss zwischen $\alpha$ und $\beta$ . . . . .	387
12.5.3	Die Suche nach weiteren Interpretationsmöglichkeiten: „post hoc-Analyse“ . . . . .	388
12.6	Vorgehensweise nach Neyman und Pearson . . . . .	388
12.6.1	Beispiel 1: Vorzeichenstest nach Neyman und Pearson . . . . .	389
12.6.2	Beispiel 2: <i>t</i> -Test nach Neyman und Pearson . . . . .	393
12.6.3	Akzeptanz des Ansatzes in Psychologie und Sozialwissenschaften . . . . .	395
12.7	Das konventionelle Verfahren: Der „Hybrid“ . . . . .	395
12.7.1	Bestandteile . . . . .	396
12.7.2	Vorgehensweise und Ergebnisinterpretation . . . . .	397
12.8	Signifikanztests: Was man noch wissen sollte . . . . .	398
12.8.1	Spezifikation von Null- und Alternativhypothese . . . . .	398
12.8.2	Wie man <i>p</i> -Werte <i>nicht</i> interpretieren sollte . . . . .	400
12.8.3	Signifikanztest und Konfidenzintervall . . . . .	401
12.8.4	Allgemeine Hinweise und Empfehlungen . . . . .	403
Kapitel 13 <i>t</i> -Tests . . . . .		
13.1	Unterschied zwischen zwei Mittelwerten . . . . .	409
13.1.1	Unabhängige Stichproben . . . . .	409
13.1.2	Abhängige Stichproben . . . . .	414

13.2	Weitere $t$ -Tests . . . . .	418
13.2.1	Korrelation . . . . .	418
13.2.2	Regression . . . . .	421
13.3	Effektgrößenberechnung aus Testergebnissen von $t$ -Tests . . . . .	421
13.3.1	Generelle Idee . . . . .	422
13.3.2	Eine Stichprobe (Mittelwert vs. vorgegebener Wert) . . . . .	422
13.3.3	Zwei unabhängige Stichproben . . . . .	423
13.3.4	Zwei abhängige Stichproben . . . . .	424
13.3.5	Korrelation und Regression . . . . .	426
14.1	Warum nicht mehrere $t$ -Tests? . . . . .	431
14.2	Die Logik der Varianzanalyse . . . . .	434
14.2.1	Zwei Wege zu einer Schätzung der Populationsvarianz . . . . .	435
14.2.2	Varianzzerlegung . . . . .	444
14.3	Voraussetzungen der einfaktoriellen Varianzanalyse . . . . .	452
14.4	Post-hoc-Tests . . . . .	453
14.5	Effektgrößen in der einfaktoriellen Varianzanalyse . . . . .	456
14.6	Power in der einfaktoriellen Varianzanalyse . . . . .	458
15.1	Mehrfaktorielle Varianzanalyse . . . . .	465
15.1.1	Varianzzerlegung in der zweifaktoriellen Varianzanalyse . . . . .	467
15.1.2	ANOVA-Tabelle . . . . .	476
15.1.3	Varianzanalysen mit mehr als zwei Faktoren . . . . .	477
15.1.4	Voraussetzungen der mehrfaktoriellen Varianzanalyse . . . . .	478
15.1.5	Mehrfaktorielle Varianzanalysen mit ungleichen Stichprobengrößen . . . . .	478
15.1.6	Effektgrößen in der mehrfaktoriellen Varianzanalyse . . . . .	479
15.1.7	Power in der mehrfaktoriellen Varianzanalyse . . . . .	482
15.2	Varianzanalyse mit abhängigen Stichproben . . . . .	485
15.2.1	Varianzzerlegung in der einfaktoriellen Varianzanalyse mit abhängigen Stichproben . . . . .	487
15.2.2	ANOVA-Tabelle . . . . .	495
15.2.3	Voraussetzungen der Varianzanalyse mit abhängigen Stichproben . . . . .	496
15.2.4	Effektgrößen in der Varianzanalyse mit abhängigen Stichproben . . . . .	498
15.2.5	Power in der Varianzanalyse mit abhängigen Stichproben . . . . .	499
15.2.6	Erweiterungen zur Varianzanalyse mit abhängigen Stichproben . . . . .	500
15.3	Der $F$ -Test in der Regressionsrechnung . . . . .	500
15.4	Weitere Varianten der Varianzanalyse . . . . .	504

