



# Forschungsmethoden und Statistik

für Psychologen und Sozialwissenschaftler

4., aktualisierte und erweiterte Auflage

**Peter Sedlmeier**  
**Frank Renkewitz**

# **Forschungsmethoden und Statistik für Psychologen und Sozialwissenschaftler**

# Forschungsmethoden und Statistik für Psychologen und Sozialwissenschaftler

## Inhaltsverzeichnis

Forschungsmethoden und Statistik für Psychologen und  
Sozialwissenschaftler

Inhaltsverzeichnis

Vorwort

Vorwort zur 4. Auflage

Teil 1: Grundlagen und Konzepte

1 Alltagswissen versus Wissenschaft: Beispiel Psychologie

1.1 Die Fallstricke der Alltagspsychologie

1.1.1 Fehler beim Wahrnehmen

1.1.2 Fehler beim Erinnern

1.1.3 Fehler beim logischen Denken

1.1.4 Fehler beim Umgang mit Wahrscheinlichkeiten

1.2 Sprachgebrauch in Alltag und Wissenschaft

1.2.1 Missverständnisse beim Verstehen von Sprache im Alltag

1.2.2 Präzisierung der Sprache in der Wissenschaft

1.3 Die wissenschaftliche Methode

1.3.1 Theorien, Hypothesen und ihre Präzisierung

1.3.2 Design

1.3.3 Durchführung von Studien

1.3.4 Datenanalyse und -interpretation

1.4 Was gewinnen wir durch die wissenschaftliche Vorgehensweise?

Weiterführende Literatur

2 Wissenschaftstheorie, Theorien und Hypothesen

2.1 Was ist die Wirklichkeit und wie können wir sie erkennen?

2.1.1 Das Leib-Seele-Problem

2.1.2 Induktion vs. Deduktion

# Inhaltsverzeichnis

## 2.2 Wissenschaftstheoretische Ansätze im Überblick

### 2.2.1 Logischer Empirismus

### 2.2.2 Kritischer Rationalismus

Fallibilismus

Falsifikationismus

Das Abgrenzungsproblem

Die Bedeutung der Kritik

Empirischer Gehalt und Theoriwahl

Probleme der Falsifizierbarkeit

### 2.2.3 Historisch-soziologische Analyse (Kuhn)

### 2.2.4 Methodologie wissenschaftlicher Forschungsprogramme (Lakatos)

### 2.2.5 Wirklichkeit als Konstruktion

Positivismus und Positivismus-Kritik

Diskursive Psychologie

## 2.3 Spezialprobleme der Psychologie

### 2.3.1 Latente Variablen

### 2.3.2 Verhältnis zwischen Forscher und „Erforschten“

## 2.4 Woher kommen Theorien?

### 2.4.1 Bed, Bathroom and Bicycle

Intuition

Induktion

Metaphern

### 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

Theorie

Forschungshypothese

Präzisierung der Forschungshypothese

Statistische Hypothesen

Weitere Vorgehensweise

### 2.5.3 Von der Theorie zur Hypothesenprüfung: Beispiele

Konformitätsdruck in Gruppen (Asch, 1955)

Die Entstehung von Emotionen (Schachter & Singer, 1962)

### 2.5.4 Hypothesenprüfung und Wissenschaftstheorie

## Weiterführende Literatur

## 3 Messen und Testen

# Inhaltsverzeichnis

## 3.1 Was ist Messen?

## 3.2 Messtheorie

### 3.2.1 Messtheoretische Probleme

Das Repräsentationsproblem

Das Eindeutigkeitsproblem

Das Bedeutsamkeitsproblem

## 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

Durchführungsobjektivität

Auswertungsobjektivität

Interpretationsobjektivität

### 3.5.2 Reliabilität

Die Retest-Methode

Die Paralleltest-Methode

Die Testhalbierungsmethode

### 3.5.3 Validität

Inhaltsvalidität

Kriteriumsvalidität

Konstruktvalidität

## Weiterführende Literatur

## 4 Datenerhebung: Befragung und Beobachtung

## 4.1 Befragung: Unterschiedliche Perspektiven

### 4.1.1 Mündlich oder schriftlich?

Mündliche Befragung

Schriftliche Befragung

Benutzen neuer Technologien

### 4.1.2 Freie oder festgelegte Antwortmöglichkeiten?

Festgelegte Antwortmöglichkeiten

Freie Antwortmöglichkeiten

### 4.1.3 Einzel- oder Gruppenbefragung?

# Inhaltsverzeichnis

Einzelbefragung

Gruppenbefragung

## 4.1.4 Wie sehr standardisieren?

Leitfaden-Interviews

Standardisierte Interviews, Fragebögen und Tests

## 4.2 Befragung: Fehlermöglichkeiten und Gegenmaßnahmen

### 4.2.1 Potenzielle Probleme bei der Gestaltung und Anordnung von Items

Soziale Erwünschtheit

Zusatzinformationen in der Frage

Gestaltung der Antwortvorgaben

Praktische Interpretation von Skalenwerten

### 4.2.2 Potenzielle Probleme bei der Durchführung der Befragung

Intervieweffekte

Antwortverweigerung

## 4.3 Befragung: Ein kurzes Resümee

### 4.3.1 Wann welche Art von Befragung?

Theoriengenerierung vs. Theorienüberprüfung

Sensible Themen

### 4.3.2 Einige abschließende Hinweise

Ablaufplanung

Probedurchgänge

Gestaltung von Fragebögen und standardisierten Interviews

## 4.4 Beobachtung: Unterschiedliche Perspektiven

### Offen oder verdeckt?

Wie sehr in die Beobachtungssituation eingreifen?

