

Hermann Singer

**Vertiefung der
Wirtschaftsmathematik
und Statistik**

4. Auflage

31. Mai 2019

Der vorliegende Kurs wird am *Lehrstuhl für angewandte Statistik und Methoden der empirischen Sozialforschung* angeboten.

Autor:
Univ.-Prof. Dr. Hermann Singer

<p>Kursergänzende Lehrbücher und Multimedia-Übungen siehe Homepage des Lehrstuhls für angewandte Statistik und Methoden der empirischen Sozialforschung (www.fernuni-hagen.de/ls_statistik/)</p>

Inhaltsverzeichnis

I	Induktive Statistik	15
1	Grundlagen der induktiven Statistik	17
1.1	Einführende Bemerkungen	17
1.2	Stichprobenraum und Stichprobenfunktionen	22
1.2.1	Mittelwert	22
1.2.2	Stichprobenvarianz	26
1.2.3	Standardabweichung	27
1.3	Stichprobenverteilungen und Testverteilungen	28
1.3.1	Die Normal (Gauß)-Verteilung	29
1.3.2	Die Chi-Quadrat-Verteilung	31
1.3.3	Die t -Verteilung	35
1.3.4	Die F -Verteilung	37
2	Punkt-, Intervallschätzung, Hypothesentests	41
2.1	Punktschätzung	41
2.2	Wünschenswerte Eigenschaften von Schätzfunktionen	49
2.2.1	Erwartungstreue	49
2.2.2	Wirksamkeit (Effizienz)	50
2.2.3	Konsistenz	51
2.3	Methoden zur Konstruktion von Schätzfunktionen	53
2.3.1	Prinzip der kleinsten Quadrate	53
2.3.2	Momentenmethode	53
2.3.3	Maximum-Likelihood(ML)-Methode	54

2.4	Konfidenzintervalle	56
2.4.1	Konfidenzintervall für μ : normalverteilte Grundgesamtheit, bekannte Varianz	57
2.4.2	Konfidenzintervall für μ : normalverteilte Grundgesamtheit, unbekannte Varianz	59
2.4.3	Konfidenzintervall für μ : beliebig verteilte Grundgesamtheit, großer Stich- probenumfang	63
2.4.4	Konfidenzintervall für den Anteilswert π : Binomialverteilte Grundgesamtheit, großer Stich- probenumfang	64
2.4.5	Konfidenzintervall für die Varianz σ^2 : Normalverteilte Grundgesamtheit	66
2.4.6	Konfidenzintervall für den Korrelationskoeffizienten ρ : Normalverteilte Grundgesamtheit	67
2.5	Statistische Überprüfung von Hypothesen: Der Signifikanztest	69
2.5.1	Hypothesen	69
2.5.2	Aufbau und Interpretation von Hypothesentests .	71
2.5.3	Fehler bei Testentscheidungen	74
2.5.4	Einseitige Hypothesen	80
2.5.5	Allgemeine Gesichtspunkte beim Hypothesentest	83
3	Tests bei speziellen Parametern	85
3.1	Gesichtspunkte bei statistischen Tests	85
3.2	Ein-Stichproben-Fall	86
3.2.1	quantitative Variable $X : H_0 : \mu = \mu_0$ Prüfung des Erwartungswerts	86
3.2.2	quantitative Variable $X : H_0 : \sigma = \sigma_0$ Prüfung der Varianz	88
3.2.3	Dichotome Variable X : Prüfung von Anteilswerten	89
3.2.4	Nichtparametrische Methoden	92

3.2.5	Prüfung auf eine bestimmte Verteilung: χ^2 -Anpassungstest	99
3.3	Zwei-Stichproben-Fall	104
3.3.1	Unabhängige Stichproben	104
3.3.2	Abhängige Stichproben	115
3.4	Zusammenhang Test-Konfidenzintervall	123
4	Zusammenhangsanalyse	127
4.1	Metrische Merkmale	127
4.1.1	Signifikanztest $\rho = 0$	128
4.1.2	Signifikanztest $\rho = \rho_0$	130
4.2	Ordinale Merkmale: Spearman's Rangkorrelation	133
4.3	Beliebiges Meßniveau: χ^2 -Unabhängigkeitstest	135
4.4	Kausalität, Abhängigkeit und Korrelation	137
5	Regressionsanalyse	141
5.1	Lineares Modell und Kleinste-Quadrate-Schätzung (KQ)	141
5.1.1	Das lineare Modell	141
5.1.2	Kleinste-Quadrate-Schätzung (KQ)	144
5.2	Konstruktion von Konfidenzintervallen und Tests	152
5.3	Konfidenzintervalle für prognostizierte Werte	155
5.3.1	Konfidenzintervall für $E[Y X] = \alpha + \beta X$ (Fall 1)	156
5.3.2	Prognoseintervall für individuelles Y_0 (Fall 2)	157
5.4	Varianz-Zerlegung, Bestimmtheitsmaß und globaler F -Test	160
5.4.1	Varianz-Zerlegung	160
5.4.2	Bestimmtheitsmaß	162
5.4.3	Globaler F -Test und ANOVA-Tafel	164

5.4.4	Residualanalyse	167
6	Varianz-Analyse	169
6.1	Multiple Vergleiche	169
6.2	Einfaktorielle Varianzanalyse	174
6.3	Einfaktorielle Varianzanalyse in Effektdarstellung	179
6.4	Varianzanalyse als Prognosemethode	181
II	Empirische Sozialforschung	185
7	Statistik und Empirische Forschung	187
8	Qualitative vs. quantitative Methoden	191
9	Der Forschungsprozeß	195
9.1	Erkundungsphase	195
9.2	Theoretische Phase	197
9.2.1	Erläuterungen zu den einzelnen Punkten:	197
9.3	Planungsphase	202
9.3.1	Variablen-Auswahl	202
9.3.2	Untersuchungs-Designs	205
9.3.3	Operationalisierung	207
9.3.4	Stichprobe	207
9.3.5	Planung der statistischen Auswertung	209
9.4	Untersuchungsphase	210
9.5	Auswertungsphase	210
9.6	Entscheidungsphase	217
10	Falsifikation und statistische Hypothesenprüfung	221
11	Hypothesenprüfung mit Prädiktionsregeln	225
11.1	Wahrheitstafeln	225
11.2	A priori-Regeln	228
11.3	A posteriori-Regeln	232

III Computergestützte Datenanalyse	241
12 Fallstudien	245
12.1 Einführung	245
12.2 Makroökonomische Daten	247
12.3 Aktienkurs-Renditen	253
12.4 Graphische Zusammenhänge zwischen Variablen	255
12.5 Deskriptive Statistik	263
12.6 Statistische Tests	265
12.7 Übungen mit SPSS und SAS/JMP	269
13 Erste Auswertungsschritte	271
13.1 Daten-Sätze und Daten-Matrizen	271
13.2 Erzeugen einer SPSS-System-Datei	272
13.2.1 Öffnen einer SPSS-System-Datei	272
13.2.2 Öffnen einer Tabelle anderer Programme	272
13.2.3 Manuelle Eingabe im SPSS-Dateneditor	275
13.3 Explorative Datenanalyse	277
13.3.1 Univariate Analyse: Histogramme und summarische Statistiken	277
13.3.2 Bivariate Analyse: Kreuztabellen und Streudiagramme	281
13.4 Einfache Unterschieds-Hypothesen	288
14 Gesichtspunkte bei statistischen Analysen	295
14.1 Übersicht	295
14.2 Univariate Statistiken nach Skalen-Niveaus	299
14.3 Bivariate Korrelationen nach Skalen-Niveaus	300
14.4 Asymmetrische Verfahren nach Skalen-Niveaus	305
14.4.1 Implikation	305
14.4.2 Kontraposition	307

14.5	Zusammenhänge in Kreuztabellen I: symmetrische Maße	314
14.6	Zusammenhänge in Kreuztabellen II: Prädiktionsregeln	315
14.7	Regressions-Modelle (Prognose)	317
14.7.1	Einfach-Regression (X und Y univariat)	320
14.7.2	Varianz-Analyse (X und Y univariat)	322
14.8	Scheinkorrelationen, optimale Prognose, Residualstreuung	329
14.8.1	diskrete Kontrollvariablen	330
14.8.2	Stetige Kontrollvariablen: Allgemeiner Fall	338
14.8.3	Stetige Kontrollvariablen: Linearer Fall	340
14.8.4	* Stetige Kontrollvariablen: Multivariate Regression	352
14.9	Reliabilität	360
14.9.1	Meßmodell	360
14.9.2	Reliabilität	360
14.9.3	Parallel-Test, Re-Test	361
14.9.4	Test-Halbierung	361
14.9.5	Allgemeine Spearman-Brown-Formel (k Test-Teile)	362
14.9.6	Cronbach's Alpha	363
14.9.7	Abschwächungs-Korrektur	365
14.10	Skalen	369
14.10.1	Likert-Skala	369
14.10.2	Item-Analyse	370
14.11	Faktoren-Analyse	372
14.11.1	Modell	372
14.11.2	Hauptkomponentenanalyse	380
14.11.3	Mathematischer Einschub: Hauptachsentransformation	385
14.11.4	Maximum-Likelihood-Methode	390
14.12	Methode der kleinsten Quadrate	397
15	Fallstudie: Filialgestaltung und Kundenzufriedenheit	399

15.1 Fragestellung	399
15.1.1 Ausgangslage	399
15.1.2 Untersuchungsziel	400
15.1.3 Stand der Forschung	401
15.1.4 Wirkungsmodell	401
15.2 Operationalisierung der Konstrukte	405
15.2.1 Wahrgenommene Filialgestaltung	405
15.2.2 Wahrgenommene Kontrolle	406
15.2.3 Nutzung und Zielerreichung	406
15.2.4 Zufriedenheit	407
15.2.5 Loyalität	407
15.2.6 Soziodemographische Fragen	408
15.2.7 Objektive Filialgestaltung	408
15.3 Durchführung der Untersuchung	408
15.4 Variablen in SPSS	412
15.5 Überblick	414
15.5.1 Univariate Analyse	414
15.5.2 Bivariate Analyse: Kreuztabellen, Korrelationen, Streudiagramme	419
15.6 Item-Analyse	428
15.6.1 Wahrgenommene Kontrolle	428
15.6.2 Nutzung und Zielerreichung	430
15.6.3 Wahrgenommene Kontrolle + Nutzung und Zielerreichung	430
15.6.4 Andere Items	432
15.7 Zusammenhänge der Konstrukte	436
15.7.1 Umcodierung	436
15.7.2 Korrelation und gruppierte Streudiagramme	436
15.7.3 Mittelwerte	442
15.7.4 Statistische Überprüfung	445
15.7.5 Regression	445
15.7.6 Mittelwerts-Vergleiche	449

IV Anhang	457
16 Aufgaben	459
16.1 Aufgaben KE 1	459
16.1.1 Punkt- und Intervallschätzung	459
16.1.2 Tests für Anteilswerte	463
16.1.3 Anpassungstests/Unabhängigkeitstests	464
16.1.4 Tests auf zentrale Tendenz	466
16.1.5 Zusammenhangsanalyse	471
16.1.6 Regressionsanalyse	472
16.1.7 Varianzanalyse	476
16.2 Aufgaben KE 2	478
16.3 Aufgaben KE 3	483
17 Tabellen	491
17.1 Binomialverteilung	492
17.2 Normalverteilung	500
17.3 χ^2 -Verteilung	504
17.4 F -Verteilung	508
17.5 Student- t -Verteilung	511
17.6 Wilcoxon-Vorzeichen-Rangtest	515
17.7 Wilcoxon-Rangsummentest	516
17.8 Zufallszahlentabelle	520
17.9 Übersicht zur Approximation von Verteilungen	522
18 Notation und Rechenregeln	525
18.1 Formeln (alphabetisch)	525
18.2 Beispiel: symmetrischer Würfel	533
18.2.1 Einmaliges Würfeln	533
18.2.2 Zweimaliges Würfeln	534

Vorwort

Der hier vorgelegte Kurs *Vertiefung der Wirtschaftsmathematik und Statistik* (Teil Statistik) ist als A-Modul im Master-Studiengang Wirtschaftswissenschaft konzipiert.

Er setzt Kenntnisse der deskriptiven Statistik sowie der Wahrscheinlichkeitsrechnung voraus, wie sie in Bachelor-Studiengängen vermittelt werden (etwa im Modul *Grundlagen der Wirtschaftsmathematik und Statistik* des Bachelor-Studiengangs Wirtschaftswissenschaft der FernUniversität).

Der Kurs besteht aus den 3 Kurseinheiten

- Induktive Statistik
- Empirische Sozialforschung sowie
- Computergestützte Datenanalyse.

Im ersten Teil wird eine Einführung in die Theorie und Praxis der statistischen Inferenz gegeben, nämlich wie aus Daten Schlüsse über Grundgesamtheiten gezogen werden können. Dies erfolgt mit Hilfe von Schätzern, Konfidenzintervallen und statistischen Hypothesentests, die für unterschiedliche Fragestellungen und Datenqualitäten formuliert werden können.

Der zweite Teil befaßt sich mit dem Ablauf von empirischen Studien (Forschungsprozeß), mit Designproblemen sowie dem Problem der Hypothesenprüfung und Falsifikation.

Im dritten Teil werden anhand von realen Datensätzen (makroökonomische Daten, Kundenzufriedenheit bei Bankkunden etc.) die Methoden mit Hilfe von Statistik-Software-Paketen (SPSS, SAS/JMP) angewandt. Dies erfolgt immer im Zusammenhang mit der statistischen Theorie und es werden Querbezüge zu eigenen Berechnungen hergestellt, die auch mit einem Taschenrechner ausgeführt werden können.

Es handelt sich also keinesfalls um ein Bedienungshandbuch der entsprechenden Software.

Im Rahmen einer Fallstudie wird gezeigt, wie ausgehend von einem theoretischen Konzept eine Operationalisierung von Konstrukten, deren empirische Analyse und Zusammenhangsanalysen durchgeführt werden können. Die entsprechenden Berechnungen können (und sollen) von den Studierenden anhand der beigefügten Datensätze selbst ausgeführt werden. Nur dann ergibt sich ein Transfer des theoretischen Stoffs auf praktische Handlungsmöglichkeiten, wie sie in vielen Berufsfeldern verlangt werden.

Die Multimedia-Ausstattung des Kurses wird einerseits durch Standardsoftware wie SAS/JMP und SPSS, andererseits durch vom Lehrstuhl entwickelte interaktive Mathematica-Applets realisiert, die mit dem frei erhältlichen Mathematica-Player (bzw. Wolfram CDF-Player) in Realzeit abgespielt werden können (siehe http://www.fernuni-hagen.de/ls_statistik/lehre/).

Dies verbindet die Kombination von berufsbefähigenden Softwarekenntnissen (Marketing, Banken, etc.) mit speziellen Lern-Applets (etwa Scheinkorrelationen, Simulation von normalverteilten Daten etc.).

Hinweis: Kapitel, die mit einem * gekennzeichnet sind, enthalten weiterführendes Material, das beim ersten Lesen übersprungen werden kann.

Ich danke meinen Mitarbeitern vom Lehrstuhl für angewandte Statistik und Methoden der empirischen Sozialforschung, Frau Daniela Doliwa, Frau Anja Bittner, Frau Marina Lorenz und Herrn Dr. Thomas Mazzoni für vielfältige Unterstützung und kritische Lektüre des Manuskripts.

Hagen, im Juni 2009

Vorwort zur 2. Auflage

Ich danke Frau Dipl. Stat. Anja Bittner sowie Herrn Dr. Thomas Mazzoni für die sorgfältige Lektüre des Manuskripts und das Auffinden einiger Schreib- und Rechenfehler.

Ansonsten wurden bei der 2. Auflage kleine Ergänzungen vorgenommen, etwa zu Konfidenzintervallen im Rahmen der Bayes-Statistik und zur Umrechnung des Thompson-Schätzers in der Faktoren-Analyse.

Hagen, im Juli 2011

Vorwort zur 3. Auflage

In der dritten Auflage wurde der McNemar-Test zur Prüfung von Anteilswerten bei abhängigen Stichproben hinzugenommen. Weiterhin habe ich die Notation von Vektoren und Matrizen auf Fettdruck umgestellt sowie diverse Schreibfehler korrigiert. Frau M.Sc. Zulfiya Davidova danke ich für die Erstellung der Musterlösungen.

Hagen, im Juli 2014

Vorwort zur 4. Auflage

In der 4. Auflage wurden verschiedene Korrekturen vorgenommen und Hinweise zur Benutzung der Tabellen und applets eingefügt. Auch wurde auf die aktuelle Diskussion zum p -Wert vertieft eingegangen. Ich danke Herrn M.Sc. Oliver Old für die sorgfältige Lektüre des Manuskripts und diverse Hinweise.

Hagen, im April 2019

Teil I

Induktive Statistik

Kapitel 1

Grundlagen der induktiven Statistik

1.1 Einführende Bemerkungen

Deskriptive (beschreibende) Statistik ist in einem gewissen Sinne *theorielos*, indem nur bestimmte statistische Maßzahlen wie z.B. der Mittelwert, die Standardabweichung oder Korrelationskoeffizienten aus Daten berechnet werden.

Auf der anderen Seite ist die Wahrscheinlichkeitsrechnung *theoretisch* orientiert, indem Wahrscheinlichkeiten axiomatisch definiert sowie bestimmte Relationen zwischen Ereignissen betrachtet werden. Weiterhin werden bestimmte Verteilungen für Zufallsvariablen betrachtet, etwa die Normal (Gauß)-Verteilung, die Binomialverteilung etc.

Die Frage ist nun, wie empirische Daten und die Theorie zufälliger Ereignisse in einem einheitlichen Rahmen behandelt werden können.

Dies leistet die **induktive (schließende) Statistik**, welche für die aus Stichproben berechneten statistischen Maßzahlen (sog. ‘Statistiken’) zugrundeliegende theoretische Größen (Parameter) postuliert.

induktive Statistik

Für die empirischen Größen werden dann Wahrscheinlichkeitsaussagen hergeleitet.

Es findet also ein Rückschluß von den Daten auf die zugrundeliegenden theoretischen Größen (etwa Wahrscheinlichkeiten) statt.

Beispiel 1.1 (Mittelwert und Erwartungswert)**Mittelwert**

Der Mittelwert (Durchschnittswert) ist die wohl bekannteste Statistik. Sie dient zur Schätzung des theoretischen Mittelwerts, der auch als Erwartungswert bezeichnet wird.

Etwa ist die mittlere Körpergröße der Personen einer Stichprobe eine Schätzung des Mittelwerts aller Personen einer Population (etwa der gesamten deutschen Bevölkerung).

**Beispiel 1.2 (Relative Häufigkeit und Wahrscheinlichkeit)**

Befragt man $N = 1000$ Personen nach Ihrem Wahlverhalten bei der nächsten Bundestagswahl, so ist die relative Häufigkeit $f(A) = 26.7\% = 267/1000$ der 267 Personen, die für die A -Partei stimmen wollen, eine Schätzung für die (unbekannte) Wahrscheinlichkeit $P(A)$, daß eine zufällig ausgewählte Person die A -Partei wählen würde (oder die A -Partei tatsächlich einen Stimmenanteil von $P(A)$ erhält).

Definiert man die Zufalls-Variable (Indikator)

$$X_n = \begin{cases} 1 & \text{wenn } A \\ 0 & \text{sonst} \end{cases} \quad (1.1)$$

(1, wenn die n -te Person die A -Partei wählt, 0 sonst), so ist der Mittelwert

$$\bar{X} = \frac{1}{N} \sum_{n=1}^N X_n = f(A) \quad (1.2)$$

relative Häufigkeit

gleich der relativen Häufigkeit und ein Schätzer für die unbekannte Wahrscheinlichkeit $P(A)$.



Die Idee ist also, theoretische Größen, etwa Verteilungen, deren Parameter oder Wahrscheinlichkeiten für bestimmte Ereignisse zu postulieren und deren numerische Größe durch Daten (Stichproben) abzuschätzen. Es ist unmittelbar klar, daß (aus praktischen oder ökonomischen Gründen) die Stichprobe nur ein kleiner Teil der Grundgesamtheit (GG) sein kann.

Definition 1.1 (Grundgesamtheit (GG), Population)**Grundgesamtheit**

Die Grundgesamtheit (Population) Ω ist die Menge aller relevanten Merkmalsträger.

Definition 1.2 (Stichprobe)

Als Stichprobe $S = \{\omega_1, \dots, \omega_N\}$ (Stichprobenumfang N) wird eine Untermenge $S \subset \Omega$ der Grundgesamtheit (Population) Ω bezeichnet.

Stichprobe

Bei einer **reinen (uneingeschränkten) Zufallsauswahl** hat jedes Element $\omega \in \Omega$ die gleiche Chance, ausgewählt zu werden.

Das zufällig ausgewählte Element ω wird nun einem numerischen Merkmalswert $X(\omega)$ zugeordnet und man betrachtet die **Verteilung dieses Merkmals in der Grundgesamtheit**, d.h.

$$F(x) = P(X \leq x). \quad (1.3)$$

Im allgemeinen sind jedoch die Parameter $\mu = E[X]$, $\text{Var}(X) = \sigma^2$ etc. des Merkmals X (Körpergröße, Wahlverhalten, Einkommen etc.) unbekannt und müssen durch \bar{X} , S^2 etc. geschätzt werden.

Die Stichprobe wird durch folgendes wahrscheinlichkeitstheoretisches Modell beschrieben:

$$\begin{aligned} \text{Stichprobenvariablen : } & X_1, \dots, X_N \quad (\text{Zufallsvariablen}) \\ \text{Realisation : } & x_1, \dots, x_N \quad (\text{konkrete Werte}), \end{aligned}$$

**Stichprobenvariablen
Realisation**

wobei alle Stichprobenvariablen X_n genauso wie das Merkmal X (identisch) verteilt sind

$$F(x) = P(X_n \leq x), n = 1, \dots, N. \quad (1.4)$$

Das Merkmal wird also N -fach repliziert.

Im allgemeinen geht man davon aus, daß bei einer uneingeschränkten Zufallsauswahl die Stichprobenvariablen X_n auch voneinander **unabhängig** sind.

Unabhängigkeit

Man spricht dann von einer einfachen Stichprobe.

**einfache
Stichprobe**

Dies ist nicht immer erfüllt, etwa bei Meßwiederholungen der gleichen Person oder bei Zeitreihen, etwa Finanzdaten (Aktienkurse) oder makroökonomischen Variablen, z.B. BIP (Bruttoinlandsprodukt).

In diesem Fall muß die **Abhängigkeit bei Berechnungen berücksichtigt** werden.

Auch beim Ziehen ohne Zurücklegen ist zu beachten, daß die Stichprobenvariablen voneinander abhängig sind, da ja bereits gezogene Einheiten nicht mehr zur Verfügung stehen.

Etwa werden bereits befragte Personen nicht noch einmal interviewt.

Man unterscheidet daher das

- Modell ohne Zurücklegen (M.o.Z.): X_n abhängig
- Modell mit Zurücklegen (M.m.Z.): X_n unabhängig (einfache Stichprobe).

Konkrete Stichproben

Konkrete Stichproben können folgendermaßen gezogen werden:

- Originalverfahren:
Zufallszahlen: Tabellen, Computer: pseudo-zufällig
- Ersatzverfahren:
Buchstabenverfahren, Geburtstagsverfahren: pseudo-zufällig

Zufallszahlen

Übung 1.1 (Zufallszahlen)

- Warum sind auf dem Computer erzeugte Zufallszahlen pseudo-zufällig?
- Wie könnte man 'echte' Zufallszahlen erzeugen?
- Was sind Buchstabenverfahren, Geburtstagsverfahren?



Arten von Grundgesamtheiten

Man kann auch verschiedene **Arten von Grundgesamtheiten (GG)** unterscheiden:

- konkret vorliegende GG: alle Studenten im Hörsaal (alle Beleger des Kurses)
- teilweise konkret vorliegende GG: alle Studenten

Wann sind Umfragen repräsentativ?

Eine Umfrage ist dann repräsentativ, wenn jeder in der so genannten Grundgesamtheit - beim Politbarometer sind das alle Wahlberechtigten - die gleiche oder eine angebbare und damit korrigierbare Chance hat, befragt zu werden. Dies wird gewährleistet durch eine Zufallsauswahl: Für jede Umfrage werden aus allen eingetragenen privaten Telefonanschlüssen zufällig Telefonnummern ausgewählt. Die letzte Ziffer wird dann durch eine zufällig generierte Ziffer ersetzt. Dieses Verfahren (RLD-Verfahren: randomize last digit) gewährleistet, dass auch solche Haushalte, die nicht im Telefonbuch verzeichnet sind, in die Auswahl gelangen können.

Das Zufallsprinzip wird bei der Auswahl der Personen innerhalb eines Haushaltes noch mal angewendet (Last-Birthday-Methode: ausgewählt wird, wer zuletzt Geburtstag hatte). Insgesamt ist diese strenge Zufallsauswahl also der Garant dafür, dass die im Politbarometer ermittelten Ergebnisse repräsentativ sind für die wahlberechtigte Bevölkerung in Deutschland.

Abbildung 1.1: Text auf der homepage des ZDF (Politbarometer).

- hypothetische GG: alle möglichen Ziehungen im Lotto, alle Wirtschaftssubjekte.

Übung 1.2 (Stichprobenziehung in der Praxis)

Quelle:

<http://politbarometer.zdf.de/ZDFde/inhalt/12/0,1872,2337900,00.html>

- Diskutieren Sie die Aussagen in Abb. 1.1.
- Hat jeder Bundesbürger wirklich die gleiche Chance, in die Stichprobe zu gelangen?
- Handelt es sich um eine einfache Stichprobe?



1.2 Stichprobenraum und Stichprobenfunktionen

Hat man eine Stichprobe gezogen, etwa die Körpergröße oder das Einkommen von $N = 10$ Probanden ermittelt, so ergibt sich das Stichprobenergebnis (Realisation)

$$\textit{Realisation} : \quad x_1, \dots, x_N \quad (\text{konkrete Werte}),$$

Dagegen sind die zugeordneten Stichprobenvariablen Zufallsvariablen:

$$\textit{Stichprobenvariablen} : \quad X_1(\omega), \dots, X_N(\omega) \quad (\text{Zufallsvariablen}).$$

Um die Abhängigkeit vom Zufall auszudrücken, wird manchmal das Argument ω (Zufall, Element der Ergebnismenge Ω) der Zufallsvariablen $X_n(\omega)$ explizit angeschrieben.

Definition 1.3 (Stichprobenfunktion, Statistik)

Eine Funktion der Stichprobenvariablen

$$T(\omega) = g(X_1(\omega), \dots, X_N(\omega)) \quad (1.5)$$

heißt Stichprobenfunktion ('Statistik').

Statistik

Die Idee ist, die Stichprobe zu einfachen, aussagefähigen Größen zusammenzufassen.

1.2.1 Mittelwert

Beispiel 1.3 (Mittelwert)

Der Mittelwert

$$\bar{X} = \frac{1}{N} \sum_{n=1}^N X_n \quad (1.6)$$

Mittelwert

ist der Durchschnitt aller Stichprobenvariablen.

Er hat die Eigenschaften:

$$E[\bar{X}] = \frac{1}{N} \sum_{n=1}^N E[X_n] = \frac{1}{N} \sum_{n=1}^N \mu = \mu \quad (1.7)$$

$$\text{Var}[\bar{X}] = \frac{1}{N^2} \sum_{n=1}^N \text{Var}[X_n] = \frac{1}{N^2} \sum_{n=1}^N \sigma^2 = \frac{\sigma^2}{N} \quad (1.8)$$

Die Wichtigkeit dieser Formeln kann gar nicht überschätzt werden:

- Der Erwartungswert von \bar{X} ist gleich μ , also dem Erwartungswert des Merkmals X .

Dies wird als **Erwartungstreue** bezeichnet.

- Die Varianz von \bar{X} ist gleich $\frac{\sigma^2}{N}$, also der Streuung des Merkmals X geteilt durch die Stichprobengröße N .

Dies bedeutet, daß **Mittelwerte eine geringere Schwankung aufweisen** als die Daten selbst.

Durch Mittelung kann also aus zufälligen Daten ein näherungsweise konstanter Wert errechnet werden.

Dies ist ein wichtiges Grundprinzip der Statistik.

- Voraussetzung ist allerdings, daß die Stichprobenvariablen **von-einander unabhängig** sind, da ansonsten $\text{Var}[\bar{X}]$ nicht klein wird oder nur langsamer als $\frac{\sigma^2}{N}$ gegen Null geht.

Die Unabhängigkeit der Stichprobenvariablen bewirkt einen **Mittelungseffekt**, d.h. Schwankungen einer Variablen nach oben werden durch Ausschläge einer anderen Variablen nach unten ausgeglichen.

Korrelierte Stichprobenvariablen**Übung 1.3 (Korrelierte Stichprobenvariablen)**

Wenn die Stichprobenvariablen voneinander abhängig sind, etwa $\text{Cov}(X_i, X_j) = \sigma^2 \rho, i \neq j$, so ergibt sich:

$$\text{Var}(\bar{X}) = \frac{1}{N^2} \sum_{i=1, j=1}^N \text{Cov}(X_i, X_j) \quad (1.9)$$

$$= \frac{1}{N^2} \left(\sum_{i=1}^N \sigma^2 + \sum_{i, j=1, i \neq j}^N \sigma^2 \rho \right) \quad (1.10)$$

$$= \frac{\sigma^2}{N} + \frac{N^2 - N}{N^2} \sigma^2 \rho \quad (1.11)$$

$$= \sigma^2 \left(\frac{1}{N} + \frac{N-1}{N} \rho \right) \quad (1.12)$$

$$\rightarrow \sigma^2 \rho. \quad (1.13)$$

Hierbei wurde die Formel

$$\text{Var}(X + Y) = \text{Var}(X) + 2 \text{Cov}(X, Y) + \text{Var}(Y) \quad (1.14)$$

benutzt (vgl. Kap. 18).

Die Abhängigkeit der Varianz des Mittelwerts von der Stichprobengröße N ist in Abb. 1.2 für die Korrelationen $\rho = -1/19, 0, .25, .5, 1$ aufgetragen (von unten).

Für $\rho \neq 0$ erhält man also in großen Stichproben eine Streuung von $\sigma^2 \rho$ anstatt σ^2/N für $\rho = 0$.

In diesem Fall strebt der Mittelwert nicht gegen den konstanten Wert μ , sondern schwankt stark um diese Größe.

Bei stark korrelierten Daten sind also keine Mittelungseffekte zu verzeichnen. Bei perfekter positiver Korrelation $\rho = 1$ schwankt der Mittelwert genauso stark wie die einzelnen Stichprobenvariablen (mit σ^2).

**Bemerkung:*

Bei negativen Korrelationen ist die Bedingung $\rho \geq -\frac{1}{N-1}$ zu beachten. Dies ist notwendig, damit die Korrelationsmatrix der Stichprobenvariablen positiv semidefinit ist. Dann ist auch $\frac{1}{N} + \frac{N-1}{N} \rho \geq 0$, wie es für eine Varianz notwendig ist.



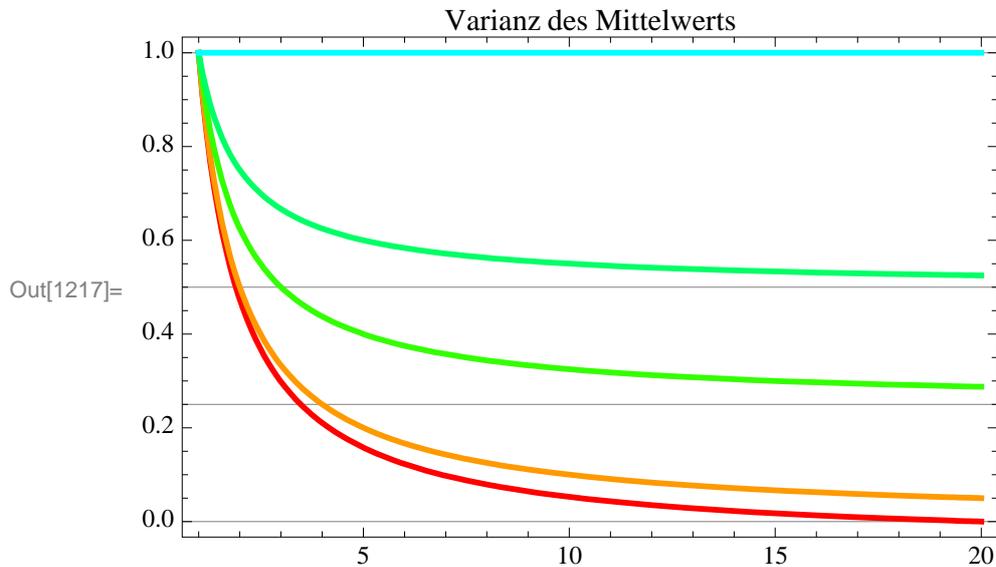


Abbildung 1.2: Varianz des Mittelwerts $\text{Var}[\bar{X}] = \sigma^2(\frac{1}{N} + \frac{N-1}{N}\rho)$ bei korrelierten Daten als Funktion der Stichprobengröße ($\sigma^2 = 1, \rho = -1/19, 0, .25, .5, 1$; von unten). Die Kurven streben gegen $\sigma^2\rho$.

Beispiel 1.4 (Dichotome Variablen (Indikatoren))

Wenn die Stichprobenvariablen nur die Werte 0 und 1 annehmen, also

$$X_n = \begin{cases} 1 & \text{wenn } A \\ 0 & \text{sonst} \end{cases} \quad (1.15)$$

so gilt

$$E[X_n] = \pi = P(A), \text{Var}[X_n] = \pi(1 - \pi) \quad (1.16)$$

(Bernoulli-Variablen $X_n \sim B(1; \pi)$).

Dann ist der Mittelwert (Stichprobenfunktion)

$$\bar{X} = \frac{1}{N} \sum_{n=1}^N X_n \quad (1.17)$$

ein Schätzer für die unbekannte Wahrscheinlichkeit $P(A) = \pi$.

Es gilt für die binomialverteilte Größe $\bar{X} \sim B(N, \pi)$

**Dichotome
Variablen
(Indikatoren)**

**Bernoulli-
Variablen**

Binomialverteilung

$$E[\bar{X}] = \frac{1}{N} \sum_{n=1}^N E[X_n] = \frac{1}{N} \sum_{n=1}^N \pi = \pi \quad (1.18)$$

$$\begin{aligned} \text{Var}[\bar{X}] &= \frac{1}{N^2} \sum_{n=1}^N \text{Var}[X_n] = \frac{1}{N^2} \sum_{n=1}^N \pi(1 - \pi) \\ &= \frac{\pi(1 - \pi)}{N} \end{aligned} \quad (1.19)$$

Befragt man $N = 1000$ Personen nach Ihrem Wahlverhalten bei der nächsten Bundestagswahl, so ergibt sich die relative Häufigkeit

$$\begin{aligned} \bar{x} &= f(A) = p = (1 + 0 + 1 + 1 + \dots + 1 + 0)/1000 \\ &= 26.7\% = 267/1000 \end{aligned}$$

der 267 Personen, die für die A -Partei stimmen wollen (berechnet mit den Realisationen x_1, \dots, x_N : kleines x).

Dies ist eine Schätzung für die (unbekannte) Wahrscheinlichkeit $P(A)$, daß eine zufällig ausgewählte Person die A -Partei wählen würde, oder daß die A -Partei tatsächlich einen Stimmenanteil von $P(A)$ erhält.

Die Stichprobenvarianz $\frac{p(1-p)}{N} = \frac{.267(1-.267)}{1000} = 0.000196$ ist dann eine Schätzung für die Varianz $\frac{\pi(1-\pi)}{N}$ der Stichprobenfunktion \bar{X} .

Die Standardabweichung $\sqrt{\frac{p(1-p)}{N}} = 0.014 = 1.4\%$ ist ein Maß für die Schwankung (Unsicherheit) des geschätzten Prozentwerts $p = .267 = 26.7\%$ der A -Partei.



1.2.2 Stichprobenvarianz

Die quadratische Abweichung vom Mittelwert

Stichprobenvarianz

$$S^2 = \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})^2 \quad (1.20)$$

wird als Stichprobenvarianz bezeichnet.

Die etwas seltsame Division durch $N - 1$ hat den Grund, daß dadurch die Stichprobenvarianz **unverzerrt (erwartungstreu)** ist, d.h. $E[S^2] = \sigma^2$. Dies ist eine wünschenswerte Eigenschaft von Schätzfunktionen.

Hätte man die naheliegende Größe (mittlere quadratische Abweichung; Maximum-Likelihood-Schätzer Kap. 2.3.3)

$$S_*^2 = \frac{1}{N} \sum_{n=1}^N (X_n - \bar{X})^2 = \frac{N-1}{N} S^2 \quad (1.21)$$

definiert, so gilt $E[S_*^2] = \frac{N-1}{N} E[S^2] = \frac{N-1}{N} \sigma^2 < \sigma^2$. Dieser Schätzer ist also verzerrt.

Beispiel 1.5 (Erwartungswert von S^2)

Mit Hilfe des Verschiebungssatzes

$$\sum_{n=1}^N (X_n - c)^2 = \sum_{n=1}^N (X_n - \bar{X})^2 + N(\bar{X} - c)^2 \quad (1.22)$$

Verschiebungssatz

kann man schreiben ($c = \mu$):

$$\sum_{n=1}^N (X_n - \bar{X})^2 = \sum_{n=1}^N (X_n - \mu)^2 - N(\bar{X} - \mu)^2. \quad (1.23)$$

Daher gilt

$$E\left[\sum_{n=1}^N (X_n - \bar{X})^2\right] = \sum_{n=1}^N E(X_n - \mu)^2 - NE(\bar{X} - \mu)^2 \quad (1.24)$$

$$= N\sigma^2 - N(\sigma^2/N) = (N-1)\sigma^2 \quad (1.25)$$

also $E[S^2] = \sigma^2$ (Erwartungstreue).



1.2.3 Standardabweichung

Zieht man die Wurzel aus S^2 , so ergibt sich die **Standardabweichung**

Standardabweichung

$$S = \sqrt{\frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})^2}. \quad (1.26)$$

Die Standardabweichung hat die gleiche Dimension wie der Mittelwert (etwa cm, Euro etc.) und dient zur Angabe von Intervallen, etwa $\bar{X} \pm S$ (Mittelwert plus/minus eine Standardabweichung). Dieser Bereich überdeckt in etwa 2/3 der Fälle den unbekanntem Erwartungswert μ .

Beispiel 1.6 (Verzerrung von S)

Interessant ist, daß S verzerrt ist, obwohl S^2 als erwartungstreue Größe definiert wurde.

Es muß nämlich für eine Varianz immer gelten

$$\text{Var}(S) = E[S^2] - (E[S])^2 = \sigma^2 - (E[S])^2 \geq 0 \quad (1.27)$$

oder

$$E[S] \leq \sigma. \quad (1.28)$$

Hierbei wurde die Rechenregel

$$E[X] = \mu \quad (1.29)$$

$$\text{Var}(X) = E[X - \mu]^2 = E[X^2] - \mu^2 \geq 0 \quad (1.30)$$

für Varianzen benutzt.



1.3 Stichprobenverteilungen und Testverteilungen

Im vorigen Abschnitt wurden bestimmte Stichprobenfunktionen (Statistiken)

$$T(\omega) = g(X_1(\omega), \dots, X_N(\omega)), \quad (1.31)$$

etwa der Mittelwert

$$\bar{X} = \frac{1}{N} \sum_{n=1}^N X_n \quad (1.32)$$

oder die Stichprobenvarianz

$$S^2 = \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})^2 \quad (1.33)$$

eingeführt.

Diese dienen der Zusammenfassung der Stichprobe und zur Abschätzung unbekannter Parameter, etwa dem Erwartungswert oder der Varianz eines Merkmals X .

Zur statistischen Beurteilung dieser Zufallsvariablen T interessiert man sich für deren Erwartungswerte $E[T]$, Varianzen $\text{Var}[T]$ oder allgemeiner für deren Wahrscheinlichkeitsverteilungen

$$F(t) = P\{T \leq t\} = P\{\omega \mid T(\omega) \leq t\}. \quad (1.34)$$

**Stichproben-
verteilungen**

Insbesondere bei statistischen Tests (etwa Prüfung der Hypothese $\mu = 2600$ Euro = mittleres Einkommen) sind sogenannte Testverteilungen notwendig, die Wahrscheinlichkeitsaussagen über T erlauben.

1.3.1 Die Normal (Gauß)-Verteilung

Die Normal (Gauß)-Verteilung ist die wohl wichtigste Wahrscheinlichkeitsverteilung, die in unzähligen Kontexten auftritt. Addiert man sehr viele Zufallsgrößen, so ist die Verteilung des Summenwerts ungefähr normalverteilt (zentraler Grenzwertsatz).