Kapitel 16 Kontrastanalyse

16.1 Kontraste vs. „Omnibus-Hypothesen“ . . . . .

    16.1.1 Die Problematik von Omnibus-Hypothesen . . . . .

    16.1.2 Kontraste als präzise Hypothesen . . . . .

16.2 Kontrastanalyse für unabhängige Stichproben. . . . .

    16.2.1  $F_{Kontrast}$  und  $t_{Kontrast}$  . . . . .

    16.2.2 Orthogonale Kontraste . . . . .

    16.2.3 Effektgrößen bei der Kontrastanalyse für unabhängige  
Stichproben . . . . .

    16.2.4 Poweranalyse bei der Kontrastanalyse für unabhängige  
Stichproben . . . . .

    16.2.5 Kontrastanalyse für unabhängige Stichproben bei  
komplexen Fragestellungen. . . . .

16.3 Kontrastanalyse für abhängige Stichproben. . . . .

    16.3.1 Bestimmen der zusammengefassten Werte. . . . .

    16.3.2  $t$ -Test für die Kontrastanalyse bei abhängigen Stichproben. . . . .

    16.3.3 Effektgrößen bei der Kontrastanalyse für abhängige Stichproben

    16.3.4 Poweranalyse bei der Kontrastanalyse für abhängige Stichproben

Kapitel 17 Verfahren zur Analyse nominalskaliertter Daten:  
Chi-Quadrat ( $\chi^2$ -)Tests

17.1 Der  $\chi^2$ -Test für eine Variable . . . . .

    17.1.1 Die Gleichverteilungsannahme als Nullhypothese . . . . .

    17.1.2 Der  $\chi^2$ -Wert . . . . .

    17.1.3  $\chi^2$ -Verteilung und Freiheitsgrade . . . . .