Teilnehmend oder nicht?

Wie sehr standardisieren?

Selbstbeobachtung

Non-reaktive Beobachtung

## 4.5 Beobachtung: Fehlermöglichkeiten und Gegenmaßnahmen

### Ungenaue Beobachtungen

Reaktivität der Beobachteten

Observer-Bias

Interpretationsfehler

Gedächtnisfehler

## 4.6 Beobachtung: Ein kurzes Resümee

### 4.6.1 Wann welche Form von Beobachtung?

### 4.6.2 Einige abschließende Hinweise

# Inhaltsverzeichnis

## 4.7 Generalisierbarkeit von Befragungs- und Beobachtungsergebnissen

### 4.7.1 Auswahl der Situation

### 4.7.2 Auswahl der Studienteilnehmer

## Weiterführende Literatur

## 5 Experimentelle Designs

### 5.1 Warum werden Experimente durchgeführt?

### 5.2 Die Logik des Experiments

#### 5.2.1 Grundlage für Kausalschlüsse

Kovariation

Zeitliche Präzedenz

Ausschluss von Alternativerklärungen

#### 5.2.2 Interne Validität

### 5.3 Kontrolltechniken

#### 5.3.1 Kontrolle personengebundener Störvariablen

Randomisieren

Parallelisieren

#### 5.3.2 Kontrolle von Störvariablen in der Versuchssituation

Konstanthalten und Eliminieren

Balancieren

Randomisieren

Kontrolle von Erwartungseffekten

### 5.4 Externe Validität

#### 5.4.1 Wie wichtig ist die externe Validität?

#### 5.4.2 Wie kann die externe Validität erhöht werden?

### 5.5 Within-Subjects-Designs

#### 5.5.1 Warum werden Within-Subjects-Designs eingesetzt?

#### 5.5.2 Positionseffekte und ihre Kontrolle

Block-Randomisierung

ABBA-Balancierung

Vollständiges Ausbalancieren

Lateinisches Quadrat

#### 5.5.3 Carry-Over-Effekte

### 5.6 Mehrfaktorielle Designs

#### 5.6.1 Haupteffekte und Interaktionen in 2 x 2-Designs

Grafische Darstellung der Ergebnisse aus 2 x 2-Designs

#### 5.6.2 Komplexere Designs

# Inhaltsverzeichnis

5.6.3 Interaktionen und externe Validität

5.7 Quasi-Experimente

Weiterführende Literatur

## Teil 2: Deskriptive und explorative Datenanalyse

### 6 Lage- und Streuungsmaße

6.1 Warum brauchen wir Streuungsmaße?

6.2 Lage und Streuung auf einen Blick

6.2.1 Stamm-Blatt-Diagramme

6.2.2 Box-Plots

6.3 Lagemaße im Detail

6.3.1 Arithmetisches Mittel

6.3.2 Median und Quantile

6.3.3 Modalwert

6.3.4 Weitere Lagemaße

6.4 Streuungsmaße im Detail

6.4.1 Standardabweichung und Varianz

6.4.2 Interquartilsabstand und andere Quantilsabstände

6.4.3 Weitere Streuungsmaße

6.5 Wann welches Maß?

6.5.1 Skalenniveau

6.5.2 Form der Verteilung

6.6 Standardisierung: z-Werte

6.7 Population vs. Stichprobe

Weiterführende Literatur

### 7 Korrelation

7.1 Die grafische Darstellung von Korrelationen: Streudiagramme

7.2 Korrelationsmuster

7.2.1 Lineare und kurvilineare Zusammenhänge

7.2.2 Richtung und Stärke von Zusammenhängen

7.2.3 Die Bedeutung des Korrelationsmusters für die weitere Analyse

7.3 Der Produkt-Moment-Korrelationskoeffizient

7.3.1 z-Werte und der Produkt-Moment-Korrelationskoeffizient



# Inhaltsverzeichnis

## 7.4 Verzerrungen des Produkt-Moment-Korrelationskoeffizienten

### 7.4.1 Ausreißerwerte

### 7.4.2 Einschränkungen der Variabilität

### 7.4.3 Zusammenfassung von heterogenen Untergruppen

## 7.5 Korrelation und Kausalität

## 7.6 Partialkorrelation

## 7.7 Andere Zusammenhangsmaße

### 7.7.1 Korrelation zweier dichotomer Merkmale – der Phi-Koeffizient

Berechnung des Phi-Koeffizienten

### 7.7.2 Korrelation zweier ordinalskalierten Merkmale – Kendalls Tau

## Weiterführende Literatur

# 8 Lineare Regression

## 8.1 Grundbegriffe der Regressionsrechnung

### 8.1.1 Prädiktor und Kriterium

### 8.1.2 Deterministische Zusammenhänge und die Geradengleichung

### 8.1.3 Stochastische Zusammenhänge und die Regressionsgerade

### 8.1.4 Das Kriterium der kleinsten Quadrate

### 8.1.5 Bestimmung der Regressionsgeraden

### 8.1.6 Die Beziehung zwischen der Korrelation und dem Regressionsgewicht $b$

### 8.1.7 Regression mit z-standardisierten Variablen

### 8.1.8 Der Regressionseffekt

### 8.1.9 Die Vorhersage von X aus Y

## 8.2 Die Güte der Vorhersage

### 8.2.1 Varianzzerlegung

### 8.2.2 Der Determinationskoeffizient $r^2$

### 8.2.3 Der Standardschätzfehler

## 8.3 Probleme und Verzerrungen in der Regressionsrechnung

## 8.4 Ein Ausblick auf die multiple Regression

### 8.4.1 Multiple Regression mit z-standardisierten Variablen

### 8.4.2 Eine Illustration mit zwei Prädiktoren

Bestimmung der standardisierten Regressionsgewichte