Empirische Größen resultieren oft aus der Ansammlung vieler Einflußfaktoren und sind daher approximativ normalverteilt. Auch Teststatistiken ergeben sich oft als Summen von Stichprobenvariablen.

Die Dichtefunktion einer standardnormalverteilten Zufallsvariablen $Z \sim N(0, 1)$ mit Erwartungswert $E[Z] = 0$ und Varianz $\text{Var}[Z] = 1$ ist in Abb. 1.3 dargestellt.

**Normal
(Gauß)-Verteilung**

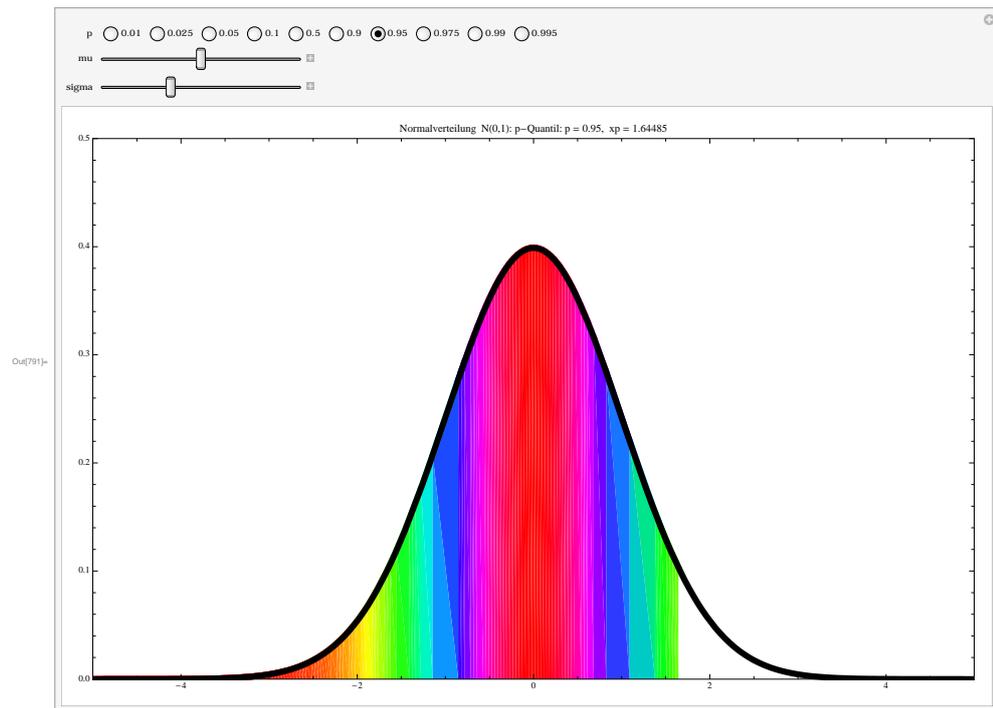


Abbildung 1.3: Dichtefunktion der $N(0, 1)$ -Zufallsvariable Z mit einem .95-Quantil $z(.95) = 1.645$. Die schraffierte Fläche ist die Wahrscheinlichkeit $p = .95 = P(-\infty < Z \leq z(.95))$. *Mathematica-Applets finden Sie auf der Homepage des Lehrstuhls:*

http://www.fernuni-hagen.de/ls_statistik/lehre/

Die Fläche unter der Dichtefunktion $f(z)$ gibt die Wahrscheinlichkeit an, daß die Zufallsvariable Z in ein bestimmtes Intervall $A = [a, b]$ fällt, also

$$P(A) = \int_a^b f(z) dz = \int_A f(z) dz. \quad (1.35)$$

Definition 1.4 (Normalverteilung)

Die Dichtefunktion in der Abbildung ist durch die **Gauß-Glocke**

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} := \phi(z) \quad (1.36)$$

gegeben.

Die Fläche unter der Kurve im Intervall $A = (-\infty, z]$, d.h.

$$F(z) = \int_{-\infty}^z f(t) dt \quad (1.37)$$

$$= \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \quad (1.38)$$

$$:= \Phi(z) = P(-\infty < Z \leq z) \quad (1.39)$$

ist die sogenannte **kumulierte Verteilungsfunktion** $F(z)$. Sie wird im Fall der Normalverteilung oft als $\Phi(z)$ bezeichnet.

**kumulierte
Verteilungsfunktion**

Gibt man sich eine Wahrscheinlichkeit p vor, so gilt für das $z(p)$ -Quantil die Bedingung

$$\Phi(z(p)) = p, \quad (1.40)$$

d.h. man sucht einen Wert $z(p)$ auf der z -Achse, sodaß die Fläche bis zu diesem Wert gleich p ist.

Quantil

Wichtige z -Quantile sind:

$$z(.95) = 1.65, z(.975) = 1.96, z(.99) = 2.33, z(.995) = 2.58.$$

Sie ergeben sich bei Tests mit den Niveaus $\alpha = .05, .01$ aus den Formeln $p = 1 - \alpha$ bzw. $p = 1 - \alpha/2$ (siehe später).

Übung 1.4 (z -Quantile)

Prüfen Sie dies bitte anhand des Applets bzw. einer Normalverteilungstabelle (Anhang) nach.

Sie können das Applet von der Homepage des Lehrstuhls herunterladen, sollten jedoch auch in der Lage sein, die Quantile aus Tabellen zu entnehmen (Klausur).



1.3.2 Die Chi-Quadrat-Verteilung

1.3.2.1 Definition

Stichprobenvarianzen involvieren Quadrate von Stichprobenvariablen. Sind diese normalverteilt, so führt dies auf folgende Quadratsumme:

Definition 1.5 (Chi-Quadrat (χ^2)-Verteilung)**Annahme:**

Die Stichprobenvariablen X_1, \dots, X_N seien unabhängig und jeweils $N(0, 1)$ -standardnormalverteilt.

Dann heißt die Summenvariable

 χ^2 -verteilt

$$\chi^2 = \sum_{n=1}^N X_n^2 \quad (1.41)$$

Chi-Quadrat (χ^2)-verteilt mit N Freiheitsgraden: $\chi^2(N)$.

Bezeichnung:

Freiheitsgrade

Der Parameter N heißt Zahl der Freiheitsgrade (englisch: df = degrees of freedom).

Meistens wird die Zahl der Freiheitsgrade in die Bezeichnung der Zufallsvariablen mit aufgenommen, d.h. $\chi^2(N)$.

Man interessiert sich vor allem für das **p -Quantil der Verteilung**, also einen Wert $\chi^2(p, N)$ auf der x -Achse (Abszisse), unterhalb dem sich $100p\%$, etwa 95% der Wahrscheinlichkeit befindet.

Man schreibt

$$P\{\chi^2(N) \leq \chi^2(p, N)\} = p, \quad (1.42)$$

wobei dies mißverständlich ist:

Sowohl das Quantil als auch die Zufallsvariable selbst wird mit dem gleichen Buchstaben (χ^2) bezeichnet.

In Abb. 1.4 ist die Dichtefunktion der $\chi^2(N)$ -Zufallsvariable dargestellt. Das $\chi^2(.95, 6)$ -Quantil = 12.5916 ist in der Graphik angegeben.

Für die $\chi^2(N)$ -Variable gelten folgende Eigenschaften:

1. $E[\chi^2(N)] = N$
2. $\text{Var}[\chi^2(N)] = 2N$
3. $Z = \frac{\chi^2(N) - N}{\sqrt{2N}} \stackrel{a}{\approx} N(0, 1)$

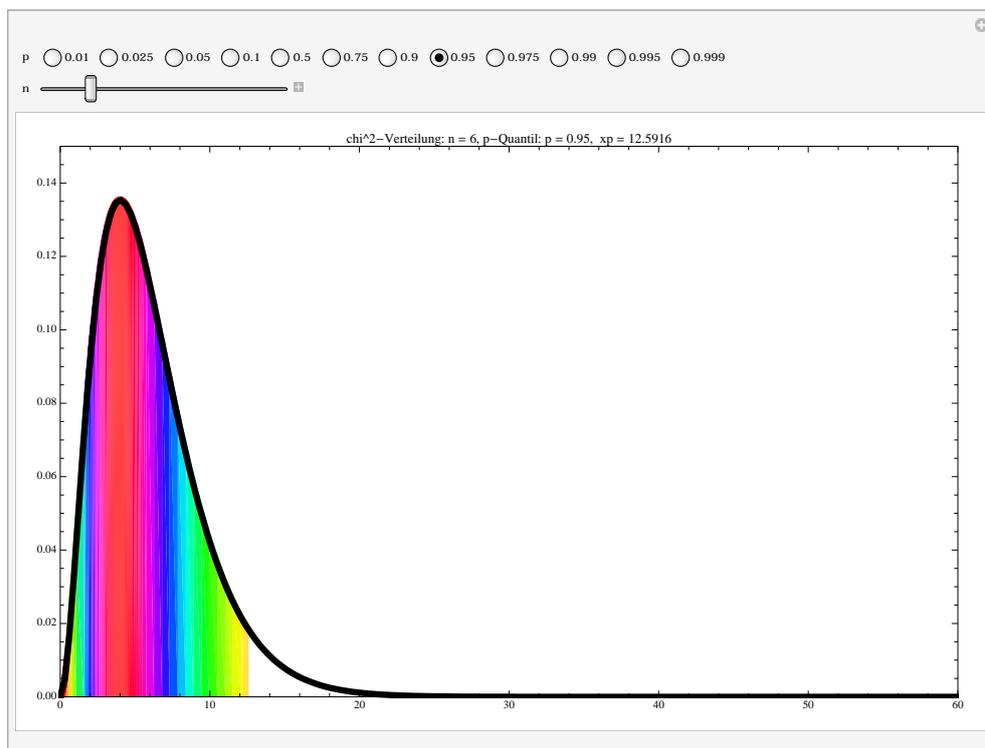


Abbildung 1.4: Dichtefunktion der $\chi^2(N)$ -Zufallsvariable mit einem $\chi^2(.95, 6)$ -Quantil.

Übung 1.5 (Eigenschaften)

Veranschaulichen Sie sich diese Eigenschaften bitte an Abb. 1.4.

Die Freiheitsgrade sind $N = 6$ und der Erwartungswert liegt bei $E[\chi^2(6)] = 6$. Der Erwartungswert stimmt nicht mit dem Maximum der Kurve überein (sie ist schief).

Die Varianz ist $\text{Var}[\chi^2(6)] = 12$, d.h. die Standardabweichung ist $\sqrt{12} = 3.4641$. Dies stimmt ungefähr mit der Intuition überein (zeichnen Sie einen Bereich 6 ± 3.4641 in die Graphik ein).

Die dritte Eigenschaft bedeutet, daß man für großes N die χ^2 -Variable nach einer Transformation als normalverteilte Größe ausdrücken kann ($\stackrel{a}{\sim}$ = asymptotisch verteilt als).



1.3.2.2 Anwendung

Nimmt man an, daß die Stichprobenvariablen $X_n \sim N(\mu, \sigma^2)$ normalverteilt sind, so kann man durch **Standardisierung**

Standardisierung

$$Z_n := \frac{X_n - \mu}{\sigma} \sim N(0, 1) \quad (1.43)$$

die oben definierte $\chi^2(N)$ -Variable verwenden:

1.

$$\frac{1}{\sigma^2} \sum_{n=1}^N (X_n - \mu)^2 = \sum_{n=1}^N \left(\frac{X_n - \mu}{\sigma} \right)^2 = \chi^2(N) \quad (1.44)$$

2.

$$\frac{1}{\sigma^2} \sum_{n=1}^N (X_n - \bar{X})^2 = \sum_{n=1}^N \left(\frac{X_n - \bar{X}}{\sigma} \right)^2 = \chi^2(N - 1) \quad (1.45)$$

Aufgrund der geschätzten Größe \bar{X} sind die einzelnen Terme nicht unabhängig und man hat einen Freiheitsgrad weniger. Der Beweis ist in Rao, 1973, Kap. 3, Abs. 3b.3 zu finden.

Aus der Eigenschaft

$$E[\chi^2(N - 1)] = N - 1 = \frac{1}{\sigma^2} E\left[\sum_{n=1}^N (X_n - \bar{X})^2\right]$$

ergibt sich die Erwartungstreue der Stichprobenvarianz S^2 .

3. Für $N > 30$ kann die Formel

$$z(p) = \sqrt{2\chi^2(N, p)} - \sqrt{2N - 1}$$

zur approximativen Berechnung des Quantils herangezogen werden ($z(p)$ nachschlagen und nach $\chi^2(N, p)$ auflösen).

Übung: $p = 0.95$: berechnen Sie $z(p)$ und $\chi^2(N, p)$.

Stichprobenvarianzen sind χ^2 -verteilt,
wenn die Stichprobenvariablen normalverteilt sind.

1.3.3 Die t -Verteilung

1.3.3.1 Definition

Bei Testproblemen für Mittelwerte ergeben sich Größen der Form

$$\frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{N}}}, \quad (1.46)$$

oder, wenn σ nicht bekannt ist und geschätzt werden muß

$$\frac{\bar{X} - \mu}{\sqrt{\frac{S^2}{N}}}. \quad (1.47)$$

Daher betrachtet man Quotienten folgender Form:

Definition 1.6 (t -Verteilung)

Man nennt die Zufallvariable

$$T = \frac{X}{\sqrt{\frac{Y}{N}}}, \quad (1.48)$$

t -Verteilung

t -verteilt mit N Freiheitsgraden (df), wobei

- $X \sim N(0, 1)$
- $Y \sim \chi^2(N)$.

Bezeichnung: $T \sim t(N)$.

Die Dichtefunktion ist in Abb. 1.5 dargestellt

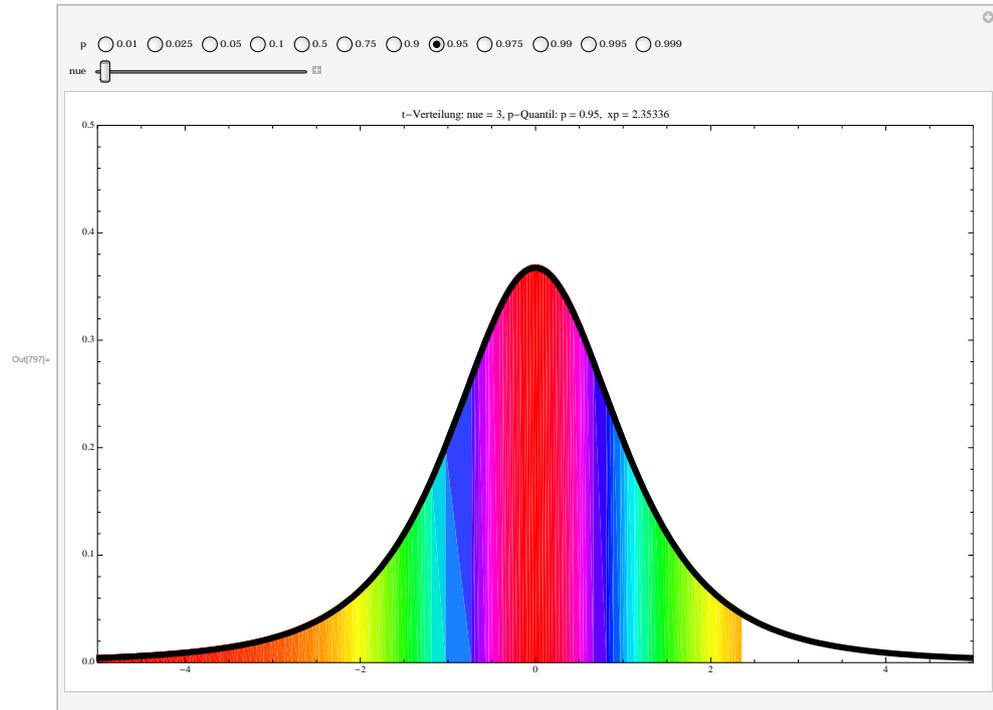


Abbildung 1.5: Dichtefunktion $f_T(t)$ der $T(N)$ -Zufallsvariable mit einem $t(0.95, 3)$ -Quantil.

1.3.3.2 Eigenschaften

1. $E[T] = 0; N > 1$
2. $\text{Var}[T] = \frac{N}{N-2}; N > 2$
3. $T \rightarrow Z$ für $N \rightarrow \infty$, wobei $Z \sim N(0, 1)$.

Für $N > 30$ kann die Standardnormalverteilung benutzt werden.

Die in der Abbildung gezeigte Kurve weicht stark von der Normalverteilung ab ($N = 3$).

Übung: Worin genau besteht die Abweichung in der Form ?

1.3.3.3 Anwendung: t -Test

Oft ist in Anwendungen die Varianz der Stichprobenvariablen nicht bekannt und muß daher geschätzt werden.

Dann ergeben sich durch Standardisierung Größen der Form

$$\frac{\bar{X} - \mu}{\sqrt{\frac{S^2}{N}}} \sim t(N - 1) \quad (1.49)$$

Ersetzt man S durch σ , so wäre dies standardnormalverteilt $N(0, 1)$, da $\text{Var}(\bar{X}) = \frac{\sigma^2}{N}$.

Nun wird aber eine normalverteilte Zufallsvariable durch die Wurzel einer $\chi^2(N - 1)$ -verteilten Zufallsvariable geteilt. Die resultierende Zufallsvariable ist dann Student- $t(N - 1)$ -verteilt.¹

Es gilt nämlich folgende Umrechnung:

$$\frac{\bar{X} - \mu}{\sqrt{\frac{S^2}{N}}} = \frac{(\bar{X} - \mu)/(\sigma/\sqrt{N})}{\sqrt{\frac{S^2}{N} \frac{N}{\sigma^2}}} \quad (1.50)$$

$$= \frac{Z}{\sqrt{\frac{(N-1)S^2}{\sigma^2} \frac{1}{N-1}}} \quad (1.51)$$

$$= \frac{Z}{\sqrt{\frac{Y}{N-1}}} \quad (1.52)$$

$$\sim t(N - 1) \quad (1.53)$$

$$Y = \frac{(N - 1)S^2}{\sigma^2} \sim \chi^2(N - 1) \quad (1.54)$$

Für großes N strebt S gegen σ und der Quotient gegen eine normalverteilte Größe.

1.3.4 Die F -Verteilung

1.3.4.1 Definition

Bei Testproblemen für Varianzen ergeben sich Größen der Gestalt

$$\frac{S_M^2}{S_N^2}, \quad (1.55)$$

¹Die Verteilung wurde unter dem Pseudonym *Student* publiziert: Student, The Probable Error of a Mean, Biometrika, Vol. 6, No. 1 (March 1908), pp. 1-25.

<http://www.jstor.org/pss/2331554>

**Student- t -
Verteilung**

Daher betrachtet man Quotienten folgender Form:

Definition 1.7 (*F*-Verteilung)

Man nennt die Zufallsvariable

$$F = \frac{X/m}{Y/n}, \quad (1.56)$$

***F*-Verteilung**

F-verteilt mit Freiheitsgraden (df) m und n (Zähler- und Nenner-Freiheitsgrade dürfen nicht vertauscht werden!), wobei

- $X \sim \chi^2(m)$.
- $Y \sim \chi^2(n)$.

Bezeichnung: $F \sim F(m, n)$.

Die Dichtefunktion ist in Abb. 1.6 dargestellt

1.3.4.2 Eigenschaften

1. $E[F] = \frac{n}{n-2}; n \geq 3$
2. $\text{Var}[F] = \frac{2n^2(n+m-2)}{m(n-4)(n-2)^2}; n \geq 5$
3. Quantile $f(p, m, n) = 1/f(1-p, n, m)$
4. $F(1, n) = \frac{\chi^2(1)}{\chi^2(n)/n} = T(n)^2$
da $T(n) = Z/\sqrt{\chi^2(n)/n}$ und $Z^2 \sim \chi^2(1)$.
Etwa ist $\left(\frac{\bar{X}-\mu}{S/\sqrt{n}}\right)^2 \sim F(1, n-1)$ -verteilt.

1.3.4.3 Anwendung: *F*-Test

Möchte man Varianzen σ_1^2, σ_2^2 auf Gleichheit testen, so ergibt sich der Quotient (Stichprobengröße N, M)

$$\frac{(S_2^2/\sigma_2^2)}{(S_1^2/\sigma_1^2)}, \quad (1.57)$$

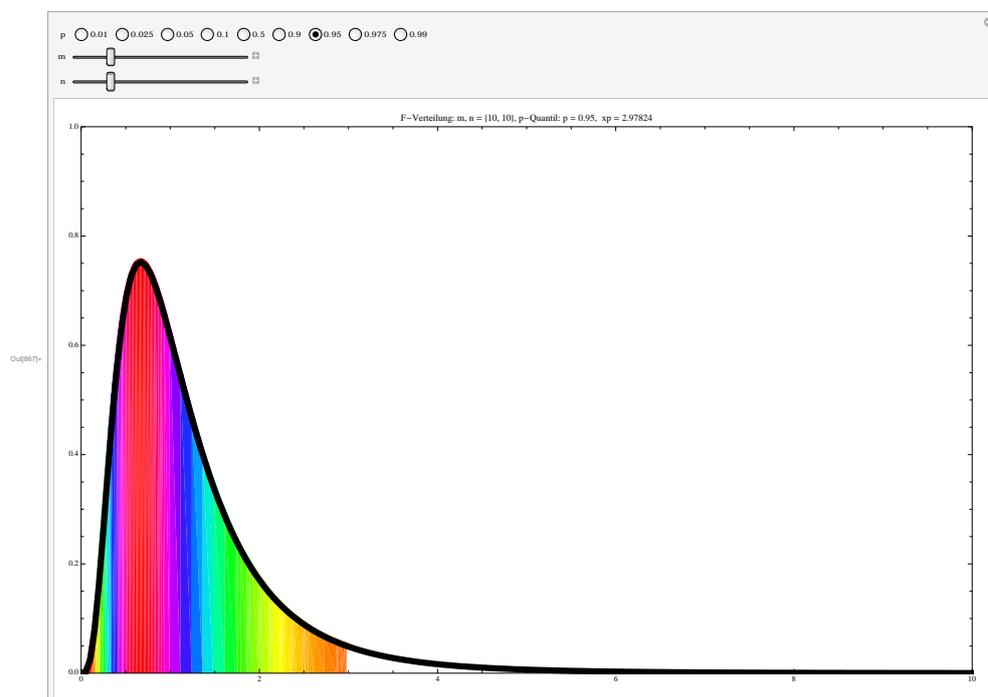


Abbildung 1.6: Dichtefunktion $f_F(f)$ der $F(m, n)$ -Zufallsvariable mit einem $f(.95, 10, 10)$ -Quantil.

der F -verteilt ist mit Freiheitsgraden $M - 1, N - 1$.

Kapitel 2

Punktschätzung, Intervallschätzung, Hypothesentests

2.1 Punktschätzung

Ziel der Punktschätzung ist die Abschätzung unbekannter Parameter θ (Punkte) der Merkmalsverteilung mit Hilfe von Daten aus einer Stichprobe.

Punktschätzung

Beispiele:

- Stimmenanteil der A -Partei bei der nächsten Wahl: $\pi = P(A)$
- Lebenserwartung von Rauchern: $\tau = E[T]$
($T =$ Zufallsvariable der Lebensdauer; $\tau =$ mittlere Lebenserwartung)
- Zusammenhang $\rho = \text{Corr}(X, Y)$ zwischen Abiturnote X und Studienerfolg Y
- Rendite μ und Risiko σ einer Geldanlage (Erwartungswert und Standardabweichung der zufälligen Rendite R)

Mit Hilfe einer geeigneten Stichprobenfunktion g kann ein Schätzwert sowie ein Schätzer für den gesuchten Parameter θ berechnet werden:

Schätzwert

$$\text{Schätzwert: } \hat{\theta} = g(x_1, \dots, x_N) \quad (2.1)$$

(konkreter Wert)

Schätzer

$$\text{Schätzer: } \hat{\Theta}(\omega) = g(X_1(\omega), \dots, X_N(\omega)) \quad (2.2)$$

(Zufallsvariable).

Beim Schätzwert wurde das Stichprobenergebnis

$$x_1, \dots, x_N \quad (\text{konkrete Werte})$$

in die Stichprobenfunktion eingesetzt.

Dagegen ist der Schätzer $\hat{\Theta}(\omega)$ eine Zufallsgröße (groß geschrieben), da die eingesetzten Stichprobenvariablen Zufallsvariablen sind:

$$X_1(\omega), \dots, X_N(\omega) \quad (\text{Zufallsvariablen}).$$

$\hat{\theta}$ ist also eine Realisation von $\hat{\Theta}$

Mit *einer* Stichprobe erhält man nur einen einzigen konkreten Schätzwert $\hat{\theta} = g(x_1, \dots, x_N)$.

Wie kann also ein Schätzer zufällig schwanken?

Man stellt sich am besten vor, daß M Stichproben gezogen wurden. Die Ergebnisse werden als x_{mn} , $m = 1, \dots, M$; $n = 1, \dots, N$ notiert und in einer Tabelle eingetragen (Tab. 2.1).

Die Schätzungen $\hat{\theta}_1, \dots, \hat{\theta}_M$ sind dann ein Maß für die Variabilität des Schätzers $\hat{\Theta}$.

1. Stichprobe	x_{11}	x_{12}	\dots	x_{1N}	$\hat{\theta}_1$
2. Stichprobe	x_{21}	x_{22}	\dots	x_{2N}	$\hat{\theta}_2$
\vdots	\vdots	\vdots	\dots	\dots	\vdots
M -te Stichprobe	x_{M1}	x_{M2}	\dots	x_{MN}	$\hat{\theta}_M$

Tabelle 2.1: Daten in M Stichproben.

Etwa kann man die Stichproben-Varianz der Schätzungen $\hat{\theta}_1, \dots, \hat{\theta}_M$

$$s^2(\hat{\theta}) = \frac{1}{M-1} \sum_{m=1}^M (\hat{\theta}_m - \bar{\hat{\theta}})^2 \quad (2.3)$$

$$\bar{\hat{\theta}} = \frac{1}{M} \sum_m \hat{\theta}_m \quad (2.4)$$

ausrechnen und so eine Schätzung für $\text{Var}(\hat{\Theta})$ gewinnen. Informativ ist auch die Berechnung eines Histogramms aus den Schätzungen.

In der Praxis zieht man jedoch meist nur eine Stichprobe.

Beispiel 2.1 (Simulierte Daten)

Es wurden $M = 20$ Datensätze der Größe $N = 10$ aus der Normalverteilung $N(\mu = 10, \sigma^2 = 1)$ gezogen (Abb. 2.1).

Die Mittelwerte $\bar{x}_m; m = 1, \dots, M$ (letzte Spalte) sind Schätzungen des Erwartungswerts $\mu = 10$, der in Anwendungen unbekannt ist. In jeder Stichprobe ergibt sich ein anderer Mittelwert.

Diese Mittelwerte schwanken jedoch nicht so stark wie die Originaldaten. Es gilt $E(\bar{X}_m) = \mu$ und $\text{Var}(\bar{X}_m) = \sigma^2/N = 1/10$ für jede der M Stichproben.

Die empirischen Mittelwerte sind $\bar{\bar{x}} = 9.991$ und $s^2(\bar{\bar{x}}) = 0.1426$ (Mittelwert der Mittelwerte und Stichprobenvarianz der Mittelwerte).

In Abb. 2.2 ist das Histogramm der $M = 20$ Mittelwerte mit der theoretischen Verteilung $\bar{X}_m \sim N(\mu, \sigma^2/N) = N(10, 1/10)$ (oben) sowie das Histogramm der $MN = 200$ Datenpunkte mit der theoretischen Verteilung $X_{mn} \sim N(\mu, \sigma^2) = N(10, 1)$ aufgetragen (unten).

■

	1	2	3	4	5	6	7	8	9	10	thetadach
1	12.22	10.10	9.57	11.21	10.04	10.11	9.96	9.48	10.28	8.86	10.18
2	8.78	8.73	8.78	10.12	10.07	10.01	10.24	9.06	12.85	9.67	9.83
3	10.02	8.82	11.26	10.36	10.36	10.43	11.99	11.69	10.49	8.67	10.41
4	8.02	9.85	9.43	9.56	10.15	9.73	8.59	9.71	10.09	11.18	9.63
5	10.25	9.16	9.68	10.08	9.73	9.46	10.42	9.66	10.14	10.78	9.94
6	9.75	10.78	10.06	12.44	9.09	8.58	9.43	9.15	12.06	8.44	9.98
7	8.76	10.01	10.27	10.71	10.26	10.31	10.15	9.35	9.21	8.72	9.78
8	12.07	10.50	9.41	11.70	11.41	9.58	8.94	11.19	9.95	9.58	10.43
9	10.68	9.58	11.77	11.35	9.25	9.38	9.81	9.71	9.25	10.87	10.17
10	10.13	12.46	10.95	9.56	10.49	9.37	10.27	8.67	10.17	9.09	10.12
11	11.49	9.78	11.47	10.76	10.57	11.21	10.85	11.41	10.69	10.57	10.88
12	8.97	10.01	10.28	10.34	8.13	10.95	8.02	9.61	10.51	10.61	9.74
13	9.77	9.69	9.14	10.25	10.45	9.64	9.93	8.98	11.38	11.68	10.09
14	11.02	8.34	9.21	9.33	8.75	10.19	8.58	7.78	9.62	10.26	9.31
15	10.46	11.78	11.22	10.14	10.53	10.26	8.97	7.97	10.66	9.59	10.16
16	8.74	11.02	10.00	9.46	8.11	9.66	9.34	9.86	9.07	11.53	9.68
17	9.46	8.34	10.49	10.70	9.12	10.81	9.15	8.79	10.06	12.13	9.90
18	10.16	11.23	10.73	9.53	9.23	11.34	8.91	11.56	11.11	10.65	10.44
19	10.60	11.00	8.72	9.52	8.24	8.56	8.56	8.80	10.27	10.13	9.44
20	9.31	9.48	11.04	9.92	11.27	8.19	9.71	9.93	8.38	9.80	9.70

Abbildung 2.1: $M = 20$ Datensätze der Größe $N = 10$ aus der Normalverteilung $N(10, 1)$ und zeilenweise Mittelwerte $\hat{\theta}_m = \bar{x}_m; m = 1, \dots, M$.

Beispiel 2.2 (Weitere Beispiele für Punktschätzer)

- $\hat{\mu} = \bar{X} = \frac{1}{N} \sum_{n=1}^N X_n$ (Mittelwert)
- $\hat{\pi} = \bar{X} = \frac{1}{N} \sum_{n=1}^N X_n$ (dichotome Variablen (0,1), Indikatoren)
- $S^2 = \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})^2$ (Stichproben-Varianz)
- $X_{med} = \text{Median} = \begin{cases} X_{(\frac{N+1}{2})} & N \text{ ungerade} \\ \frac{1}{2}[X_{(\frac{N}{2})} + X_{(\frac{N}{2}+1)}] & N \text{ gerade} \end{cases}$

$X_{(n)}$: n -ter Datenpunkt in der geordneten Liste.

- $R(X, Y) = \frac{S(X, Y)}{S(X)S(Y)}$ = Korrelationskoeffizient
 $S(X, Y) = \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})(Y_n - \bar{Y})$ = Stichproben-Kovarianz
 $S(X)^2 = \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})^2 = S(X, X)$ = Stichproben-Varianz
- Multinomialverteilung $M(N, (\pi_1, \dots, \pi_k))$
 (k Kategorien A_1, \dots, A_k)
 Spezialfall: Binomialverteilung $B(N, \pi) = M(N, (\pi, 1 - \pi))$

Stichproben-Variablen $X_{nj}, n = 1, \dots, N, j = 1, \dots, k$

$$X_{nj} = \begin{cases} 1 & \text{falls Kategorie } A_j \\ 0 & \text{sonst} \end{cases}$$

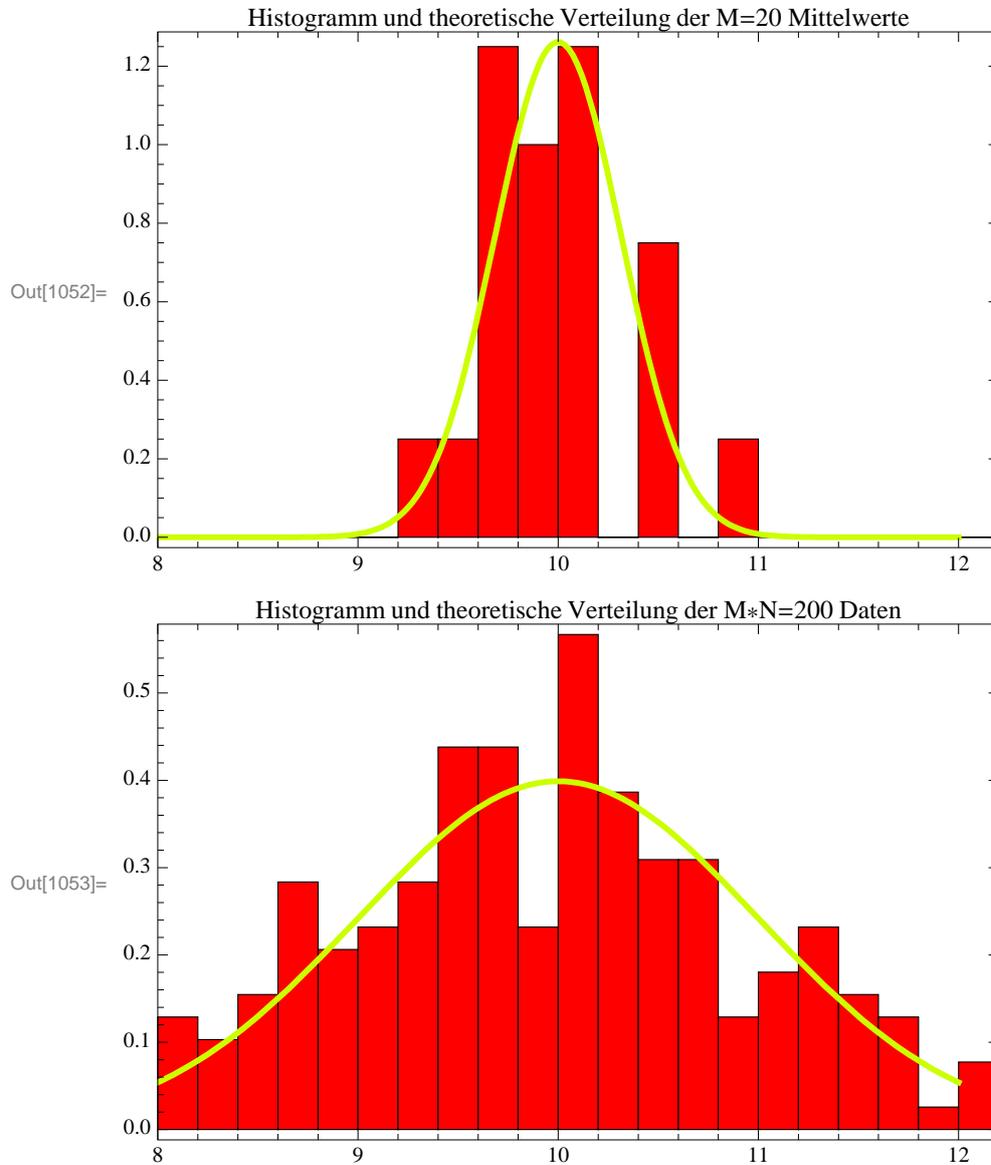


Abbildung 2.2:

Histogramm der $M = 20$ Mittelwerte \bar{x}_m mit der theoretischen Verteilung $\bar{X}_m \sim N(\mu, \sigma^2/N) = N(10, 1/10)$ (oben).

Histogramm der $MN = 200$ Datenpunkte x_{mn} mit der theoretischen Verteilung $X_{mn} \sim N(\mu, \sigma^2) = N(10, 1)$ (unten).

(Person n wählt Kategorie A_j)

$$E[X_{nj}] = \pi_j$$

$X_j = \sum_n X_{nj}$ (Zahl der Personen, die Kategorie A_j wählen)

$\hat{\pi}_j = \frac{X_j}{N}$ (relative Häufigkeit der Kategorie A_j)

$$E[\hat{\pi}_j] = \pi_j$$

$$\text{Var}[\hat{\pi}_j] = \frac{1}{N} \pi_j (1 - \pi_j)$$

Kovarianz der relativen Häufigkeiten:

$$\begin{aligned} \text{Cov}[\hat{\pi}_j, \hat{\pi}_{j'}] &= \frac{1}{N^2} \sum_n \text{Cov}[X_{nj}, X_{nj'}] \\ &= \frac{1}{N^2} \sum_n (E[X_{nj} X_{nj'}] - E[X_{nj}] E[X_{nj'}]) \\ &= \frac{1}{N^2} \sum_n (\pi_j \delta_{jj'} - \pi_j \pi_{j'}) \\ &= \frac{1}{N} (\pi_j \delta_{jj'} - \pi_j \pi_{j'}), \end{aligned}$$

**Kronecker
Delta-Symbol**

$$\delta_{jj'} = \begin{cases} 1 & j = j' \\ 0 & \text{sonst} \end{cases}$$

(Kronecker Delta-Symbol).

Hierbei wurde die Unabhängigkeit der Stichprobenvariablen für $n \neq n'$ sowie die Rechenregel

$$E[X] = \mu \tag{2.5}$$

$$\text{Var}(X) = E[X - \mu]^2 = E[X^2] - \mu^2 \geq 0 \tag{2.6}$$

für Varianzen benutzt.



Beispiel 2.3 (Multinomialverteilung für Wahlumfrage)

Quelle (siehe Abb. 2.3):

<http://politbarometer.zdf.de/ZDFde/inhalt/17/0,1872,7550801,00.html?dr=1>

Zur Modellierung der Daten dient eine Multinomialverteilung $M(N, (\pi_1, \dots, \pi_k))$ mit $k = 6$ Parteien.

- Kategorien:

$$A_1 = \text{CDU/CSU}, A_2 = \text{SPD}, A_3 = \text{FDP},$$

$$A_4 = \text{Grüne}, A_5 = \text{Linkspartei}, A_6 = \text{Andere}$$

- Stichproben-Vektor (Realisation):

$$x_{nj} = (0, 0, 0, 0, 0, 1): \text{ Person } n \text{ wählt Andere } (x_{n6} = 1)$$

$$x_6 = \sum_n x_{n6} \text{ (Zahl der Personen, die Andere wählen)}$$

$$N = 1245 \text{ (siehe Abb. 2.3, unten)}$$

$$(x_1, \dots, x_6) = (498., 311.25, 174.3, 99.6, 112.05, 49.8)$$

$$\approx (498, 311, 174, 100, 112, 50) \text{ (Prozentwerte sind gerundet)}$$

- Prozentwerte

$$p_j = \hat{\pi}_j = \frac{x_j}{N} = (40, 25, 14, 8, 9, 4)\%$$

$$= (0.4, 0.25, 0.14, 0.08, 0.09, 0.04)$$

$$\text{(relative Häufigkeit der Parteien; } \sum_j \hat{\pi}_j = 100\% = 1)$$

- Standardfehler

$$\text{Var}[\hat{\pi}_j] = \frac{1}{N} \pi_j (1 - \pi_j)$$

Schätzung:

$$\widehat{\text{Var}}[\hat{\pi}_j] = \frac{1}{N} p_j (1 - p_j)$$

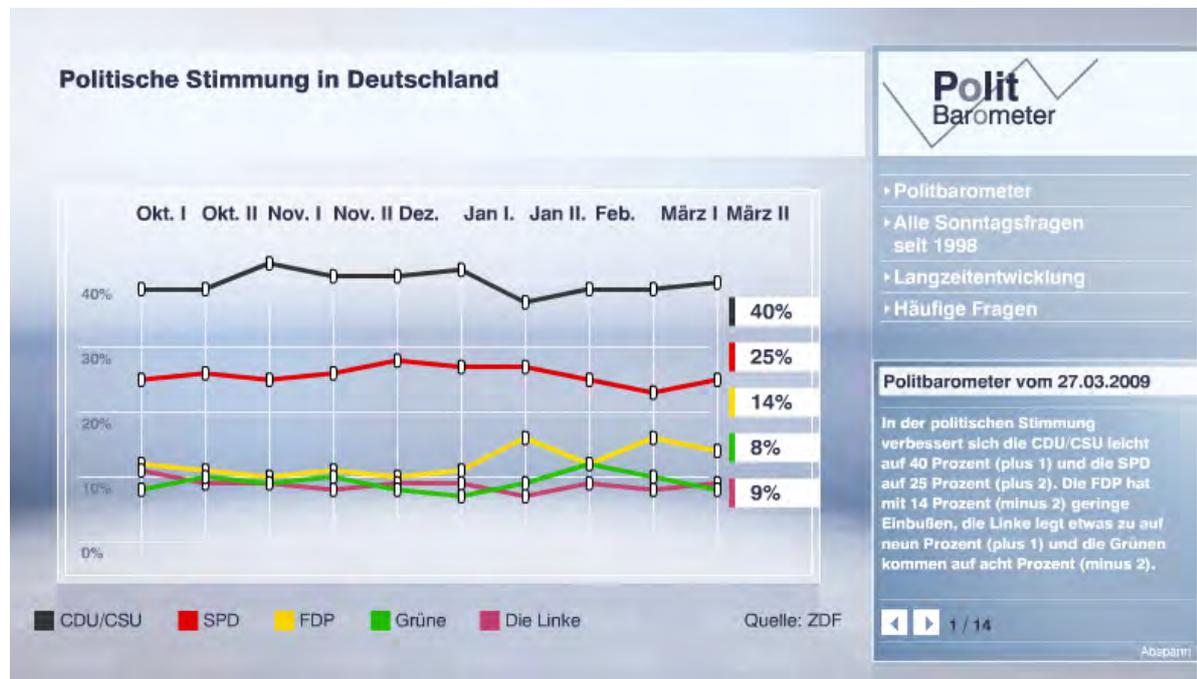
2 Standardabweichungen (approximatives 95%-Konfidenz-Intervall):

$$2 \text{ std} = 2\sqrt{\widehat{\text{Var}}[\hat{\pi}_j]} = (2.78, 2.45, 1.97, 1.54, 1.62, 1.11)\%$$

Dies entspricht etwa den Angaben aus dem Text:

... Die Fehlertoleranz bei den großen Parteien beträgt 2,7 Prozentpunkte, bei den kleineren 1,4 Prozentpunkte. ...