    17.1.4 Andere Verteilungsannahmen als Nullhypothese . . . . .

    17.1.5 Effektgrößen . . . . .

    17.1.6 Power . . . . .

17.2 Der  $\chi^2$ -Test für zwei Variablen . . . . .

    17.2.1 Die Unabhängigkeitsannahme als Nullhypothese . . . . .

    17.2.2 Berechnung des  $\chi^2$ -Werts . . . . .

    17.2.3 Freiheitsgrade und Signifikanzprüfung . . . . .

    17.2.4 Effektgrößen . . . . .

    17.2.5 Power . . . . .

17.3 Voraussetzungen der  $\chi^2$ -Tests . . . . .

Kapitel 18 Verfahren zur Analyse ordinalskaliertter Daten

18.1 Voraussetzungsverletzungen in parametrischen Tests. . . . .

18.2 Der  $U$ -Test . . . . .

    18.2.1 Zuordnung der Rangplätze . . . . .

    18.2.2 Null- und Alternativhypothese . . . . .

    18.2.3 Der  $U$ -Wert . . . . .

    18.2.4 Signifikanzprüfung in kleinen Stichproben . . . . .

    18.2.5 Signifikanzprüfung in großen Stichproben. . . . .

Kapitel 16	Kontrastanalyse	509
16.1	Kontraste vs. „Omnibus-Hypothesen“	511
16.1.1	Die Problematik von Omnibus-Hypothesen	511
16.1.2	Kontraste als präzise Hypothesen	512
16.2	Kontrastanalyse für unabhängige Stichproben	516
16.2.1	$F_{Kontrast}$ und $t_{Kontrast}$	516
16.2.2	Orthogonale Kontraste	522
16.2.3	Effektgrößen bei der Kontrastanalyse für unabhängige Stichproben	524
16.2.4	Poweranalyse bei der Kontrastanalyse für unabhängige Stichproben	530
16.2.5	Kontrastanalyse für unabhängige Stichproben bei komplexen Fragestellungen	533
16.3	Kontrastanalyse für abhängige Stichproben	533
16.3.1	Bestimmen der zusammengefassten Werte	534
16.3.2	$t$ -Test für die Kontrastanalyse bei abhängigen Stichproben	535
16.3.3	Effektgrößen bei der Kontrastanalyse für abhängige Stichproben	539
16.3.4	Poweranalyse bei der Kontrastanalyse für abhängige Stichproben	540
Kapitel 17	Verfahren zur Analyse nominalskaliert Chi-Quadrat ( $\chi^2$ )-Tests	543
17.1	Der $\chi^2$ -Test für eine Variable	546
17.1.1	Die Gleichverteilungsannahme als Nullhypothese	546
17.1.2	Der $\chi^2$ -Wert	548
17.1.3	$\chi^2$ -Verteilung und Freiheitsgrade	549
17.1.4	Andere Verteilungsannahmen als Nullhypothese	551
17.1.5	Effektgrößen	553
17.1.6	Power	555
17.2	Der $\chi^2$ -Test für zwei Variablen	556
17.2.1	Die Unabhängigkeitsannahme als Nullhypothese	558
17.2.2	Berechnung des $\chi^2$ -Werts	561
17.2.3	Freiheitsgrade und Signifikanzprüfung	561
17.2.4	Effektgrößen	563
17.2.5	Power	567
17.3	Voraussetzungen der $\chi^2$ -Tests	567
Kapitel 18	Verfahren zur Analyse ordinalskaliert Daten	571
18.1	Voraussetzungsverletzungen in parametrischen Tests	573
18.2	Der $U$ -Test	574
18.2.1	Zuordnung der Rangplätze	575
18.2.2	Null- und Alternativhypothese	576
18.2.3	Der $U$ -Wert	577
18.2.4	Signifikanzprüfung in kleinen Stichproben	579
18.2.5	Signifikanzprüfung in großen Stichproben	579
18.2.6	Rangbindungen	581

18.3	Der Wilcoxon-Test . . . . .	581
18.3.1	Durchführung des Wilcoxon-Tests . . . . .	582
18.3.2	Eine Voraussetzung des Wilcoxon-Tests . . . . .	584
18.4	Powerbestimmung im <i>U</i> -Test und Wilcoxon-Test . . . . .	584

## Kapitel 19: Resampling-Verfahren

19.1	Konventionelle Inferenzstatistik versus Resampling-Verfahren . . . . .	589
19.2	Resampling-Verfahren: Warum und wie? . . . . .	589
19.2.1	Zwei wesentliche Vorteile . . . . .	590
19.2.2	Die Stichprobe als repräsentatives Abbild der Population . . . . .	591
19.2.3	Resampling-Stichprobenverteilungen . . . . .	591
19.3	Bootstrap: Konfidenz nach Münchhausen-Art . . . . .	594
19.3.1	Wie funktioniert der Bootstrap? . . . . .	594
19.3.2	Bootstrap: Anwendungsbeispiele . . . . .	595
19.4	Randomisierungstests . . . . .	599
19.4.1	Wie funktionieren Randomisierungstests? . . . . .	600
19.4.2	Randomisierungstests: Anwendungsbeispiele . . . . .	601
19.4.3	Besonderheiten bei Randomisierungstests . . . . .	605
19.5	Resampling-Verfahren im Kontext . . . . .	607
19.5.1	Bootstrappen oder Randomisieren? . . . . .	607
19.5.2	Weitere Resampling-Verfahren . . . . .	607
19.5.3	Resampling-Verfahren versus traditionelle Inferenzstatistik . . . . .	608
19.5.4	Praktische Vorgehensweise . . . . .	608