Bestimmung der unstandardisierten Regressionskoeffizienten

### 8.4.3 Gütemaße in der multiplen Regression

Multiple Korrelation und der multiple Determinationskoeffizient

# Inhaltsverzeichnis

Der Standardschätzfehler

Weiterführende Literatur

## 9 Effektgrößen

9.1 Was sind Effektgrößen?

9.2 Abstandsmaße

9.3 Zusammenhangsmaße

9.4 Effektgrößen aus Effektgrößen

9.4.1 Abstandsmaße aus Abstandsmaßen

9.4.2 Korrelationen aus Abstandsmaßen

9.4.3 Abstandsmaße aus Korrelationen

9.5 Wie bedeutsam ist eine Effektgröße?

9.6 Weitere Effektgrößen-Maße

9.6.1 Relatives Risiko

9.6.2 Odds Ratio

9.6.3 Mehr zu Effektgrößen in diesem Buch

Weiterführende Literatur

## Teil 3: Inferenzstatistik

### 10 Grundlagen der Inferenzstatistik

10.1 Wahrscheinlichkeiten, kurz gefasst

10.1.1 Was ist Wahrscheinlichkeit?

10.1.2 Wahrscheinlichkeit von Konjunktionen und bedingte Wahrscheinlichkeiten

10.2 Von der Population über Stichproben zur Stichprobenverteilung

10.2.1 Simulationsbeispiel für Anteile

10.2.2 Simulationsbeispiel für Mittelwerte

10.2.3 Die tatsächliche Vorgehensweise: Von der Stichprobe zur Population

10.3 Stichprobenverteilung für Anteile

10.3.1 Binomialverteilung „per Hand“

10.3.2 Binomialverteilung mit Binomialformel

10.4 Lage- und Streuungsmaße von Stichprobenverteilungen

10.4.1 Binomialverteilung

10.4.2 Stichprobenverteilungen für Mittelwerte

10.5 Der Einfluss der Stichprobengröße auf die Stichprobenverteilung

# Inhaltsverzeichnis

10.5.1 Empirisches Gesetz der großen Zahlen

10.5.2 Zentraler Grenzwertsatz

Normalverteilung und Standardnormalverteilung

Demonstration des zentralen Grenzwertsatzes

10.6 Rekapitulation und Ausblick

Weiterführende Literatur

## 11 Konfidenzintervalle

11.1 Was ist ein Konfidenzintervall?

11.1.1 Wahrscheinlichkeitsintervalle: Ein Gedankenexperiment

11.1.2 Konfidenzintervalle für Anteile

11.1.3 Auswirkungen der Höhe der Konfidenz und der Stichprobengröße

11.1.4 Die Berechnung von Konfidenzintervallen

11.2 Konfidenzintervalle für Mittelwerte

11.3 Konfidenzintervalle für Mittelwertsunterschiede

11.3.1 Unabhängige Messungen

11.3.2 Abhängige (gepaarte) Messungen

Positive Korrelation

Negative Korrelation

Unabhängige Messungen

11.4 Die Interpretation von Konfidenzintervallen

Weiterführende Literatur

## 12 Signifikanztests

12.1 Wie funktioniert ein Signifikanztest?

12.2 Vorgehensweise nach R. A. Fisher

12.2.1 Beispiel 1: Vorzeichentest

12.2.2 Beispiel 2: t-Test für Mittelwert

12.2.3 Probleme mit der Vorgehensweise nach Fisher

12.3 Neymans & Pearsons Verbesserungsvorschläge

12.3.1 Warum braucht man die Alternativhypothese und wie wird sie bestimmt?

12.3.2 Fehler erster und zweiter Art ( und )

12.3.3 Die „Verhaltensinterpretation“ des Signifikanztestergebnisses

12.4 Welche Faktoren beeinflussen das Ergebnis eines  
Signifikanztests?

12.4.1 Populations-Effektgröße

# Inhaltsverzeichnis

12.4.2 Stichprobengröße

12.4.3 Abwägung der Fehler erster und zweiter Art

12.4.4 Minimierung des „experimentellen Fehlers“

12.4.5 Homogenität der Population(en)

## 12.5 Poweranalyse

12.5.1 Die Suche nach der Stichprobengröße: „A priori-Analyse“

12.5.2 Die Suche nach einem Kompromiss zwischen und

12.5.3 Die Suche nach weiteren Interpretationsmöglichkeiten: „post hoc-Analyse“

## 12.6 Vorgehensweise nach Neyman und Pearson

12.6.1 Beispiel 1: Vorzeichentest nach Neyman und Pearson

12.6.2 Beispiel 2: t-Test nach Neyman und Pearson

12.6.3 Akzeptanz des Ansatzes in Psychologie und Sozialwissenschaften

## 12.7 Das konventionelle Verfahren: Der „Hybrid“

12.7.1 Bestandteile

12.7.2 Vorgehensweise und Ergebnisinterpretation

## 12.8 Signifikanztests: Was man noch wissen sollte

12.8.1 Spezifikation von Null- und Alternativhypothese

Einseitige vs. zweiseitige Tests

Spezifizieren der Alternativhypothese

12.8.2 Wie man p-Werte nicht interpretieren sollte

Warum ist p kein Indikator für die Effektgröße?