INFOBOX

Die Umfragen zum Politbarometer wurden ...

... wie immer von der Mannheimer Forschungsgruppe Wahlen durchgeführt. Die Interviews wurden in der Zeit vom 24. bis 26. März 2009 bei 1245 zufällig ausgewählten Wahlberechtigten telefonisch erhoben. Die Befragung ist repräsentativ für die wahlberechtigte Bevölkerung in ganz Deutschland. Die Fehlertoleranz bei den großen Parteien beträgt 2,7 Prozentpunkte, bei den kleineren 1,4 Prozentpunkte. Das nächste Politbarometer sendet das ZDF am Freitag, 24. April, nach dem heute-journal.

Abbildung 2.3: Politbarometer vom 27.3.2009.

2.2 Wünschenswerte Eigenschaften von Schätzfunktionen

2.2.1 Erwartungstreue

Definition 2.1 (Erwartungstreue)

Der Schätzer

$$\hat{\Theta} = g(X_1(\omega), \dots, X_N(\omega)) \quad (2.7)$$

heißt **erwartungstreu für Θ** , falls

$$E[\hat{\Theta}] = \theta, \quad (2.8)$$

d.h. im Mittel soll $\hat{\Theta}$ mit θ übereinstimmen.

Beispiel 2.4 (Erwartungstreue Schätzer)

Im folgenden wird angenommen, daß $\mu = E[X]$ und $\sigma^2 = \text{Var}(X)$ in der Grundgesamtheit ist.

- Mittelwert \bar{X}
- Stichprobenvarianz

$$S^2 = \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})^2 \quad (2.9)$$

- verzerrte Stichprobenvarianz

$$S_*^2 = \frac{1}{N} \sum_{n=1}^N (X_n - \bar{X})^2 \quad (2.10)$$

ist **nicht erwartungstreu** ($E[S_*^2] = \frac{N-1}{N}\sigma^2$).

- Die Standardabweichung $S = \sqrt{S^2}$ ist **nicht erwartungstreu**.



Definition 2.2

Der **Schätzfehler** ist die Abweichung vom wahren Wert $\hat{\Theta} - \theta$.

Die **Verzerrung (bias)** ist der erwartete Schätzfehler $E[\hat{\Theta}] - \theta$.

Ein erwartungstreuer Schätzer ist demnach unverzerrt, d.h. $E[\hat{\Theta}] - \theta = \theta - \theta = 0$.

Definition 2.3 (Mittlerer quadratischer Fehler)

Der mittlere quadratische Fehler (mean squared error) ist die erwartete quadrierte Abweichung vom wahren Wert, d.h.

$$\text{MSE} = E[\hat{\Theta} - \theta]^2 = E[\text{Schätzfehler}^2] \quad (2.11)$$

$$E[\hat{\Theta} - \theta]^2 = E[\hat{\Theta} - E[\hat{\Theta}] + E[\hat{\Theta}] - \theta]^2 \quad (2.12)$$

$$= E[\hat{\Theta} - E[\hat{\Theta}]]^2 + [E[\hat{\Theta}] - \theta]^2. \quad (2.13)$$

$$\text{MSE} = \text{Var}[\hat{\Theta}] + \text{bias}^2 \quad (2.14)$$

Der MSE setzt sich aus der Streuung und der Verzerrung des Schätzers zusammen.

Bei einem unverzerrten Schätzer stimmen also MSE und Varianz überein.

2.2.2 Wirksamkeit (Effizienz)**Definition 2.4 (Relative Effizienz)**

Ein unverzerrter Schätzer $\hat{\Theta}_1$ heißt wirksamer (effizienter) als $\hat{\Theta}_2$, wenn er eine kleinere Varianz aufweist, d.h. $\text{Var}(\hat{\Theta}_1) < \text{Var}(\hat{\Theta}_2)$.

Definition 2.5 (Effizienz)

Ein unverzerrter Schätzer $\hat{\Theta}$ heißt effizient (wirksamst), wenn er die kleinste Varianz unter allen unverzerrten Schätzern für θ aufweist.

Beispiel 2.5

Daten X_1, \dots, X_N ,

$$E[X_n] = \mu, \text{Var}(X_n) = \sigma^2$$

Schätzer:

$$\hat{\Theta}_1 = \bar{X} = \frac{1}{N} \sum X_n, N > 3$$

$$\hat{\Theta}_2 = \frac{1}{2}(X_1 + X_2)$$

$$\hat{\Theta}_3 = \frac{1}{3}(X_1 + X_2 + X_3)$$

Es gilt: $E[\hat{\Theta}_{1,2,3}] = \mu$ (Erwartungstreue), jedoch

$$\text{Var}(\hat{\Theta}_1) = \frac{\sigma^2}{N} < \frac{\sigma^2}{3} < \frac{\sigma^2}{2}$$

$\hat{\Theta}_1$ ist effizienter als $\hat{\Theta}_2$ und $\hat{\Theta}_3$.



Im allgemeinen gilt:

Satz 2.1

Die wirksamste **lineare** Schätzfunktion für μ ist der Mittelwert \bar{X} .

Beweis:

Minimiere $\text{Var}(\sum_n \alpha_n X_n)$ unter der Nebenbedingung $\sum_n \alpha_n = 1$.

Hinweis: Verwenden Sie eine Lagrange-Funktion.

2.2.3 Konsistenz

Bei Erhöhung des Stichprobenumfangs sollte sich auch die Schätzgenauigkeit erhöhen, etwa indem die Streuung des Schätzers immer kleiner wird.

Der Schätzfehler $\hat{\Theta} - \theta$ ist eine Zufallsgröße, sodaß man Wahrscheinlichkeitsaussagen machen muß, etwa daß die Wahrscheinlichkeit einer Abweichung vom wahren Wert bei großem Stichprobenumfang immer geringer wird.

Die Wahrscheinlichkeit einer Abweichung des Schätzers vom wahren Wert um mehr als die Zahl c kann man in der Form $P\{|\hat{\Theta} - \theta| \geq c\}$ schreiben.

Definition 2.6 (Konsistenz)

Eine Folge von Schätzfunktionen $\hat{\Theta}_N$ heißt *konsistent für θ* , wenn für jede beliebige Zahl $c > 0$ die Abweichungswahrscheinlichkeit gegen Null geht, d.h.

$$P\{|\hat{\Theta}_N - \theta| \geq c\} \rightarrow 0 \text{ für } N \rightarrow \infty \quad (2.15)$$

**Konvergenz in
Wahrscheinlichkeit**

Man schreibt (Konvergenz in Wahrscheinlichkeit)

$$p\text{-}\lim_{N \rightarrow \infty} \hat{\Theta}_N = \theta \quad (2.16)$$

Wie schon oben erwähnt, sollte die Varianz des Schätzers gegen Null gehen. Dies ist in der Tat hinreichend für die Konvergenz in Wahrscheinlichkeit:

$$P\{|\hat{\Theta}_N - \theta| \geq c\} \leq \text{Var}(\hat{\Theta}_N)/c^2 \rightarrow 0, \quad (2.17)$$

wenn $\text{Var}(\hat{\Theta}_N) \rightarrow 0$ strebt. Die umgekehrte Richtung gilt nicht.

**Tschebyscheff-
Ungleichung**

Hierbei wurde die **Tschebyscheff-Ungleichung** und $E[\hat{\Theta}_N] = \theta$ benutzt:

$$P\{|X - E[X]| \geq c\} \leq \text{Var}(X)/c^2 \quad (2.18)$$

Beispiel 2.6 (Konsistenz des Mittelwerts)

Daten X_1, \dots, X_N ,
 $E[X_n] = \mu, \text{Var}(X_n) = \sigma^2$.

Schätzer:

$$\hat{\Theta}_N = \bar{X}_N = \frac{1}{N} \sum X_n$$

$$E[\hat{\Theta}_N] = \frac{1}{N} \sum E[X_n] = \mu \text{ (Erwartungstreue)}$$

$$\text{Var}(\bar{X}_N) = \frac{\sigma^2}{N}$$

$$P\{|\bar{X}_N - \mu| \geq c\} \leq \frac{\sigma^2}{Nc^2} \rightarrow 0 \tag{2.19}$$

oder

$$p\text{-}\lim_{n \rightarrow \infty} \bar{X}_N = \mu \tag{2.20}$$

■

2.3 Methoden zur Konstruktion von Schätzfunktionen

Bisher wurden die Schätzfunktionen mehr oder weniger erraten oder plausibel gemacht. Es gibt jedoch auch systematische Prinzipien zur Ableitung solcher Schätzer. Einige sollen hier kurz diskutiert werden.

2.3.1 Prinzip der kleinsten Quadrate

siehe Regressionsanalyse (Kap. 5.1.2).

2.3.2 Momentenmethode

In vielen Fällen sind Schätzer durch sogenannte Momente der Verteilung motiviert, etwa durch das k -te Moment

$$\mu_k = \int x^k f(x) dx. \tag{2.21}$$

Ersetzt man dies durch das k -te Stichprobenmoment

**k -tes
Stichprobenmoment**

$$M_k = \frac{1}{N} \sum_n X_n^k = \hat{\mu}_k, \quad (2.22)$$

so ist dies ein erwartungstreuer Schätzer, da

$$E[M_k] = \frac{1}{N} \sum_n E[X_n^k] = \mu_k. \quad (2.23)$$

Im allgemeinen sind jedoch Momentenschätzer wenig effizient.

2.3.3 Maximum-Likelihood(ML)-Methode

**Likelihood-
Funktion**

Hier wird die sogenannte Likelihood-Funktion

$$L(\theta) := f(X_1, \dots, X_N | \theta) \quad (2.24)$$

betrachtet, wobei die gemeinsame Dichtefunktion aller Daten X_1, \dots, X_N als Funktion des Parameters θ aufgefaßt wird.

**Maximum-
Likelihood(ML)-
Schätzer**

Man bestimmt nun das Maximum der Likelihood und nennt den maximierenden zufälligen θ -Wert Maximum-Likelihood-Schätzer $\hat{\Theta}_{ML}$.

Der Begriff *Likelihood* wurde von Fisher eingeführt, da f als Funktion von θ keine Wahrscheinlichkeit ist.

Die a posteriori-Wahrscheinlichkeit (vgl. 18.1)

$$f(\theta | X_1, \dots, X_N) = \frac{f(X_1, \dots, X_N | \theta) f(\theta)}{f(X_1, \dots, X_N)}$$

**Maximum-a
posteriori(MAP)-
Schätzer**

führt auf den Bayes-Schätzer (Maximum-a posteriori(MAP)-Schätzer). Bei einer nichtinformativen a priori-Verteilung $f(\theta) = \text{const.}$ ergibt sich als Spezialfall der ML-Schätzer. Allerdings wird bei der Bayes-Schätzung dem Parameter eine Verteilung zugeordnet (welche die Information über den Parameter widerspiegelt), während bei der ML-Schätzung ein fixer, aber unbekannter Parameter θ auftritt.

Beispiel 2.7 (ML-Schätzer bei Normalverteilung)

Annahme:

Die Daten X_1, \dots, X_N sind unabhängig und identisch normalverteilt

$$X_n \sim N(\mu, \sigma^2). \quad (2.25)$$

Die gemeinsame Dichtefunktion aller Daten ist somit:

$$f(x_1, \dots, x_N | \mu, \sigma^2) = \phi(x_1; \mu, \sigma^2) \dots \phi(x_N; \mu, \sigma^2) \quad (2.26)$$

$$\phi(x; \mu, \sigma^2) = (2\pi\sigma^2)^{-1/2} \exp \left\{ -\frac{1}{2} \frac{(x - \mu)^2}{\sigma^2} \right\} \quad (2.27)$$

Die Likelihood-Funktion

$$L(\mu, \sigma^2) = (2\pi\sigma^2)^{-N/2} \exp \left\{ -\frac{1}{2} \sum_n \frac{(X_n - \mu)^2}{\sigma^2} \right\} \quad (2.28)$$

$$= (2\pi\sigma^2)^{-N/2} \exp \left\{ -\frac{N}{2\sigma^2} (S_*^2 + (\bar{X} - \mu)^2) \right\} \quad (2.29)$$

ist die Dichtefunktion $f(X_1, \dots, X_N | \mu, \sigma^2)$, aufgefaßt als Funktion der Parameter und mit eingesetzten Daten.

Hier wurde der Mittelwert \bar{X} und die Stichprobenvarianz (mittlere quadratische Abweichung)

$$S_*^2 = \frac{1}{N} \sum_{n=1}^N (X_n - \bar{X})^2 = \frac{N-1}{N} S^2 \quad (2.30)$$

**Stichprobenvarianz
(ML)**

substituiert (Übung).

Die Likelihood hängt also von den Daten nur über die (suffizienten) Statistiken \bar{X}, S_*^2 ab.

Das Maximum von L ist durch die Parameter-Schätzwerte

$$\hat{\mu} = \bar{X} \quad (2.31)$$

$$\hat{\sigma}^2 = S_*^2 \quad (2.32)$$

gegeben (Übung). Man erhält also i.A. verzerrte Schätzer.



Die Bedeutung der ML-Methode wird durch folgende Eigenschaften belegt:

Maximum-Likelihood (ML)-Schätzer**Satz 2.2 (Maximum-Likelihood-Schätzer)**

Maximum-Likelihood-Schätzer sind
(unter verschiedenen technischen Bedingungen)

- konsistent
- asymptotisch effizient
- asymptotisch normalverteilt

Dies bedeutet, daß man in großen Stichproben Tests und Konfidenzintervalle für ML-Schätzer mit Hilfe der Normalverteilung berechnen kann und die (nicht unbedingt normalverteilten) Daten optimal ausnutzt. Allerdings muß das Maximum von L in vielen Fällen numerisch berechnet werden (vgl. Kap. 14.11.4).

2.4 Konfidenzintervalle

Bei der Punktschätzung stellt sich ein Problem: der Schätzwert $\hat{\Theta}$ weicht immer vom wahren Wert θ ab, d.h. man hat einen Schätzfehler $\hat{\Theta} - \theta$. Dieser ist unbekannt, da man den wahren Parameter nicht kennt.

Es wird also ein Maß, etwa $\text{Var}(\hat{\Theta})$, für die Streuung des Schätzers benötigt.

Dann ist $\hat{\Theta} \pm \sqrt{\text{Var}(\hat{\Theta})}$ ein Bereich, in dem mit hoher Wahrscheinlichkeit der wahre Parameter enthalten sein wird.

Man sucht ein Intervall, das mit einer (hohen) Wahrscheinlichkeit den wahren (aber unbekannt) Parameter überdeckt. Anstatt eines Punkts erhält man also ein Intervall (untere, obere Grenze) als Schätzergebnis.

Definition 2.7 (Konfidenz-Intervall (KI))Ein **zufälliges** Intervall

$$[U, O] = [g_1(X_1, \dots, X_N), g_2(X_1, \dots, X_N)], \quad (2.33)$$

das als Funktion der Daten geschrieben werden kann und für das gilt

$$P\{U \leq \theta \leq O\} = 1 - \alpha \quad (2.34)$$

heißt **Konfidenz(Vertrauens)-Intervall** zum Konfidenz-Niveau $1 - \alpha$ (z.B. 95%, 99%).**Wichtig:**

Die Wahrscheinlichkeitsaussage ist **nicht** für den Parameter formuliert (dieser ist keine Zufallsvariable), sondern für die zufälligen Intervallgrenzen $U(\omega), O(\omega)$.

Hat man konkrete Daten x_1, \dots, x_N , so ergeben sich konkrete Intervallgrenzen und ein realisiertes Intervall $[u, o]$. Man weiß jedoch nicht, ob der wahre Parameter in diesem Intervall wirklich enthalten ist. Auch ist die Aussage

$$P\{u \leq \theta \leq o\} = 1 - \alpha \quad (2.35)$$

nicht sinnvoll, da θ keine Zufallsgröße ist.¹

Im Mittel überdeckt jedoch $[U, O]$ in $100(1 - \alpha)\%$ der Fälle den wahren Wert, etwa in M Replikationen der Stichprobe.

2.4.1 Konfidenzintervall für μ : normalverteilte Grundgesamtheit, bekannte Varianz

Annahmen:

¹Im Fall des Bayes-Schätzers (siehe Abs. 2.3.3) kann man dagegen schreiben

$$P(u \leq \theta \leq o | X_1, \dots, X_N) = 1 - \alpha,$$

da hier der Parameter eine Verteilung aufweist. Dies ist die a posteriori-Wahrscheinlichkeit, gegeben die Daten, daß der Parameter in einem bestimmten Intervall liegt.

- Daten $X_1, \dots, X_N, X_n \sim N(\mu, \sigma^2)$
- Punktschätzer $\bar{X} = \frac{1}{N} \sum X_n$
- Erwartungswert $E(\bar{X}) = \mu$
- Varianz $\text{Var}(\bar{X}) = \sigma^2/N$

Dann gilt die Wahrscheinlichkeitsaussage [$z = z(1 - \alpha/2) =$ Quantil der Normalverteilung]

$$P\left\{\mu - z \frac{\sigma}{\sqrt{N}} \leq \bar{X} \leq \mu + z \frac{\sigma}{\sqrt{N}}\right\} = 1 - \alpha, \quad (2.36)$$

die aber noch kein Konfidenzintervall ist, da der Schätzer (und nicht der Parameter) in der Mitte steht.

Umstellen nach μ ergibt

$$P\left\{\bar{X} - z \frac{\sigma}{\sqrt{N}} \leq \mu \leq \bar{X} + z \frac{\sigma}{\sqrt{N}}\right\} = 1 - \alpha, \quad (2.37)$$

$$z = z(1 - \alpha/2).$$

Übung 2.1

Stellen Sie bitte $\mu - a \leq \bar{X} \leq \mu + a$ nach μ um (in die Mitte der Ungleichung).

Hinweis: Betrachten Sie die Teile der Ungleichung einzeln.



Eigenschaften des Konfidenzintervalls:

- Länge des Konfidenzintervalls:
 $l = O - U = 2 z(1 - \alpha/2) \frac{\sigma}{\sqrt{N}} =$ feste Größe, obwohl O und U zufällig sind.
- Je größer σ ist, desto länger ist das Intervall (Streuung der Daten).
- Je größer N ist, desto kürzer ist das Intervall (mehr Information bei größerer Stichprobe).

- Je größer $1 - \alpha$ ist, desto länger ist das Intervall (größere Sicherheit).

Beispiel 2.8 (Bruttoinlandsprodukt BIP)

Das nominale Wachstum des Bruttoinlandsprodukts im Jahr 2007 für verschiedene OECD-Länder hat folgende Werte (Abb. 2.4; Datensatz Regression.jmp). Aus Abb. 2.5 entnimmt man folgende Informationen:

$$\bar{x} = 5.53, s = 1.86, N = 32.$$

Unter der Annahme, daß s mit der tatsächlichen Varianz σ übereinstimmt, sowie der Normalverteilung der Daten (dies ist nicht unplausibel; siehe Anpassungstests) ergibt sich folgendes 95%-Konfidenzintervall für den unbekanntem Erwartungswert μ :

$$z = z(1 - \alpha/2) = z(1 - .05/2) = z(.975) = 1.96,$$

$$[\bar{x} - z \frac{\sigma}{\sqrt{N}}, \bar{x} + z \frac{\sigma}{\sqrt{N}}] =$$

$$[5.53 - 1.96 \frac{1.86}{\sqrt{32}}, 5.53 + 1.96 \frac{1.86}{\sqrt{32}}] =$$

$$[5.53 - 0.65, 5.53 + 0.65] = [4.89, 6.18]$$

■

2.4.2 Konfidenzintervall für μ : normalverteilte Grundgesamtheit, unbekannte Varianz

Falls σ unbekannt ist, muß es zuvor geschätzt werden.

Annahmen:

- Daten $X_1, \dots, X_N, X_n \sim N(\mu, \sigma^2)$
- Punktschätzer

$$\bar{X} = \frac{1}{N} \sum X_n,$$

$$S^2 = \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})^2$$
- Erwartungswert $E(\bar{X}) = \mu, E(S^2) = \sigma^2$

Dann ist die Stichprobenfunktion

$$T = \frac{\bar{X} - \mu}{S/\sqrt{N}} \sim t(N - 1) \quad (2.38)$$

t -verteilt mit $N - 1$ Freiheitsgraden.

	Land	Nominales GDP- Wachstum in % 2007
1	Australia	6,59
2	Austria	4,08
3	Belgium	4,21
4	Canada	5,22
5	Czech Republic	8,32
6	Denmark	5,38
7	Finland	3,77
8	France	3,88
9	Germany	3,18
10	Greece	7,20
11	Hungary	7,30
12	Iceland	5,71
13	Ireland	8,54
14	Italy	3,54
15	Japan	2,80
16	Korea	6,33
17	Luxembourg	6,51
18	Mexico	6,92
19	Netherlands	4,13
20	New Zealand	4,29
21	Norway	5,69
22	Poland	6,45
23	Portugal	3,36
24	Slovak Rep...	8,65
25	Spain	6,78
26	Sweden	5,26
27	Switzerland	3,22
28	Turkey	10,43
29	United King...	4,54
30	United States	5,48
31	Euro area	4,18
32	Total OECD	4,94

Abbildung 2.4: Wachstum des Bruttoinlandsprodukts im Jahr 2007.

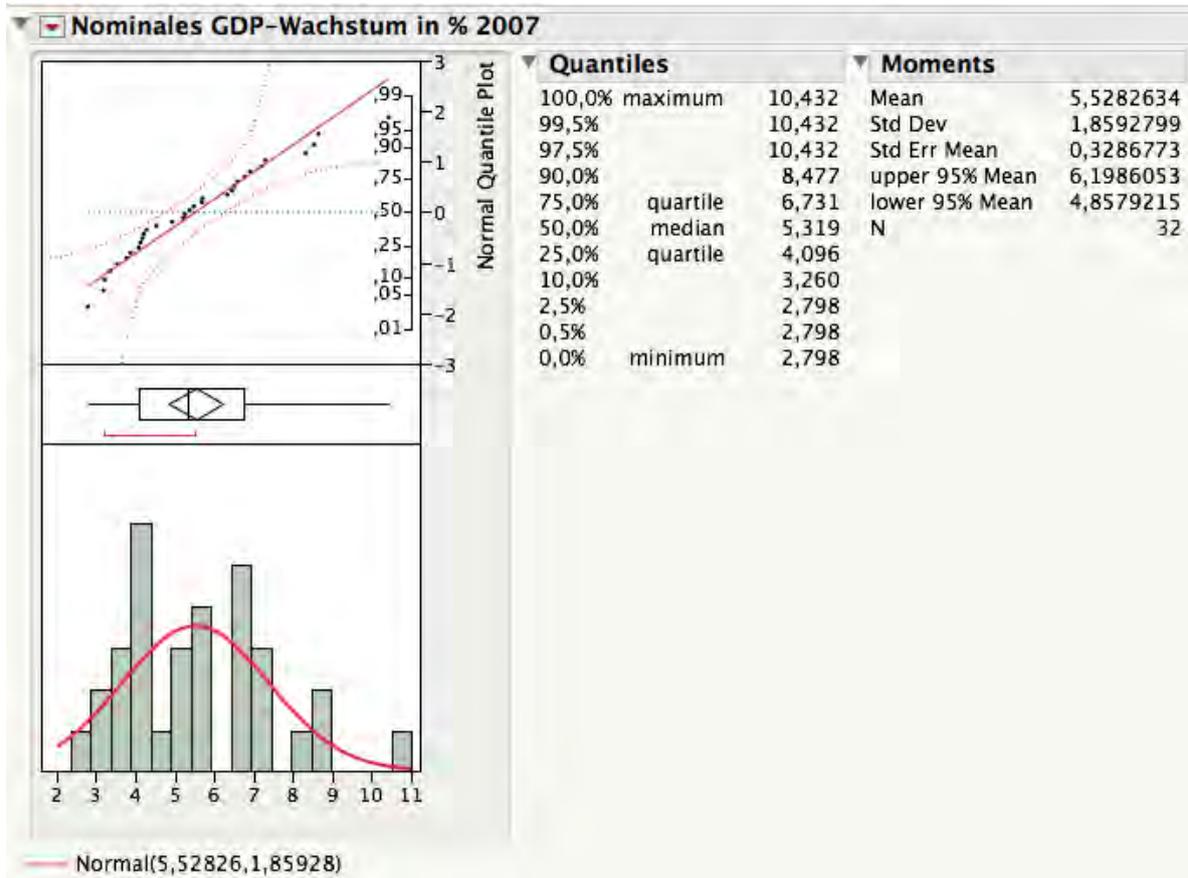


Abbildung 2.5: Wachstum des Bruttoinlandsprodukts im Jahr 2007: Verteilung und Statistiken.

Somit gilt die Wahrscheinlichkeitsaussage

$$P\{-t \leq T \leq t\} = 1 - \alpha \quad (2.39)$$

Stellt man dies wieder nach μ um, so ergibt sich das **t -Intervall**

$$P\left\{\bar{X} - t \frac{S}{\sqrt{N}} \leq \mu \leq \bar{X} + t \frac{S}{\sqrt{N}}\right\} = 1 - \alpha \quad (2.40)$$

mit dem **t -Quantil** $t(1 - \alpha/2, N - 1)$.

Es hat eine **zufällige Länge** von $l = 2 t S/\sqrt{N}$ und ist damit im Durchschnitt etwas länger als im Fall bekannter Varianz.

Beispiel:

$$t(.975, 8) = 2.306, z(.975) = 1.960.$$

Übung 2.2 (Modulevaluation)

Bei einer Modulevaluation nahmen 13 Studenten teil. Der Mittelwert der Benotung eines Moduls (Noten: 1–5) war 2.5, die Standardabweichung 3.1.

Berechnen Sie bitte ein 95%-Konfidenzintervall für den unbekanntem Erwartungswert der Modul-Note (Grundgesamtheit: alle potentiellen Teilnehmer der Evaluation).

Lösung:

$$\bar{x} = 2.5, s = 3.1, N = 13, \alpha = 0.05 \Rightarrow t(.975, 12) = 2.179.$$

Daraus ergibt sich das (ziemlich große) 95%- t -Intervall

$$\bar{x} \pm d = 2.5 \pm 1.873 = [0.627, 4.373]$$

$$\text{mit } d = 2.179 \times 3.1/\sqrt{13} = 1.873.$$

Dies heißt aber nicht, daß der unbekanntem Erwartungswert in diesem konkreten Intervall wirklich liegt. Er würde aber bei M -facher Wiederholung der Erhebung im Mittel von 95% der Intervalle überdeckt werden.



Übung 2.3 (Bruttoinlandsprodukt: t -Intervall)

Berechnen Sie ein t -Intervall für die OECD-Daten (Bsp. 2.8).



2.4.3 Konfidenzintervall für μ : beliebig verteilte Grundgesamtheit, großer Stichprobenumfang

In den bisher diskutierten Konfidenzintervallen wurde eine Verteilungsannahme (Normalverteilung) für die Stichprobenvariablen X_n gemacht. Ist dies unrealistisch, so kann in großen Stichproben der sogenannte **zentrale Grenzwertsatz** für den Mittelwert $\hat{\mu} = \bar{X} = \frac{1}{N} \sum X_n$ ausgenutzt werden: Summen von beliebig verteilten Zufallsvariablen sind nämlich approximativ normalverteilt.

Etwas genauer gilt:

Satz 2.3 (zentraler Grenzwertsatz)

Die Summe von beliebig verteilten unabhängigen Zufallsvariablen ist asymptotisch normalverteilt.

Sei $S_N = \sum_{n=1}^N X_n$; $E[S_N] = N\mu$, $\text{Var}(S_N) = N\sigma^2$

Dann ist die standardisierte Variable

$$Z_N = \frac{S_N - N\mu}{\sqrt{N\sigma^2}} \underset{a}{\sim} N(0, 1) \text{ für } N \rightarrow \infty \quad (2.41)$$

asymptotisch normalverteilt.

**zentraler
Grenzwertsatz**

Man erhält also die approximative Wahrscheinlichkeitsaussage für den Mittelwert \bar{X}

$$P \left\{ \bar{X} - z \frac{\sigma}{\sqrt{N}} \leq \mu \leq \bar{X} + z \frac{\sigma}{\sqrt{N}} \right\} \approx 1 - \alpha, \quad (2.42)$$

$N > 30$; $z = z(1 - \alpha/2)$.

(approximatives Konfidenzintervall für $\mu = E[X_n]$ bei beliebig verteilten X_n).

2.4.4 Konfidenzintervall für den Anteilswert π : Binomialverteilte Grundgesamtheit, großer Stichprobenumfang

Betrachtet man die Indikatorvariable

$$X_n = \begin{cases} 1 & \text{wenn } A \\ 0 & \text{sonst} \end{cases} \quad (2.43)$$

$$E[X_n] = \pi \quad (2.44)$$

für das Merkmal A (etwa Person n wählt die A -Partei), so ist ein Vertrauensintervall für die unbekannte Wahrscheinlichkeit π von großer praktischer Relevanz (Genauigkeit von Wahlumfragen).

Als Punkt-Schätzer von π nimmt man den Mittelwert (Prozentanteil P)

$$\hat{\pi} = P = \bar{X}_N = \frac{1}{N} \sum_n X_n \quad (2.45)$$

$$E[\hat{\pi}] = \pi \quad (2.46)$$

$$\text{Var}[\hat{\pi}] = \frac{1}{N} \pi(1 - \pi) \quad (2.47)$$

Für große Stichproben kann man wieder ausnutzen, daß

$$Z_N = \frac{\hat{\pi} - \pi}{\sqrt{\frac{1}{N} \pi(1 - \pi)}} \stackrel{a}{\sim} N(0, 1) \text{ für } N \rightarrow \infty \quad (2.48)$$

asymptotisch normalverteilt ist.

Man fordert, daß $N\pi \geq 5$, $N(1 - \pi) \geq 5$, bzw. $N\hat{\pi} \geq 5$, $N(1 - \hat{\pi}) \geq 5$.

Ersetzt man auch noch die Varianz $\frac{1}{N} \pi(1 - \pi)$ durch den konsistenten Schätzer $\frac{1}{N} \hat{\pi}(1 - \hat{\pi})$, so ergibt sich das approximative Vertrauensintervall

$$P \left\{ \hat{\pi} - z \sqrt{\frac{1}{N} \hat{\pi}(1 - \hat{\pi})} \leq \pi \leq \hat{\pi} + z \sqrt{\frac{1}{N} \hat{\pi}(1 - \hat{\pi})} \right\} \approx 1 - \alpha, \quad (2.49)$$

$$z = z(1 - \alpha/2)$$

(approximatives Konfidenzintervall für $\pi = E[X_n]$ bei dichotomen X_n (Indikatoren)).

Man kann auch mit Hilfe der Binomialverteilung von $\hat{\pi}$ ein exaktes KI konstruieren.

Beispiel 2.9 (Wahlumfrage)

In Beispiel 2.3 wurden die Prozentwerte der Parteien als Multinomialverteilung modelliert. Nimmt man der Einfachheit halber nur 2 Parteien (etwa CDU/CSU und Andere), so erhält man als Spezialfall $k = 2$ eine Binomialverteilung $B(N, \pi)$:

- Kategorien:

$$A = \text{CDU/CSU}, \bar{A} = \text{Andere}$$

- Stichproben-Vektor (Realisation):

$$x_n = \begin{cases} 1 & \text{Person } n \text{ wählt CDU/CSU} \\ 0 & \text{Person } n \text{ wählt Andere} \end{cases}$$

$$n = 1, \dots, N = 1245$$

- Prozentwerte

$$\hat{\pi} = p = 0.4$$

(relative Häufigkeit der Parteien; $p + (1 - p) = 1$)

- Standardfehler

$$\text{Var}[\hat{\pi}] = \frac{1}{N}\pi(1 - \pi)$$

Schätzung:

$$\widehat{\text{Var}}[\hat{\pi}] = \frac{1}{N}p(1 - p) = 0.000193$$

$$\text{std} = \sqrt{\frac{1}{N}p(1 - p)} = 0.0139$$

$$z(.975) = 1.96$$

- approximatives 95%-Konfidenz-Intervall:

$$0.4 \pm 1.96 \times 0.0139 = (40 \pm 2.78)\%$$



2.4.5 Konfidenzintervall für die Varianz σ^2 : Normalverteilte Grundgesamtheit

Die Stichprobenvarianz

$$S^2 = \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})^2 \quad (2.50)$$

ist ein erwartungstreuer Schätzer für die Varianz σ^2 . Teilt man durch diese und multipliziert mit dem Vorfaktor, so ergibt sich eine $\chi^2(N-1)$ -verteilte Zufallsgröße mit $N-1$ Freiheitsgraden

$$Y := (N-1)S^2/\sigma^2 = \frac{1}{\sigma^2} \sum_{n=1}^N (X_n - \bar{X})^2 \sim \chi^2(N-1). \quad (2.51)$$

Es gilt also

$$P \{ \chi^2(\alpha/2, N-1) \leq Y \leq \chi^2(1-\alpha/2, N-1) \} = 1 - \alpha \quad (2.52)$$

und durch Umstellen nach σ^2

$$P \left\{ \frac{(N-1)S^2}{\chi^2(1-\alpha/2, N-1)} \leq \sigma^2 \leq \frac{(N-1)S^2}{\chi^2(\alpha/2, N-1)} \right\} = 1 - \alpha \quad (2.53)$$

(Konfidenzintervall für die Varianz σ^2).

Übung 2.4

Stellen Sie bitte die Ungleichung

$$a \leq (N-1)S^2/\sigma^2 \leq b \quad (2.54)$$

nach σ^2 um.

Hinweis: behandeln Sie die Ungleichungen einzeln.



Beispiel 2.10 (Bruttoinlandsprodukt: Fortsetzung)

$$s = 1.86, N = 32, \alpha = .05$$

$$a = \chi^2(\alpha/2, N - 1) = 17.539$$

$$b = \chi^2(1 - \alpha/2, N - 1) = 48.232$$

KI für σ^2 : [2.224, 6.115]

KI für σ : [1.491, 2.473] (Wurzel ziehen)

(bitte nachrechnen).



2.4.6 Konfidenzintervall für den Korrelationskoeffizienten ρ : Normalverteilte Grundgesamtheit

Der Korrelationskoeffizient zwischen 2 Variablen X, Y ist ein Maß für den linearen Zusammenhang, bei normalverteilten Größen folgt sogar die Unabhängigkeit aus der Unkorreliertheit der Merkmale.²

Die Unkorreliertheit impliziert jedoch nicht unbedingt die Unabhängigkeit, etwa bei einem parabelförmigen Zusammenhang ($Y = X^2$; unkorreliert, jedoch perfekt abhängig).

Zunächst ist die Korrelation ganz allgemein die normierte Kovarianz der Variablen, d.h.

$$\rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} \quad (2.55)$$

Ein Punktschätzer für ρ ist durch die Produkt-Moment-Korrelation

$$R = \frac{S(X, Y)}{S(X)S(Y)} \quad (2.56)$$

²da bei einer diagonalen Kovarianzmatrix die Dichtefunktion in Produkte faktorisiert. Siehe Kap. 18

gegeben, wobei die Stichproben(ko)varianzen durch

$$S(X, Y) = \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})(Y_n - \bar{Y}) \quad (2.57)$$

$$S(X)^2 = \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})^2 = S(X, X) \quad (2.58)$$

$$S(Y)^2 = \frac{1}{N-1} \sum_{n=1}^N (Y_n - \bar{Y})^2 = S(Y, Y) \quad (2.59)$$

definiert sind.

Explizit kann man auch

Produkt-Moment-Korrelation

$$R = \frac{\sum_{n=1}^N (X_n - \bar{X})(Y_n - \bar{Y})}{\sqrt{\sum_{n=1}^N (X_n - \bar{X})^2 \sum_{n=1}^N (Y_n - \bar{Y})^2}} \quad (2.60)$$

schreiben.

2.4.6.1 Fishers Z-Transformation

Wenn die Variablen (X_n, Y_n) bivariat normalverteilt sind, so ist die transformierte Variable

$$Z = \frac{\sqrt{N-3}}{2} \left(\ln \frac{1+R}{1-R} - \ln \frac{1+\rho}{1-\rho} \right) \quad (2.61)$$

für große Stichproben $N(0, 1)$ -normalverteilt (**Fishers Z-Transformation**).

Somit ergibt sich das Intervall $[z = z(1 - \alpha/2)]$

$$P \{-z \leq Z \leq z\} = 1 - \alpha, \quad (2.62)$$

das man wie immer nach ρ umformen kann

$$P \left\{ \frac{e^A - 1}{e^A + 1} \leq \rho \leq \frac{e^B - 1}{e^B + 1} \right\} = 1 - \alpha, \quad (2.63)$$

wobei

$$A = \ln \frac{1 + R}{1 - R} - \frac{2z}{\sqrt{N - 3}} \quad (2.64)$$

$$B = \ln \frac{1 + R}{1 - R} + \frac{2z}{\sqrt{N - 3}}. \quad (2.65)$$

Beispiel 2.11

Zwei Variablen korrelieren zu $r = 0.475$ bei einer Stichprobengröße von $N = 80$. Gesucht ist ein 95%-Vertrauensintervall für den unbekanntem Korrelationskoeffizienten ρ .

97.5%-Quantil: $z(0.975) = 1.96$

$$a = \ln \frac{1 + 0.475}{1 - 0.475} - \frac{2 \times 1.96}{\sqrt{77}} = \ln(2.809) - .447 \quad (2.66)$$

$$= 1.033 - .447 = 0.586 \quad (2.67)$$

$$b = 1.033 + .447 = 1.48. \quad (2.68)$$

$$\frac{e^a - 1}{e^a + 1} = \frac{1.797 - 1}{1.797 + 1} = \frac{0.797}{2.797} = 0.285 \quad (2.69)$$

$$\frac{e^b - 1}{e^b + 1} = \frac{4.393 - 1}{4.393 + 1} = \frac{3.393}{5.393} = 0.629 \quad (2.70)$$

Daraus findet man das 95%-Konfidenzintervall (KI)
[0.285, 0.629].



2.5 Statistische Überprüfung von Hypothesen: Der Signifikanztest

2.5.1 Hypothesen

Im Rahmen der Berechnung von Konfidenzintervallen wurden keine Hypothesen über die unbekanntem Parameter formuliert, sondern zufällige

Bereiche berechnet, die den (unbekannten) Parameter mit einer bestimmten Wahrscheinlichkeit überdecken. Formuliert man nun eine Hypothese über den Parameter, etwa

- *Der Zusammenhang zwischen Gewicht und Größe ist $\rho = 0.65$,*

so würde ein Konfidenzintervall, das diesen Wert nicht überdeckt, etwa $[0.285, 0.628]$ im obigen Beispiel, dieser Hypothese widersprechen. Es genügt also nicht, daß der berechnete Wert (Punktschätzer) $r = 0.475$ etwas neben der Hypothese liegt (er liegt sogar mit Wahrscheinlichkeit 1 daneben), sondern das gesamte Intervall muß die Hypothese verfehlen. In diesem Fall würde man die Hypothese: $\rho = 0.65$ ablehnen.

Diese Ablehnung kann aber auch falsch sein, denn wenn der wahre Parameter tatsächlich $\rho = 0.65$ ist, gibt es durchaus Konfidenzintervalle, die diesen Wert nicht überdecken.

Diese Wahrscheinlichkeit (Fehler 1. Art) ist aber nur $\alpha = P\{\rho \notin KI\}$ (ρ wird nicht vom Konfidenzintervall überdeckt).