## Teil IV Inferenzstatistik: Praktische Probleme und alternative Sichtweisen 611

### Kapitel 20: Replizierbarkeit, Reproduzierbarkeit und die Replizationskrise 611

20.1	Replizierbarkeit in der Psychologie . . . . .	615
20.1.1	Das Reproducibility Project: Psychology . . . . .	615
20.1.2	Andere Befunde zur Replizierbarkeit in der Psychologie . . . . .	622
20.2	Ursachen der Replikationskrise . . . . .	624
20.2.1	Probleme bei der Interpretation des Signifikanztests . . . . .	624
20.2.2	Probleme in der Praxis: Publikationsbias, HARKing und <i>p</i> -Hacking . . . . .	632
20.3	Problemlösungen . . . . .	644

### Kapitel 21: Replikation, Preregistrierung, Open Science 649

21.1	Replikation . . . . .	651
21.1.1	Typen von Replikationsstudien . . . . .	653
21.1.2	Wann ist eine Replikation erfolgreich? . . . . .	658
21.1.3	Was ist eine gute Replikationsstudie? . . . . .	664

18.3	Der Wilcoxon-Test . . . . .	581
18.3.1	Durchführung des Wilcoxon-Tests . . . . .	582
18.3.2	Eine Voraussetzung des Wilcoxon-Tests . . . . .	584
18.4	Powerbestimmung im <i>U</i> -Test und Wilcoxon-Test . . . . .	584

**Kapitel 19 Resampling-Verfahren** 587

19.1	Konventionelle Inferenzstatistik versus Resampling-Verfahren . . . . .	589
19.2	Resampling-Verfahren: Warum und wie? . . . . .	589
19.2.1	Zwei wesentliche Vorteile . . . . .	590
19.2.2	Die Stichprobe als repräsentatives Abbild der Population . . . . .	591
19.2.3	Resampling-Stichprobenverteilungen . . . . .	591
19.3	Bootstrap: Konfidenz nach Münchhausen-Art . . . . .	594
19.3.1	Wie funktioniert der Bootstrap? . . . . .	594
19.3.2	Bootstrap: Anwendungsbeispiele . . . . .	595
19.4	Randomisierungstests . . . . .	599
19.4.1	Wie funktionieren Randomisierungstests? . . . . .	600
19.4.2	Randomisierungstests: Anwendungsbeispiele . . . . .	601
19.4.3	Besonderheiten bei Randomisierungstests . . . . .	605
19.5	Resampling-Verfahren im Kontext . . . . .	607
19.5.1	Bootstrappen oder Randomisieren? . . . . .	607
19.5.2	Weitere Resampling-Verfahren . . . . .	607
19.5.3	Resampling-Verfahren versus traditionelle Inferenzstatistik . . . . .	608
19.5.4	Praktische Vorgehensweise . . . . .	608

**Teil IV Inferenzstatistik: Praktische Probleme und alternative Sichtweisen** 611

**Kapitel 20 Probleme der Reproduzierbarkeit in der Forschungspraxis** 615

20.1	Replizierbarkeit in der Psychologie . . . . .	615
20.1.1	Das Reproducibility Project: Psychology . . . . .	615
20.1.2	Andere Befunde zur Replizierbarkeit in der Psychologie . . . . .	622
20.2	Ursachen der Replikationskrise . . . . .	624
20.2.1	Probleme bei der Interpretation des Signifikanztests . . . . .	624
20.2.2	Probleme in der Praxis: Publikationsbias, HARKing und <i>p</i> -Hacking . . . . .	632
20.3	Problemlösungen . . . . .	644