Warum kann man mit p die Wahrscheinlichkeit der Hypothese nicht abschätzen?

Warum ist p nicht die Irrtumswahrscheinlichkeit?

Warum ist p kein Indikator für Replizierbarkeit?

Wie sollte man p-Werte interpretieren?

12.8.3 Signifikanztest und Konfidenzintervall

12.8.4 Allgemeine Hinweise und Empfehlungen

## Weiterführende Literatur

## 13 t-Tests

### 13.1 Unterschied zwischen zwei Mittelwerten

13.1.1 Unabhängige Stichproben

Generelle Vorgehensweise

Poweranalyse

Voraussetzungen für die Durchführung

13.1.2 Abhängige Stichproben

# Inhaltsverzeichnis

Generelle Vorgehensweise  
Poweranalyse  
Voraussetzungen für die Anwendung

## 13.2 Weitere t-Tests

### 13.2.1 Korrelation

Generelle Vorgehensweise  
Poweranalyse  
Voraussetzungen für die Durchführung

### 13.2.2 Regression

## 13.3 Effektgrößenberechnung aus Testergebnissen von t-Tests

### 13.3.1 Generelle Idee

### 13.3.2 Eine Stichprobe (Mittelwert vs. vorgegebener Wert)

### 13.3.3 Zwei unabhängige Stichproben

### 13.3.4 Zwei abhängige Stichproben

### 13.3.5 Korrelation und Regression

## Weiterführende Literatur

## 14 Der F-Test in der einfaktoriellen Varianzanalyse

### 14.1 Warum nicht mehrere t-Tests?

### 14.2 Die Logik der Varianzanalyse

#### 14.2.1 Zwei Wege zu einer Schätzung der Populationsvarianz

Schätzung der Populationsvarianz aufgrund der Variation innerhalb der Stichproben  
Schätzung der Populationsvarianz aufgrund der Variation zwischen den Stichprobenmittelwerten  
F-Wert und F-Verteilung