Zunächst wird die (statistische) Hypothese formuliert:

Hypothesen

Definition 2.8 (Hypothesen)

Statistische Hypothesen sind Annahmen über die Wahrscheinlichkeitsverteilung eines oder mehrerer Merkmale bzw. deren Parameter.

Beispiele:

- Der Anteil der A-Partei bei der kommenden Bundestagswahl ist $\pi \geq 0.40$ ('Vierzig plus X')
- Männer sind (im Mittel) 10 cm größer als Frauen ($\mu_y - \mu_x = 10$).
- Der Zusammenhang zwischen Gewicht und Größe ist $\rho = 0.65$
- Die Daten sind normalverteilt.

Hypothesen betreffen also stets die Parameter in der Grundgesamtheit, also

- $\mu_y - \mu_x = 10$, **nicht die Schätzer** $\bar{Y} - \bar{X} = 10$.

2.5.2 Aufbau und Interpretation von Hypothesentests

Hypothesentests (Signifikanztests) sind im allgemeinen folgendermaßen aufgebaut:

1. Festlegung einer sogenannten Nullhypothese H_0 und einer Alternativhypothese H_1 .
2. Festlegung eines Signifikanz-Niveaus α :
 $\alpha = P\{H_0 \text{ ablehnen} | H_0 \text{ richtig}\} = \alpha\text{-Fehler}$
3. Bestimmung einer Testfunktion (Teststatistik)
 $T = T(X_1, \dots, X_N)$.
Die Verteilung der Teststatistik muß von der Hypothese abhängen und wird unter H_0 berechnet.
4. Festlegung eines Verwerfungs (Ablehnungs)-Bereichs R mit
 $P\{T \in R | H_0\} = \alpha$.
5. Entscheidungsregel:
wenn $T \in R$ wird H_0 abgelehnt (Realisation $T = t$ einsetzen).

Aufbau eines Hypothesentests

Bemerkungen:

- Die Notation $P\{A|H_0\}$ bedeutet, daß die Wahrscheinlichkeit des Ereignisses A unter Gültigkeit der Hypothese H_0 berechnet wird. Es handelt sich nicht um eine bedingte Wahrscheinlichkeit, da H_0 keine Ereignis ist.
- Die Wahrscheinlichkeit, H_0 abzulehnen, wenn H_0 richtig ist, soll klein sein (z.B. $\alpha = 5\%, 1\%$).
- Die Hypothese selbst hat keine Wahrscheinlichkeit. Man kann also nicht sagen: H_1 ist zu 95% richtig, etc.
- Der Ablehnungsbereich (rejection region) R ist eine Untermenge des Parameter-Raums (Wertebereich des Parameters, etwa $-1 \leq \rho \leq 1, -\infty < \mu < \infty$).

Da der Test für den Korrelationskoeffizienten etwas komplizierter ist (es wird eine transformierte Testgröße benutzt), hier zuerst der Test für den Erwartungswert μ :

Beispiel 2.12 (Test des Erwartungswerts μ ; σ bekannt)

Man betrachtet das Wachstum des Bruttoinlandsprodukts 2007 in Prozent. Es wird unterstellt, daß $\sigma = s = 1.86$ bekannt ist (ansonsten führt man einen t -Test durch). Der Mittelwert des Wachstums war $\bar{x} = 5.53$.

Ablauf des Tests:

1. Festlegung einer Nullhypothese $H_0 : \mu = \mu_0 = 6$ (Prozent) und einer Alternativhypothese $H_1 : \mu \neq \mu_0 = 6$.

2. Festlegung eines Signifikanz-Niveaus $\alpha = 0.05$:
 $\alpha = P\{H_0 \text{ ablehnen} | H_0 \text{ richtig}\} = \alpha\text{-Fehler}$

3. Bestimmung einer Testfunktion (Teststatistik)

$$T = T(X_1, \dots, X_N) = \bar{X}$$

(Mittelwert).

Es gilt unter H_0 :

$$E[T | H_0] = E[\bar{X} | H_0] = \mu_0 = 6$$

Die Verteilung der Teststatistik hängt somit, wie gefordert, von der Null-Hypothese ab.

4. Verwerfungsbereich R festlegen:

$$P\{T \in R | H_0\} = P\{\bar{X} \notin [\mu_0 - z \sigma / \sqrt{N}, \mu_0 + z \sigma / \sqrt{N}] | H_0\} = \alpha$$

bzw. für die standardisierte Teststatistik $Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{N}}$

$$P\{|Z| > z | H_0\} = \alpha$$

$$[\text{Quantil } z = z(1 - \alpha/2) = 1.96].$$

Ablehnungsbereich:

$$\begin{aligned} R &= (-\infty, \mu_0 - z \sigma / \sqrt{N}) \cup (\mu_0 + z \sigma / \sqrt{N}, \infty) \\ &= (-\infty, 6 - 1.96 \times 1.86 / \sqrt{32}) \cup (6 + 1.96 \times 1.86 / \sqrt{32}, \infty) \\ &= (-\infty, 5.36) \cup (6.65, \infty) \end{aligned}$$

5. Entscheidungsregel: wenn $t = \bar{x} \in R$, wird H_0 abgelehnt.

Da $5.53 \in [5.36, 6.65] = \bar{R}$ wird $H_0 : \mu = \mu_0 = 6$ beibehalten (kann nicht abgelehnt werden).

Entsprechend war das Konfidenzintervall $[4.89, 6.18]$, es überdeckte also die Nullhypothese (keine Ablehnung) ■

Beispiel 2.13 (Test des Korrelationskoeffizienten ρ)

1. Festlegung einer Nullhypothese $H_0 : \rho = \rho_0 = 0.65$

und einer Alternativhypothese $H_1 : \rho \neq 0.65$.

2. Festlegung eines Signifikanz-Niveaus $\alpha = 0.05$:

$\alpha = P\{H_0 \text{ ablehnen} | H_0 \text{ richtig}\} = \alpha\text{-Fehler}$

3. Bestimmung einer Testfunktion (Teststatistik)

$$T = T(X_1, \dots, X_N) = R$$

(empirischer Korrelationskoeffizient).

Die transformierte approximativ normalverteilte Test-Variable Z (die man hier benutzt, da die Verteilung von $T = R$ unbekannt ist; vgl. Abs. 2.4.6)

$$Z = \frac{\sqrt{N-3}}{2} \left(\ln \frac{1+R}{1-R} - \ln \frac{1+\rho_0}{1-\rho_0} \right) \quad (2.71)$$

enthält den Korrelations-Koeffizienten ρ_0 .

Die Verteilung der Teststatistik hängt somit von der Null-Hypothese ab.

4. Verwerfungsbereich R festlegen:

$$P\{T \in R | H_0\} = P\{Z \in R_Z | H_0\} = P\{|Z| > z | H_0\} = \alpha$$

[Quantil $z = z(1 - \alpha/2)$] d.h.

$$R_Z = (-\infty, -z) \cup (z, \infty) \quad (2.72)$$

$$= (-\infty, -1.96) \cup (1.96, \infty) \quad (2.73)$$

(Verwerfungsbereich für Z).

5. Entscheidungsregel: wenn $z \in R_Z$ wird H_0 abgelehnt.

Berechne

$$z = \frac{\sqrt{77}}{2} \left(\ln \frac{1 + .475}{1 - .475} - \ln \frac{1 + .65}{1 - .65} \right) = -2.27088 \quad (2.74)$$

Da $-2.27088 < -1.96$ wird $H_0 : \rho = \rho_0 = 0.65$ abgelehnt.

■

2.5.3 Fehler bei Testentscheidungen

Fällt man die Testentscheidung: H_0 ablehnen, oder H_0 beibehalten, so ist dies immer mit einem Fehler verbunden:

- H_0 ablehnen, obwohl sie richtig ist, oder
- H_0 beibehalten, obwohl sie falsch ist.

Die entsprechenden Wahrscheinlichkeiten sind

α -Fehler

- $\alpha = P(H_0 \text{ ablehnen, obwohl richtig})$
 $= P(H_0 \text{ ablehnen} | H_0 \text{ richtig})$

β -Fehler

- $\beta = P(H_0 \text{ beibehalten, obwohl falsch})$
 $= P(H_0 \text{ beibehalten} | H_1 \text{ richtig})$

In Form einer Tabelle ergibt sich übersichtlich (Tab. 2.2).

Der α -Fehler kann innerhalb der Testprozedur vorab auf einen kleinen Wert fixiert und der Ablehnungsbereich (konkret für den Gauß-Test)

$$P\{T \in R | H_0\} = P\{\bar{X} \notin [\mu_0 - z \sigma / \sqrt{N}, \mu_0 + z \sigma / \sqrt{N}] | H_0\} = \alpha$$

entsprechend gewählt werden ($z = z(1 - \alpha/2)$).

Andererseits ist

$$\beta = P(H_0 \text{ beibehalten, obwohl falsch}) \quad (2.75)$$

$$= P(H_0 \text{ beibehalten} | H_1 \text{ richtig}) = P(T \notin R | H_1) \quad (2.76)$$

unter Gültigkeit der Alternativ-Hypothese H_1 zu berechnen. Der konkrete Wert hängt aber vom unbekanntem wahren Parameter $\mu \in H_1$ ab, also $\beta = \beta(\mu)$.

		wahrer Zustand	
		H_0 richtig	H_0 falsch
Entscheidung	H_0 beibehalten	richtige Entscheidung	β -Fehler (Fehler 2. Art)
	H_0 ablehnen	α -Fehler (Fehler 1. Art)	richtige Entscheidung

Tabelle 2.2: Fehlentscheidungen bei statistischen Tests.

Ist beispielsweise die unspezifische Alternative $H_1 : \mu \neq \mu_0 = 6$ gegeben (wie im Beispiel), so gehören alle Erwartungswerte $\mu \neq 6$ zur Alternative und es gibt unendlich viele $\beta(\mu)$, die sehr groß, maximal $\beta(\mu) < 1 - \alpha = 0.95$ werden können.

Daher spricht man auch vorsichtig von
 H_0 **beibehalten (nicht ablehnen)**.

Der β -Fehler könnte ja maximal 95% sein, daher ist die H_0 **nicht bewiesen**.

Nur bei einer einfachen Alternative, etwa $H_1 : \mu = 7$, läßt sich $\beta(\mu)$ konkret berechnen.

Es ist wichtig, sich diese Sachverhalte graphisch zu verdeutlichen (Abb. 2.6). Im oberen Bild ist die Verteilung der Teststatistik unter $H_0 : \mu = 6$ (rote Kurve, links) und unter dem Parameterwert $\mu = 7 \in H_1$ (gelbe Kurve, rechts) dargestellt. Der β -Fehler ist die Wahrscheinlichkeit, daß die Teststatistik unter Gültigkeit von H_1 nicht in den Ablehnungsbereich fällt (gelb unterlegter Bereich), also H_0 beigehalten wird.

Im unteren Bild ist der Parameterwert $\mu = 6.5 \in H_1$. Daher ist der β -Fehler wesentlich größer (0.67). Er kann sogar fast $1 - \alpha$ werden, wenn μ fast mit dem Wert $\mu_0 = 6$ aus der H_0 übereinstimmt.

Dagegen ist die rot unterlegte Fläche in beiden Fällen gleich, d.h. der α -Fehler kann vorab fixiert werden.

Der β -Fehler kann mit Hilfe der kritischen Werte und der Verteilung der Teststatistik unter $\mu \in H_1$ wie folgt berechnet werden (konkret für den

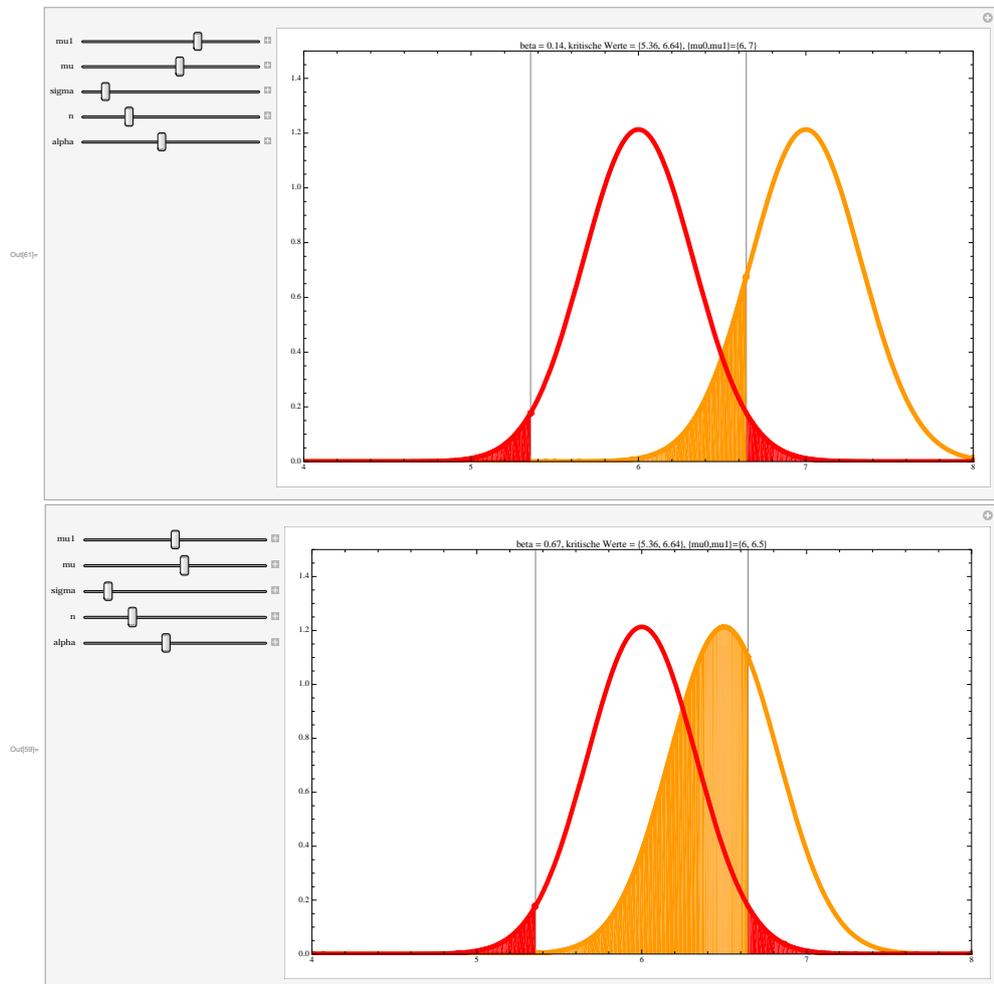


Abbildung 2.6: β -Fehler beim Gauß-Test:

oben: $H_0 : \mu = 6$ und wahrer Parameter $\mu = 7 \in H_1$. Kritische Werte: $\mu_0 \pm z \times \sigma / \sqrt{N} = 6 \pm .65 = [5.36, 6.65]$ (senkrechte Linien). Der β -Fehler ist 0.14 (gelb unterlegt).

unten: $H_0 : \mu = 6$ und wahrer Parameter $\mu = 6.5 \in H_1$. β -Fehler = 0.67.

http://www.fernuni-hagen.de/ls_statistik/lehre/

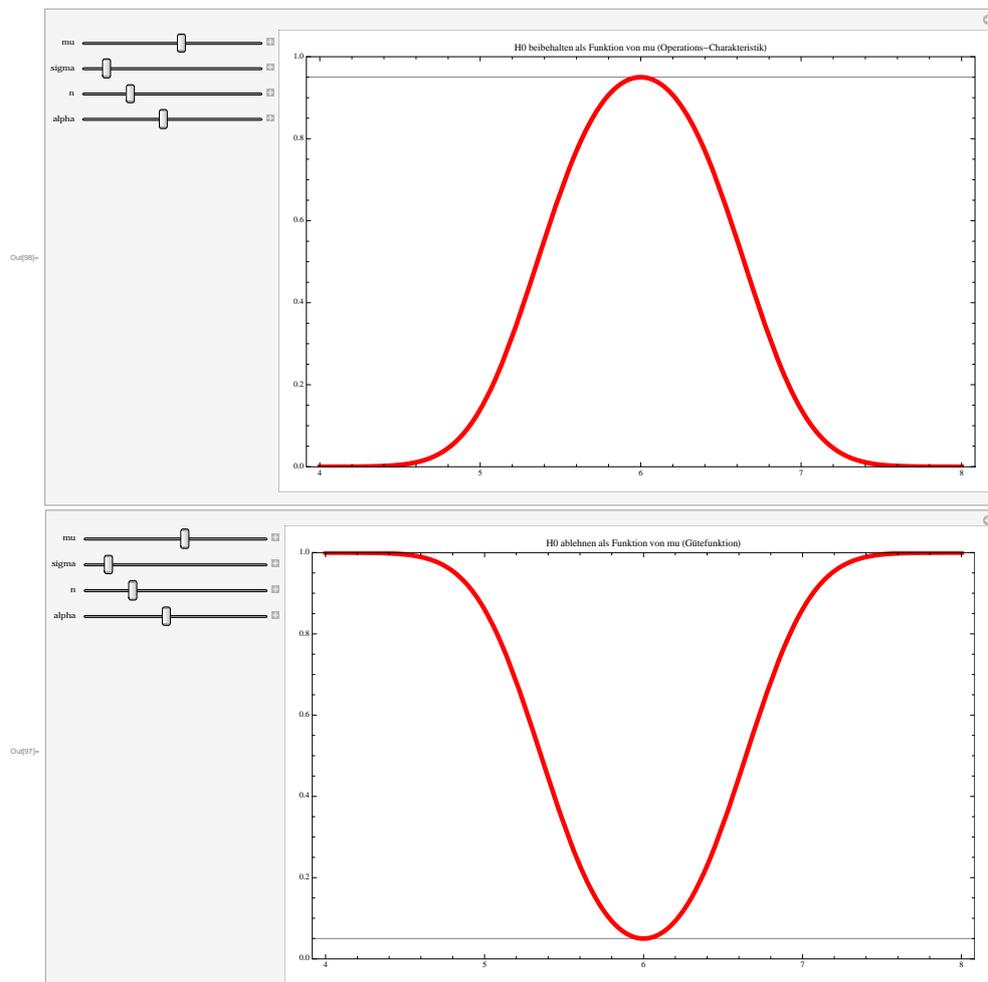


Abbildung 2.7:

oben: Wahrscheinlichkeit, H_0 beizubehalten als Funktion des wahren Parameters μ (Operations-Charakteristik).

unten: Wahrscheinlichkeit, H_0 abzulehnen als Funktion des wahren Parameters μ (Güte-Funktion).

http://www.fernuni-hagen.de/ls_statistik/lehre/

Gauß-Test).

Setzt man die kritischen Werte $[\mu_0 - z \sigma/\sqrt{N}, \mu_0 + z \sigma/\sqrt{N}] = [x_1, x_2]$, so gilt

$$\beta = P(T \in \bar{R} | \mu \in H_1) \quad (2.77)$$

β -Fehler

$$= \Phi(x_2; \mu, \sigma^2/N) - \Phi(x_1; \mu, \sigma^2/N) \quad (2.78)$$

$$= \Phi((x_2 - \mu)/(\sigma/\sqrt{N}); 0, 1) - \Phi((x_1 - \mu)/(\sigma/\sqrt{N}); 0, 1) \quad (2.79)$$

$$= \Phi(z_2 - \mu_z) - \Phi(z_1 - \mu_z) \quad (2.80)$$

Hierbei sind $z_1 = z(\alpha/2) = -z$, $z_2 = z(1 - \alpha/2) = z$ die standardisierten Quantile, $\mu_z = (\mu - \mu_0)/(\sigma/\sqrt{N})$ ist ein standardisierter Parameterwert und $\Phi(x; \mu, \sigma^2)$ ist die kumulierte Verteilungsfunktion der Normalverteilung.

Dies ist aber genau die gelb unterlegte Fläche in Abb. 2.6. Da der wahre Wert μ in Anwendungen unbekannt ist, kann man den Schätzwert $\hat{\mu} = \bar{X}$ einsetzen und erhält so $\hat{\mu}_z = (\bar{X} - \mu_0)/(\sigma/\sqrt{N})$ (standardisierte Teststatistik beim Gauß-Test).

Trägt man den β -Fehler gegen den wahren Wert des Parameters μ auf, so erhält man die sog. **Operations-Charakteristik OC** des Tests: die Wahrscheinlichkeit, H_0 beizubehalten als Funktion von μ . Außerhalb von H_0 ist dies gerade $\beta(\mu)$, innerhalb von H_0 ist es die Wahrscheinlichkeit einer richtigen Entscheidung.

Trägt man dagegen die Wahrscheinlichkeit, H_0 abzulehnen, als Funktion von μ auf, so ist dies die sog. **Güte-Funktion**

Güte-Funktion

$$g(\mu) = P(H_0 \text{ ablehnen} | \mu) \quad (2.81)$$

$$= 1 - P(H_0 \text{ beibehalten} | \mu) = 1 - \text{OC}(\mu) \quad (2.82)$$

Innerhalb von H_0 ist dies also $\alpha(\mu)$, außerhalb die Wahrscheinlichkeit einer richtigen Entscheidung (nämlich H_0 abzulehnen, wenn μ nicht in H_0 liegt). Die Kurven sind in Abb. 2.7 dargestellt. Bei Erhöhung des Stichprobenumfangs geht die Güte-Funktion schneller gegen 1, die Macht (power) des Tests wird größer. Dies ist auch sinnvoll, da ja Abweichungen der Daten vom hypothetischen Wert bei großen Stichproben schneller detektiert werden sollten (Abb. 2.8). Auf der anderen Seite führt die Verkleinerung des Fehlers 1. Art (kleines α) zu einer Vergrößerung der kriti-

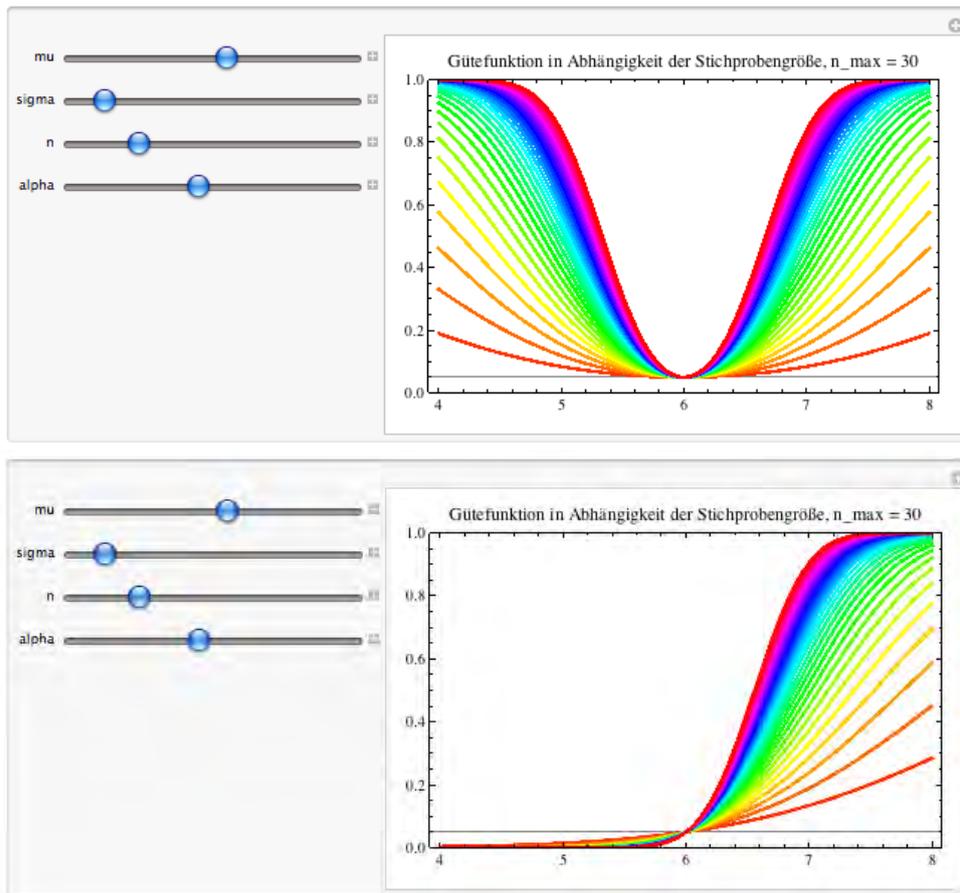


Abbildung 2.8: Güte-Funktion für verschiedene Stichprobenumfänge $n = 1, \dots, N$. oben: Zweiseitiger Test. unten: Einseitiger Test. Die Wahrscheinlichkeit, H_0 abzulehnen, steigt mit der Größe der Stichprobe.

http://www.fernuni-hagen.de/lis_statistik/lehre/

schen Werte $\mu_0 \pm z(1 - \alpha/2) \times \sigma / \sqrt{N}$ und somit bei gleicher Konstellation zu einem vergrößerten Fehler 2. Art (H_0 wird also eher beibehalten).

Man hat also das Dilemma (vgl. Tab. 2.2):

Je kleiner α gewählt wird, desto geringer ist die Wahrscheinlichkeit der Fehlentscheidung, H_0 abzulehnen, obwohl sie richtig ist (Fehler 1. Art). Andererseits steigt die Gefahr, H_0 beizubehalten, obwohl sie falsch ist (Fehler 2. Art)

2.5.4 Einseitige Hypothesen

Bisher wurden nur zweiseitige Hypothesen diskutiert, etwa $H_0 : \mu = \mu_0$ gegen $H_1 : \mu \neq \mu_0$. In manchen Fällen interessiert man sich für die einseitigen Hypothesenpaare (oder umgekehrt)

$$H_0 : \mu \leq \mu_0 \quad (2.83)$$

$$H_1 : \mu > \mu_0, \quad (2.84)$$

etwa

“die A-Partei erhält bei der nächsten Wahl $40 + x$ %”

Dann würde man für die unbekannte Wahrscheinlichkeit $\pi = P(\text{Person wählt die die A-Partei})$ formulieren:

$$H_0 : \pi \leq 0.4 \quad (2.85)$$

$$H_1 : \pi > 0.4 \quad (2.86)$$

Bei Ablehnung der Nullhypothese entscheidet man sich für die Aussage “die A-Partei erhält bei der nächsten Wahl mehr als 40%” und hat nur eine Irrtumswahrscheinlichkeit von α (etwa 5%).

Die α - und β -Fehler sowie Verläufe der Gütefunktion sind (für das BIP-Beispiel 2.12) in den Abb. 2.9–2.10 dargestellt. Der Ablehnungsbereich (rot unterlegt) ist nur einseitig (von der Nullhypothese abgewandt). Da die Nullhypothese eine größere Information beinhaltet, sind die β -Fehler bei gleichem α etwas kleiner als bei der zweiseitigen Alternative.

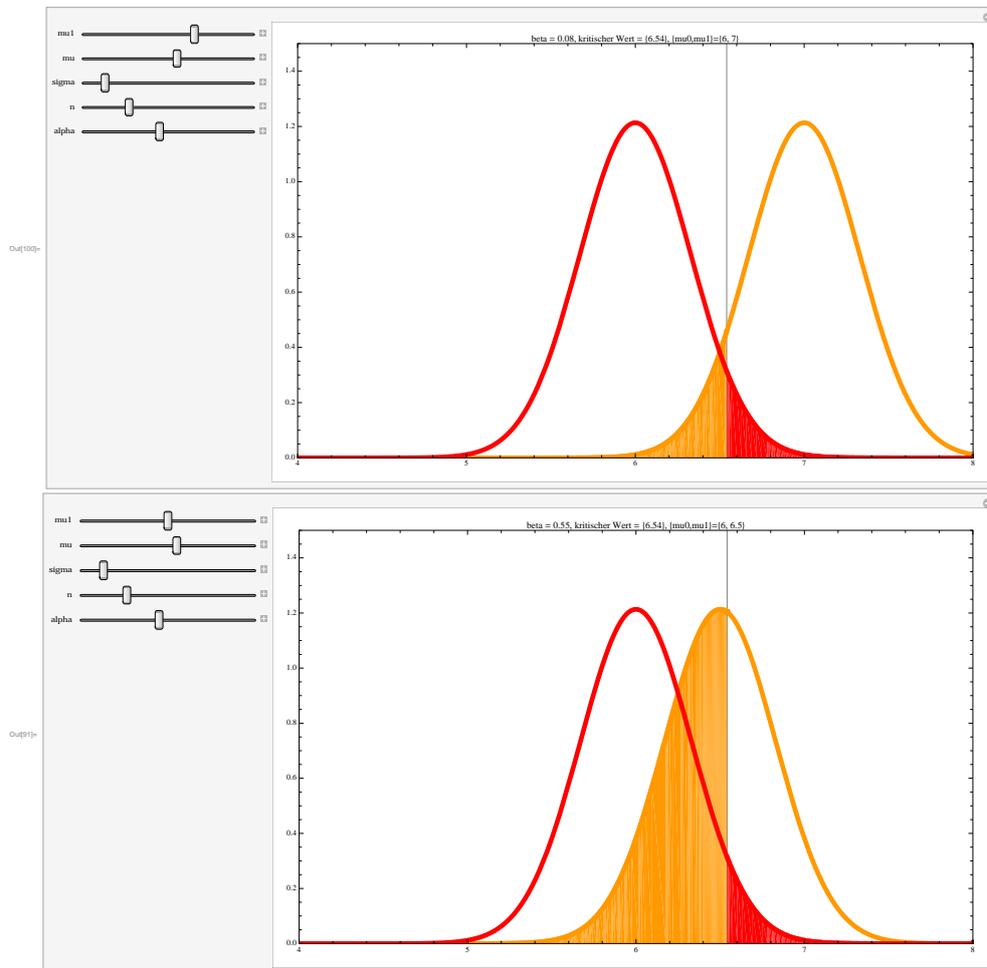


Abbildung 2.9: β -Fehler beim einseitigen Gauß-Test:

oben: $H_0 : \mu \leq 6$ und wahrer Parameter $\mu = 7 \in H_1$. Kritischer Wert: $\mu_0 + z(1 - \alpha) \times \sigma / \sqrt{N} = 6.54$ (senkrechte Linie). Der β -Fehler ist 0.08 (gelb unterlegt).

unten: $H_0 : \mu \leq 6$ und wahrer Parameter $\mu = 6.5 \in H_1$. β -Fehler = 0.55.

http://www.fernuni-hagen.de/ls_statistik/lehre/

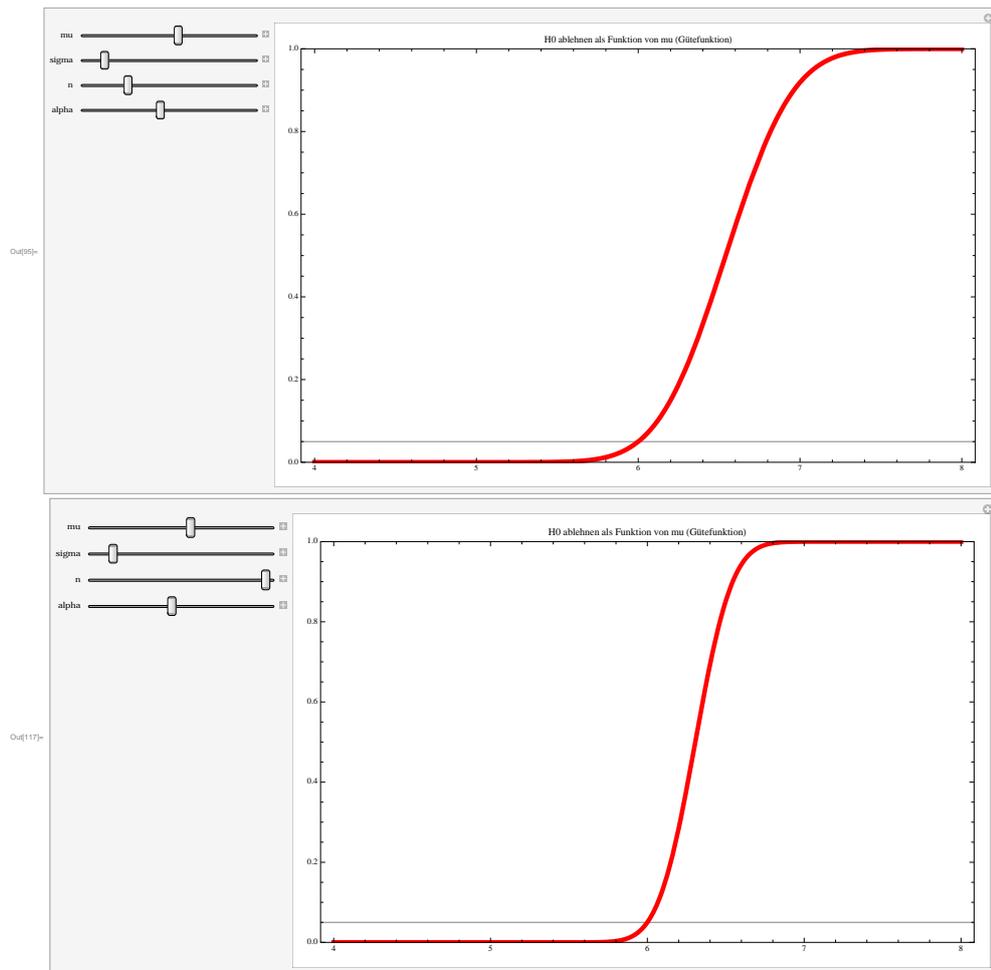


Abbildung 2.10: Gütefunktion (einseitiger Gauß-Test):
oben: Wahrscheinlichkeit, H_0 abzulehnen als Funktion des wahren Parameters μ (Güte-Funktion $g(\mu)$).
unten: Güte-Funktion für $N = 100$.

2.5.5 Allgemeine Gesichtspunkte beim Hypothesentest

Hier wird wieder allgemein θ als Name für den Parameter verwendet.

- zunächst wird der Fehler 1. Art (α) festgelegt (5%, 1%,...)
- Die Teststatistik T und der Verwerfungsbereich R werden so gewählt, daß der β -Fehler möglichst klein wird.
- Wenn sich die Nullhypothese aus mehreren Werten zusammensetzt (zusammengesetzte Hypothese, etwa $\theta \leq \theta_0$), gilt immer $\alpha(\theta) = P(H_1 \text{ annehmen} | \theta) \leq \alpha$.
- Jedoch kann der β -Fehler bei Nichtablehnung von H_0 sehr groß werden ($\beta < 1 - \alpha$).
- Daher besteht eine **Asymmetrie** zwischen H_0 und H_1 : α und β lassen sich nicht simultan fixieren.
- Die Entscheidung: $T \in \bar{R} \Rightarrow H_0$ beibehalten (nicht ablehnen) bedeutet nicht, daß H_0 bestätigt ist, sondern nur, daß die Daten zur Ablehnung nicht ausreichen ("im Zweifel für den Angeklagten").

**zusammengesetzte
Hypothese**

**Asymmetrie beim
Test**

Praktiker der empirischen Wirtschafts- und Sozialforschung wählen diejenige Hypothese als H_1 , die sie bestätigen wollen (H_0 als diejenige, die sie verwerfen wollen).

Ein Anhänger von Karl Popper müßte seine Hypothese als H_0 wählen, da er sie falsifizieren will (der Fehler 1. Art ist dann klein).

- Häufig werden nur "signifikante" Ergebnisse veröffentlicht, die anderen (bei denen "nichts herausgekommen" ist) werden vernachlässigt.

Kapitel 3

Tests bei speziellen Parametern

Nach den grundsätzlichen Überlegungen des letzten Kapitels werden nun eine Reihe von speziellen Testproblemen behandelt, wobei sich die eher theoretischen Gesichtspunkte konkretisieren.

3.1 Gesichtspunkte bei statistischen Tests

Zunächst können Testprobleme nach verschiedenen Gesichtspunkten klassifiziert werden.

- Zahl der Stichproben
(1 Stichprobe, 2 Stichproben (Männer, Frauen; Kontrollgruppe, Behandlungsgruppe), mehrere Stichproben (Designs))
- Unabhängige vs. abhängige Stichproben
etwa Meßwiederholungen: BIP(2007), BIP(2008)
- Meßniveau der Variablen (quantitativ vs. qualitativ)
- parametrisch vs. nichtparametrisch
ist der Verteilungstyp bekannt (bis auf Parameter)
oder nicht: nichtparametrische Tests
- Zusammenhangsanalyse: mehrdimensionale Merkmale
Test auf Unabhängigkeit

Kovarianz und Korrelation

Regressionsanalyse: Einfluß von $X \rightarrow Y: Y = \alpha + \beta X + \epsilon$

Test: $\beta = 0$: kein Einfluß

3.2 Ein-Stichproben-Fall

3.2.1 quantitative Variable $X : H_0 : \mu = \mu_0$ Prüfung des Erwartungswerts

Voraussetzungen	Prüfgröße	Prüfverteilung	Bezeichnung
Normalverteilung $X_n \sim N(\mu, \sigma^2)$ σ bekannt	$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{N}}$	$Z \sim N(0, 1)$	Ein-Stichproben- Gauß-Test
$X_n \sim N(\mu, \sigma^2)$ σ unbekannt	$T = \frac{\bar{X} - \mu_0}{S/\sqrt{N}}$	$T \sim t(N - 1)$	Ein-Stichproben- t - Test
keine Verteilungs- annahme für X_n σ bekannt, $N > 30$	$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{N}}$	$Z \sim N(0, 1)$ (approximativ)	approximativer Gauß-Test
keine Verteilungs- annahme für X_n σ unbekannt, $N > 30$	$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{N}}$	$Z \sim N(0, 1)$ (approximativ)	approximativer Gauß-Test

Tabelle 3.1: Ein-Stichproben-Tests für den Erwartungswert.

Der Hypothesentest erfolgt wie bereits besprochen durch Berechnung der kritischen Grenzen, die durch die z - oder t -Quantile gegeben sind (vgl. Abb. 3.1): **2-seitiger Test:**

**Test des
Erwartungswerts**

$$H_0 : \mu = \mu_0 \quad (3.1)$$

$$H_1 : \mu \neq \mu_0 \quad (3.2)$$

$$\text{berechne } z = z(1 - \alpha/2) \text{ oder } t = t(1 - \alpha/2, N - 1) \quad (3.3)$$

$$\text{falls } |Z| > z \text{ bzw. } |T| > t : H_0 \text{ ablehnen} \quad (3.4)$$

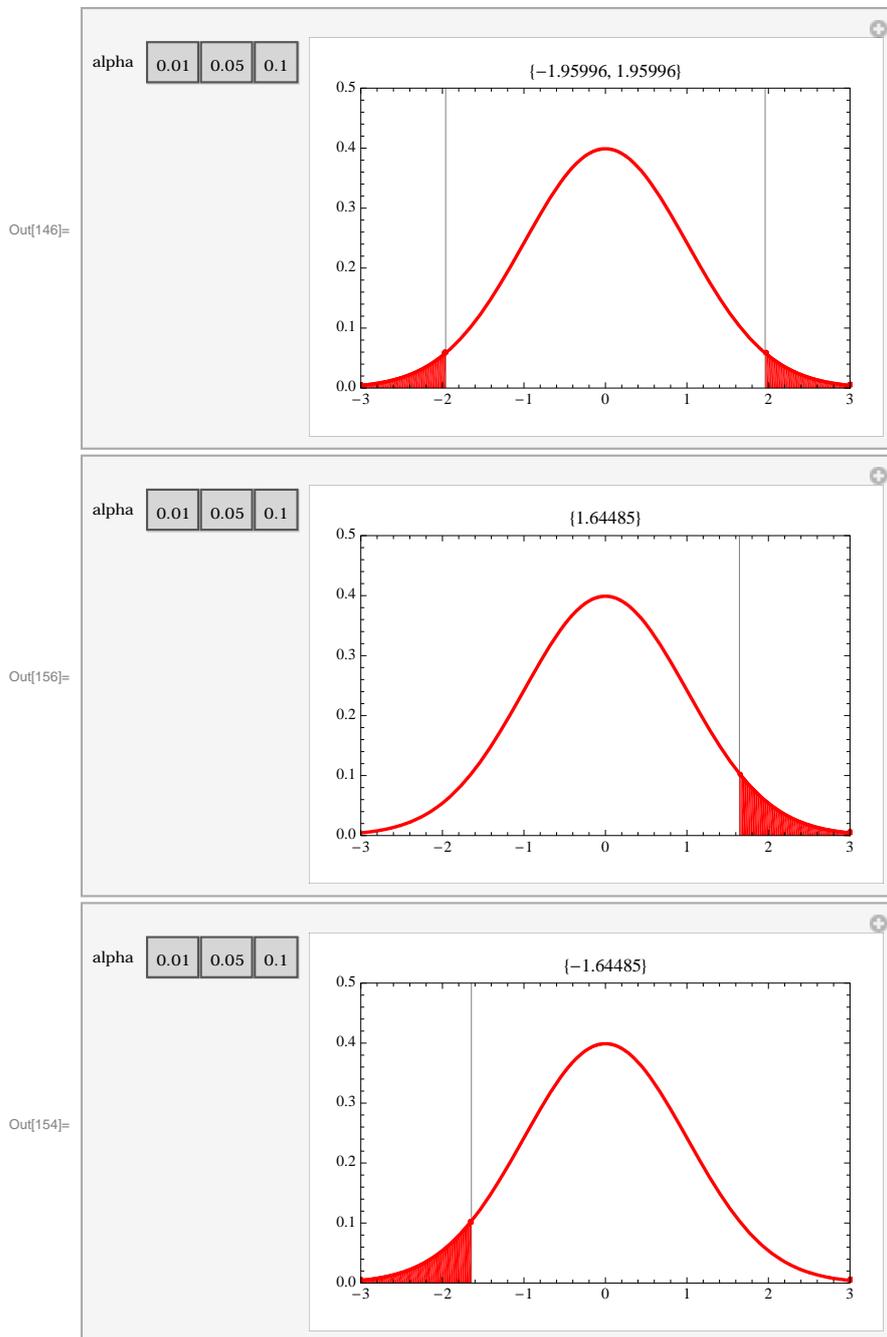


Abbildung 3.1: Ablehnungsbereiche für den Mittelwertstest ($\alpha = 0.05$, Gauß-Test).