**Kapitel 21 Replikation, Präregistrierung, Open Science** 649

21.1	Replikation . . . . .	651
21.1.1	Typen von Replikationsstudien . . . . .	653
21.1.2	Wann ist eine Replikation erfolgreich? . . . . .	658
21.1.3	Was ist eine gute Replikationsstudie? . . . . .	664

21.2	Präregistrierung .....	668
21.3	Open Science.....	674

**Kapitel 22 Bayesianische Statistik** 679

22.1	Die Revision von Wahrscheinlichkeiten.....	681
22.1.1	Das Bayes-Theorem .....	682
22.2	Bayesianische Wahrscheinlichkeiten .....	685
22.3	Priors, Likelihoods und Posteriors .....	689
22.3.1	Priorverteilung .....	689
22.3.2	Likelihoods .....	690
22.3.3	Posteriorverteilung .....	692
22.4	Stetige Priorverteilungen und konjugierte Priors.....	694
22.5	Einflussgrößen auf die Posteriorverteilung .....	699
22.5.1	Auswirkungen der Priorverteilung .....	699
22.5.2	Auswirkungen der Stichprobengröße .....	701
22.6	Bayes-Faktor.....	703
22.7	Klassisch vs. Bayesianisch .....	707

**Teil V Das Allgemeine Lineare Modell** 711

**Kapitel 23 Das Allgemeine Lineare Modell** 713

23.1	Was ist das Allgemeine Lineare Modell? .....	715
23.2	Der <i>t</i> -Test als Spezialfall der einfachen Regression .....	717
23.3	Varianzanalyse mit zwei Gruppen als Spezialfall der einfachen Regression .....	724
23.4	Varianzanalyse mit mehr als zwei Gruppen als Spezialfall der multiplen Regression.....	728

**Kapitel 24 Regressionsrechnung, Ergänzungen und Erweiterungen** 733

24.1	Multiple Regression: Ergänzungen .....	735
24.1.1	Schrittweise Regression.....	735
24.1.2	Effektgrößen bei der multiplen Regression .....	739
24.1.3	Inferenzstatistik bei der multiplen Regression.....	742
24.1.4	Analyse nichtlinearer Beziehungen .....	746
24.2	Kontrastanalyse mittels Regressionsrechnung.....	747
24.3	Kovarianzanalyse mittels Regressionsrechnung .....	748
24.4	Moderatoranalyse: Die generelle Behandlung von Interaktionen .....	752
24.4.1	Interaktion als multiplikative Komponente .....	752
24.4.2	Zentrieren der Prädiktorvariablen.....	755
24.4.3	Interaktion zwischen zwei nominalskalierten Variablen.....	756
24.4.4	Interaktion zwischen einer nominal- und einer intervallskalierten Variable .....	758
24.4.5	Interaktion zwischen zwei intervallskalierten Variablen.....	759
24.4.6	Interaktion in komplexeren Fällen .....	762

<b>Kapitel 25</b>	<b>Exploratorische Datenanalyse und die Pfadanalyse</b>	<b>767</b>
25.1	Pfadanalyse	767
25.1.1	Zusammenhang zwischen Regressionsrechnung und Pfadanalyse	767
25.1.2	Pfadanalyse mit Mediatorvariable	771
25.2	Strukturgleichungsmodelle	773
25.2.1	Identifizierbarkeit	776
25.2.2	Mess- und Strukturmodelle	778
25.2.3	Schätzen der freien Parameter	780
25.2.4	Die Überprüfung des Modells: Gütemaße	783
25.2.5	Anwendungsvoraussetzungen	785
25.3	Exploratorische Faktorenanalyse	785
25.3.1	Datenbeispiel	787
25.3.2	Fundamentaltheorem der Faktorenanalyse	788
25.3.3	Extraktionsverfahren	789
25.3.4	Ladungen, Kommunalitäten, Eigenwerte	792
25.3.5	Faktorauswahl	794
25.3.6	Rotation und Interpretation	795
25.4	Mehrebenenanalyse	797
25.4.1	Warum Mehrebenenanalyse?	798
25.4.2	Regressionsgleichung für ein einfaches Mehrebenenmodell	800
25.4.3	Feste versus zufällige Effekte	802
25.4.4	Theoriegeleitete Analyse	803
25.4.5	Explorative Vorgehensweise	804
25.4.6	Maße zur Beurteilung der Ergebnisse	806
25.4.7	Mehrebenenanalyse als Metaanalyseprozedur	811
25.4.8	Möglichkeiten und Grenzen der Mehrebenenanalyse	812