#### 14.2.2 Varianzzerlegung

ANOVA-Tabelle

### 14.3 Voraussetzungen der einfaktoriellen Varianzanalyse

### 14.4 Post-hoc-Tests

### 14.5 Effektgrößen in der einfaktoriellen Varianzanalyse

### 14.6 Power in der einfaktoriellen Varianzanalyse

## 15 Weitere F-Tests

### 15.1 Mehrfaktorielle Varianzanalyse

#### 15.1.1 Varianzzerlegung in der zweifaktoriellen Varianzanalyse

Quadratsummen  
Freiheitsgrade, Varianzschätzungen und F-Werte

#### 15.1.2 ANOVA-Tabelle

# Inhaltsverzeichnis

15.1.3 Varianzanalysen mit mehr als zwei Faktoren

15.1.4 Voraussetzungen der mehrfaktoriellen Varianzanalyse

15.1.5 Mehrfaktorielle Varianzanalysen mit ungleichen Stichprobengrößen

15.1.6 Effektgrößen in der mehrfaktoriellen Varianzanalyse

15.1.7 Power in der mehrfaktoriellen Varianzanalyse

## 15.2 Varianzanalyse mit abhängigen Stichproben

15.2.1 Varianzzerlegung in der einfaktoriellen Varianzanalyse mit abhängigen Stichproben

Die Idee der Varianzzerlegung bei abhängigen Stichproben

Quadratsummen

Freiheitsgrade, Varianzschätzungen und F-Wert

15.2.2 ANOVA-Tabelle

15.2.3 Voraussetzungen der Varianzanalyse mit abhängigen Stichproben

15.2.4 Effektgrößen in der Varianzanalyse mit abhängigen Stichproben

15.2.5 Power in der Varianzanalyse mit abhängigen Stichproben

15.2.6 Erweiterungen zur Varianzanalyse mit abhängigen Stichproben

## 15.3 Der F-Test in der Regressionsrechnung

## 15.4 Weitere Varianten der Varianzanalyse

Varianzanalyse mit zufälligen Effekten

Kovarianzanalyse

Multivariate Varianzanalyse

## Weiterführende Literatur

## 16 Kontrastanalyse

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

16.1.1 Die Problematik von Omnibus-Hypothesen

16.1.2 Kontraste als präzise Hypothesen

Kontraste: Lambdagewichte aus spezifischen Vorhersagen

Arten von Kontrasten

### 16.2 Kontrastanalyse für unabhängige Stichproben

16.2.1 F-Kontrast und t-Kontrast

Quantitative Hypothese

Ordinale Fragestellungen

Qualitative Fragestellungen

16.2.2 Orthogonale Kontraste

16.2.3 Effektgrößen bei der Kontrastanalyse für unabhängige Stichproben

# Inhaltsverzeichnis

Berechnung von effect size aus den Rohwerten

Berechnung von  $\eta^2$

Effektgrößen aus Signifikanztestergebnissen

Die Effektgrößenmaße im Vergleich

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

Weiterführende Literatur

## 17 Verfahren zur Analyse nominalskalierten Daten: Chi-Quadrat (2-)Tests

17.1 Der 2-Test für eine Variable

17.1.1 Die Gleichverteilungsannahme als Nullhypothese

17.1.2 Der 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 2-Test für zwei Variablen

17.2.1 Die Unabhängigkeitsannahme als Nullhypothese

17.2.2 Berechnung des 2-Werts

17.2.3 Freiheitsgrade und Signifikanzprüfung

17.2.4 Effektgrößen

17.2.5 Power

17.3 Voraussetzungen der 2-Tests

Weiterführende Literatur

## 18 Verfahren zur Analyse ordinalskalierten Daten

18.1 Voraussetzungsverletzungen in  $\chi^2$ -parametrischen -Tests

18.2 Der U-Test

18.2.1 Zuordnung der Rangplätze

# Inhaltsverzeichnis

- 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
- 18.2.6 Rangbindungen

## 18.3 Der Wilcoxon-Test

- 18.3.1 Durchführung des Wilcoxon-Tests
- 18.3.2 Eine Voraussetzung des Wilcoxon-Tests

## 18.4 Powerbestimmung im U-Test und Wilcoxon-Test

Weiterführende Literatur

## 19 Resampling-Verfahren

### 19.1 Konventionelle Inferenzstatistik versus Resampling-Verfahren

### 19.2 Resampling-Verfahren: Warum und wie?

- 19.2.1 Zwei wesentliche Vorteile
- 19.2.2 Die Stichprobe als repräsentatives Abbild der Population
- 19.2.3 Resampling-Stichprobenverteilungen

### 19.3 Bootstrap: Konfidenz nach Münchhausen-Art

- 19.3.1 Wie funktioniert der Bootstrap?
- 19.3.2 Bootstrap: Anwendungsbeispiele

### 19.4 Randomisierungstests

- 19.4.1 Wie funktionieren Randomisierungstests?
- 19.4.2 Randomisierungstests: Anwendungsbeispiele
- 19.4.3 Besonderheiten bei Randomisierungstests

Bestimmung der Nullhypothese

Rolle der Alternativhypothese

Effektgrößen

### 19.5 Resampling-Verfahren im Kontext

- 19.5.1 Bootstrappen oder Randomisieren?
- 19.5.2 Weitere Resampling-Verfahren
- 19.5.3 Resampling-Verfahren versus traditionelle Inferenzstatistik
- 19.5.4 Praktische Vorgehensweise

Weiterführende Literatur

## Teil 4: Inferenzstatistik: Praktische Probleme und



# Inhaltsverzeichnis

## alternative Sichtweisen

### 20 Probleme der klassischen Inferenzstatistik in der Forschungspraxis

#### 20.1 Replizierbarkeit in der Psychologie

##### 20.1.1 Das Reproducibility Project: Psychology

Methode

Ergebnisse

Schlussfolgerungen

##### 20.1.2 Andere Befunde zur Replizierbarkeit in der Psychologie

#### 20.2 Ursachen der Replikationskrise

##### 20.2.1 Probleme bei der Interpretation des Signifikanztests

Wie häufig sind falsch positive Resultate?

Wie sollten wir signifikante Ergebnisse interpretieren?

##### 20.2.2 Probleme in der Praxis: Publikationsbias, HARKing und p-Hacking

Publikationsbias

HARKing

p-Hacking

Wie verbreitet ist p-Hacking?

#### 20.3 Problemlösungen

#### Weiterführende Literatur

### 21 Replikation, Präregistrierung, Open Science

#### 21.1 Replikation

##### 21.1.1 Typen von Replikationsstudien

Direkte Replikationen

Konzeptuelle Replikationen

##### 21.1.2 Wann ist eine Replikation erfolgreich?

War die Originalstudie aussagekräftig?

Bestätigt die Replikation eher das Ergebnis der Originalstudie oder die Nullhypothese?

Wann welches Verfahren?

##### 21.1.3 Was ist eine gute Replikationsstudie?

#### 21.2 Präregistrierung

#### 21.3 Open Science

#### Weiterführende Literatur

### 22 Bayesianische Statistik

#### 22.1 Die Revision von Wahrscheinlichkeiten

##### 22.1.1 Das Bayes-Theorem

# Inhaltsverzeichnis

22.2 Bayesianische Wahrscheinlichkeiten

22.3 Priors, Likelihoods und Posteriors

22.3.1 Priorverteilung

22.3.2 Likelihoods

22.3.3 Posteriorverteilung

22.4 Stetige Priorverteilungen und konjugierte Priors

22.5 Einflussgrößen auf die Posteriorverteilung

22.5.1 Auswirkungen der Priorverteilung

22.5.2 Auswirkungen der Stichprobengröße

22.6 Bayes-Faktor

22.7 Klassisch vs. Bayesianisch

Weiterführende Literatur

## Teil 5: Das Allgemeine Lineare Modell

### 23 Das Allgemeine Lineare Modell

23.1 Was ist das Allgemeine Lineare Modell?

23.2 Der t-Test als Spezialfall der einfachen Regression

23.3 Varianzanalyse mit zwei Gruppen als Spezialfall der einfachen Regression

23.4 Varianzanalyse mit mehr als zwei Gruppen als Spezialfall der multiplen Regression

Weiterführende Literatur

### 24 Regressionsrechnung: Ergänzungen und Erweiterungen

24.1 Multiple Regression: Ergänzungen

24.1.1 Schrittweise Regression

Theoriegeleitete Vorgehensweise

Reihenfolge nach Erklärungskraft der Prädiktoren

Wann was?