1-seitiger Test; kritischer Wert rechts:

$$H_0 : \mu \leq \mu_0 \quad (3.5)$$

$$H_1 : \mu > \mu_0 \quad (3.6)$$

$$\text{berechne} \quad z = z(1 - \alpha) \text{ oder } t = t(1 - \alpha, N - 1) \quad (3.7)$$

$$\text{falls} \quad Z > z \text{ bzw. } T > t : H_0 \text{ ablehnen} \quad (3.8)$$

1-seitiger Test; kritischer Wert links:

$$H_0 : \mu \geq \mu_0 \quad (3.9)$$

$$H_1 : \mu < \mu_0 \quad (3.10)$$

$$\text{berechne} \quad z = z(\alpha) \text{ oder } t = t(\alpha, N - 1) \quad (3.11)$$

$$\text{falls} \quad Z < z \text{ bzw. } T < t : H_0 \text{ ablehnen} \quad (3.12)$$

3.2.2 quantitative Variable $X : H_0 : \sigma = \sigma_0$ Prüfung der Varianz

Meistens interessieren sich Praktiker nur für den Mittelwert einer Variable. Genausogut ist aber auch der Streubereich von Interesse ($s = 1.86$, Bereich des BIP-Wachstums; vgl. Abb. 2.5).

Dann ist ein Test für σ gefragt. Als Teststatistik bietet sich die Stichprobenvarianz S^2 an, die χ^2 -verteilt ist (Tab. 3.2). Die Ablehnungsbereiche sind Abb. 3.2 zu entnehmen.

Test der Varianz

Ausführlich lauten die Hypothesenpaare:

2-seitiger Test:

$$H_0 : \sigma = \sigma_0 \quad (3.13)$$

$$H_1 : \sigma \neq \sigma_0 \quad (3.14)$$

$$\text{berechne} \quad c_1 = \chi^2(\alpha/2, N - 1), c_2 = \chi^2(1 - \alpha/2, N - 1) \quad (3.15)$$

$$\text{falls} \quad \chi^2 < c_1 \text{ bzw. } \chi^2 > c_2 : H_0 \text{ ablehnen} \quad (3.16)$$

1-seitiger Test; kritischer Wert oben:

$$H_0 : \sigma \leq \sigma_0 \quad (3.17)$$

$$H_1 : \sigma > \sigma_0 \quad (3.18)$$

$$\text{berechne} \quad c_2 = \chi^2(1 - \alpha, N - 1) \quad (3.19)$$

$$\text{falls} \quad \chi^2 > c_2 : H_0 \text{ ablehnen} \quad (3.20)$$

1-seitiger Test; kritischer Wert unten:

$$H_0 : \sigma \geq \sigma_0 \quad (3.21)$$

$$H_1 : \sigma < \sigma_0 \quad (3.22)$$

$$\text{berechne } c_1 = \chi^2(\alpha, N - 1) \quad (3.23)$$

$$\text{falls } \chi^2 < c_1 : H_0 \text{ ablehnen} \quad (3.24)$$

Voraussetzungen	Prüfgröße	Prüfverteilung	Bezeichnung
$X_n \sim N(\mu, \sigma^2)$	$\chi^2 = \frac{1}{\sigma_0^2} \sum (X_n - \bar{X})^2$ $= \frac{(N-1)S^2}{\sigma_0^2}$	$\chi^2(N - 1)$	Ein-Stichproben- Chi-Quadrat-Test

Tabelle 3.2: Ein-Stichproben-Test für die Varianz.

Beispiel 3.1 (BIP-Wachstum: Varianz-Test)

Die BIP-Wachstumsraten 2007 aus Bsp. 3.10 weisen eine Standardabweichung von $s = 1.86\%$ auf ($N = 32$).

Man kann nun die Null-Hypothese prüfen, ob der wahre Streubereich $\sigma = \sigma_0 = 2\%$ ist.

Teststatistik:

$$\chi^2 = \frac{1}{\sigma_0^2} \sum (X_n - \bar{X})^2 = \frac{(N-1)S^2}{\sigma_0^2} = \frac{(32-1)1.86^2}{2^2} = 26.8119$$

Quantile ($\alpha = 0.05$, $df = 31$):

$c_1 = 17.53$; $c_2 = 48.2319$ (Abb. 3.2, oben)

Da $26.8119 \in [17.53, 48.2319]$ kann die Null-Hypothese $H_0 : \sigma = \sigma_0 = 2$ nicht abgelehnt werden.

■

3.2.3 Dichotome Variable X : Prüfung von Anteilswerten

In diesem Fall werden als Stichprobenvariablen die Indikatoren X_n betrachtet, die das Vorliegen der Eigenschaft A anzeigen:

$$X_n = \begin{cases} 1 & \text{wenn } A \\ 0 & \text{sonst} \end{cases} \quad (3.25)$$

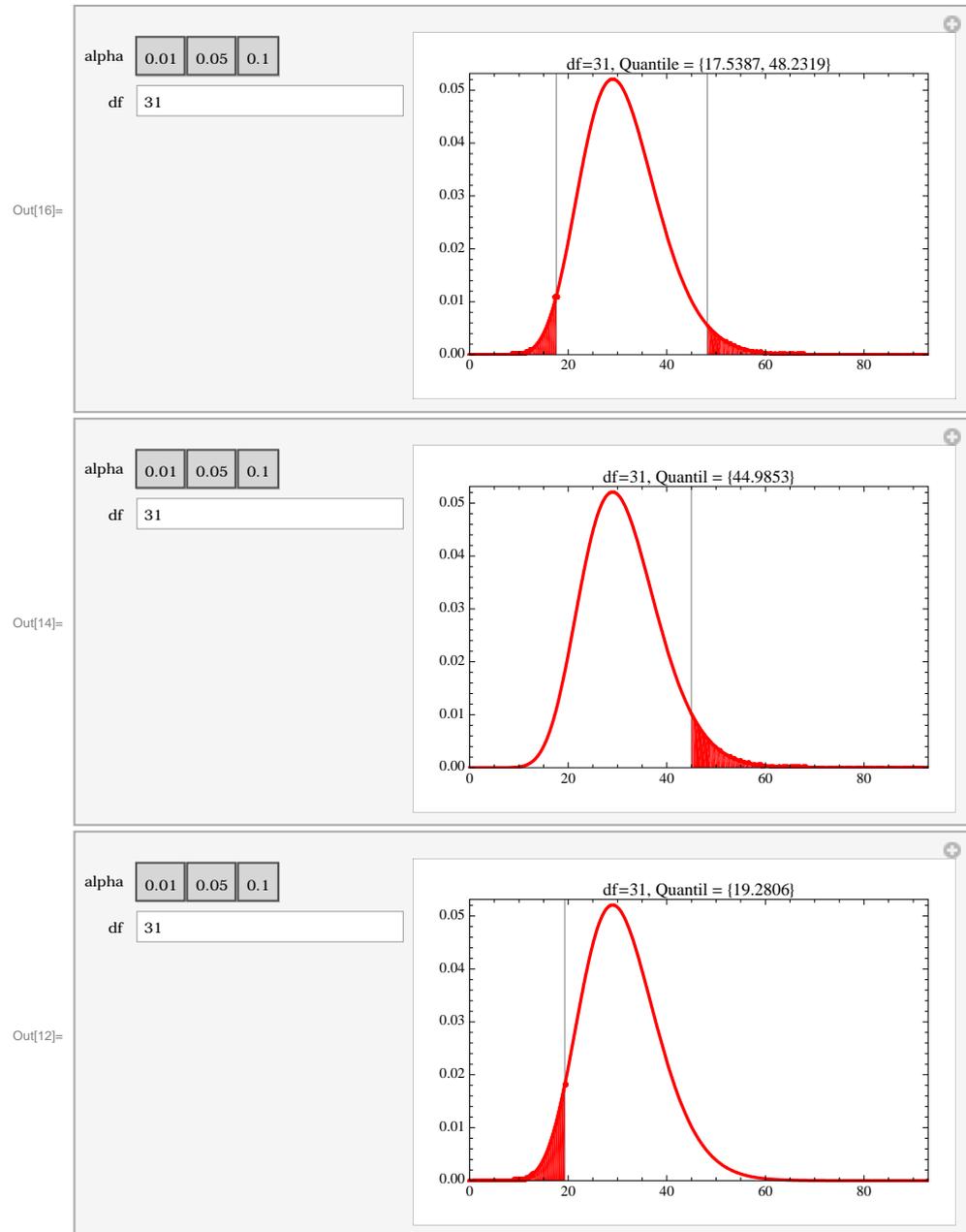


Abbildung 3.2: Ablehnungsbereiche für den Varianz-Test ($\alpha = 0.05$). Testverteilung $\chi^2(N - 1)$.

(Bsp.: $X_n = 1$, wenn die n -te Person die A -Partei wählt, 0 sonst; vgl. Bsp. 1.2).

Der Mittelwert

$$\bar{X} = \frac{1}{N} \sum_{n=1}^N X_n = f(A) = \hat{\pi} \quad (3.26)$$

ist gleich der relativen Häufigkeit der Eigenschaft A und ein Schätzer für die unbekannte Wahrscheinlichkeit $\pi = P(A) = E[\hat{\pi}]$.

Auch hier kann man zweiseitige und einseitige Hypothesen formulieren:

2-seitiger Test:

$$H_0 : \pi = \pi_0 \quad (3.27)$$

$$H_1 : \pi \neq \pi_0 \quad (3.28)$$

1-seitiger Test; kritischer Wert rechts:

$$H_0 : \pi \leq \pi_0 \quad (3.29)$$

$$H_1 : \pi > \pi_0 \quad (3.30)$$

1-seitiger Test; kritischer Wert links:

$$H_0 : \pi \geq \pi_0 \quad (3.31)$$

$$H_1 : \pi < \pi_0. \quad (3.32)$$

Da \bar{X} als Summe von Indikatoren binomialverteilt ist, kann ein exakter Test formuliert werden. Bei genügend großen Stichproben kann aber die Binomialverteilung durch die Normalverteilung approximiert werden.

Für $N\pi_0 \geq 5$ und $N(1 - \pi_0) \geq 5$ gilt näherungsweise (asymptotisch verteilt $\stackrel{a}{\sim}$):

$$\bar{X} = \frac{1}{N} \sum_{n=1}^N X_n = \hat{\pi} \stackrel{a}{\sim} N\left(\pi, \frac{1}{N}\pi(1 - \pi)\right). \quad (3.33)$$

Daher ist unter H_0 :

$$Z = \frac{\hat{\pi} - \pi_0}{\sqrt{\frac{1}{N}\pi_0(1 - \pi_0)}} \stackrel{a}{\sim} N(0, 1) \quad (3.34)$$

**Test des
Anteilswerts**

Die kritischen Bereiche sind wie beim Gauß-Test (vgl. Abb. 3.1).

Beispiel 3.2 (Politbarometer)

Ein Politiker behauptet, daß seine Partei bei der nächsten Wahl mehr als 30% der Stimmen erhält. Bei einer Wahlumfrage ergab sich aber ein Prozentwert von 25% (Abb. 2.3).

Will man diese Behauptung widerlegen, erhält man das Hypothesenpaar
1-seitiger Test; kritischer Wert links:

$$H_0 : \pi \geq 0.3 \quad (3.35)$$

$$H_1 : \pi < 0.3. \quad (3.36)$$

Die Teststatistik ist somit

$$z = \frac{.25 - .3}{\sqrt{\frac{1}{1245} \cdot .3(1 - .3)}} = -3.85 \quad (3.37)$$

Das Quantil $z(0.05) = -1.65$ (Abb. 3.1) ist größer als die Teststatistik, daher muß die Aussage des Politikers auf dem 5%-Niveau abgelehnt werden. Da $z(.01) = -2.33$ ist, wäre die Hypothese auch auf dem 1%-Niveau abgelehnt worden.

In der Tat ist der p -Wert (Überschreitungswahrscheinlichkeit)

$$p = P(Z < z) = P(Z < -3.85) = 1 - \Phi(3.85) \quad (3.38)$$

$$= 1 - 0.999941 = 0.0000590925 < \alpha. \quad (3.39)$$

Man hätte also auch noch kleinere Signifikanzniveaus nehmen können. α muß aber **vor dem Test** auf einem bestimmten Wert fixiert werden.

Übung:

Wie müßte der Politiker die Hypothesen formulieren, um seine Behauptung zu prüfen ($\alpha = 0.05$). Führen Sie den Test durch.



3.2.4 Nichtparametrische Methoden

Bisher wurden für die Stichprobenvariablen X bestimmte Verteilungen vorausgesetzt, z.B. $N(\mu, \sigma^2)$, $B(N, \pi)$, die Multinomialverteilung oder ein großer Stichprobenumfang angenommen (zentraler Grenzwertsatz: \bar{X} approximativ normalverteilt).

Wenn der Verteilungstyp der Stichprobenvariablen nicht bekannt ist, bestimmt man eine Testfunktion T , deren Verteilung nicht von der Verteilung von X abhängt.¹

Dies führt auf die sogenannten **nichtparametrischen** (verteilungsfreien) Verfahren.

**nichtparametrische
Verfahren**

3.2.4.1 Vorzeichentest

stetiges Merkmal X : keine Bindungen (ties)

Für den Median $x(0.5)$ der unbekanntenen Verteilung $f(x)$ des Merkmals X gilt die Aussage

$$P(X < x(0.5)) = 0.5 \quad (3.40)$$

$$P(X > x(0.5)) = 0.5 \quad (3.41)$$

$$P(X = x(0.5)) = 0, \quad (3.42)$$

da es sich um ein stetiges Merkmal handelt (Punkte haben Wahrscheinlichkeit 0).

Eine Nullhypothese kann für den unbekanntenen Wert des Medians formuliert werden:

2-seitiger Test:

$$H_0 : x(0.5) = \delta_0 \quad (3.43)$$

$$H_1 : x(0.5) \neq \delta_0 \quad (3.44)$$

**Test des Medians
(Vorzeichentest)**

1-seitiger Test; kritischer Wert rechts:

$$H_0 : x(0.5) \leq \delta_0 \quad (3.45)$$

$$H_1 : x(0.5) > \delta_0 \quad (3.46)$$

1-seitiger Test; kritischer Wert links:

$$H_0 : x(0.5) \geq \delta_0 \quad (3.47)$$

$$H_1 : x(0.5) < \delta_0. \quad (3.48)$$

Aus der Stichprobe X_1, \dots, X_N wird die Teststatistik

¹Nur in diesem Sinne ist der Test verteilungsfrei.

$$T = \text{„Anzahl der Werte rechts von } \delta_0\text{“} = \sum_n Z_n \quad (3.49)$$

$$Z_n = \begin{cases} 1 & \text{wenn } X_n > \delta_0 \\ 0 & \text{sonst} \end{cases} \quad (3.50)$$

berechnet.

Dann sind Z_n Indikatoren und die Summenvariable T ist unter H_0 binomialverteilt $B(N, 0.5)$, da die Wahrscheinlichkeit, links oder rechts vom Median zu liegen, gleich ist, also

$$P(Z_n = 1) = P(X_n > \delta_0) = 0.5. \quad (3.51)$$

Für genügend großes N ist T näherungsweise normalverteilt mit Parametern

$$E[T] = \sum_n E[Z_n] = N/2 \quad (3.52)$$

$$\text{Var}[T] = \sum_n \text{Var}[Z_n] = N \cdot 0.5(1 - 0.5) = N/4 \quad (3.53)$$

Daher ist unter H_0

$$Z = \frac{T - N/2}{\sqrt{N/4}} \stackrel{a}{\sim} N(0, 1). \quad (3.54)$$

Die kritischen Bereiche sind wie beim Gauß-Test (vgl. Abb. 3.1).

beliebiges Merkmal X : mit Bindungen (ties)

Bei beliebigen Merkmalen, die etwa ordinalskaliert sind, kann

$$P(X = x(0.5)) \neq 0 \quad (3.55)$$

sein. Im obigen Abschnitt trat dieser Fall nur mit Wahrscheinlichkeit 0 auf. Entsprechend kommen die Werte mit $X_n = \delta_0$ in der Teststatistik (3.49) nicht vor.

Man kann nun so vorgehen, daß die Werte $X_n = \delta_0$ weggelassen werden. Dann wird das Signifikanzniveau nicht überschritten und es gilt

$$P(T \in R) \leq \alpha \quad (3.56)$$

(vgl. Bosch, 1992, Kap. XI.2.2). Es handelt sich daher um einen sog. **konservativen Test**.

Beispiel 3.3 (Bruttoinlandsprodukt BIP)

Das nominale Wachstum des Bruttoinlandsprodukts in Prozent im Jahr 2007 für verschiedene OECD-Länder hat folgende Werte (Abb. 2.4; Datensatz `Regression.jmp`). Es war $\bar{x} = 5.53$, $s = 1.86$, $N = 32$. Der Median sowie die unteren und oberen Quantile lauten: $x(0.5) = 5.32$, $x(0.25) = 4.10$, $x(0.75) = 6.73$.

Unter Normalverteilungsannahme wurde in Bsp. 2.12 ein Gauß-Test für die Hypothese $H_0 : \mu = 6$ durchgeführt. Die Nullhypothese mußte beibehalten werden ($\alpha = 0.05$).

Wenn man keine Verteilungsannahmen machen will, kann mit Hilfe des Vorzeichentests die Hypothese

$$H_0 : x(0.5) = \delta_0 = 6 \quad (3.57)$$

$$H_1 : x(0.5) \neq \delta_0 = 6 \quad (3.58)$$

für den Median überprüft werden. Man muß nur feststellen, wieviele Daten oberhalb von 6% liegen. Eine Inspektion von Abb. 2.4 ergibt $t = 12$. Daraus berechnet man die Teststatistik

$$z = \frac{12 - 32/2}{\sqrt{32/4}} = -1.41 \quad (3.59)$$

Die kritischen Werte $z(\alpha/2)$, $z(1 - \alpha/2)$ lauten -1.96 , 1.96 (Beibehaltungsbereich $\bar{R} = [-1.96, 1.96]$). Also muß auch hier die H_0 beibehalten werden.

Der p -Wert lautet: $p = P(|Z| > |z|) = P(|Z| > 1.41) = 2(1 - \Phi(|z|)) = 0.157$, da $\Phi(|z|) = .921$.

Man berechnet also die Wahrscheinlichkeit, daß die (betragsmäßige) Teststatistik außerhalb des konkreten Testwerts $|z| = 1.41$ liegt. Wenn $p < \alpha$, wird die H_0 abgelehnt.



Übung 3.1 (Bruttoinlandsprodukt BIP)

Prüfen Sie bitte die einseitige Hypothese

$$H_0 : x(0.5) \geq 6 \quad (3.60)$$

$$H_1 : x(0.5) < 6 \quad (3.61)$$

mit dem Gauß-Test und dem Vorzeichentest und berechnen Sie die p -Werte (beim einseitigen Test $P(Z < z)$).

**3.2.4.2 Wilcoxon-Vorzeichen-Rangtest (Symmetrietest)**

Beim Vorzeichentest wird nur der Median getestet, d.h. es wird nur geprüft, ob die Fläche unter der Dichtefunktion $f(x)$ links und rechts vom hypothetischen Median δ_0 gleich ($= 0.5$) ist. Die Form der Verteilung kann auf beiden Hälften völlig verschieden aussehen.

Beim Wilcoxon-Vorzeichen-Rangtest wird dagegen geprüft:

$$H_0 : f(x) \text{ ist symmetrisch zum Median } x(0.5) = \delta_0 \quad (3.62)$$

Auch hier müssen wieder die Fälle mit und ohne Bindungen unterschieden werden.

stetiges Merkmal X : keine Bindungen (ties)

Für den Median $x(0.5)$ der unbekanntten Verteilung $f(x)$ des Merkmals X gilt die Aussage

$$P(X < x(0.5)) = 0.5 \quad (3.63)$$

$$P(X > x(0.5)) = 0.5 \quad (3.64)$$

$$P(X = x(0.5)) = 0, \quad (3.65)$$

da es sich um ein stetiges Merkmal handelt (Punkte haben Wahrscheinlichkeit 0).

Die Testgröße nach Wilcoxon lautet

$$W^+ = \sum_n R_n Z_n \quad (3.66)$$

wobei

$$R_n = \text{rg}|X_n - \delta_0| \quad (3.67)$$

$$Z_n = \begin{cases} 1 & \text{wenn } X_n > \delta_0 \\ 0 & \text{sonst} \end{cases} \quad (3.68)$$

**Wilcoxon-
Vorzeichen-
Rangtest**

Hierbei versteht man unter $\text{rg}(x)$ den Rang des Datenpunkts (Index in der geordneten Liste). Da in der Teststatistik nur die Werte rechts von δ_0 gezählt werden, ergeben sich große oder kleine Werte bei einer asymmetrischen Verteilung.

Für den Erwartungswert und die Varianz der Teststatistik findet man unter H_0 :

$$E[W^+] = \sum_n n E[Z_n] \quad (3.69)$$

$$= \frac{1}{2} \sum_n n = \frac{1}{2} \frac{N(N+1)}{2} = \frac{N(N+1)}{4}. \quad (3.70)$$

In der oberen Zeile wurde ausgenutzt, daß alle Ränge zwischen $1, \dots, N$ in der Summe vorkommen und die Summenformel von Gauß verwendet (Übung). Außerdem gilt $E[Z_n] = \frac{1}{2}$ bei Gültigkeit der Nullhypothese.

Mit dem gleichen Argument findet sich ($Z_n^2 = Z_n$)

$$\text{Var}[W^+] = \sum_n n^2 \text{Var}[Z_n] = \sum_n n^2 (E[Z_n^2] - E[Z_n]^2) \quad (3.71)$$

$$= \frac{1}{4} \sum_n n^2 = \frac{N(N+1)(2N+1)}{24}. \quad (3.72)$$

Für genügend großes $N > 20$ (zentraler Grenzwertsatz) ist also

$$Z = \frac{W^+ - E[W^+]}{\sqrt{\text{Var}[W^+]}} \stackrel{a}{\sim} N(0, 1) \quad (3.73)$$

Der kritische Bereich ist also $|Z| > z(1 - \alpha/2)$ beim zweiseitigen Test. Ausgedrückt in der Wilcoxon-Teststatistik gilt also die Ablehnungsbedingung

$$|W^+| > E[W^+] + z(1 - \alpha/2)\sqrt{\text{Var}[W^+]} \quad (3.74)$$

Mit kombinatorischen Methoden kann auch eine exakte Verteilung für W^+ hergeleitet werden (vgl. Bosch, Kap. XI.5).

Beim exakten Wilcoxon-Test benutzt man die Quantile

$$w^+(\alpha/2) \quad (3.75)$$

$$w^+(1 - \alpha/2) = \frac{N(N+1)}{2} - w^+(\alpha/2) \quad (3.76)$$

(Tabelle 17.6) mit den Ablehnungsbereichen

$$W^+ < w^+(\alpha/2) \text{ oder } W^+ > w^+(1 - \alpha/2) \quad (3.77)$$

$$W^+ < w^+(\alpha) \quad (3.78)$$

$$W^+ > w^+(1 - \alpha) \quad (3.79)$$

Beispiel 3.4 (Intelligenztest)

Im Rahmen einer Marketing-Studie wurde auch ein Intelligenztest an einer Gruppe von Probanden durchgeführt. Es ergaben sich folgende Resultate (Tab. 3.3). Es soll die Hypothese geprüft werden, daß in der Grund-

n	1	2	3	4	5	6	7
x_n	112	168	97	119	187	136	109
z_n	0	1	0	0	1	1	0
$ x_n - \delta_0 $	8	48	23	1	67	16	11
$r_n = \text{rg} x_n - \delta_0 $	2	6	5	1	7	4	3

Tabelle 3.3: IQ-Daten und Wilcoxon-Hilfsgrößen.

gesamtheit der Probanden eine Intelligenz mit Median $x(0.5) = 120$ (d.h. $P(X < 120) = P(X > 120) = 0.5$) besteht. Daraus errechnet man die Teststatistik

$$w^+ = \sum_n r_n z_n = 6 + 7 + 4 = 17. \quad (3.80)$$

Aus der Tabelle (Anhang, Tab. 17.6) ergeben sich die kritischen Werte für $\alpha = 0.05, N = 7$:

$$w^+(\alpha/2) = 3, w^+(1 - \alpha/2) = N(N + 1)/2 - w^+(\alpha/2) = 28 - 3 = 25.$$

Damit muß H_0 beibehalten werden.

Zum Vergleich soll auch der approximative (asymptotische) Test durchgeführt werden (jedoch nur $N = 7 < 25$):

$$z(1 - \alpha/2) = 1.96,$$

$$E[W^+] = N(N + 1)/4 = 14,$$

$$\text{Var}[W^+] = \frac{N(N+1)(2N+1)}{24} = 35.$$

Daraus findet man die kritischen Grenzen

$$E[W^+] \pm z(1 - \alpha/2)\sqrt{\text{Var}[W^+]} = 14 \pm 1.96\sqrt{35} = 14 \pm 11.6$$

Da $w^+ = 17$ innerhalb dieser Grenzen liegt (also in \bar{R}), behält man H_0 bei.

■

beliebiges Merkmal X : mit Bindungen (ties)

In diesem Fall können keine eindeutigen Ränge vergeben werden, da gleiche Werte vorkommen. Man geht so vor, daß Durchschnittsränge für gleiche Werte vergeben werden und Daten, die auf den (hypothetischen) Median fallen, weggelassen werden (vgl. Bosch, 1992, Kap. XI, S. 696).

3.2.5 Prüfung auf eine bestimmte Verteilung: χ^2 -Anpassungstest

Wir betrachten Daten mit beliebigem Meßniveau (nominal, ordinal, metrisch; vgl. Kap. 14). Diese werden in l verschiedene Klassen eingeteilt (Tabelle 3.4). Etwa würde man bei einer quantitativen Variable wie

Klasse	1	2	3	4	...	l
Anzahl	n_1	n_2	n_3	n_4	...	n_l

Tabelle 3.4: Klasseneinteilung der Daten ($\sum_i n_i = N$).

beim Zeichnen eines Histogramms die x -Achse in Bereiche (Klassen)

$K_i = (x_i, x_{i+1}]$ oder $K_i = [x_i, x_{i+1})$ unterteilen und die Anzahl der Datenpunkte in jedem Intervall zählen.

Zur Prüfung auf die Verteilung $F(x) = P(X \leq x)$ formuliert man folgendes Hypothesenpaar:

Anpassungstest

$$H_0 : F = F_0 \quad (3.81)$$

$$H_1 : F \neq F_0. \quad (3.82)$$

Man vergleicht nun die theoretische Wahrscheinlichkeit (unter H_0)

$$\pi_i = P(X \in \text{Klasse } i) = P(X \in (x_i, x_{i+1}]) \quad (3.83)$$

$$= P(x_i < X \leq x_{i+1}) \quad (3.84)$$

mit den empirischen Häufigkeiten $f_i = n_i/N$ (bzw. n_i mit $N\pi_i$).

Die χ^2 -Teststatistik lautet

$$\chi^2 = \sum_{i=1}^l \frac{(n_i - N\pi_i)^2}{N\pi_i} = N \sum_{i=1}^l \frac{(f_i - \pi_i)^2}{\pi_i} \quad (3.85)$$

($E[n_i] = N\pi_i =$ theoretische Klassenhäufigkeit unter H_0). Die Abweichungen werden also quadriert und aufsummiert.

Es gilt für große Stichproben:

- Die χ^2 -Teststatistik ist asymptotisch $\chi^2(l-1)$ -verteilt,
- wobei gelten muß:
 - $N\pi_i \geq 1$
 - $N\pi_i \geq 5$ für mindestens 80% der Klassen.
- Falls dies nicht erfüllt ist: Klassen zusammenfassen oder N vergrößern.

Bemerkung zur Zahl der Freiheitsgrade:

- Da $\sum_{i=1}^l n_i = N$ hat $\chi^2(l-1)$ nur $l-1$ Freiheitsgrade (nur $l-1$ Häufigkeiten sind frei).

- wenn F unbekannte Parameter enthält, die geschätzt werden müssen (etwa μ und σ in der Normalverteilung $N(\mu, \sigma^2)$), so gilt für die Freiheitsgrade:

$$df = l - 1 - \text{Zahl der geschätzten Parameter}$$

Besonderheit beim Anpassungstest:

- Üblicherweise wird die Hypothese, die man bestätigen möchte, als die Alternative H_1 formuliert, da bei Ablehnung der H_0 nur ein Fehler 1. Art der Größe α zu befürchten ist.
- Hier will man aber die H_0 bestätigen, daß die Daten wie F_0 verteilt sind. Etwa möchte man herausfinden, ob die Daten normalverteilt sind, um eine auf dieser Verteilung beruhende Testprozedur zu benutzen (etwa ein F -Test, etc.)
- Bei Beibehaltung der H_0 muß mit einem großen β -Fehler gerechnet werden, der aufgrund der unspezifischen Alternative $F \neq F_0$ jedoch nicht berechnet werden kann.

Beispiel 3.5 (BIP: Test auf Normalverteilung)

Das nominale Wachstum des Bruttoinlandsprodukts im Jahr 2007 für verschiedene OECD-Länder hat folgende Werte (Abb. 2.4; Datensatz Regression.jmp). Ein Histogramm mit angepaßter Gauß-Verteilung sowie einem Stamm-Blatt-Diagramm sind in Abb. 3.3 zu sehen. Die Frage ist, ob die Normalverteilung eine gute Approximation darstellt. Mit Hilfe einer Klasseneinteilung kann der χ^2 -Anpassungstest für die Hypothese

$$H_0 : F = F_0 = N(\mu, \sigma^2) = \Phi(x; \mu, \sigma^2) \quad (3.86)$$

$$H_1 : F \neq F_0. \quad (3.87)$$

wie folgt angewendet werden:

Aus dem stem-leaf-Diagramm entnimmt man die Häufigkeiten für Klassen mit der Breite 1 (%) sowie den Randklassen $(-\infty, 3)$, $[9, \infty)$, also $j = 8$ Klassen. Daraus ergeben sich die Klassenhäufigkeiten (Tab. 3.5). Zur Berechnung der theoretischen Häufigkeiten benötigt man die Wahrscheinlichkeiten

$$\pi_i = P(X \in \text{Klasse } i) = P(X \in [x_i, x_{i+1})) \quad (3.88)$$

$$= P(x_i \leq X < x_{i+1}) = \Phi(x_{i+1}) - \Phi(x_i). \quad (3.89)$$

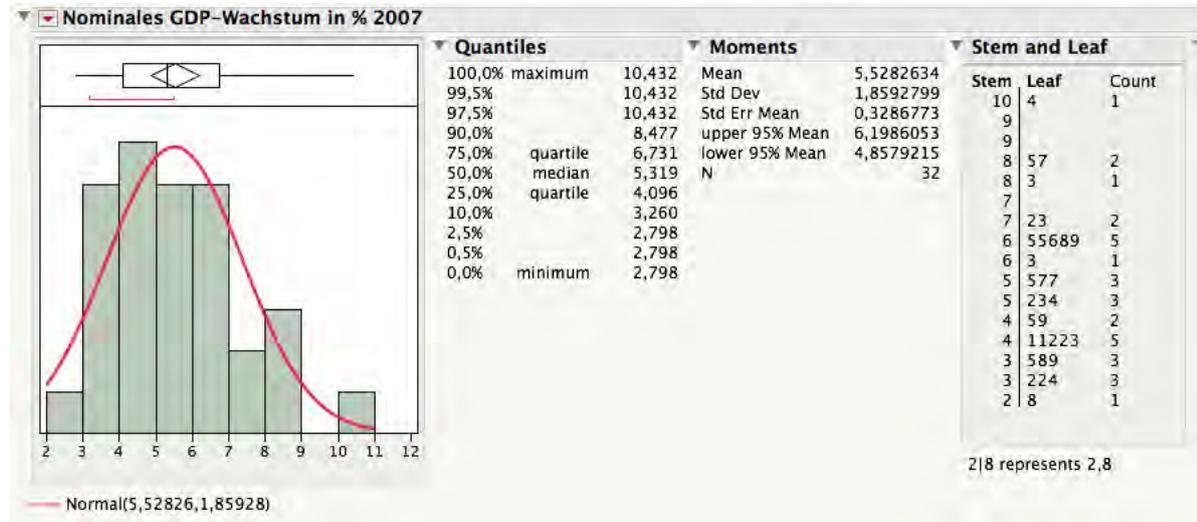


Abbildung 3.3: Wachstum des Bruttoinlandsprodukts im Jahr 2007: Verteilung und Statistiken.

Klasse K_i	$(-\infty, 3)$	$[3,4)$	$[4,5)$	$[5,6)$	$[6,7)$	$[7,8)$	$[8,9)$	$[9, \infty)$
Anzahl n_i	1	6	7	6	6	2	3	1

Tabelle 3.5: Klasseneinteilung der BIP-Daten ($\sum_i n_i = 32$).

Da die Normalverteilung nur für die standardisierte Variable $Z = (X - \mu)/\sigma$ tabelliert ist, müssen die Klassengrenzen in z -Werte umgerechnet werden. Nimmt man statt der unbekannt Parameter deren Schätzungen $\bar{x} = 5.53, s = 1.86$, so ergibt sich (alles gerundet auf 2 Stellen):

Klassengrenzen

$$x_i = \{-\infty, 3, 4, 5, 6, 7, 8, 9, \infty\}$$

standardisierte Klassengrenzen

$$z_i = (x_i - \bar{x})/s = \{-\infty, -1.36, -0.82, -0.28, 0.25, 0.79, 1.33, 1.87, \infty\}$$

kumulierte Normalverteilung

$$\Phi(z_i) = \{0, 0.09, 0.21, 0.39, 0.60, 0.79, 0.91, 0.97, 1.00\}$$

theoretische Wahrscheinlichkeiten

$$\pi_i = \{0.09, 0.12, 0.18, 0.21, 0.19, 0.12, 0.06, 0.03\}$$

erwartete Häufigkeiten

$$N\pi_i = \{2.88, 3.84, 5.76, 6.72, 6.08, 3.84, 1.92, 0.96\}$$

empirische Häufigkeiten

$$n_i = \{1, 6, 7, 6, 6, 2, 3, 1\}$$

Differenzen

$$n_i - N\pi_i = \{-1.88, 2.16, 1.24, -0.72, -0.08, -1.84, 1.08, 0.04\}$$

quadrierte normierte Differenzen

$$\frac{(n_i - N\pi_i)^2}{N\pi_i} = \{1.23, 1.22, 0.27, 0.08, 0, 0.88, 0.61, 0\}$$

Chi-Quadrat-Statistik

$$\chi^2 = 4.29$$

Quantil

$$\chi^2(8 - 1 - 2, 0.95) = 11.0705$$

Damit muß die Nullhypothese der Normalverteilung der BIP-Differenzen beibehalten werden ($\alpha = 5\%$).

Offene Probleme:

- $N\pi_i > 5$ nur für 3 Klassen
- Breite (gleich oder ungleich) und Lokalisation der Klasseneinteilung
- Anzahl der Klassen



3.3 Zwei-Stichproben-Fall

3.3.1 Unabhängige Stichproben

3.3.1.1 Quantitative Merkmale: Vergleich von Erwartungswerten

In vielen Anwendungen kann die Stichprobe in zwei oder mehrere *unabhängige* Untergruppen aufgeteilt werden.

Beispiele:

- Zufriedenheit in Bankfilialen mit unterschiedlichem Gestaltungskonzept
- Umsätze mit/ohne Werbemaßnahme
- Arbeitslosenzahlen in alten/neuen Bundesländern
- BIP-Wachstum Europa/Übersee
- Unterschiede im Kaufverhalten zwischen Männern/Frauen
- Änderung von Prozentwerten bei Wahlumfragen.

Auch hier ergeben sich wieder unterschiedliche Test-Hypothesen für die Erwartungswerte der Merkmale X (Stichprobe 1) und Y (Stichprobe 2):

2-seitiger Test:

**Test des
Erwartungswerts
(2-Stichproben-
fall)**

$$H_0 : \mu_x - \mu_y = \delta_0 \quad (3.90)$$

$$H_1 : \mu_x - \mu_y \neq \delta_0 \quad (3.91)$$

1-seitiger Test; kritischer Wert rechts:

$$H_0 : \mu_x - \mu_y \leq \delta_0 \quad (3.92)$$

$$H_1 : \mu_x - \mu_y > \delta_0 \quad (3.93)$$

1-seitiger Test; kritischer Wert links:

$$H_0 : \mu_x - \mu_y \geq \delta_0 \quad (3.94)$$

$$H_1 : \mu_x - \mu_y < \delta_0. \quad (3.95)$$

Die kritischen Bereiche sind wiederum Abb. 3.1 zu entnehmen.

Als Prüfgröße ist die Mittelwertsdifferenz

$$\bar{X} - \bar{Y} - \delta_0 \quad (3.96)$$

naheliegend, wobei

$$\bar{X} = \frac{1}{N} \sum X_n \quad (3.97)$$

$$\bar{Y} = \frac{1}{M} \sum Y_m \quad (3.98)$$

die Mittelwerte in Stichprobe 1 und 2 sind (Stichprobenumfänge N und M).

Wie im Einstichprobenfall lassen sich verschiedene Konstellationen unterscheiden:

1. $X \sim N(\mu_x, \sigma_x^2), Y \sim N(\mu_y, \sigma_y^2)$
 σ_x^2 und σ_y^2 bekannt

$$\text{Varianzterm: } S = \sqrt{\frac{\sigma_x^2}{N} + \frac{\sigma_y^2}{M}}$$

2. $X \sim N(\mu_x, \sigma_x^2), Y \sim N(\mu_y, \sigma_y^2)$
 $\sigma_x^2 = \sigma_y^2$, aber unbekannt

$$\text{Varianzterm: } S = \sqrt{\left(\frac{1}{N} + \frac{1}{M}\right) \frac{(N-1)S_x^2 + (M-1)S_y^2}{N+M-2}}$$

3. $X \sim N(\mu_x, \sigma_x^2), Y \sim N(\mu_y, \sigma_y^2)$
 $\sigma_x^2 \neq \sigma_y^2$ **und** unbekannt (Behrens-Fisher-Problem)

**Behrens-Fisher-
Problem**

$$\text{Varianzterm: } S = \sqrt{\frac{S_x^2}{N} + \frac{S_y^2}{M}}$$

$$\text{Freiheitsgrade: } k \approx \frac{\left(\frac{S_x^2}{N} + \frac{S_y^2}{M}\right)^2}{\left(\frac{S_x^2}{N}\right)^2 / (N-1) + \left(\frac{S_y^2}{M}\right)^2 / (M-1)} \quad (\text{gerundet})$$

4. X, Y beliebig verteilt, $N, M > 30$

$$\text{Varianzterm: } S = \sqrt{\frac{S_x^2}{N} + \frac{S_y^2}{M}}$$

Die Prüfgröße hat die einheitliche Form

$$\frac{\bar{X} - \bar{Y} - \delta_0}{S}, \quad (3.99)$$

wobei die Varianzterme oben aufgeführt wurden.

Fall	Prüfgröße	Prüfverteilung
1	$Z = \frac{\bar{X} - \bar{Y} - \delta_0}{S}$	$Z \sim N(0, 1)$
2	$T = \frac{\bar{X} - \bar{Y} - \delta_0}{S}$	$T \sim t(N + M - 2)$
3	$T = \frac{\bar{X} - \bar{Y} - \delta_0}{S}$	$T \sim t(k)$
4	$T = \frac{\bar{X} - \bar{Y} - \delta_0}{S}$	$T \stackrel{a}{\sim} N(0, 1)$ für $N, M > 30$

Tabelle 3.6: 2-Stichproben-Tests für Erwartungswerte.