**Teil VI Weitere Verfahren in der Datenerhebung und Datenanalyse** **815**

<b>Kapitel 26</b>	<b>Explorative Datenanalyse und die Quantil-Methode</b>	<b>817</b>
26.1	Robustheit von EDA-Verfahren: Box-Plots	819
26.2	Varianten von Streuungsdiagrammen	820
26.2.1	Streuungsdiagramme mit Box-Plots	820
26.2.2	Influence-Plot	821
26.2.3	Bubble-Plot	822
26.3	„Aufspüren“ und „Geradebiegen“ nichtlinearer Zusammenhänge	823
26.3.1	Lowess	823
26.3.2	Potenzleiter	826
26.4	Multivariate Zusammenhänge auf einen Blick: Die Streuungsdiagramm-Matrix	830



26.5	Mehrdimensionale grafische Klassifikation von Personen oder Objekten. . . . .	832
26.5.1	Rechteck-Icons. . . . .	832
26.5.2	Histogramm- und Profilplots. . . . .	833
26.5.3	Star-Plots . . . . .	833
26.5.4	Chernoff-Gesichter. . . . .	834
26.6	EDA im Kontext. . . . .	835
Kapitel 27 Effektgrößen Erweberungen und Ergänzungen		839
27.1	Populations- versus Stichprobeneffektgrößen . . . . .	841
27.2	Effektgrößenschätzung bei unvollständigen Angaben. . . . .	843
27.2.1	Nur $p$ -Werte und Stichprobengröße(n) angegeben. . . . .	843
27.2.2	Nur „globale“ Angaben . . . . .	845
27.3	Die Vergleichbarkeit von Effektgrößen . . . . .	845
27.3.1	Effektgrößen aus Rohdaten vs. Signifikanztestergebnissen . . . . .	846
27.3.2	Die Vergleichbarkeit von unterschiedlichen korrelativen Maßen	846
27.3.3	Abstandsmaße vs. korrelative Maße . . . . .	847
27.3.4	Unabhängige vs. abhängige Stichproben . . . . .	847
27.3.5	Signifikanztest auf Unterschied zweier Effektgrößen. . . . .	848
27.4	Konfidenzintervalle für $r$ und $g$ . . . . .	849
27.4.1	Approximative Konfidenzintervalle für $r$ und $g$ . . . . .	850
27.4.2	Bootstrap-Konfidenzintervalle . . . . .	852
27.4.3	Exakte Konfidenzintervalle . . . . .	856
27.5	Konfidenzintervalle für weitere Effektgrößen . . . . .	861
27.5.1	Konfidenzintervalle für Anteile . . . . .	862
27.5.2	Konfidenzintervalle für Relative Risiken (RR) und Odds Ratios (OR) . . . . .	863
Kapitel 28 Metaanalyse		867
28.1	Metaanalyse in Grundzügen . . . . .	869
28.1.1	Empirische Stichprobenverteilungen als Ausgangsbasis. . . . .	870
28.1.2	Metaanalyse versus „Signifikanz-Zählen“ . . . . .	871
28.1.3	Annahmen über Populationseffekte: „Fixed effects“ versus „random effects“. . . . .	871
28.1.4	Wichtige Einflussgrößen . . . . .	872
28.2	Praktische Durchführung . . . . .	874
28.2.1	Suche nach passenden Studien. . . . .	874
28.2.2	Auswahl von Studien: Kriterien . . . . .	875
28.2.3	Berechnung und Kombination von Effektgrößen . . . . .	876
28.2.4	Analyse potenzieller Moderatorvariablen. . . . .	879
28.3	Varianten von Metaanalysen. . . . .	881
28.3.1	„Äpfel und Birnen“: Psychometrische Metaanalyse . . . . .	882
28.3.2	„Normalverteilte Apfelsorten“: Das HO-Modell. . . . .	886
28.3.3	„Fehlen manche Äpfel systematisch?“: $p$ -Curve . . . . .	893
28.4	Weitere Ansätze zur Diagnose und Kontrolle potenzieller Probleme. . . . .	897
28.4.1	Fail-safe $N$ . . . . .	898