24.1.2 Effektgrößen bei der multiplen Regression

Gesamteffekt und Zuwachs an Erklärungskraft

Auf einzelne Prädiktoren zurückführbare Effekte

Das Problem der Multikollinearität

24.1.3 Inferenzstatistik bei der multiplen Regression

Gesamteffekt und Zuwachs an Erklärungskraft

Auf einzelne Prädiktoren zurückführbare Effekte

24.1.4 Analyse nichtlinearer Beziehungen

# Inhaltsverzeichnis

24.2 Kontrastanalyse mittels Regressionsrechnung

24.3 Kovarianzanalyse mittels Regressionsrechnung

24.4 Moderatoranalyse: Die generelle Behandlung von Interaktionen

24.4.1 Interaktion als multiplikative Komponente

24.4.2 Zentrieren der Prädiktorvariablen

24.4.3 Interaktion zwischen zwei nominalskalierten Variablen

24.4.4 Interaktion zwischen einer nominal- und einer intervallskalierten Variable

24.4.5 Interaktion zwischen zwei intervallskalierten Variablen

24.4.6 Interaktion in komplexeren Fällen

Weiterführende Literatur

## 25 Indirekte Effekte, latente Variablen und multiple Analyseebenen

### 25.1 Pfadanalyse

25.1.1 Zusammenhang zwischen Regressionsrechnung und Pfadanalyse

Regressions-Lösung

Pfadanalytische Lösung

25.1.2 Pfadanalyse mit Mediatorvariable

### 25.2 Strukturgleichungsmodelle

25.2.1 Identifizierbarkeit

25.2.2 Mess- und Strukturmodelle

25.2.3 Schätzen der freien Parameter

25.2.4 Die Überprüfung des Modells: Gütemaße

25.2.5 Anwendungsvoraussetzungen

### 25.3 Exploratorische Faktorenanalyse

25.3.1 Datenbeispiel

25.3.2 Fundamentaltheorem der Faktorenanalyse

25.3.3 Extraktionsverfahren

25.3.4 Ladungen, Kommunalitäten, Eigenwerte

25.3.5 Faktorauswahl

25.3.6 Rotation und Interpretation

### 25.4 Mehrebenenanalyse

25.4.1 Warum Mehrebenenanalyse?

25.4.2 Regressionsgleichung für ein einfaches Mehrebenenmodell

Ebene 1

Ebene 2

# Inhaltsverzeichnis

Ebene 1 und Ebene 2

25.4.3 Feste versus zufällige Effekte

25.4.4 Theoriegeleitete Analyse

25.4.5 Explorative Vorgehensweise

Schritt 1: Intercept-only-Modell

Schritte 2 und 3: Fixieren der Prädiktoren auf Ebene 1 und Ebene 2

Schritt 4: Variieren die Steigungen der Ebene-1-Prädiktoren?

Schritt 5: Gibt es eine Interaktion zwischen den Ebenen?