Begründung für die Varianzterme:

- $\text{Var}(\bar{X} - \bar{Y} - \delta_0) = \text{Var}(\bar{X}) + \text{Var}(\bar{Y}) = \frac{\sigma_x^2}{N} + \frac{\sigma_y^2}{M}$,

da δ_0 eine Konstante ist und die Stichproben unabhängig sind (ansonsten wäre bei abhängigen Stichproben noch ein Kovarianzterm $-2 \text{Cov}(\bar{X}, \bar{Y})$ zu berücksichtigen).

- falls $\sigma_x^2 = \sigma_y^2 = \sigma^2$, aber unbekannt:

durch gepooltes (alle Daten zusammenfassen)

$$S^2 = [(N - 1)S_x^2 + (M - 1)S_y^2]/(N + M - 2) \text{ schätzen.}$$

$$SSQ = (N - 1)S_x^2 + (M - 1)S_y^2 = \sum_n (X_n - \bar{X})^2 + \sum_m (Y_m - \bar{Y})^2,$$

daher ist $SSQ/\sigma^2 \sim \chi^2(N + M - 2)$ -verteilt.

Normiert man den Nenner $\bar{X} - \bar{Y} - \delta_0$ mit seiner Varianz (Ziffer 1, gleiche Varianzen), so ergibt sich die t -Variable

$$T = \frac{(\bar{X} - \bar{Y} - \delta_0) / \sqrt{\frac{\sigma_x^2}{N} + \frac{\sigma_y^2}{M}}}{\sqrt{\frac{(N-1)S_x^2 + (M-1)S_y^2}{\sigma^2(N+M-2)}}} \sim t(N+M-2)$$

Dies ist aber der Varianzterm aus Fall 2 (oben). Bitte nachrechnen.

3. Wenn die Varianzen unbekannt und auch nicht gleich sind, so ist das Testproblem für die Differenzen nur näherungsweise lösbar [eine exakte Lösung ist im Bayesianischen Paradigma möglich (Behrens-Fisher-Verteilung; vgl. Box und Tiao 1973, Kap. 2.5)]. Eine Approximation der Verteilung der Prüfgröße durch die t -Verteilung ist oben angegeben.
4. Schließlich sind bei großen Stichprobenumfängen die Mittelwerte \bar{X}, \bar{Y} approximativ normalverteilt (zentraler Grenzwertsatz) und die Varianz-Schätzungen S_x^2, S_y^2 streben gegen die Parameter σ_x^2, σ_y^2 . Daher ist die Teststatistik approximativ normalverteilt.

Beispiel 3.6 (Werbemaßnahme)

In einem Teil einer Supermarktkette ($X, N = 20$) wurde eine Werbemaßnahme durchgeführt. Nun soll die Wirksamkeit der Maßnahme anhand eines Vergleichs der Umsätze untersucht werden. Es ergaben sich folgende Statistiken (in TEuro):

$$\bar{x} = 12.34, \bar{y} = 3.38$$

$$s_x = 5.08, s_y = 2.07$$

$$N = 20, M = 20.$$

Die Werbemaßnahme erhöht also offenbar stark den Umsatz, jedoch sind auch die Streuungen sehr unterschiedlich.

Es ist nicht bekannt, ob die Varianzen gleich sind.

Daher ist Fall 3 relevant.

$$\text{Varianzterm: } s = \sqrt{\frac{s_x^2}{N} + \frac{s_y^2}{M}} = \sqrt{\frac{5.08^2}{20} + \frac{2.07^2}{20}} = 1.23$$

$$\text{Teststatistik: } t = \frac{\bar{x} - \bar{y} - \delta_0}{s} = 7.30$$

$$\text{Freiheitsgrade: } k \approx \frac{\left(\frac{s_x^2}{N} + \frac{s_y^2}{M}\right)^2}{\left(\frac{s_x^2}{N}\right)^2 / (N-1) + \left(\frac{s_y^2}{M}\right)^2 / (M-1)} = 25.14 \approx 25$$

Hypothesen:

$$H_0 : \mu_x \leq \mu_y$$

$$H_1 : \mu_x > \mu_y \text{ (Werbung stärkt den Umsatz)}$$

$$\text{kritischer Wert: } t(1 - \alpha, k) = t(.95, 25) = 1.708$$

Testentscheidung: da $t > t(1 - \alpha, k)$, wird H_1 angenommen (Werbung stärkt den Umsatz).



3.3.1.2 Dichotome Merkmale:

Vergleich von Anteilswerten

Häufig müssen Prozentwerte einer Eigenschaft in verschiedenen Populationen verglichen werden, etwa $A =$ Zugehörigkeit zur Mittelschicht im Vergleich Hauptschule/Gymnasium oder Prozentwerte von Parteien in den alten/neuen Bundesländern.

In diesem Fall werden als Stichprobenvariablen die Indikatoren X_n, Y_m betrachtet, die das Vorliegen der Eigenschaft A in Stichprobe (Gruppe) 1 oder 2 anzeigen:

$$X_n = \begin{cases} 1 & \text{wenn } A \text{ in Stichprobe 1} \\ 0 & \text{sonst} \end{cases} \quad (3.100)$$

$$Y_m = \begin{cases} 1 & \text{wenn } A \text{ in Stichprobe 2} \\ 0 & \text{sonst} \end{cases} \quad (3.101)$$

(Bsp.: $X_n = 1$, wenn die n -te Person die A -Partei in Westdeutschland wählt, 0 sonst).

Als Hypothesen formuliert man für die unbekanntes Wahrscheinlichkeiten $\pi_x = P(A|1) = E[X]$, $\pi_y = P(A|2) = E[Y]$:

2-seitiger Test:

$$H_0 : \pi_x - \pi_y = \delta_0 \quad (3.102)$$

$$H_1 : \pi_x - \pi_y \neq \delta_0 \quad (3.103)$$

1-seitiger Test; kritischer Wert rechts:

$$H_0 : \pi_x - \pi_y \leq \delta_0 \quad (3.104)$$

$$H_1 : \pi_x - \pi_y > \delta_0 \quad (3.105)$$

1-seitiger Test; kritischer Wert links:

$$H_0 : \pi_x - \pi_y \geq \delta_0 \quad (3.106)$$

$$H_1 : \pi_x - \pi_y < \delta_0. \quad (3.107)$$

Die kritischen Bereiche sind Abb. 3.1 zu entnehmen.

Die Testgröße ist die Differenz der empirischen Anteilswerte

$$\hat{\pi}_x = P_x = \bar{X} = \frac{1}{N} \sum_n X_n \quad (3.108)$$

$$\hat{\pi}_y = P_y = \bar{Y} = \frac{1}{M} \sum_m Y_m. \quad (3.109)$$

Die Streuung der Anteilswerte kann durch

$$S_x^2 = \frac{\hat{\pi}_x(1 - \hat{\pi}_x)}{N} \quad (3.110)$$

$$S_y^2 = \frac{\hat{\pi}_y(1 - \hat{\pi}_y)}{M} \quad (3.111)$$

geschätzt werden.

Die Testgröße

$$Z = \frac{\hat{\pi}_x - \hat{\pi}_y - \delta_0}{\sqrt{S_x^2 + S_y^2}} \stackrel{a}{\sim} N(0, 1), N, M > 30 \quad (3.112)$$

**Test von
Anteilswerten (2-
Stichprobenfall)**

ist für genügend große Stichproben asymptotisch normalverteilt.

Beispiel 3.7 (Politbarometer, Veränderung)

Die Prozentwerte der Parteien für den April 2009 sind in Abb. 3.4 zu sehen. Gegenüber der letzten Umfrage vom März 2009 (Abb. 2.3) ergeben sich Unterschiede.

Sind diese aber auch relevant bzw. signifikant (statistisch bedeutsam)?

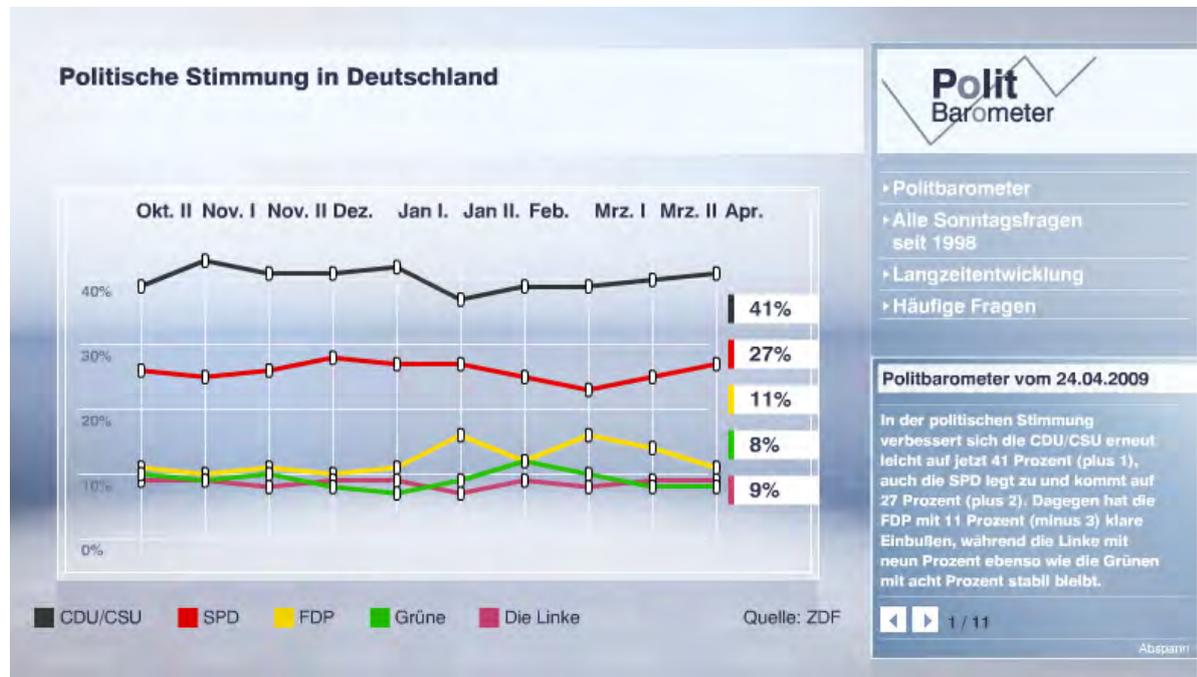
In der Teststatistik und dem kritischen Wert mischen sich Effektstärke (Unterschiede), Stichprobengröße und Signifikanzniveau. Praktische Relevanz bzw. statistische Signifikanz müssen daher strikt unterschieden werden (etwa Signifikanz durch höhere Fallzahl bei gleicher Differenz).

Die Prozentwerte im März waren:

$$\hat{\pi}_y = \{0.4, 0.25, 0.14, 0.08, 0.09, 0.04\}$$

Im April ergab sich

$$\hat{\pi}_x = \{0.41, 0.27, 0.11, 0.08, 0.09, 0.04\}$$



INFOBOX

Die Umfragen zum Politbarometer ...

... wurden wie immer von der Mannheimer Forschungsgruppe Wahlen durchgeführt. Die Interviews wurden in der Zeit vom 21. bis 23. April 2009 bei 1445 zufällig ausgewählten Wahlberechtigten telefonisch erhoben. Die Befragung ist repräsentativ für die wahlberechtigte Bevölkerung in ganz Deutschland. Der Fehlerbereich beträgt bei einem Parteianteil von 40 Prozent rund drei Prozentpunkte und bei einem Parteianteil von zehn Prozent rund zwei Prozentpunkte. Das nächste Politbarometer sendet das ZDF am Freitag, den 8. Mai.

Abbildung 3.4: Politbarometer vom 24.4.2009.

Differenz:

$$\hat{\pi}_x - \hat{\pi}_y = \{0.01, 0.02, -0.03, 0., 0., 0.\}$$

(FDP verlor 3 Prozentpunkte)

Diese Differenzen müssen aber in Relation zu den Schätzfehlern gesehen werden:

$$s_x = \{0.0129, 0.0117, 0.0082, 0.0071, 0.0075, 0.0052\}$$

$$s_y = \{0.0139, 0.0123, 0.0098, 0.0077, 0.0081, 0.0056\}$$

$$s = \sqrt{s_x^2 + s_y^2} = \{0.019, 0.0169, 0.0128, 0.0105, 0.0111, 0.0076\}$$

Die Schätzfehler liegen somit im Prozentbereich, etwa bei CDU/CSU 1.9%. Daher ist eine Steigerung von 1% vermutlich nicht statistisch bedeutsam.

Hypothesen (2-seitiger Test)

$$H_0 : \pi_x - \pi_y = 0$$

$$H_1 : \pi_x - \pi_y \neq 0$$

Teststatistik:

$$z = \frac{\hat{\pi}_x - \hat{\pi}_y}{\sqrt{s_x^2 + s_y^2}} = \{0.5269, 1.1806, -2.3393, 0, 0, 0\}$$

Kritischer Wert: $z(.975) = 1.96$ ($z(.995) = 2.58$).

Auf dem 5%-Niveau ist nur die Veränderung des FDP-Werts (-0.03) signifikant.



3.3.1.3 Stetige Merkmale: Vergleich von Varianzen

Meistens wird bei Gruppenvergleichen nur der Mittelwertsunterschied getestet. Jedoch können sich auch andere Parameter der Verteilungen $f_x(x)$, $f_y(y)$ in den Gruppen geändert haben (etwa die Streuung oder die Kurtosis bei Renditen vor und nach der Finanzkrise).

Auch können Interventionseffekte (Werbemaßnahmen, Trainingseffekte) auch auf der Ebene der Streuungen sichtbar sein (größere Heterogenität durch die Maßnahme).

Dies führt auf das Testproblem

**Test von
Varianzen
(2-Stichproben-
fall)**

(Homoskedastizität vs. Heteroskedastizität)

$$H_0 : \sigma_x^2 = \sigma_y^2 \quad (3.113)$$

$$H_1 : \sigma_x^2 \neq \sigma_y^2 \quad (3.114)$$

(oder einseitig).

Da die Stichprobenvarianzen χ^2 -verteilt sind, kann der Quotient S_y^2/S_x^2 mit Hilfe eines F -Tests geprüft werden (Kap. 1.3.4). Sind die Varianzen gleich (unter H_0), so gilt für die Prüfgröße

$$F = \frac{S_y^2}{S_x^2} \sim F(M-1, N-1). \quad (3.115)$$

Die kritischen Bereiche sind dann

$$F < f(\alpha/2, M-1, N-1) \text{ und } F > f(1-\alpha/2, M-1, N-1).$$

Bemerkung zur Teststatistik:

Es gilt

$$\frac{1}{\sigma_x^2} \sum (X_n - \bar{X})^2 = \frac{N-1}{\sigma_x^2} S_x^2 \sim \chi^2(N-1) \quad (3.116)$$

$$\frac{1}{\sigma_y^2} \sum (Y_m - \bar{Y})^2 = \frac{M-1}{\sigma_y^2} S_y^2 \sim \chi^2(M-1) \quad (3.117)$$

Daher ist der Quotient

$$\frac{\frac{1}{\sigma_y^2} S_y^2}{\frac{1}{\sigma_x^2} S_x^2} = \frac{S_y^2}{S_x^2} \quad (3.118)$$

$F(M-1, N-1)$ -verteilt unter H_0 (die Varianzen kürzen sich).

Beispiel 3.8 (Werbemaßnahme)

In einem Teil einer Supermarktkette ($X, N = 20$) wurde eine Werbemaßnahme durchgeführt. Nun soll die Wirksamkeit der Maßnahme anhand

eines Vergleichs der Umsätze untersucht werden. Es ergaben sich folgende Statistiken (in TEuro):

$$\bar{x} = 12.34, \bar{y} = 3.38$$

$$s_x = 5.08, s_y = 2.07$$

$$N = 20, M = 20.$$

In Bsp. 3.6 wurde der Mittelwertsunterschied getestet. Offenbar ergaben sich aber auch starke Unterschiede in den Streuungen.

$$f = \frac{s_y^2}{s_x^2} = 2.07^2 / 5.08^2 = 0.166. \quad (3.119)$$

Die untere und obere kritische Grenze ist dann [$F \sim F(M - 1, N - 1)$, $\alpha = 0.05$]

$$f(\alpha/2, M - 1, N - 1) = 0.396 \text{ und}$$

$$f(1 - \alpha/2, M - 1, N - 1) = 2.526.$$

Daher sind die Varianzen ungleich. Die Quantile können dem applet zur F -Verteilung entnommen werden.

http://www.fernuni-hagen.de/lis_statistik/lehre/



3.3.1.4 Nichtparametrische Methode: Wilcoxon-Rangsummen-Test

Hier wird die Idee, die ungleiche Verteilung von Rängen zu betrachten (vgl. Kap. 3.2.4.2), auf den 2-Stichprobenfall angewandt.

Für die Daten

X_1, \dots, X_N (1. Stichprobe)

Y_1, \dots, Y_M (2. Stichprobe)

werden Ränge in der zusammengefaßten (gepoolten) Stichprobe

$X_1, \dots, X_N, Y_1, \dots, Y_M$

vergeben, sodaß sich Lokalisationsunterschiede in den Rängen widerspiegeln.

Die Teststatistik

$$T_W = \sum_n \text{rg}(X_n) \quad (3.120)$$

zählt nur die Ränge der 1. Stichprobe (die aber in der gemeinsamen Stichprobe vergeben wurden).

Test des Erwartungswerts (2-Stichprobenfall)

Mit Hilfe von T_W wird das Hypothesenpaar

2-seitiger Test:

$$H_0 : \mu_x = \mu_y \quad (3.121)$$

$$H_1 : \mu_x \neq \mu_y \quad (3.122)$$

Wilcoxon-Rangsummen-Test

oder einseitige Alternativen geprüft.

Beispiel 3.9 (Ränge in gepoolter Stichprobe)

Die Datensätze

{1.0, 4.1, 6.9} (1. Stichprobe)

{2.5, 4.5} (2. Stichprobe)

führen zu den gepoolten Rängen

{1, 3, 5} (1. Stichprobe)

{2, 4} (2. Stichprobe)

Somit ist $T_W = 9$, in Stichprobe 2 ist die Summe $2+4=6$. Insgesamt summieren sich die Ränge auf $5 * 6/2 = 15$. Die Asymmetrie in den Mittelwerten $\bar{x} = 4, \bar{y} = 3.5$ zeigt sich auch in den gemeinsamen Rängen.

■

Satz 3.1 (Erwartungswert und Varianz von T_W)

Für die Teststatistik gilt (unter H_0):

$$E[T_W] = N(N + M + 1)/2 \quad (3.123)$$

$$\text{Var}[T_W] = NM(N + M + 1)/12. \quad (3.124)$$

Beweis:

$$T_W = \sum_n \text{rg}(X_n) = \sum_{k=1}^{N+M} \text{rg}(Z_k) V_k \quad (3.125)$$

wobei $Z = \{X_1, \dots, X_N, Y_1, \dots, Y_M\}$ (gepoolte Stichprobe)

$$V_k = \begin{cases} 1 & \text{falls } Z_k \in \text{Stichprobe 1} \\ 0 & \text{sonst.} \end{cases}$$

$$T_W = \sum_{k=1}^{N+M} k V_k, \text{ da alle Ränge vorkommen.} \quad (3.126)$$

$$E[T_W] = \sum_{k=1}^{N+M} k \underbrace{E[V_k]}_{N/(N+M)} \quad (3.127)$$

$$= \frac{N}{N+M} \frac{(N+M)(N+M+1)}{2} \quad (3.128)$$

Der Beweis für $\text{Var}(T_W)$ findet sich in Bosch, 1992, S. 685 und S. 713 f.

■

Mit Hilfe dieser Größen läßt sich ein asymptotischer Test formulieren ($N, M > 25$)

$$Z = \frac{T_W - E[T_W]}{\sqrt{\text{Var}[T_W]}} \underset{a}{\sim} N(0, 1) \quad (3.129)$$

Der Test wird, wie nun schon Routine, mit Hilfe der z -Quantile zweiseitig oder einseitig durchgeführt.

Mit Hilfe kombinatorischer Methoden kann auch die exakte Verteilung von T_W hergeleitet werden (Bosch, a.a.O., S. 714).

Die kritischen Werte $w(\alpha)$, $w(1-\alpha) = N(N+M+1) - w(\alpha)$ (einseitiger Test) sind im Tabellenanhang zu finden.

3.3.2 Abhängige Stichproben

3.3.2.1 Quantitative Merkmale

Bisher wurde angenommen, daß die Stichproben

X_1, \dots, X_N (1. Stichprobe)

Y_1, \dots, Y_M (2. Stichprobe)

voneinander *unabhängig* waren, d.h. es handelte sich um verschiedene Personen (allgemein: statistische Einheiten) in beiden Stichproben. Dann

sind die Kovarianzen $\text{Cov}(X_n, Y_m) = 0$ und die Kovarianzterme der Mittelwerte $\text{Cov}(\bar{X}, \bar{Y}) = 0$ verschwinden ebenfalls.

Es gibt jedoch Fälle, wo die Stichproben abhängig sind, etwa beim Bruttoinlandsprodukt (BIP) für die Jahre 2006 und 2007 (Panel-Daten X_{nt}). Dann muß in der Varianz der Mittelwertsdifferenz

$$\text{Var}(\bar{X} - \bar{Y}) = \text{Var}(\bar{X}) - 2 \text{Cov}(\bar{X}, \bar{Y}) + \text{Var}(\bar{Y}) \quad (3.130)$$

der Kovarianzterm mitberücksichtigt werden. Wenn die Daten positiv korreliert sind (Stabilität), verringert dieser Kovarianzterm die Streuung der Mittelwerts-Differenz, die Teststatistik wird also größer als bei unabhängigen Datensätzen.

Somit können Mittelwertsunterschiede leichter nachgewiesen werden, etwa wenn der Erfolg von Trainings- oder Werbemaßnahmen untersucht wird. Daher wird häufig in solchen Studien mit Meßwiederholungen gearbeitet, d.h. die gleichen Personen werden vor und nach der Intervention untersucht (pre-test/post-test).

Man kann die direkte Berechnung des Kovarianzterms umgehen, indem Differenzwerte

Meßwiederholungen pre-test/post-test

$$D_n = X_n - Y_n \quad (3.131)$$

gebildet werden. Damit wird der 2-Stichprobenfall auf den 1-Stichprobenfall reduziert und die entsprechenden Tests für D können angewandt werden (Gauß-Test, Vorzeichen-Test, Vorzeichen-Rang-Test). Die Testgröße ist der Mittelwert (vgl. Tabelle 3.1)

$$\bar{D} = \frac{1}{N} \sum_n D_n \quad (3.132)$$

mit Standardabweichung der Differenzwerte

$$S_D^2 = \frac{1}{N-1} \sum_n (D_n - \bar{D})^2 \quad (3.133)$$

$$= \frac{1}{N-1} \sum_n [(X_n - \bar{X}) - (Y_n - \bar{Y})]^2 \quad (3.134)$$

$$= S_x^2 + S_y^2 - 2S_{xy} = S_x^2 + S_y^2 - 2S_x S_y R_{xy}. \quad (3.135)$$

Somit ergibt sich die Prüfgröße

$$T = \frac{\bar{D} - \delta_0}{S_D/\sqrt{N}} \sim t(N - 1) \quad (3.136)$$

Test des Erwartungswerts (abhängige Stichproben)

die unter $H_0 : \mu_x - \mu_y = \delta_0$ t -verteilt mit $N - 1$ Freiheitsgraden ist.

Beispiel 3.10 (Bruttoinlandsprodukt BIP 2006/2007)

Das nominale Wachstum des Bruttoinlandsprodukts für verschiedene OECD-Länder hat folgende Werte für die Jahre 2006 und 2007 (Abb. 3.5; Datensatz Regression.jmp). Aus Abb. 3.6 entnimmt man folgende Informationen (gerundet):

$$\bar{x} = 5.89, s_x = 2.47.$$

$$\bar{y} = 5.53, s_y = 1.86.$$

$$\bar{d} = \bar{x} - \bar{y} = 0.36, s_d = 1.33, N = 32.$$

Die Standardabweichung der Differenzen $s_d = 1.33 = s_{x-y}$ ist wesentlich geringer als die Summe der Standardfehler $\sqrt{s_x^2 + s_y^2} = 3.09$, da im Differenzwert die Korrelation $r_{xy} = 0.85$ (Abb. 3.7) zwischen den BIP-Werten berücksichtigt wurde.

$$\text{In der Tat gilt } s_d = 1.33 = s_{x-y} = \sqrt{s_x^2 + s_y^2 - 2s_x s_y r_{xy}}$$

(bitte nachrechnen!).

Die Standardabweichung der Differenzen berücksichtigt also automatisch die Korrelation der Stichproben.

Führt man für die Differenzen einen t -Test der Form

$$T = \frac{\bar{D}}{S_D/\sqrt{N}} \quad (3.137)$$

für die Hypothese $H_0 : \delta_0 = 0$ durch, so ergibt sich:

$$t = \frac{0.36}{1.33/\sqrt{32}} = 1.54 \quad (3.138)$$

Die kritischen Werte der T -Statistik sind

$$t(\alpha/2, N - 1) = t(.025, 31) = -2.039,$$

$$t(1 - \alpha/2, N - 1) = t(.975, 31) = 2.039.$$

Verwenden Sie bitte das applet zur t -Verteilung (Abb. 1.5) oder die Tabelle im Anhang (Näherung).

Nominales GDP- Wachstum in % 2006	Nominales GDP- Wachstum in % 2007	GDP-Differenz
7,24	6,59	0,65
4,29	4,08	0,21
4,29	4,21	0,07
6,22	5,22	1,00
7,15	8,32	-1,17
5,08	5,38	-0,29
3,88	3,77	0,11
3,72	3,88	-0,15
2,24	3,18	-0,94
7,25	7,20	0,05
6,16	7,30	-1,14
8,44	5,71	2,74
8,41	8,54	-0,13
2,79	3,54	-0,75
2,14	2,80	-0,66
6,26	6,33	-0,08
8,80	6,51	2,29
9,71	6,92	2,79
4,18	4,13	0,05
3,27	4,29	-1,02
10,90	5,69	5,21
5,96	6,45	-0,49
3,59	3,36	0,23
8,52	8,65	-0,13
7,17	6,78	0,40
5,57	5,26	0,31
3,19	3,22	-0,03
11,56	10,43	1,13
4,55	4,54	0,01
6,71	5,48	1,22
3,79	4,18	-0,38
5,35	4,94	0,42

Abbildung 3.5: Veränderung des Bruttoinlandsprodukts im Jahr 2006 und 2007 (Daten).

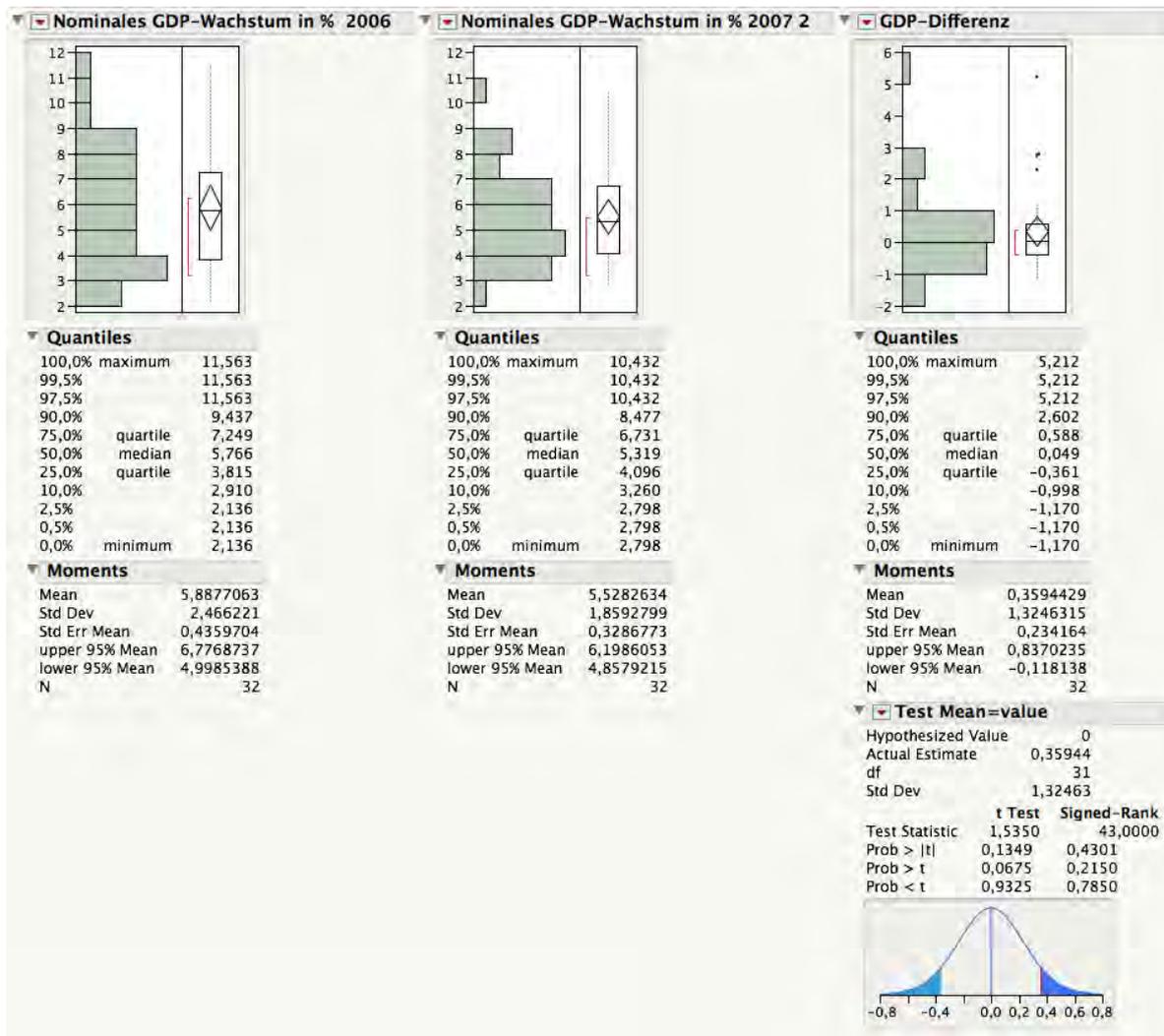


Abbildung 3.6: Veränderung des Bruttoinlandsprodukts im Jahr 2006 und 2007.

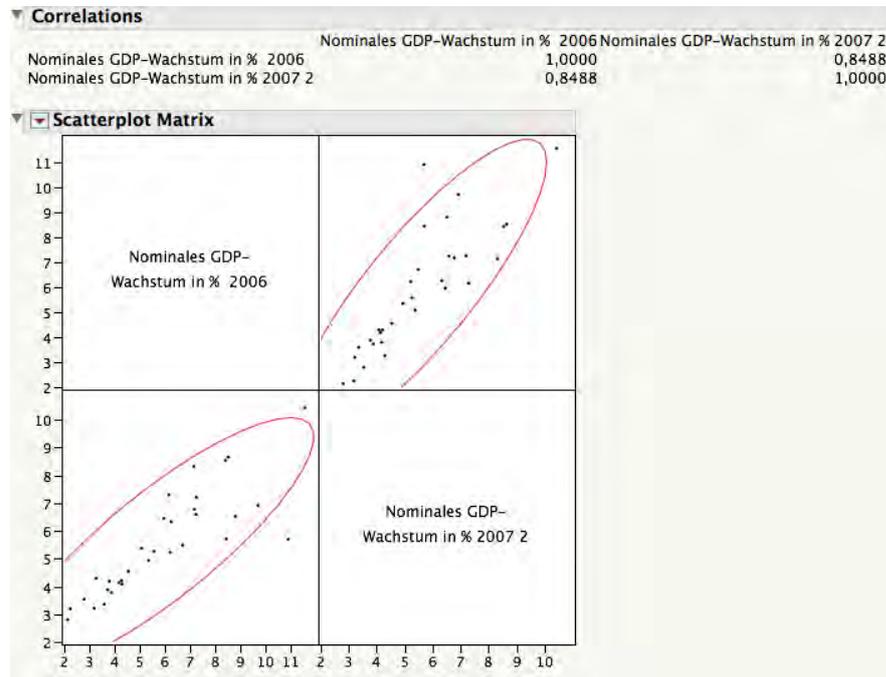


Abbildung 3.7: Korrelation $r = 0.8488$ der BIP-Änderungen im Jahr 2006 und 2007.

Damit muß die $H_0 : \delta_0 = \mu_x - \mu_y = 0$ beibehalten werden.

Bei unabhängigen Stichproben würde man durch $\sqrt{s_x^2 + s_y^2} = 3.09$ teilen (noch kleinere t -Statistik).

p -Wert

Der p -Wert ist entsprechend $p = P(|T| > |t|) = 2(1 - P(T \leq 1.54)) = 0.1349$ (siehe Abb. 3.6, rechts unten).

Man müßte also das Signifikanzniveau über 14% wählen, um ein signifikantes Resultat zu erhalten.

Allerdings muß α **vor** dem Test fixiert werden!

Die Entscheidungsregel lautet: falls $p < \alpha : H_0$ ablehnen.

Etwas genauer gilt: wenn $F(t) = P(T \leq t)$ die Verteilungsfunktion der Teststatistik unter H_0 ist, dann ergibt sich für die blau gefärbte Fläche p in Abb. 3.6 (oberhalb von $|T|$ und unterhalb von $-|T|$) der Wert $p = F(-|T|) + (1 - F(|T|)) = 2(1 - F(|T|))$ (Formel bei symmetrischer Verteilung). Setzt man nun die Teststatistik $T(\omega)$ ein, so ist auch $p = p(\omega)$ eine Zufallsgröße.

Die Entscheidungsregel ist nun: falls $p(\omega) < \alpha : H_0$ ablehnen.

Im vorliegenden Fall ist die Realisierung $t = 1.54$, $p = 0.1349$, $p > 0.05$,

Entscheidung: H_0 beibehalten. Es handelt sich also beim p -Wert um eine Transformation der Teststatistik, nicht um ein Signifikanzniveau. Die Angewohnheit von Praktikern, zuerst bei p zu schauen und dann das Signifikanzniveau zu wählen (passend zum eigenen Interesse), ist falsch (vergleiche ausführlich Greenland et al. (2016); Kuffner und Walker (2018)). Der β -Fehler kann geschätzt werden, wenn man den Mittelwert \bar{d} als Schätzwert für das unbekannte $E[D] = \mu_x - \mu_y = \delta_0$ einsetzt.

Dann gilt (vgl. 2.77):

$$\hat{\beta} = P(T \in \bar{R} | \delta_0 = \bar{d}) \quad (3.139)$$

$$= P\{T \leq t(1 - \alpha/2) - t\} - P\{T \leq t(\alpha/2) - t\} \quad (3.140)$$

$$= 0.6907 \quad (3.141)$$

wobei $t = 1.54$ und $t(1 - \alpha/2, N - 1) = t(.975, 31) = 2.039$, $t(\alpha/2, N - 1) = t(.025, 31) = -2.039$ (kritische Werte; vgl. Abb. 2.6).

Leider sind die Werte

$$P\{T \leq 0.5045\} = 0.6913$$

$$P\{T \leq -3.5745\} = 0.0006$$

nicht tabelliert, sodaß man auf die Normalverteilungs-Näherung (für $df > 30$) zurückgreifen kann:

$$z(0.025) = -1.95996$$

$$z(0.975) = 1.95996$$

$$P\{Z \leq 0.425\} = 0.6646$$

$$P\{Z \leq -3.495\} = 0.0002$$

$$\hat{\beta} = 0.6644.$$

In der Tabelle sind nur die Werte für 0.42 und 0.43 hinterlegt. Für negative Werte gilt $P(Z \leq z) = 1 - P(Z \leq -z)$. Sie können aber auch freie Software wie R benutzen, um die Werte zu berechnen.

■

Übung 3.2 (Wilcoxon-Vorzeichen-Rang-Test)

Führen Sie mit den Differenzwerten d_n (Abb. 3.5) den Wilcoxon-Vorzeichen-Rang-Test der Hypothese

$$H_0 : f(d) \text{ ist symmetrisch zum Median } d(0.5) = \delta_0 = 0 \quad (3.142)$$

durch.

■

3.3.2.2 Dichotome Merkmale: McNemar-Test

Will man Anteilswerte bei dichotomen Merkmalen in abhängigen Gruppen vergleichen (Einstellung zu Regierung/Opposition bei Wiederholungsbefragung), so kann man die Variablen $(X_n, Y_n), n = 1, \dots, N$ in Form einer Kreuztabelle auftragen:

	$Y_n = 0$	$Y_n = 1$	
$X_n = 0$	h_{00}	h_{01}	$N\hat{\pi}_x$
$X_n = 1$	h_{10}	h_{11}	
	$N\hat{\pi}_y$		

Hierbei sind $h_{ij}, i, j = 1, 2$ die absoluten Häufigkeiten, $\sum_{ij} h_{ij} = N$, und $\pi_x = P(X_n = 1), \pi_y = P(Y_n = 1)$.

Es wird angenommen, daß die Paare $(X_n, Y_n), n = 1, \dots, N$ unabhängig und identisch verteilt sind.

Getestet wird das Hypothesen-Paar:

$$H_0 : \pi_x = \pi_y \text{ gegen } \pi_x \neq \pi_y. \quad (3.143)$$

Unter der Nullhypothese ist $\pi_x = \pi_{10} + \pi_{11} = \pi_y = \pi_{01} + \pi_{11}$, d.h. $\pi_{10} = \pi_{01}$. Daher haben die Nichtdiagonalen gleiche Wahrscheinlichkeiten und für die fixierten Häufigkeiten $h_{10} + h_{01}$ gilt:

Teststatistik:

$$T = h_{01} \text{ oder } T = h_{10} \quad (3.144)$$

$$T \sim B(h_{10} + h_{01}, 1/2) \text{ unter } H_0 \quad (3.145)$$

(Binomialverteilung).

Ein approximativer Test kann so abgeleitet werden:

Die Differenzen $D_n = X_n - Y_n$ haben den Mittelwert

$$\bar{D} = \bar{X} - \bar{Y} = \hat{\pi}_x - \hat{\pi}_y \quad (3.146)$$

$$= (h_{10} + h_{11} - h_{01} - h_{11})/N = (h_{10} - h_{01})/N \quad (3.147)$$

und die Varianz lautet

$$\text{Var}(\bar{D}) = \text{Var}(\bar{X}) + \text{Var}(\bar{Y}) - 2\text{Cov}(\bar{X}, \bar{Y}) \quad (3.148)$$

$$= \frac{1}{N}[\pi_x(1 - \pi_x) + \pi_y(1 - \pi_y) - 2(\pi_{11} - \pi_x\pi_y)] \quad (3.149)$$

$$= \frac{1}{N}[\pi_{10} + \pi_{01}] \quad (3.150)$$

unter H_0 . Dies wird geschätzt durch $S^2(\bar{D}) = \frac{1}{N^2}[h_{10} + h_{01}]$. In obiger Berechnung wurde die Formel $\text{Cov}(\bar{X}, \bar{Y}) = 1/N^2 \sum_{nn'} \text{Cov}(X_n, Y_{n'}) = 1/N(\pi_{11} - \pi_x \pi_y)$ benutzt.

Für große Stichproben ist \bar{D} approximativ normalverteilt. Man erhält daher die Teststatistik

$$\chi^2 = \frac{\bar{D}^2}{S^2(\bar{D})} = \frac{(h_{10} - h_{01})^2}{h_{10} + h_{01}} \stackrel{a}{\sim} \chi^2(1) \quad (3.151)$$

**McNemar-Test
(Approximation)**

Der Ablehnungsbereich ist $\chi^2 > \chi^2(1 - \alpha, 1)$.

3.4 Zusammenhang Test-Konfidenzintervall

Beim Test wird eine Hypothese H_0 für den unbekannt Parameter θ formuliert, sodaß unter H_0 die Wahrscheinlichkeitsaussage

$$P(T \in R | H_0) = \alpha \quad (3.152)$$

$$P(T \in \bar{R} | H_0) = 1 - \alpha \quad (3.153)$$

gilt. Hierbei ist $R = R_\alpha$ ein Ablehnungsbereich mit Wahrscheinlichkeit α .

Andererseits kann man zufällige Intervalle $KI = [U, O]$ konstruieren, die den wahren Parameter mit einer bestimmten Wahrscheinlichkeit überdecken, also

$$P(\theta \in KI) = 1 - \alpha \quad (3.154)$$

$$P(\theta \in \overline{KI}) = \alpha \quad (3.155)$$

Der Test ist eine Wahrscheinlichkeitsaussage für die Prüfgröße T unter der Nullhypothese, während beim KI die zufälligen Intervallgrenzen (die von den Daten abhängen), den Parameter mit einer bestimmten Wahrscheinlichkeit überdecken.