28.4.2	Funnel-Plot mit Stichprobengrößen . . . . .	899
28.4.3	Trim-and-fill . . . . .	901
28.5	Metaanalyse im Kontext . . . . .	902
28.5.1	Weitere Varianten von Metaanalysen . . . . .	902
28.5.2	Verhältnis von Einzelstudien und Metaanalysen . . . . .	903
28.5.3	Die Aussagekraft von gemittelten Effektgrößen . . . . .	903
<b>Kapitel 29 Besonderheiten bei Datenerhebung</b> . . . . .		907
29.1	Die Problematik fehlender Daten (missing data) . . . . .	907
29.1.1	Fehlende Daten: drei unterschiedliche Fälle . . . . .	908
29.1.2	Diagnosemöglichkeiten: Fehlen die Daten zufällig? . . . . .	909
29.1.3	„Traditioneller“ (suboptimaler) Umgang mit fehlenden Daten. . . . .	909
29.1.4	Empfehlenswerte Ersetzungsverfahren . . . . .	911
29.1.5	Der Umgang mit fehlenden Daten: Rekapitulation . . . . .	913
29.2	Verfälschte Stichproben . . . . .	913
29.2.1	Selektive Stichproben . . . . .	913
29.2.2	„Nonsampling Error“: Verfälschung durch „Nichtziehen“ . . . . .	916
29.2.3	Ziehen nach Ergebnis . . . . .	919
29.3	Unverfälschte Antworten bei sensiblen Fragen: Randomized Response . . . . .	922
29.3.1	Randomized Response für Anteile I . . . . .	922
29.3.2	Randomized Response für Anteile II . . . . .	924
29.3.3	Randomized Response für Mittelwerte . . . . .	926

**Teil VII Alternative Vorgehensweisen** . . . . . **929**

**Kapitel 30 Experimentelles Design bei Interventionen** . . . . . **933**

30.1	Grundlegende Aspekte . . . . .	933
30.1.1	Die Rolle der Baselines . . . . .	933
30.1.2	Variation der Bedingungen . . . . .	936
30.1.3	Potenzielle Probleme des Standarddesigns . . . . .	938
30.2	Multiple-Baseline-Designs . . . . .	939
30.2.1	Multiple-Baselines über Personen . . . . .	939
30.2.2	Multiple Baselines über Verhaltensweisen . . . . .	940
30.2.3	Multiple-Baselines über Situationen . . . . .	941
30.3	Alternating-Treatment-Designs . . . . .	943
30.3.1	Das Prinzip . . . . .	943
30.3.2	Ein Beispiel . . . . .	944
30.4	Gütekriterien in experimentellen Einzelfallanalysen . . . . .	945
30.4.1	Interne Validität . . . . .	945
30.4.2	Externe Validität . . . . .	946
30.5	Statistische Analyse . . . . .	946
30.5.1	Signifikanztests . . . . .	947
30.5.2	Effektgrößen . . . . .	949
30.5.3	Metaanalyse . . . . .	950