25.4.6 Maße zur Beurteilung der Ergebnisse

Signifikanztests

Effektgrößen

25.4.7 Mehrebenenanalyse als Metaanalyseprozedur

25.4.8 Möglichkeiten und Grenzen der Mehrebenenanalyse

Weiterführende Literatur

## Teil 6: Weitere Verfahren in der Datenerhebung und Datenanalyse

### 26 Explorative Datenanalyse (EDA): Weitere Verfahren

26.1 Robustheit von EDA-Verfahren: Box-Plots

26.2 Varianten von Streuungsdiagrammen

26.2.1 Streuungsdiagramme mit Box-Plots

26.2.2 Influence-Plot

26.2.3 Bubble-Plot

26.3 „Aufspüren“ und „Geradebiegen“ nichtlinearer Zusammenhänge

26.3.1 Lowess

26.3.2 Potenzleiter

26.4 Multivariate Zusammenhänge auf einen Blick:

Die Streuungsdiagramm-Matrix

26.5 EDA im Kontext

Weiterführende Literatur

### 27 Visuelle Klassifikation und Clusteranalyse

27.1 Mehrdimensionale grafische Klassifikation von Fällen

27.1.1 Histogramm-Plots

27.1.2 Star-Plots

27.1.3 Chernoff-Gesichter

# Inhaltsverzeichnis

## 27.2 Clusteranalyse im Überblick

## 27.3 Proximitätsmaße

27.3.1 Ähnlichkeit versus Distanz

27.3.2 Ähnlichkeitsmatrizen

27.3.3 Distanzmatrizen

27.3.4 Weitere Möglichkeiten zur Erstellung von Proximitätsmatrizen

## 27.4 Clusteralgorithmen

27.4.1 Hierarchisch-agglomerative Vorgehensweise

27.4.2 Hierarchisch-divisive Vorgehensweise

27.4.3 Partitionierende Vorgehensweise

## 27.5 Clusteranzahl

27.5.1 Illustration für ein hierarchisches (agglomeratives) Verfahren

27.5.2 Illustration für ein partitionierendes Verfahren

## 27.6 Weitere Aspekte bei der Durchführung von Clusteranalysen

27.6.1 Indizes zur Auswahl der besten Methode

27.6.2 Empfehlungen für die Vorgehensweise

## Weiterführende Literatur

## 28 Effektgrößen: Erweiterungen und Ergänzungen

### 28.1 Populations- versus Stichprobeneffektgrößen

### 28.2 Effektgrößenschätzung bei unvollständigen Angaben

28.2.1 Nur p-Werte und Stichprobengröße(n) angegeben

requisit und t-Verteilung

Effektgrößenberechnung mithilfe der Standardnormalverteilung

28.2.2 Nur „globale“ Angaben

### 28.3 Die Vergleichbarkeit von Effektgrößen

28.3.1 Effektgrößen aus Rohdaten vs. Signifikanztestergebnissen

28.3.2 Die Vergleichbarkeit von unterschiedlichen korrelativen Maßen

28.3.3 Abstandsmaße vs. korrelative Maße

28.3.4 Unabhängige vs. abhängige Stichproben

28.3.5 Signifikanztest auf Unterschied zweier Effektgrößen

### 28.4 Konfidenzintervalle für r und g

28.4.1 Approximative Konfidenzintervalle für r und g

Konfidenzintervall für r

Konfidenzintervall für g

# Inhaltsverzeichnis

## 28.4.2 Bootstrap-Konfidenzintervalle

Konfidenzintervall für  $r$

Konfidenzintervall für  $g$

Wann Bootstrap-Konfidenzintervalle?

## 28.4.3 Exakte Konfidenzintervalle

Bestimmung exakter Konfidenzintervalle mithilfe der  $t$ -Verteilung

Exaktes Konfidenzintervall für  $g$

## 28.5 Konfidenzintervalle für weitere Effektgrößen

### 28.5.1 Konfidenzintervalle für Anteile

### 28.5.2 Konfidenzintervalle für Relative Risiken (RR) und Odds Ratios (OR)

Relatives Risiko (RR)

Odds Ratio (OR)

## Weiterführende Literatur

## 29 Metaanalyse

### 29.1 Metaanalyse in Grundzügen

#### 29.1.1 Empirische Stichprobenverteilungen als Ausgangsbasis

#### 29.1.2 Metaanalyse versus „Signifikanzen-Zählen“

#### 29.1.3 Annahmen über Populationseffekte: „Fixed effects“ versus „random effects“

#### 29.1.4 Wichtige Einflussgrößen

Stichprobengrößen in den Einzelstudien

Methodische Qualität der Studie

Inhaltliche Unterschiede

### 29.2 Praktische Durchführung

#### 29.2.1 Suche nach passenden Studien

#### 29.2.2 Auswahl von Studien: Kriterien

#### 29.2.3 Berechnung und Kombination von Effektgrößen

Abhängige vs. unabhängige Effektgrößen

Aggregation von Effektgrößen

Korrektur von Effektgrößen

#### 29.2.4 Analyse potenzieller Moderatorvariablen

Korrelation und Regression

Konfidenzintervalle für aggregierte Effektgrößen

### 29.3 Varianten von Metaanalysen

#### 29.3.1 „Äpfel und Birnen“: Psychometrische Metaanalyse

#### 29.3.2 „Normalverteilte Apfelsorten“: Das HO-Modell

#### 29.3.3 „Fehlen manche Äpfel systematisch?“: $p$ -Curve

# Inhaltsverzeichnis

## 29.4 Weitere Ansätze zur Diagnose und Kontrolle potenzieller Probleme

29.4.1 Fail-safe N

29.4.2 Funnel-Plot mit Stichprobengrößen

29.4.3 Trim-and-fill

## 29.5 Metaanalyse im Kontext

29.5.1 Weitere Varianten von Metaanalysen

29.5.2 Verhältnis von Einzelstudien und Metaanalysen

29.5.3 Die Aussagekraft von gemittelten Effektgrößen

## Weiterführende Literatur

## 30 Besonderheiten der Datenerhebung

### 30.1 Die Problematik fehlender Daten (missing data)

30.1.1 Fehlende Daten: drei unterschiedliche Fälle

Missing completely at random (MCAR)

Missing at random (MAR)

Not missing at random (NMAR)