Daher kann man folgende **Entscheidungsregel** konstruieren:

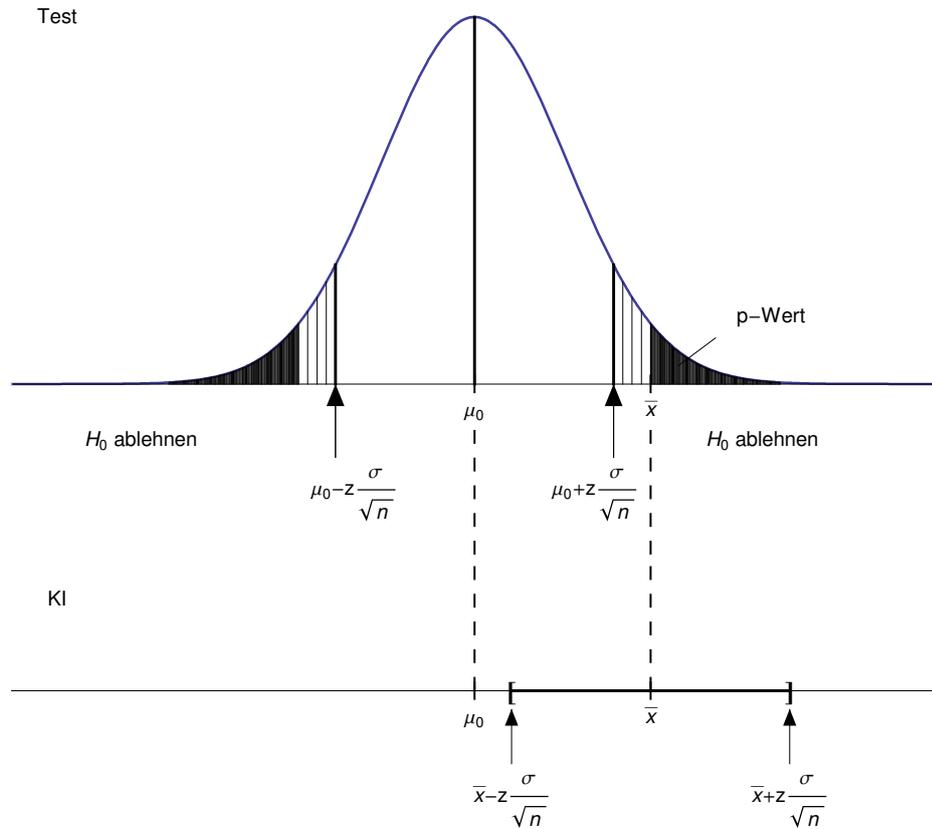


Abbildung 3.8: Zweiseitiger Gauß-Test und Konfidenzintervall.

$$\begin{aligned}
 \text{KI überdeckt den Wert } \theta_0 &\Leftrightarrow H_0 \text{ beibehalten} \\
 P(\theta_0 \in \text{KI}) &= 1 - \alpha \\
 &= P(T \notin R | H_0)
 \end{aligned}$$

Betrachten wir konkret den Gauß-Test $H_0 : \mu = \mu_0$ bei normalverteilten Daten $X_n \sim N(\mu, \sigma^2)$.

Die Teststatistik und das Komplement des Ablehnungsbereichs lauten:

$$T = \frac{\bar{X} - \mu}{\sigma / \sqrt{N}} \quad (3.156)$$

$$\bar{R} = [-z, z]; \quad z := z(1 - \alpha/2) \quad (3.157)$$

Daher gilt

$$T \notin R \Leftrightarrow T \in \bar{R}$$

$$\Leftrightarrow \left| \frac{\bar{X} - \mu}{\sigma/\sqrt{N}} \right| \leq z$$

(Test für μ , H_0 beibehalten) (3.158)

$$\Leftrightarrow \bar{X} - z \sigma/\sqrt{N} \leq \mu \leq \bar{X} + z \sigma/\sqrt{N}$$

(KI für μ) (3.159)

**Test \Leftrightarrow
Konfidenzintervall**

Dieser Sachverhalt ist in Abb. 3.8 graphisch verdeutlicht. Dargestellt ist die Verteilung von \bar{X} unter $H_0 : \mu = \mu_0$. Die Wahrscheinlichkeit außerhalb der gemessenen Teststatistik t , also $p = P(|T| > t)$, ist der sog. p -Wert (Überschreitungswahrscheinlichkeit). Bei $p < \alpha$ wird die H_0 abgelehnt. Äquivalent hierzu ist, daß das Konfidenzintervall den Wert μ_0 nicht überdeckt.

**Überschreitungs-
wahrscheinlichkeit
(p -Wert)**

Etwas genauer gilt: wenn $F(t) = P(T \leq t)$ die Verteilungsfunktion der Teststatistik unter H_0 ist, dann ergibt sich für die schwarz gefärbte Fläche in Abb. 3.8 (oberhalb von $|\bar{X}|$ und unterhalb von $-|\bar{X}|$) der Wert $p(T) = F(-|T|) + (1 - F(|T|))$. Setzt man die Teststatistik $T(\omega)$ ein, so ist auch $p = p(T(\omega))$ eine Zufallsgröße.

Die Entscheidungsregel lautet nun: falls $p < \alpha : H_0$ ablehnen.

Es handelt sich also beim p -Wert um eine Transformation der Teststatistik, nicht um ein Signifikanzniveau. Anstatt mit der Teststatistik T kann der Test auch mit $p(T)$ durchgeführt werden.

Die Angewohnheit von Praktikern, zuerst bei p zu schauen und dann das Signifikanzniveau zu wählen (passend zum eigenen Interesse), ist falsch (vergleiche ausführlich Wasserstein et al. (2016); Greenland et al. (2016); Kuffner und Walker (2018)).

Im allgemeinen ergibt sich der p -Wert durch den kleinsten Ablehnungsbereich R_p , der die Teststatistik noch enthält, d.h. $T \in R_p$. Die Ablehnungsbereiche müssen hierbei verschachtelt sein: $R_p \subset R_{p'}$ für $p < p'$.

Kapitel 4

Zusammenhangsanalyse

In Abschnitt 2.4.6 und Bsp. 2.13 wurden Konfidenzintervalle und Tests für Korrelationen ρ bereits diskutiert. Aufgrund der hohen Bedeutung solcher Analysen wird dem Thema ein eigener Abschnitt gewidmet.

4.1 Metrische Merkmale

Beispiele:

- Zusammenhang zwischen Abiturnote und Studienerfolg
- Korrelation der Renditen von Wertpapieren (Portfolio-Theorie von Markowitz)
- Korrelation der BIP-Zuwächse 2006, 2007 (Abb. 3.7)

Zunächst einige Definitionen:

Korrelationskoeffizient in der Population

$$\rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} \quad (4.1)$$

**Korrelations-
koeffizient
(Population)**

Ein Punktschätzer für ρ ist durch die **Produkt-Moment-Korrelation (Pearson-Korrelation)**

$$R = \frac{S(X, Y)}{S(X)S(Y)} \quad (4.2)$$

gegeben, wobei die Stichproben(ko)varianzen durch

$$S(X, Y) = \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})(Y_n - \bar{Y}) \quad (4.3)$$

$$S(X)^2 = \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})^2 = S(X, X) \quad (4.4)$$

$$S(Y)^2 = \frac{1}{N-1} \sum_{n=1}^N (Y_n - \bar{Y})^2 = S(Y, Y) \quad (4.5)$$

definiert sind.

Explizit kann man auch schreiben

**Korrelations-
koeffizient
(Stichprobe)**

$$R = \frac{\sum_{n=1}^N (X_n - \bar{X})(Y_n - \bar{Y})}{\sqrt{\sum_{n=1}^N (X_n - \bar{X})^2 \sum_{n=1}^N (Y_n - \bar{Y})^2}}. \quad (4.6)$$

Es wird bei den folgenden Korrelationstests angenommen, daß die Wertepaare $(X_n, Y_n), n = 1, \dots, N$, voneinander **unabhängig und identisch verteilt** sind wie (X, Y) . Außerdem wird eine bivariate Normalverteilung unterstellt.

4.1.1 Signifikanztest $\rho = 0$

Bei bivariat normalverteilten Merkmalen (Abb. 4.1)

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim N \left(\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \begin{bmatrix} \sigma_x^2 & \sigma_x \sigma_y \rho \\ \sigma_x \sigma_y \rho & \sigma_y^2 \end{bmatrix} \right). \quad (4.7)$$

ist unter

$$H_0 : \rho = 0 \quad (4.8)$$

$$H_1 : \rho \neq 0 \quad (4.9)$$

die Teststatistik

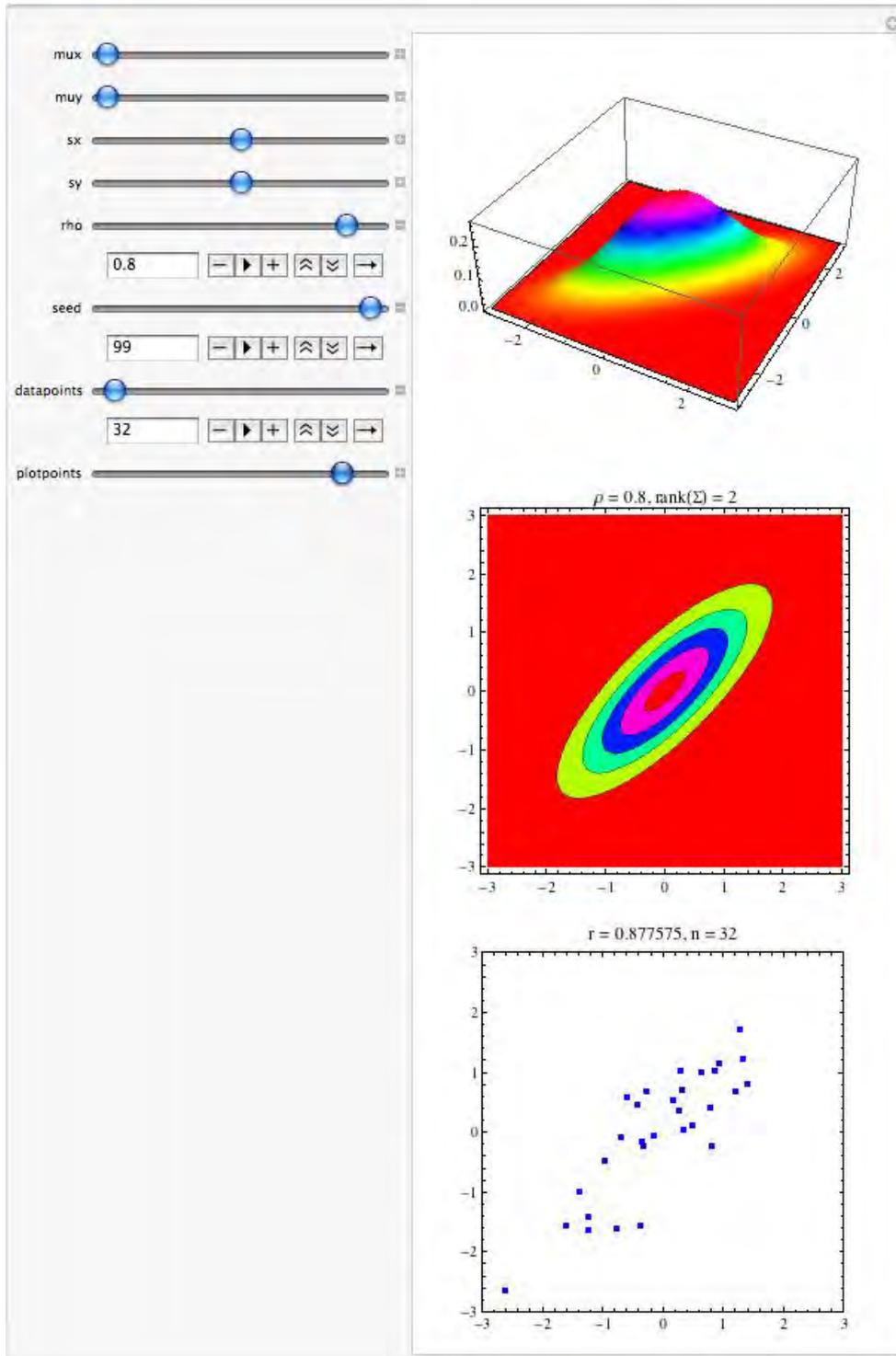


Abbildung 4.1: Bivariate Normalverteilung mit $\rho = 0.8$ und simulierten Daten (Streudiagramm; $r = .877$, $N = 32$, $\text{seed} = 99$).

http://www.fernuni-hagen.de/ls_statistik/lehre/

Test des Korrelationskoeffizienten

$$T = \sqrt{N-2} \frac{R}{\sqrt{1-R^2}} \sim t(N-2). \quad (4.10)$$

t -verteilt mit $N-2$ Freiheitsgraden.

Beispiel 4.1 (Test auf $\rho = 0$)

In Abb. 4.1 ist $r = .877$, $N = 32$, $seed = 99$.¹ Daraus ergibt sich

$$t = \sqrt{32-2} \frac{.877}{\sqrt{1-.877^2}} = 9.99. \quad (4.11)$$

Das $t(30, .975)$ -Quantil ist 2.04. Daher muß $H_0 : \rho = 0$ abgelehnt werden ($\alpha = 0.05$).

■

4.1.2 Signifikanztest $\rho = \rho_0$

Will man auf einen beliebigen Korrelationswert ρ_0 testen, so muß zunächst auf Fishers Z -Statistik

Fisher- Z -Statistik

$$Z = \frac{\sqrt{N-3}}{2} \left(\ln \frac{1+R}{1-R} - \ln \frac{1+\rho_0}{1-\rho_0} \right) \quad (4.12)$$

transformiert werden. Sie ist für große Stichproben $N(0,1)$ -normalverteilt.

Beispiel 4.2 (BIP 2006/2007: Test auf $\rho = 0.8$)

In Abb. 3.7 war das Streudiagramm der BIP-Zuwächse 2006/2007 aufgetragen ($r = .849$, $N = 32$). Hier soll die Nullhypothese $\rho = \rho_0 = 0.8$ getestet werden:

$$z = \frac{\sqrt{N-3}}{2} \left(\ln \frac{1+r}{1-r} - \ln \frac{1+\rho_0}{1-\rho_0} \right) = 0.829. \quad (4.13)$$

¹die sog. *random seed* determiniert den Startwert des Zufallsgenerators. Man erhält dann andere Stichproben und Korrelationen r .

Das $z(.975)$ -Quantil ist 1.96. Daher muß H_0 beibehalten werden. Die Nullhypothese $\rho = 0$ wird jedoch abgelehnt (bitte nachrechnen).



Beispiel 4.3 (Scheinkorrelation)

Die Annahme, daß die Datenpaare identisch verteilt sind, bedeutet insbesondere, daß kein 'Trend' in den Daten vorliegen sollte, etwa $X_n = n + \epsilon_n$, $\epsilon_n \sim N(0, 1)$, $Y_n = n + \delta_n$, $\delta_n \sim N(0, 1)$, da dann $E[X_n] = n$, $E[Y_n] = n$ von n abhängt und somit die Daten (Trend + Rauschen) nicht identisch wie (X, Y) verteilt sind.

Dies hätte den Effekt, daß für unkorrelierte² Rausch-Variablen mit $\text{Cov}(\epsilon_n, \delta_n) = 0$ die Daten zwar keine Kovarianz

$$\text{Cov}(X_n, Y_n) = \text{Cov}(\epsilon_n, \delta_n) = 0$$

aufweisen, jedoch die Stichprobenkovarianz

$$S_N(X, Y) = \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})(Y_n - \bar{Y})$$

keine sinnvolle Teststatistik darstellt. Es gilt nämlich (X_n, Y_n) sind unabhängig)

$$\begin{aligned} E[S_N(X, Y)] &= \frac{1}{N-1} \sum_{n=1}^N E(X_n - \bar{X})E(Y_n - \bar{Y}) \\ &= \frac{1}{N-1} \sum_{n=1}^N (n - (N+1)/2)^2 \\ &= \frac{1}{N-1} \sum_{n=1}^N (n^2 - n(N+1) + (N+1)^2/4) \\ &= \frac{N(N^2-1)}{12(N-1)} \\ &\propto N^2/12 \text{ für große } N \end{aligned}$$

mit den Summenformeln $\sum_{n=1}^N n^2 = N(N+1)(1+2N)/6$ und $\sum_{n=1}^N n = N(N+1)/2$. Dies bedeutet, daß die Stichprobenkovarianz nicht gegen den Populationsparameter $\text{Cov}(X, Y) = 0$ streben kann.³ Wie

²aufgrund der Normalverteilung sind die Datenpaare auch unabhängig.

³Zur Konvergenz im Quadratmittel gegen die Konstante c , d.h. $E[S_N - c]^2 = \text{Var}(S_N) + (E[S_N] - c)^2 \rightarrow 0$ müßte gelten: $E[S_N] \rightarrow c$, $\text{Var}[S_N] \rightarrow 0$. Dies würde auch die Konvergenz in Wahrscheinlichkeit bedeuten (vgl. Shiryaev, 1995, S. 256).

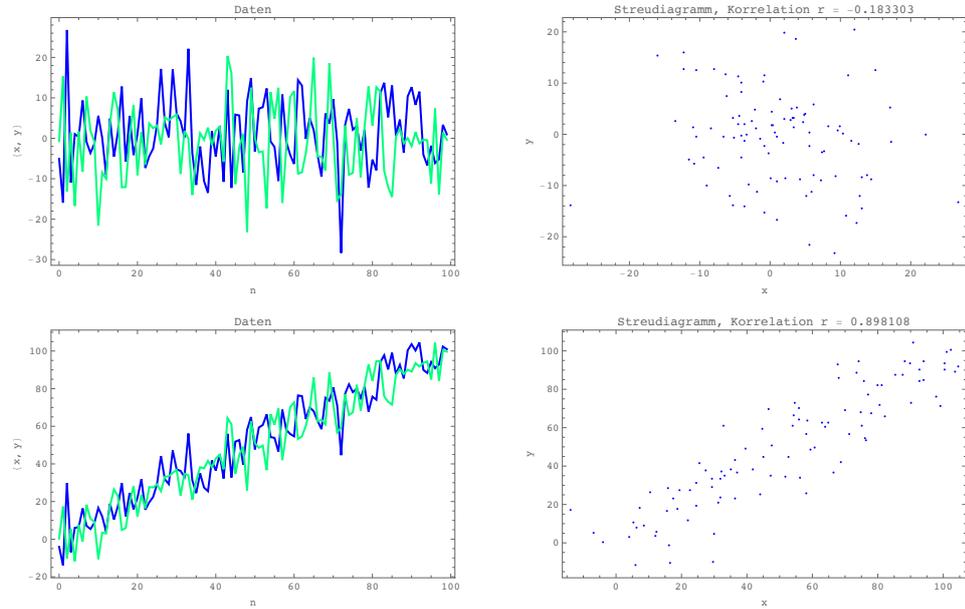


Abbildung 4.2: Scheinkorrelation. Linke Spalte: Statistisch unabhängige Daten ohne Trend (oben), mit Trend (unten). Der empirische Korrelationskoeffizient ist vorwiegend durch den Trend erzeugt (rechte Spalte, unten).

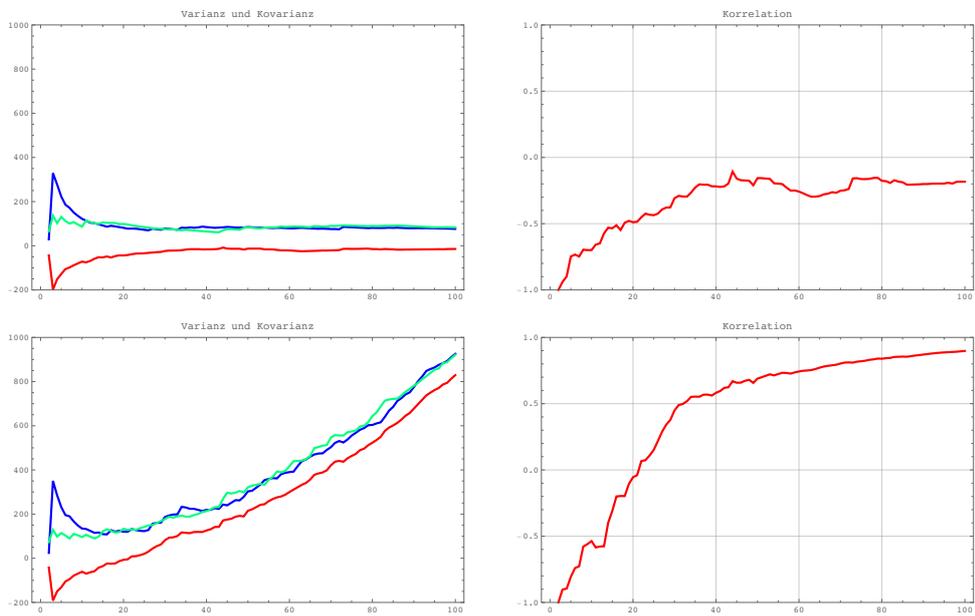


Abbildung 4.3: Scheinkorrelation. Linke Spalte: Varianzen und Kovarianz (rot) als Funktion der Stichprobengröße. Rechte Spalte: empirischer Korrelationskoeffizient als Funktion der Stichprobengröße. Obere Zeile: ohne Trend, untere Zeile: mit Trend.

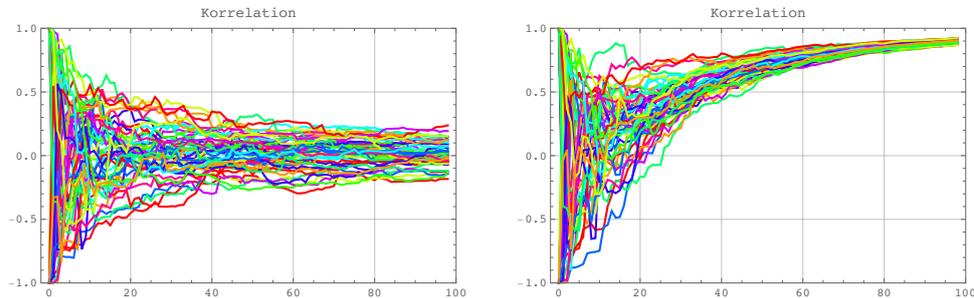


Abbildung 4.4: Korrelation in 50 Replikationen als Funktion der Stichprobengröße. Links: ohne Trend, rechts: mit Trend.

den Abb. 4.2–4.4 zu entnehmen ist, wird durch den hinzugefügten deterministischen Trend eine empirische Korrelation erzeugt⁴ Für große Stichproben geht diese sogar gegen 1, obwohl die wirkliche Korrelation $\rho = \text{Cov}(X, Y) / \sqrt{\text{Var}(X)\text{Var}(Y)} = 0$ ist. Vgl. Davidson et al. (2004, Kap. 14.2).

■

4.2 Ordinale Merkmale: Spearman's Rangkorrelation

Wenn die Daten nur ordinalskaliert sind, also nur Vergleiche $X_i < X_j$ erklärt sind (keine Produkte oder Summen), so kann man von den Urdaten auf Rangdaten $\text{rg}(X_1), \dots, \text{rg}(X_N), \text{rg}(Y_1), \dots, \text{rg}(Y_N)$ übergehen und mit diesen Rängen den Produkt-Moment-Korrelationskoeffizienten R_{Sp} berechnen

$$R_{Sp} = \frac{\sum_{n=1}^N (\text{rg}(X_n) - \overline{\text{rg}X})(\text{rg}(Y_n) - \overline{\text{rg}Y})}{\sqrt{\sum_{n=1}^N (\text{rg}(X_n) - \overline{\text{rg}X})^2 \sum_{n=1}^N (\text{rg}(Y_n) - \overline{\text{rg}Y})^2}}. \quad (4.14)$$

**Spearman's
Rangkorrelation**

⁴es wurde $\text{Var}(\epsilon_n) = \text{Var}(\delta_n) = 100$ gewählt, um gut sichtbare Schwankungen zu erzeugen.

Hierbei sind

$$\overline{\text{rg}X} = \frac{1}{N} \sum_{n=1}^N \text{rg}(X_n) \quad (4.15)$$

$$= \frac{1}{N} \sum_{n=1}^N n = \frac{N(N+1)}{2N} = \frac{N+1}{2} \quad (4.16)$$

(entsprechend für Y) die Rangmittelwerte. Diese hängen gar nicht von den Daten ab. Der Rangkorrelations-Koeffizient ist invariant gegenüber streng monoton wachsenden Transformationen $\tilde{X} = g(X)$, $\tilde{Y} = h(Y)$, da die Ränge unverändert bleiben.

Führt man die Differenzwerte

$$D_n = \text{rg}(X_n) - \text{rg}(Y_n) \quad (4.17)$$

ein, so gilt ohne Bindungen (gleiche Ränge) die Formel

$$R_{Sp} = 1 - 6 \frac{\sum_n D_n^2}{N(N^2 - 1)}. \quad (4.18)$$

Bei Vorliegen von Bindungen werden Durchschnittsränge vergeben und Glg. 4.14 benutzt. Als Hypothesenpaar wird getestet:

$$H_0 : X, Y \text{ sind unabhängig} \quad (4.19)$$

$$H_1 : X, Y \text{ sind korreliert} \quad (4.20)$$

oder 1-seitig ($H_1 : X, Y$ sind negativ (positiv) korreliert).

Für große Stichproben gilt:

$$T = \sqrt{N-2} \frac{R_{Sp}}{\sqrt{1-R_{Sp}^2}} \sim t(N-2) \quad (4.21)$$

ist t -verteilt mit $N-2$ Freiheitsgraden.

Für kleine Stichproben ($N \leq 10$) kann ein exakter Test für die Quadratsumme $D = \sum_n D_n^2$ ausgeführt werden (Tabellen-Anhang).

Beispiel 4.4 (Tanzturnier)

Bei einem Tanzturnier vergaben 2 Wertungsrichter folgende Ränge:

Paar	1	2	3	4	5	6	7	8	9	10
rg(x_n)	8	10	6	1	4	9	5	7	3	2
rg(y_n)	6	10	8	3	2	7	1	9	5	4
d_n	2	0	-2	-2	2	2	4	-2	-2	-2

Tabelle 4.1: Ränge der Wertungsrichter.

Da keine Bindungen vorliegen, gilt

$$r_{Sp} = 1 - 6 \frac{\sum_n d_n^2}{N(N^2 - 1)} = 1 - (6 \times 48)/(10 \times 99) = 0.709.$$

Der gleiche Wert ergibt sich bei Berechnung mit Glg. 4.14.

Als asymptotische Teststatistik findet man

$$t = \sqrt{8} \frac{.709}{\sqrt{1 - .709^2}} = 2.83. \tag{4.22}$$

mit $t(8, 0.975) = 2.306$. Daher muß $H_0 : X, Y$ sind unabhängig, auf dem Signifikanzniveau $\alpha = 5\%$ abgelehnt werden. Die Urteile der Preisrichter sind also korreliert (werden durch die Leistungen beeinflusst).



4.3 Beliebige Meßniveau: χ^2 -Unabhängigkeitstest

Beispiel 4.5 (Zusammenhang Geschlecht/akademische Stellung)

An einer Universität ergab sich folgende Relation zwischen akademischer Stellung $X =$ (Professor/Mittelbau) und Geschlecht $Y =$ (männlich/weiblich) (Tab. 4.2).

Rechnet man in bedingte Häufigkeiten um (spaltenweise, d.h. in den Gruppen m/w), so ergibt sich das asymmetrische Bild (Tab. 4.3).



n_{ij}	m	w	$n_{i.}$
Professor	87	12	99
Mittelbau	122	50	172
$n_{.j}$	209	62	271

Tabelle 4.2: Absolute Häufigkeiten n_{ij} . Stichprobenumfang $\sum_{ij} n_{ij} = N = 271$.

$f(i j)$	m	w
Professor	.42	.18
Mittelbau	.58	.82
$f(. j)$	1	1

Tabelle 4.3: Bedingte Häufigkeiten $f(i|j) = n_{ij}/n_{.j}$ (spaltenweise).

n_{ij}	$N\hat{\pi}_{ij}$	$\frac{(n_{ij}-N\hat{\pi}_{ij})^2}{N\hat{\pi}_{ij}}$
87	76.35	1.49
12	22.65	5.01
122	132.65	0.86
50	39.35	2.88
		10.24

Tabelle 4.4: Terme der χ^2 -Teststatistik.

Es ist nun zu prüfen, ob die Variablen X und Y voneinander unabhängig sind.

Dazu kann der χ^2 -Test herangezogen werden (vgl. Abs. 3.2.5). Man vergleicht die tatsächlichen Häufigkeiten $n_{ij}, i = 1, \dots, I; j = 1, \dots, J$ mit den erwarteten Häufigkeiten unter Unabhängigkeit

$$N\hat{\pi}_{ij} = N(n_{i.}/N)(n_{.j}/N) = n_{i.}n_{.j}/N \quad (4.23)$$

(Randsummen $n_{i.} = \sum_j n_{ij}, n_{.j} = \sum_i n_{ij}$).

Dies beruht auf der Eigenschaft $\pi_{ij} = \pi_{i.}\pi_{.j}$ für unabhängige Ereignisse [$P(A \cap B) = P(A)P(B)$]. Diese Wahrscheinlichkeiten werden durch die relativen Häufigkeiten $\hat{\pi}_{ij} = \hat{\pi}_{i.}\hat{\pi}_{.j} = (n_{i.}/N)(n_{.j}/N) = f_{i.}f_{.j}$ geschätzt.

Somit ergibt sich die χ^2 -Teststatistik für die $I \times J$ -Kreuztabelle

Unabhängigkeit

$$\chi^2 = \sum_{i,j=1}^{I,J} \frac{(n_{ij} - N\hat{\pi}_{ij})^2}{N\hat{\pi}_{ij}} \quad (4.24)$$

$$= N \sum_{i,j=1}^{I,J} \frac{(f_{ij} - \hat{\pi}_{ij})^2}{\hat{\pi}_{ij}} \sim \chi^2((I-1)(J-1)) \quad (4.25)$$

Unabhängigkeitstest

Im Beispiel 4.5 ergeben sich die Terme (Tab. 4.4). Vergleicht man mit dem $\chi^2(1, .99) = 6.635$ -Quantil, so ergibt sich ein signifikantes Ergebnis auf den 1%-Niveau.

4.4 Kausalität, Abhängigkeit und Korrelation

Abhängigkeiten bzw. Korrelationen sind notwendige, aber nicht hinreichende Bedingungen für einen kausalen Zusammenhang (aus X folgt Y , Ursache und Wirkung). Sie besagen nur, daß Veränderungen in der einen Variable mit solchen in der anderen Variable einhergehen. Außerdem wird mit Hilfe des Produkt-Moment-Korrelationskoeffizienten $\rho = \text{Cov}(X, Y) / \sqrt{\text{Var}(X)\text{Var}(Y)}$ nur ein linearer Zusammenhang erfaßt. Variablen können also durchaus abhängig sein, obwohl die (lineare) Kor-

relation verschwindet. Etwas genauer gilt: ⁵

Unabhängigkeit und Korrelation

$$\text{Unabhängigkeit} \Rightarrow \text{keine Korrelation} \quad (4.26)$$

$$\text{Korrelation} \Rightarrow \text{Abhängigkeit.} \quad (4.27)$$

Aber auch nichtlineare Zusammenhänge können durch dritte und weitere Einflüsse gegeben sein. Daher ist bei Vorliegen eines korrelativen Zusammenhangs immer zu untersuchen, ob es weitere Variablen gibt, die den Zusammenhang erklären (Scheinkorrelationen; eine ausführliche Diskussion wird in KE 3 gegeben).

In Abb. 4.5 wird ein Zusammenhang zwischen Parodontose ($X = (ja/nein)$) und Herzkrankheiten ($Y = ja/nein$) berichtet. Die (geschätzte) Wahrscheinlichkeit

$$f(\text{Herzkrankheiten} = ja | \text{Parodontose} = ja)$$

ist also um 25% höher als

$$f(\text{Herzkrankheiten} = ja | \text{Parodontose} = nein)$$

Es könnte also die Kausalkette

- *Parodontose* \rightarrow *Bakterien* \rightarrow *Herzkrankheiten*

zu dieser Abhängigkeit führen.

Andererseits könnte mangelnde Gesundheitspflege (Z) sowohl die Neigung zu Herzkrankheiten als auch die Parodontose bewirken:

- *mangelnde Gesundheitspflege* \rightarrow (*Parodontose, Herzkrankheiten*)

und somit zu einer Korrelation zwischen Parodontose und Herzkrankheiten führen. Diese Korrelation, die etwas irreführend als Scheinkorrelation bezeichnet wird, ist durch eine Drittvariable (oder weitere Variablen wie Geschlecht oder Alter) induziert.

⁵ $\text{Cov}(X, Y) = E[XY] - E[X]E[Y] = E[X]E[Y] - E[X]E[Y] = 0$ bei Unabhängigkeit, da $E[XY] = E[X]E[Y]$. Unabhängigkeit ist jedoch eine stärkere Eigenschaft: $E[f(X)g(Y)] = E[f(X)]E[g(Y)]$ für beliebige Funktionen f, g . Bei Normalverteilungen ist allerdings die Unabhängigkeit und die Unkorreliertheit gleichbedeutend, da bei einer diagonalen Kovarianzmatrix die Dichtefunktion in Produkte faktorisiert. Siehe Kap. 18

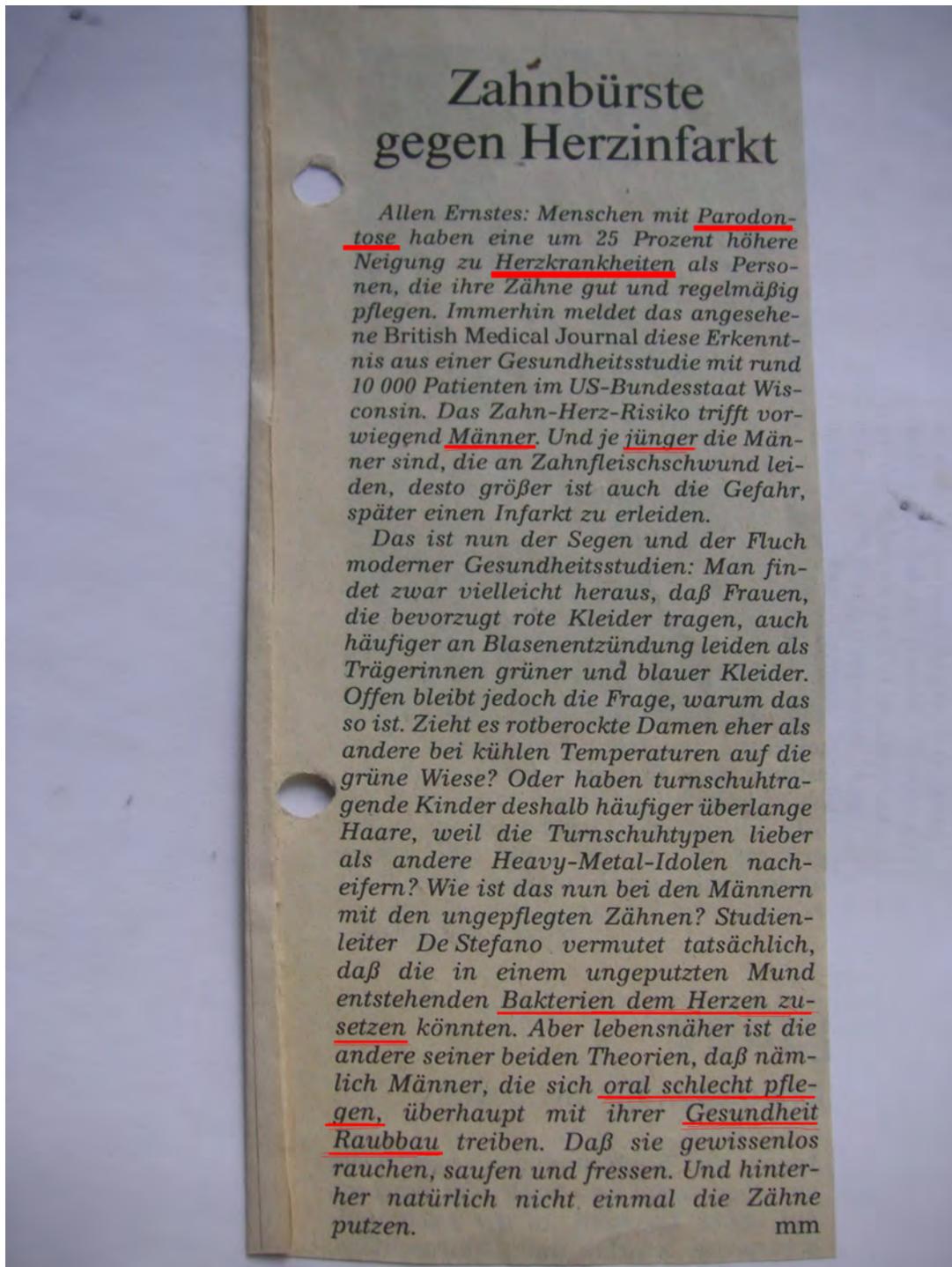


Abbildung 4.5: Zusammenhang Herzinfarkt und Parodontose.

Kapitel 5

Regressionsanalyse

5.1 Lineares Modell und Kleinste-Quadrate-Schätzung (KQ)

5.1.1 Das lineare Modell

Im Rahmen der Korrelationsanalyse wird ein symmetrischer Zusammenhang zwischen Variablen X und Y untersucht. Andererseits kann auch ein asymmetrisches Modell formuliert werden, das zumindest ansatzweise einen Kausalzusammenhang $X \rightarrow Y$ statistisch formuliert.

Der einfachste Ansatz ist das lineare Regressionsmodell (lineare Prognoseregel)

$$Y_n = \alpha + \beta X_n + \epsilon_n, n = 1, \dots, N, \quad (5.1)$$

**lineares
Regressionsmodell**

bei dem eine Gerade an die Daten angepaßt wird (Abb. 5.1, 5.2). Hierbei wird X als unabhängige Variable (UV) und Y als abhängige Variable (AV) bezeichnet.

Da in der Realität nicht alle Daten-Punkte (X_n, Y_n) auf der Geraden liegen, muß das (zu einfache) lineare Modell durch stochastische Gleichungsfehler ϵ_n ergänzt werden. Diese modellieren Spezifikationsfehler des Modells, etwa Abweichungen des Zusammenhangs vom linearen Modell oder weggelassene bzw. nicht erfaßte Variablen, die aber zumindest in stochastischer Weise berücksichtigt werden können.

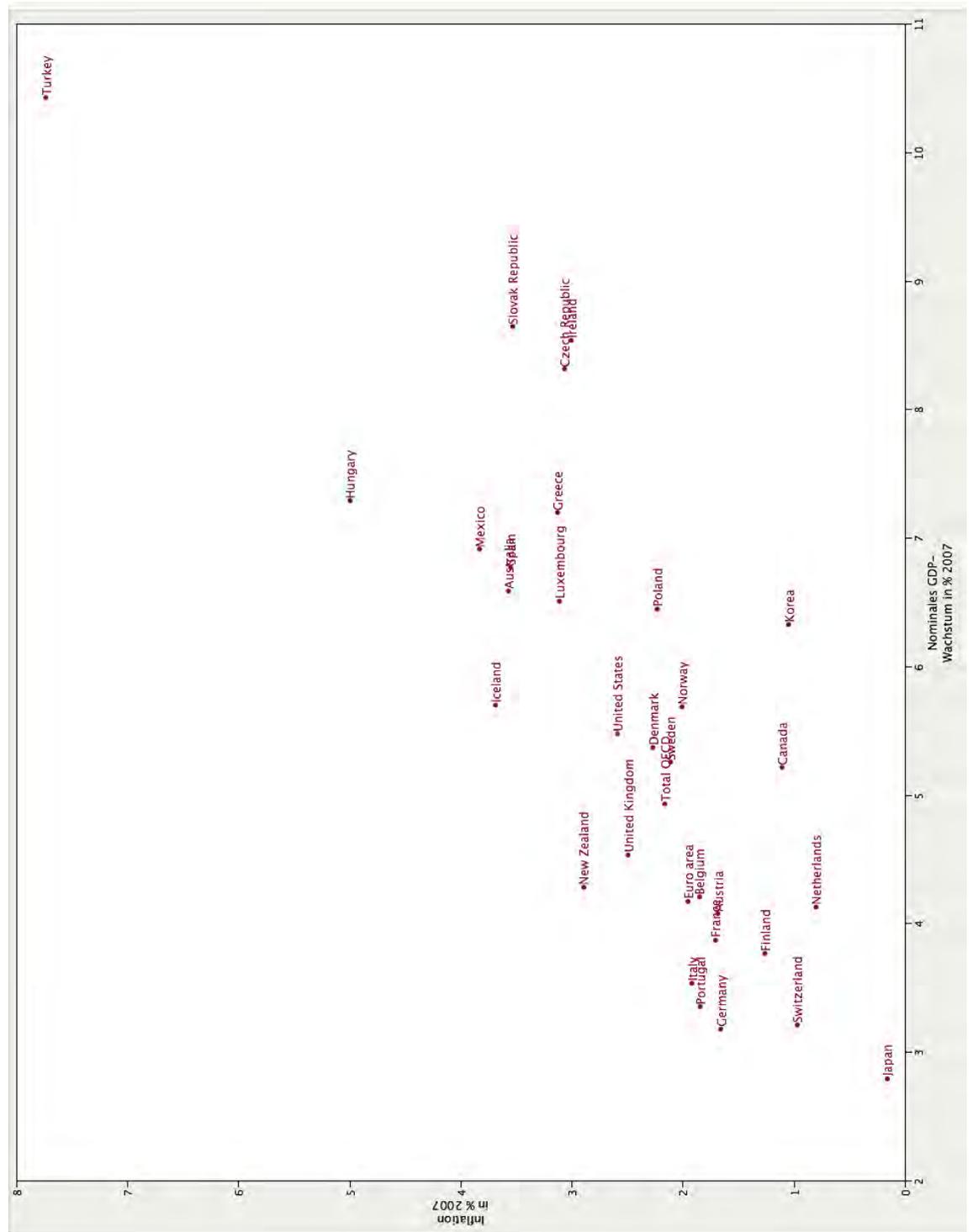


Abbildung 5.1: Streudiagramm BIP 2007 und Inflationsrate.