Kapitel 31	Computersimulation als Forschungsmethode	953
31.1	Warum Computermodellierung?	955
31.1.1	„Reichere“ Modelle	955
31.1.2	Präzisere Vorhersagen	955
31.1.3	Aufhebung künstlicher Trennungen	956
31.2	Was kann man wie modellieren?	957
31.2.1	Art der Repräsentation: Symbolisch vs. subsymbolisch	957
31.2.2	Art der modellierten Prozesse: Kognition, Sozialverhalten und Evolution	958
31.3	Produktionssysteme	959
31.3.1	Architektur und Funktionsweise	959
31.3.2	Ein spezifisches Modell: ACT-R	961
31.3.3	Wofür sind Produktionssystem-Modelle geeignet?	963
31.4	Verteilte Modelle	963
31.4.1	Architektur und Funktionsweise	964
31.4.2	Beispiele	965
31.4.3	Wofür sind einfache verteilte Modelle geeignet?	969
31.5	Neuronale Netzwerke	969
31.5.1	Architektur und Funktionsweise	970
31.5.2	Beispiele	973
31.5.3	Wofür sind neuronale Netzwerke geeignet?	978
31.6	Genetische Algorithmen	978
31.6.1	Architektur und Funktionsweise	979
31.6.2	Beispiele	981
31.6.3	Wofür sind genetische Algorithmen geeignet?	984
31.7	Praktische Vorgehensweise	985
31.7.1	Bewertung von Simulationsergebnissen	985
31.7.2	Programmierung	986
31.7.3	Simulationsumgebungen	986
31.8	Möglichkeiten und Grenzen der Computermodellierung	987
Kapitel 32	Qualitative Methoden	991
32.1	Qualitative Methoden im Überblick	993
32.1.1	Zielstellung qualitativer Forschung: Drei Sichtweisen	993
32.1.2	Die wissenschaftliche Methode: Qualitative Version	995
32.1.3	Die Vielfalt qualitativer Ansätze	997
32.2	Spezifische Ansätze: Eine Auswahl	998
32.2.1	Qualitative Inhaltsanalyse	998
32.2.2	Grounded Theory	1001
32.2.3	Diskursanalyse	1006
32.3	Der qualitative Forschungsprozess	1010
32.3.1	Datensammlung	1010
32.3.2	Datenanalyse	1011
32.3.3	Gütekriterien	1013

32.4	Qualitative Methoden: Eine kritische Bewertung . . . . .	1015
32.4.1	Qualitative „Messung“ . . . . .	1016
32.4.2	Qualitative Methoden und Falsifizierbarkeit . . . . .	1018
32.4.3	Wie man qualitative Forschung <i>nicht</i> betreiben sollte . . . . .	1019
32.4.4	Wann sind qualitative Methoden nützlich? . . . . .	1019

## **Teil VIII    Reflexion** **1023**

### **Kapitel 33    Bewährte Methoden und neue Ansätze** **1027**

33.1	Bewährte Methoden und neue Ansätze . . . . .	1027
33.1.1	Inferenzstatistik: Erweiterte Perspektiven . . . . .	1028
33.1.2	Die Rolle von experimentellen Einzelfallanalysen . . . . .	1029
33.1.3	Die Rolle von Simulationen . . . . .	1029
33.1.4	Die Rolle der qualitativen Methoden . . . . .	1030
33.2	Forschungsmethoden und Statistik als Argument . . . . .	1030
33.2.1	Die zwei Funktionen von Forschungsmethoden und Statistik . .	1031
33.2.2	Überzeugende Argumente: Die MAGIC-Kriterien . . . . .	1031
33.2.3	Die Rolle des Signifikanztests in der statistischen Argumentation. . . . .	1032
33.3	Die Methodenbrille: Sehhilfe oder Sehbehinderung? . . . . .	1035

## **Anhang Tabellen** **1040**

## **Bibliografie** **1059**

## **Stichwortverzeichnis** **1083**