30.1.2 Diagnosemöglichkeiten: Fehlen die Daten zufällig?

30.1.3 „Traditioneller“ (suboptimaler) Umgang mit fehlenden Daten

Fallweises Löschen

Paarweises Löschen

Mittelwertsimulation

Regressionsbasierte Imputation

30.1.4 Empfehlenswerte Ersetzungsverfahren

EM Algorithmus

Multiple zufallsbasierte Imputation

30.1.5 Der Umgang mit fehlenden Daten: Rekapitulation

### 30.2 Verfälschte Stichproben

30.2.1 Selektive Stichproben

30.2.2 „Nonsampling Error“: Verfälschung durch „Nichtziehen“

30.2.3 Ziehen nach Ergebnis

### 30.3 Unverfälschte Antworten bei sensiblen Fragen: Randomized Response

30.3.1 Randomized Response für Anteile I

30.3.2 Randomized Response für Anteile II

30.3.3 Randomized Response für Mittelwerte

## Weiterführende Literatur

# Inhaltsverzeichnis

## Teil 7: Alternative Vorgehensweisen

### 31 Experimentelle Einzelfallanalyse

#### 31.1 Grundlegende Aspekte

31.1.1 Die Rolle der Baselines

31.1.2 Variation der Bedingungen

31.1.3 Potenzielle Probleme des Standarddesigns

Carry-Over-Effekte

Mangelnde Kontrolle von Störvariablen

#### 31.2 Multiple-Baseline-Designs

31.2.1 Multiple-Baselines über Personen

31.2.2 Multiple Baselines über Verhaltensweisen

31.2.3 Multiple-Baselines über Situationen

#### 31.3 Alternating-Treatment-Designs

31.3.1 Das Prinzip

31.3.2 Ein Beispiel

#### 31.4 Gütekriterien in experimentellen Einzelfallanalysen

31.4.1 Interne Validität

31.4.2 Externe Validität

#### 31.5 Statistische Analyse

31.5.1 Signifikanztests

Die Problematik von Standardverfahren

Alternativen

31.5.2 Effektgrößen

31.5.3 Metaanalyse

#### Weiterführende Literatur

### 32 Qualitative Methoden

#### 32.1 Qualitative Methoden im Überblick

32.1.1 Zielstellung qualitativer Forschung: Drei Sichtweisen

Ergänzung zu konventionellen Verfahren

Gleichberechtigte Ansätze

Alternative zum konventionellen Ansatz

32.1.2 Die wissenschaftliche Methode: Qualitative Version

32.1.3 Die Vielfalt qualitativer Ansätze

#### 32.2 Spezifische Ansätze: Eine Auswahl



# Inhaltsverzeichnis

## 32.2.1 Qualitative Inhaltsanalyse

## 32.2.2 Grounded Theory

Codieren

Memos schreiben

Theoretical Sampling

Theoriebildung

Art der Theorien

## 32.2.3 Diskursanalyse

Diskursive Psychologie

Foucaultsche Diskursanalyse

Art der Theorien

## 32.3 Der qualitative Forschungsprozess

### 32.3.1 Datensammlung

### 32.3.2 Datenanalyse

Kunst oder Technik?

Generelle Vorgehensweise

### 32.3.3 Gütekriterien

Prozedurale Reliabilität

Alternative Sicht der Validität

Triangulation

Verallgemeinerung

Generelle Hilfen

## 32.4 Qualitative Methoden: Eine kritische Bewertung

### 32.4.1 Qualitative „Messung“

Das Repräsentationsproblem bei qualitativen Messungen

Die Subjektivität qualitativer Messungen

Variablen untersuchen vs. nach Variablen suchen

### 32.4.2 Qualitative Methoden und Falsifizierbarkeit

### 32.4.3 Wie man qualitative Forschung nicht betreiben sollte

### 32.4.4 Wann sind qualitative Methoden nützlich?

Weiterführende Literatur

## Teil 8: Reflexion

### 33 Methode und Inhalt

#### 33.1 Bewährte Methoden und neue Ansätze

##### 33.1.1 Inferenzstatistik: Erweiterte Perspektiven

##### 33.1.2 Die Rolle von experimentellen Einzelfallanalysen

# Inhaltsverzeichnis

33.1.3Die Rolle von Simulationen

33.1.4Die Rolle der qualitativen Methoden

## 33.2Forschungsmethoden und Statistik als Argument

33.2.1Die zwei Funktionen von Forschungsmethoden und Statistik

33.2.2Überzeugende Argumente: Die MAGIC-Kriterien

33.2.3Die Rolle des Signifikanztests in der statistischen Argumentation

Signifikanztestergebnis vs. Idealfall

Bezug zur Falsifikationsstrategie

Harte vs. weiche Evidenz

## 33.3Die Methodenbrille: Sehhilfe oder -Sehbehinderung?

Weiterführende Literatur

Bibliografie

Stichwortverzeichnis

Anhang

Tabellen

Bibliografie

Stichwortverzeichnis

Impressum

# Copyright

Daten, Texte, Design und Grafiken dieses eBooks, sowie die eventuell angebotenen eBook-Zusatzdaten sind urheberrechtlich geschützt. Dieses eBook stellen wir lediglich als **persönliche Einzelplatz-Lizenz** zur Verfügung!

Jede andere Verwendung dieses eBooks oder zugehöriger Materialien und Informationen, einschließlich

- der Reproduktion,
- der Weitergabe,
- des Weitervertriebs,
- der Platzierung im Internet, in Intranets, in Extranets,
- der Veränderung,
- des Weiterverkaufs und
- der Veröffentlichung

bedarf der **schriftlichen Genehmigung** des Verlags. Insbesondere ist die Entfernung oder Änderung des vom Verlag vergebenen Passwort- und DRM-Schutzes ausdrücklich untersagt!

Bei Fragen zu diesem Thema wenden Sie sich bitte an: **info@pearson.de**

## Zusatzdaten

Möglicherweise liegt dem gedruckten Buch eine CD-ROM mit Zusatzdaten oder ein Zugangscode zu einer eLearning Plattform bei. Die Zurverfügungstellung dieser Daten auf unseren Websites ist eine freiwillige Leistung des Verlags. **Der Rechtsweg ist ausgeschlossen.** ZugangsCodes können Sie darüberhinaus auf unserer Website käuflich erwerben.

## Hinweis

Dieses und viele weitere eBooks können Sie rund um die Uhr und legal auf unserer Website herunterladen:

**<https://www.pearson-studium.de>**