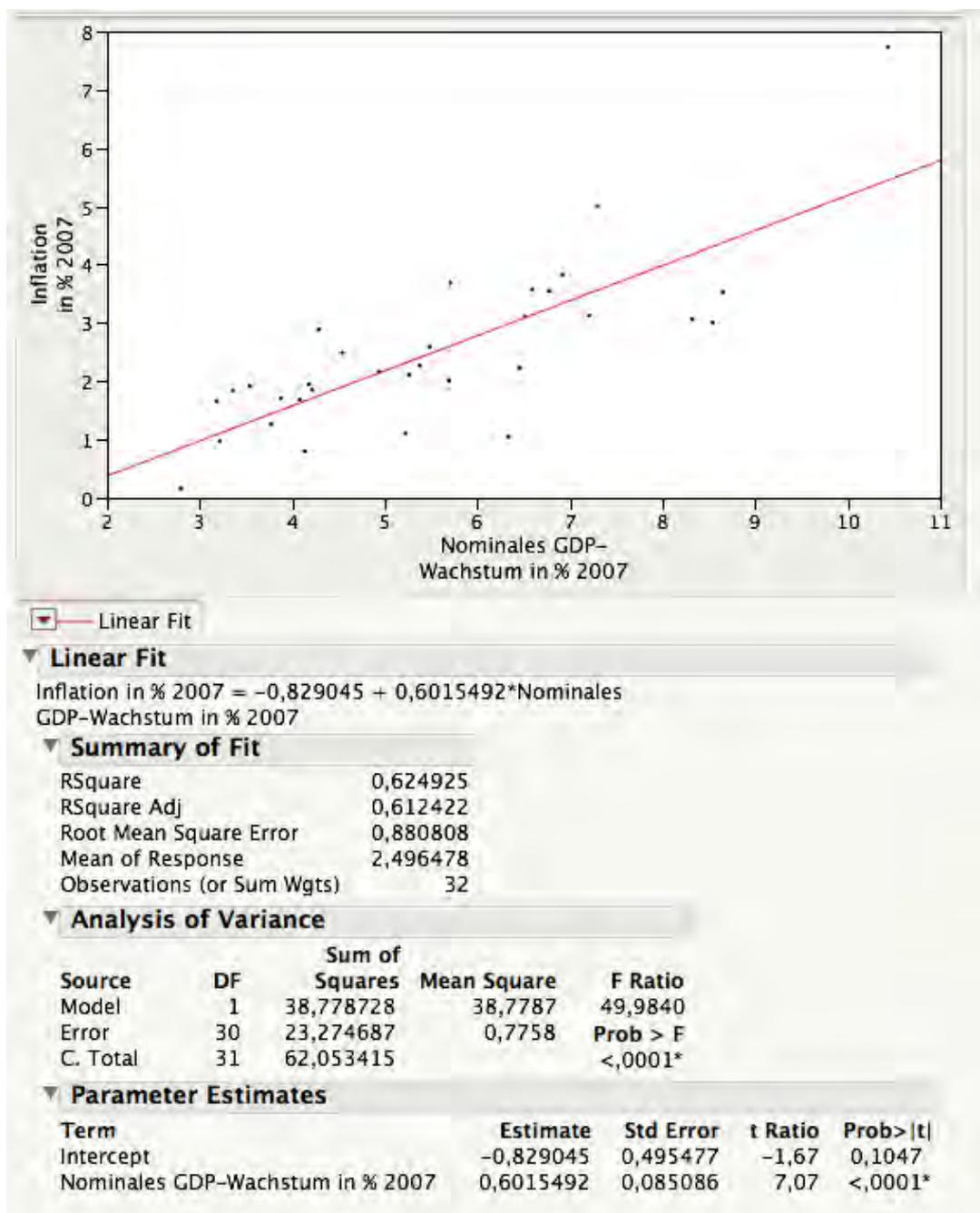


Abbildung 5.2: Regression BIP 2007 und Inflationsrate.

Die Parameter der Geraden sind:

- Achsenabschnitt α
- Steigung β
- Varianz des Gleichungsfehlers (Störgröße) $\sigma^2 = \text{Var}(\epsilon_n)$

wahre Gerade

Bei Glg. 5.1 handelt es sich allerdings um die wahre Gerade, die durch die Populationsparameter α, β und $\sigma^2 = \text{Var}(\epsilon_n)$ definiert ist. Diese können nur geschätzt werden. Im Bild sieht man die geschätzte Gerade (lineare Prognose)

$$\hat{Y} = \hat{\alpha} + \hat{\beta}X. \quad (5.2)$$

5.1.2 Kleinste-Quadrate-Schätzung (KQ)

5.1.2.1 Berechnung der Kleinste-Quadrate-Schätzer

Die Schätzer für α, β und $\sigma^2 = \text{Var}(\epsilon_n)$ können mit Hilfe des Prinzips der kleinsten Quadrate (KQ) gefunden werden. Hierbei wird α und β so gewählt, daß

$$S(\alpha, \beta) = \sum_n (Y_n - \alpha - \beta X_n)^2 = \sum_n \epsilon_n^2 \quad (5.3)$$

minimal wird, also die Gerade optimal an die Datenpunkte (X_n, Y_n) angepaßt wird.

Differenziert man $S(\alpha, \beta)$ nach α und β und setzt die Ableitungen Null, so ergeben sich die KQ-Schätzer

kleinste Quadrate-Schätzer (KQ, OLS)

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} \quad (5.4)$$

$$\hat{\beta} = \frac{\sum_n (X_n - \bar{X})(Y_n - \bar{Y})}{\sum_n (X_n - \bar{X})^2} \quad (5.5)$$

$$= R_{xy} \frac{S_y}{S_x}. \quad (5.6)$$

Hierbei ist R_{xy} der Korrelationskoeffizient und $\bar{X} = \frac{1}{N} \sum X_n$ sowie $\bar{Y} = \frac{1}{N} \sum Y_n$ sind die Mittelwerte.

Aus den geschätzten Parametern können die **geschätzten Y-Werte**

$$\hat{Y}_n = \hat{\alpha} + \hat{\beta}X_n \quad (5.7)$$

$$= \bar{Y} + \hat{\beta}(X_n - \bar{X}), \quad (5.8)$$

die Residuen (geschätzte Gleichungsfehler)

Residuen

$$\hat{\epsilon}_n = Y_n - \hat{Y}_n \quad (5.9)$$

sowie der Varianz-Schätzer

$$\hat{\sigma}^2 = \frac{1}{N-2} \sum_n \hat{\epsilon}_n^2 \quad (5.10)$$

gewonnen werden.

Der Mittelwert der geschätzten Y-Werte ist gleich dem Mittelwert

$$\bar{\hat{Y}} = \hat{\alpha} + \hat{\beta}\bar{X} = \bar{Y} \quad (5.11)$$

sodaß sich die Residuen zu Null addieren

$$\sum_n \hat{\epsilon}_n = \sum_n (Y_n - \hat{Y}_n) = N(\bar{Y} - \bar{Y}) = 0. \quad (5.12)$$

Beispiel 5.1 (BIP 2007 und Inflationsrate)

In Abb. 5.2 wurde eine Gerade an die Datenpunkte $(x_n, y_n) = (\text{BIP}_n, \text{Inflation}_n)$ angepaßt. Die Daten sind nochmals in Abb. 5.3 angegeben (gerundet auf 2 Kommastellen).

Die entsprechenden Berechnungen lauten im Detail:

Zunächst werden die Mittelwerte und Quadratsummen in

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \quad (5.13)$$

$$\hat{\beta} = \frac{\sum_n x_n y_n - N\bar{x}\bar{y}}{\sum_n x_n^2 - N\bar{x}^2} \quad (5.14)$$

$$(5.15)$$

	Land	Nominales GDP- Wachstum in % 2007	Inflation in % 2007
1	Australia	6,59	3,57
2	Austria	4,08	1,68
3	Belgium	4,21	1,85
4	Canada	5,22	1,11
5	Czech Republic	8,32	3,06
6	Denmark	5,38	2,27
7	Finland	3,77	1,26
8	France	3,88	1,71
9	Germany	3,18	1,66
10	Greece	7,20	3,13
11	Hungary	7,30	4,99
12	Iceland	5,71	3,69
13	Ireland	8,54	3,00
14	Italy	3,54	1,92
15	Japan	2,80	0,16
16	Korea	6,33	1,05
17	Luxembourg	6,51	3,11
18	Mexico	6,92	3,83
19	Netherlands	4,13	0,80
20	New Zealand	4,29	2,89
21	Norway	5,69	2,01
22	Poland	6,45	2,23
23	Portugal	3,36	1,84
24	Slovak Republic	8,65	3,53
25	Spain	6,78	3,55
26	Sweden	5,26	2,11
27	Switzerland	3,22	0,97
28	Turkey	10,43	7,73
29	United Kingdom	4,54	2,49
30	United States	5,48	2,59
31	Euro area	4,18	1,95
32	Total OECD	4,94	2,16

Abbildung 5.3: Daten BIP 2007 und Inflationsrate 2007.

	x	y	x^2	xy	y^2	\hat{y}	$\hat{\epsilon}$	$\hat{\epsilon}^2$
1	6.59	3.57	43.43	23.53	12.74	3.14	0.43	0.19
2	4.08	1.68	16.65	6.85	2.82	1.63	0.05	0
3	4.21	1.85	17.72	7.79	3.42	1.7	0.15	0.02
4	5.22	1.11	27.25	5.79	1.23	2.31	-1.2	1.44
5	8.32	3.06	69.22	25.46	9.36	4.18	-1.12	1.25
6	5.38	2.27	28.94	12.21	5.15	2.41	-0.14	0.02
7	3.77	1.26	14.21	4.75	1.59	1.44	-0.18	0.03
8	3.88	1.71	15.05	6.63	2.92	1.51	0.2	0.04
9	3.18	1.66	10.11	5.28	2.76	1.09	0.57	0.33
10	7.2	3.13	51.84	22.54	9.8	3.5	-0.37	0.14
11	7.3	4.99	53.29	36.43	24.9	3.56	1.43	2.04
12	5.71	3.69	32.6	21.07	13.62	2.61	1.08	1.17
13	8.54	3.	72.93	25.62	9.	4.31	-1.31	1.71
14	3.54	1.92	12.53	6.8	3.69	1.3	0.62	0.38
15	2.8	0.16	7.84	0.45	0.03	0.86	-0.7	0.49
16	6.33	1.05	40.07	6.65	1.1	2.98	-1.93	3.72
17	6.51	3.11	42.38	20.25	9.67	3.09	0.02	0
18	6.92	3.83	47.89	26.5	14.67	3.33	0.5	0.25
19	4.13	0.8	17.06	3.3	0.64	1.66	-0.86	0.73
20	4.29	2.89	18.4	12.4	8.35	1.75	1.14	1.29
21	5.69	2.01	32.38	11.44	4.04	2.59	-0.58	0.34
22	6.45	2.23	41.6	14.38	4.97	3.05	-0.82	0.68
23	3.36	1.84	11.29	6.18	3.39	1.19	0.65	0.42
24	8.65	3.53	74.82	30.53	12.46	4.37	-0.84	0.71
25	6.78	3.55	45.97	24.07	12.6	3.25	0.3	0.09
26	5.26	2.11	27.67	11.1	4.45	2.34	-0.23	0.05
27	3.22	0.97	10.37	3.12	0.94	1.11	-0.14	0.02
28	10.43	7.73	108.78	80.62	59.75	5.45	2.28	5.22
29	4.54	2.49	20.61	11.3	6.2	1.9	0.59	0.34
30	5.48	2.59	30.03	14.19	6.71	2.47	0.12	0.01
31	4.18	1.95	17.47	8.15	3.8	1.69	0.26	0.07
32	4.94	2.16	24.4	10.67	4.67	2.14	0.02	0
Summe	176.88	79.9	1084.83	506.07	261.45	79.9	0	23.21

Tabelle 5.1: Berechnungstabelle BIP 2007 und Inflationsrate 2007.

berechnet (Tab. 5.1). Daraus ergibt sich ($\bar{x} = 5.528, \bar{y} = 2.497$)

$$\hat{\alpha} = 2.497 - 0.601 \times 5.528 = -0.827 \quad (5.16)$$

$$\hat{\beta} = \frac{506.07 - 32 \times 5.528 \times 2.497}{1084.83 - 32 \times 5.582^2} = \frac{64.42}{107.12} = 0.601 \quad (5.17)$$

Dies stimmt mit den Berechnungen des Computerprogramms überein (Abb. 5.2, unten; intercept = Achsenabschnitt = $\hat{\alpha}$).

Die geschätzten Datenpunkte ergeben sich aus

$$\hat{y}_n = \hat{\alpha} + \hat{\beta}x_n, \quad (5.18)$$

(Spalte ydach), die Residuen (geschätzte Gleichungsfehler) durch

$$\hat{\epsilon}_n = y_n - \hat{y}_n \quad (5.19)$$

(Spalte epsdach) sowie der Varianz-Schätzer durch Summation

$$\hat{\sigma}^2 = \frac{1}{N-2} \sum_n \hat{\epsilon}_n^2 = 23.21/30 = .774. \quad (5.20)$$

Die Residuen summieren sich zu 0. Dies gilt ganz allgemein (warum?).

Die geschätzte (empirische Regressionsgerade) ist also

$$\hat{y} = -0.827 + 0.601 x \quad (5.21)$$

mit Achsenabschnitt $\hat{\alpha} = -0.827 = \hat{y}(x = 0)$.

■

Es ist wichtig, sich klarzumachen, daß die Gerade durch die Mittelwerte $\bar{x} = 5.528, \bar{y} = 2.497$ geht, denn

$$\hat{y} = \hat{\alpha} + \hat{\beta}x \quad (5.22)$$

$$= \bar{y} - \hat{\beta}\bar{x} + \hat{\beta}x \quad (5.23)$$

$$= \bar{y} + \hat{\beta}(x - \bar{x}) \quad (5.24)$$

$$\hat{y}(x = \bar{x}) = \bar{y}. \quad (5.25)$$

(Abb. 5.4). Der Zeichnung ist auch zu entnehmen, daß die Steigung durch

$$\hat{\beta} = (\bar{y} - \hat{\alpha})/\bar{x} \quad (5.26)$$

gegeben ist. Umstellen ergibt $\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$.

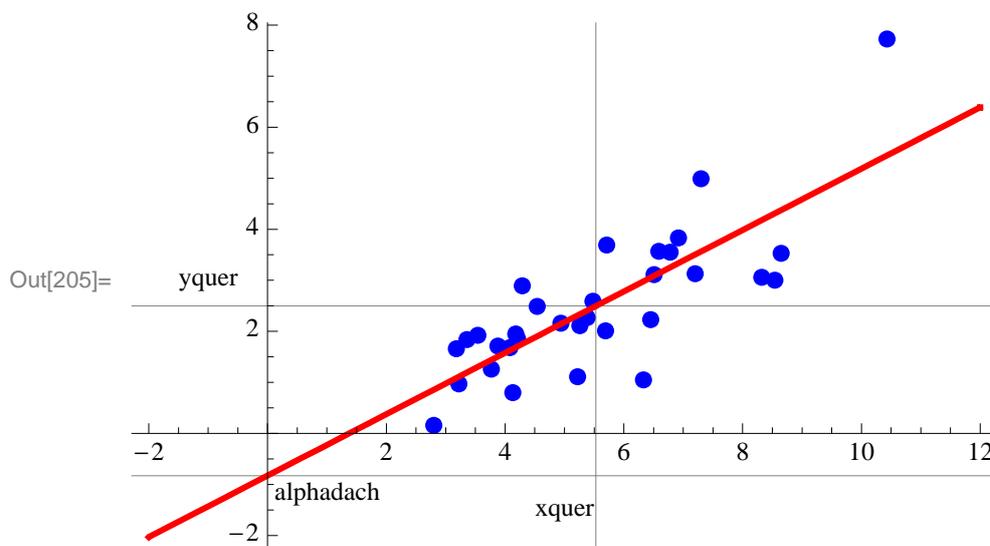


Abbildung 5.4: Geschätzte Gerade mit Achsenabschnitt $\hat{\alpha}$ und Mittelwerten \bar{x}, \bar{y} .

5.1.2.2 Eigenschaften der Kleinste-Quadrate-Schätzer

Im allgemeinen werden beim klassischen Regressionsmodell folgende **Annahmen** gemacht:

1. $\epsilon_n \sim N(0, \sigma^2)$: Normalverteilungsannahme
2. $\text{Cov}(\epsilon_n, \epsilon_m) = 0$ für $n \neq m$: Unkorreliertheit der Gleichungsfehler
3. $E[\epsilon_n] = 0$
4. $\text{Var}[\epsilon_n] = \sigma^2$: Varianzhomogenität (Homoskedastizität)
5. X_n ist deterministisch

**Annahmen beim
klassischen
Regressionsmodell**

Letztere Annahme ist im obigen Beispiel 5.5 nicht erfüllt, da das BIP-Wachstum der einzelnen Länder nicht vorgegeben werden kann, sondern faktisch vorliegt. Fixe, deterministische Regressoren lassen sich nur in einem experimentellen Design realisieren, wo die x -Werte (etwa bestimmte Dosen von Medikamenten oder andere Stimuli) vom Versuchsleiter vorgegeben werden.

Die folgenden Resultate gelten jedoch auch **bedingt** auf die gemessenen x -Werte.

**stochastische
Regressoren**

In Erweiterung des Modells können auch stochastische Regressoren eingeführt werden, deren Verteilung $f(x)$ bzw. Momente $E[X]$, $\text{Var}(X)$ zusätzlich zu schätzen sind.

Etwa enthalten Strukturgleichungsmodelle $\eta = B\eta + \Gamma\xi + \zeta$ wie LISREL stochastische Regressoren ξ und zusätzliche Parameter $\Phi = \text{Var}(\xi)$, d.h. die Kovarianzmatrix der Regressoren.

Da die Regressoren deterministisch sind (bzw. bedingt auf diese), sind somit die KQ-Schätzer

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} \quad (5.27)$$

$$\hat{\beta} = \frac{\sum_n (X_n - \bar{X})(Y_n - \bar{Y})}{\sum_n (X_n - \bar{X})^2} \quad (5.28)$$

**Normalverteilung
der KQ-Schätzer**

normalverteilt ($\hat{\sigma}^2$ ist χ^2 -verteilt). Dies folgt aus der Normalverteilung von $Y_n = \alpha + \beta X_n + \epsilon_n$ (Annahme 1) sowie der Eigenschaft, daß Summen von normalverteilten Größen wieder normalverteilt sind. Es genügt also, die Erwartungswerte und Varianzen der KQ-Schätzer zu berechnen:

Satz 5.1 (Erwartungstreue der KQ-Schätzer)

- $E[\hat{\alpha}] = \alpha$
- $E[\hat{\beta}] = \beta$
- $E[\hat{\sigma}^2] = \sigma^2$

Beweis:

$$E[\hat{\beta}] = \frac{\sum_n X_n E[Y_n] - N\bar{X}E[\bar{Y}]}{\sum_n X_n^2 - N\bar{X}^2} \quad (5.29)$$

$$= \frac{\sum_n X_n(\alpha + \beta X_n) - N\bar{X}(\alpha + \beta\bar{X})}{\sum_n X_n^2 - N\bar{X}^2} \quad (5.30)$$

$$= \frac{\beta(\sum_n X_n^2 - N\bar{X}^2)}{\sum_n X_n^2 - N\bar{X}^2} = \beta \quad (5.31)$$

$$E[\hat{\alpha}] = E[\bar{Y}] - E[\hat{\beta}]\bar{X} = \alpha + \beta\bar{X} - \beta\bar{X} = \alpha \quad (5.32)$$

$$E[\hat{\sigma}^2] = \frac{\sigma^2}{N-2} E\left[\sum_n \hat{\epsilon}_n^2 / \sigma^2\right] = \sigma^2. \quad (5.33)$$

Die letzte Zeile ergibt sich aus der Tatsache, daß $\sum_n \hat{\epsilon}_n^2 / \sigma^2 \sim \chi^2(N-2)$ und $E[\chi^2(N-2)] = N-2$ (Kap. 1.3.2).

■

Satz 5.2 (Varianzen der KQ-Schätzer)

- $\sigma_\alpha^2 = \text{Var}(\hat{\alpha}) = \sigma^2 \left(\frac{1}{N} + \frac{\bar{X}^2}{\sum_n (X_n - \bar{X})^2} \right)$
- $\sigma_\beta^2 = \text{Var}(\hat{\beta}) = \frac{\sigma^2}{\sum_n (X_n - \bar{X})^2}$

Beweis:

$$\text{Var}(\hat{\beta}) = \text{Var} \left(\frac{\sum_n (X_n - \bar{X}) Y_n}{\sum_n (X_n - \bar{X})^2} \right) \quad (5.34)$$

$$= \frac{\sum_n (X_n - \bar{X})^2 \text{Var}(Y_n)}{[\sum_n (X_n - \bar{X})^2]^2} \quad (5.35)$$

$$= \frac{\sigma^2}{\sum_n (X_n - \bar{X})^2} \quad (5.36)$$

$$\text{Var}(\hat{\alpha}) = \text{Var}(\bar{Y} - \hat{\beta} \bar{X}) \quad (5.37)$$

$$= \text{Var}(\bar{Y}) - 2\text{Cov}(\bar{Y}, \hat{\beta} \bar{X}) + \text{Var}(\hat{\beta} \bar{X}) \quad (5.38)$$

$$= \sigma^2/N + \text{Var}(\hat{\beta}) \bar{X}^2 \quad (5.39)$$

$$= \sigma^2 \left(\frac{1}{N} + \frac{\bar{X}^2}{\sum_n (X_n - \bar{X})^2} \right) \quad (5.40)$$

Bei obiger Berechnung wurde ausgenutzt, daß

$$\sum_n (X_n - \bar{X})(Y_n - \bar{Y}) = \sum_n (X_n - \bar{X}) Y_n, \quad (5.41)$$

sowie

$$\text{Cov}(\bar{Y}, \hat{\beta} \bar{X}) = \frac{\bar{X}}{\sum_n (X_n - \bar{X})^2} \sum_n \text{Cov}(\bar{Y}, Y_n)(X_n - \bar{X}) \quad (5.42)$$

$$= \frac{\bar{X}}{\sum_n (X_n - \bar{X})^2} \sum_n (X_n - \bar{X})(\sigma^2/N) = 0 \quad (5.43)$$

und $\text{Cov}(\bar{Y}, Y_n) = \sigma^2/N$.

■

5.2 Konstruktion von Konfidenzintervallen und Tests

Da die Gleichungsfehler ϵ_n als normalverteilt angenommen wurden, lassen sich für die KQ-Schätzer

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} \sim N(\alpha, \sigma_\alpha^2) \quad (5.44)$$

$$\hat{\beta} = \frac{\sum_n (X_n - \bar{X})(Y_n - \bar{Y})}{\sum_n (X_n - \bar{X})^2} \sim N(\beta, \sigma_\beta^2) \quad (5.45)$$

einfache Konfidenzintervalle und Tests ableiten.

Etwa ist $(\hat{\alpha} - \alpha)/\sigma_\alpha \sim N(0, 1)$. Da jedoch die Varianz σ^2 des Gleichungsfehlers nicht bekannt ist, müssen die t -verteilten Variablen

$$\frac{\hat{\alpha} - \alpha}{\widehat{\sigma}_\alpha} \sim t(N - 2) \quad (5.46)$$

$$\frac{\hat{\beta} - \beta}{\widehat{\sigma}_\beta} \sim t(N - 2) \quad (5.47)$$

benutzt werden. Man erhält so die t -Konfidenzintervalle

t -Konfidenzintervalle für Regressionsparameter

$$P\{\alpha \in \hat{\alpha} \pm t(1 - \alpha/2, N - 2) \widehat{\sigma}_\alpha\} = 1 - \alpha \quad (5.48)$$

$$P\{\beta \in \hat{\beta} \pm t(1 - \alpha/2, N - 2) \widehat{\sigma}_\beta\} = 1 - \alpha \quad (5.49)$$

mit

$$\widehat{\sigma}_\alpha = \hat{\sigma} \sqrt{\frac{1}{N} + \frac{\bar{X}^2}{\sum_n (X_n - \bar{X})^2}} = \hat{\sigma} \sqrt{\frac{\sum_n X_n^2}{N \sum_n (X_n - \bar{X})^2}} \quad (5.50)$$

$$\widehat{\sigma}_\beta = \frac{\hat{\sigma}}{\sqrt{\sum_n (X_n - \bar{X})^2}}. \quad (5.51)$$

Entsprechend können t -Tests für die Parameter α und β mit Hilfe der Testgrößen

$$T_\alpha = \frac{\hat{\alpha} - \alpha_0}{\widehat{\sigma}_\alpha} \sim t(N - 2) \quad (5.52)$$

$$T_\beta = \frac{\hat{\beta} - \beta_0}{\widehat{\sigma}_\beta} \sim t(N - 2) \quad (5.53)$$

**t-Tests für
Regressions-
parameter**

ausgeführt werden. Die entsprechenden Hypothesenpaare sind wie immer (entsprechend für β):

2-seitiger Test:

$$H_0 : \alpha = \alpha_0 \quad (5.54)$$

$$H_1 : \alpha \neq \alpha_0 \quad (5.55)$$

$$\text{berechne } t(1 - \alpha/2, N - 2) \quad (5.56)$$

$$\text{falls } |T_\alpha| > t : H_0 \text{ ablehnen} \quad (5.57)$$

1-seitiger Test; kritischer Wert rechts:

$$H_0 : \alpha \leq \alpha_0 \quad (5.58)$$

$$H_1 : \alpha > \alpha_0 \quad (5.59)$$

$$\text{berechne } t(1 - \alpha, N - 2) \quad (5.60)$$

$$\text{falls } T_\alpha > t : H_0 \text{ ablehnen} \quad (5.61)$$

1-seitiger Test; kritischer Wert links:

$$H_0 : \alpha \geq \alpha_0 \quad (5.62)$$

$$H_1 : \alpha < \alpha_0 \quad (5.63)$$

$$\text{berechne } t(\alpha, N - 2) \quad (5.64)$$

$$\text{falls } T_\alpha < t : H_0 \text{ ablehnen} \quad (5.65)$$

Beispiel 5.2 (BIP 2007 und Inflationsrate (KI und Tests))

Im Beispiel 5.5 war

$$\hat{\alpha} = -0.827 \quad (5.66)$$

$$\hat{\beta} = 0.601 \quad (5.67)$$

$$\widehat{\sigma}^2 = .774 \quad (5.68)$$

$$\hat{\sigma} = \sqrt{.774} = .88 \quad (5.69)$$

Weiterhin benötigt man die Hilfsgrößen

$$\sum x_n^2 = 1084.83 \quad (5.70)$$

$$\sum (x_n - \bar{x})^2 = \sum x_n^2 - N\bar{x}^2 = 107.121. \quad (5.71)$$

Daraus findet man die geschätzten Standardfehler der Schätzer $\hat{\alpha}, \hat{\beta}$

$$\hat{\sigma}_\alpha = \hat{\sigma} \sqrt{\frac{\sum_n x_n^2}{N \sum_n (x_n - \bar{x})^2}} \quad (5.72)$$

$$= .88 \sqrt{\frac{1084.83}{32 \times 107.121}} = .495 \quad (5.73)$$

$$\hat{\sigma}_\beta = \frac{\hat{\sigma}}{\sqrt{\sum_n (x_n - \bar{x})^2}} \quad (5.74)$$

$$= \frac{.88}{\sqrt{107.121}} = 0.085. \quad (5.75)$$

Ein 95%-Konfidenzintervall für α ist somit

$$\hat{\alpha} \pm t(1 - \alpha/2, N - 2) \hat{\sigma}_\alpha = -0.827 \pm 1.01092 \quad (5.76)$$

$$= [-1.838, 0.1837] \quad (5.77)$$

mit dem t -Quantil $t(.975, 30) = 2.04$.

Der t -Test für $H_0 : \alpha = 0$ ergibt:

$$t_\alpha = \frac{\hat{\alpha} - \alpha_0}{\hat{\sigma}_\alpha} \quad (5.78)$$

$$= -0.827/.495 = -1.67. \quad (5.79)$$

Damit muß die Nullhypothese beibehalten werden. Entsprechend überdeckte das KI den Wert $\alpha = 0$.

Das Konfidenz-Intervall für β lautet:

$$\hat{\beta} \pm t(1 - \alpha/2, N - 2) \hat{\sigma}_\beta = 0.601 \pm 0.173 \quad (5.80)$$

$$= [0.427804, 0.774932] \quad (5.81)$$

mit dem t -Quantil $t(.975, 30) = 2.04$.

Der t -Test für $H_0 : \beta = 0$ ergibt:

$$t_\beta = \frac{\hat{\beta} - \beta_0}{\hat{\sigma}_\beta} \quad (5.82)$$

$$= 0.601/0.084 = 7.076. \quad (5.83)$$

Damit muß die Nullhypothese abgelehnt werden (das KI überdeckte den Wert $\beta = 0$ nicht).



5.3 Konfidenzintervalle für prognostizierte Werte

Das lineare Prognose-Modell lautet

$$Y = \alpha + \beta X + \epsilon. \quad (5.84)$$

Wenn nun α und β bekannt wären, hätte man für die Prognose

$$\hat{Y} = \alpha + \beta X = E[Y|X] \quad (5.85)$$

den Schätzfehler

$$Y - \hat{Y} = \epsilon \quad (5.86)$$

mit Varianz $\text{Var}(\epsilon) = \sigma^2$ zu betrachten.

Jedoch sind im allgemeinen nur Schätzungen für α und β erhältlich, was zu einer zusätzlichen Fehlerquelle führt:

$$\hat{Y} = \hat{\alpha} + \hat{\beta}X = \hat{E}[Y|X]. \quad (5.87)$$

Bei der Einbeziehung von Unsicherheit in der Geradengleichung bzw. der Prognose müssen zwei Fragestellungen unterschieden werden:

1. Konfidenzintervall für die wahre Gerade $\alpha + \beta X_0 = E[Y|X_0]$, also ein zufälliges Intervall, das die Gerade mit Wahrscheinlichkeit $1 - \alpha$ überdeckt.
2. Vorhersage(Prognose)-Intervall für ein einzelnes (individuelles) Y_0 : Prognose $\hat{Y}_0 = \hat{\alpha} + \hat{\beta}X_0$ mit Prognose-Fehler-Varianz $E[Y_0 - \hat{Y}_0]^2$.

Unterschied:

1. Es wird für einen festen, aber unbekanntem Wert $\alpha + \beta X_0$ ein zufälliges Intervall gesucht.
2. Es wird für einen zufälligen, unbekanntem Wert $Y_0 = \alpha + \beta X_0 + \epsilon_0$ ein Vorhersage(Prognose)-Intervall gesucht.

5.3.1 Konfidenzintervall für $E[Y|X] = \alpha + \beta X$ (Fall 1)

Analog zum Vorgehen bei Parametern berechnet man zunächst einen Punktschätzer für $E[Y|X] = \alpha + \beta X$, nämlich

$$\hat{E}[Y|X] = \hat{\alpha} + \hat{\beta}X = \hat{E}. \quad (5.88)$$

Dessen Varianz ist durch

$$\text{Var}(\hat{E}) = \text{Var}(\hat{\alpha} + \hat{\beta}X) \quad (5.89)$$

$$= \text{Var}(\hat{\alpha}) + \text{Var}(\hat{\beta}X) + 2 \text{Cov}(\hat{\alpha}, \hat{\beta}X) \quad (5.90)$$

$$= \sigma^2 \left(\frac{1}{N} + \frac{(X - \bar{X})^2}{\sum_n (X_n - \bar{X})^2} \right) \quad (5.91)$$

gegeben. Dies beruht auf der Unkorreliertheit $\text{Cov}(\bar{Y}, \hat{\beta}) = 0$ (Glg. 5.42). Damit gilt $\text{Cov}(\hat{\alpha}, \hat{\beta}) = \text{Cov}(\bar{Y} - \hat{\beta}\bar{X}, \hat{\beta}) = -\bar{X}\text{Var}(\hat{\beta})$. Einsetzen der Varianzen $\text{Var}(\hat{\alpha})$, $\text{Var}(\hat{\beta})$ ergibt den gesuchten Varianzterm der geschätzten Gerade.

Ersetzt man auch noch σ^2 durch $\hat{\sigma}^2$, so ist

$$\frac{\hat{E} - E}{\sqrt{\widehat{\text{Var}}(\hat{E})}} \sim t(N - 2). \quad (5.92)$$

Daraus ergibt sich das Konfidenzintervall für die Gerade $E[Y|X] = \alpha + \beta X$

**Konfidenzintervall
für Regressions-
Gerade**

$$\hat{E}[Y|X] \pm t(1 - \alpha/2, N - 2) \sqrt{\widehat{\text{Var}}(\hat{E})} \quad (5.93)$$

mit

$$\sqrt{\widehat{\text{Var}}(\hat{E})} = \hat{\sigma} \sqrt{\frac{1}{N} + \frac{(X - \bar{X})^2}{\sum_n (X_n - \bar{X})^2}} \quad (5.94)$$

Das Konfidenzintervall hängt sowohl bei der Position als auch in der Breite von der Stelle X ab. Für große $|X|$ wird die gekrümmte Kurve zu einer Geraden.

Beispiel 5.3 (BIP 2007 und Inflationsrate (Konfidenzintervall))

Im Beispiel 5.5 war

$$\begin{aligned}\hat{\alpha} &= -0.827 \\ \hat{\beta} &= 0.601 \\ \hat{\sigma}^2 &= .774 \\ \hat{\sigma} &= \sqrt{.774} = .88 \\ \sum (x_n - \bar{x})^2 &= \sum x_n^2 - N\bar{x}^2 = 107.121.\end{aligned}$$

Daraus findet man den Streuungsterm

$$\begin{aligned}\sqrt{\widehat{\text{Var}}(\hat{E})} &= \hat{\sigma} \sqrt{\frac{1}{N} + \frac{(x - \bar{x})^2}{\sum_n (x_n - \bar{x})^2}} \\ &= 0.88 \sqrt{\frac{1}{32} + \frac{(x - 5.5275)^2}{107.121}}.\end{aligned}$$

Das 95%-KI ist damit

$$(-0.827 + 0.601 x) \pm 1.7972 \sqrt{1/32 + 0.00934(-5.5275 + x)^2}$$

mit dem Quantil $t(.975, 30) = 2.04227$. An der Stelle $x = \bar{x} = 5.5275$ ergibt sich der minimale Wert

$$2.497 \pm 0.318.$$

Abb. 5.5 zeigt das Konfidenz-Intervall als Konfidenz-Band für alle x -Werte zwischen 0 und 11.

■

**5.3.2 Prognoseintervall für individuelles Y_0
(Fall 2)**

Der individuelle zufällige Wert $Y_0 = \alpha + \beta X_0 + \epsilon_0 = E[Y_0|X_0] + \epsilon_0$ kann mit Hilfe von $\hat{Y}_0 = \hat{\alpha} + \hat{\beta} X_0$ geschätzt werden. Der dabei zu erwartende quadrierte Prognose-Fehler ist (vgl. 5.91; alles folgende bedingt auf X_0)

$$E[Y_0 - \hat{Y}_0]^2 = \sigma^2 \left(1 + \frac{1}{N} + \frac{(X_0 - \bar{X})^2}{\sum_n (X_n - \bar{X})^2} \right). \quad (5.95)$$

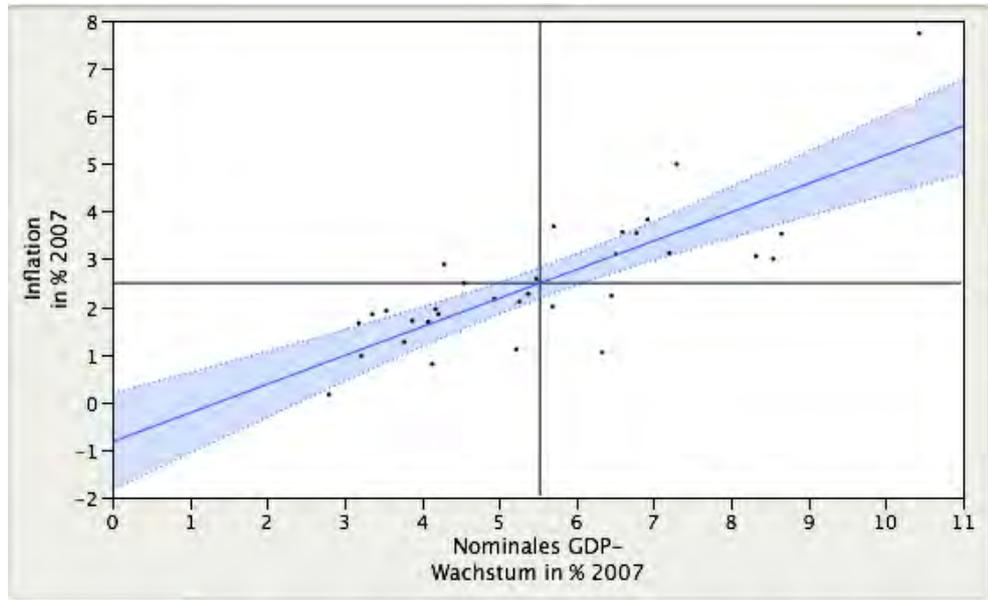


Abbildung 5.5: Geschätzte Gerade mit Konfidenzband $\hat{E}[Y|X] \pm t(1 - \alpha/2, N - 2)\sqrt{\widehat{\text{Var}}(\hat{E})}$.

Man erhält also im Vergleich zu Glg. 5.91 einen zusätzlichen Term σ^2 , der durch den Gleichungsfehler ϵ_0 im stochastischen $Y_0 = \alpha + \beta X_0 + \epsilon_0$ erzeugt wird.

Dies kann wie folgt gezeigt werden (bedingt auf X_0):

Da $E[Y_0] = \alpha + \beta X_0$ und $E[\hat{Y}_0] = \alpha + \beta X_0$ (Erwartungstreue der KQ-Schätzer) gilt

$$E[Y_0 - \hat{Y}_0]^2 = \text{Var}(Y_0 - \hat{Y}_0) = \text{Var}(Y_0) + \text{Var}(\hat{Y}_0). \quad (5.96)$$

Hierbei wurde $\text{Cov}(Y_0, \hat{Y}_0) = \text{Cov}(\alpha + \beta X_0 + \epsilon_0, \hat{\alpha} + \hat{\beta} X_0) = 0$ ausgenutzt, da die KQ-Schätzer unabhängig vom Gleichungsfehler ϵ_0 sind. Außerdem gilt ganz allgemein für Zufallsvariablen $E[Z^2] = \text{Var}(Z)$ für $E[Z] = 0$.

Setzt man noch $\text{Var}(Y_0) = \sigma^2$ und $\text{Var}(\hat{Y}_0) = \sigma^2 \left(\frac{1}{N} + \frac{(X_0 - \bar{X})^2}{\sum_n (X_n - \bar{X})^2} \right)$ (Fall 1) ein und ersetzt wieder $\sigma \rightarrow \hat{\sigma}$, so ergibt sich das gesuchte Prognose-Intervall

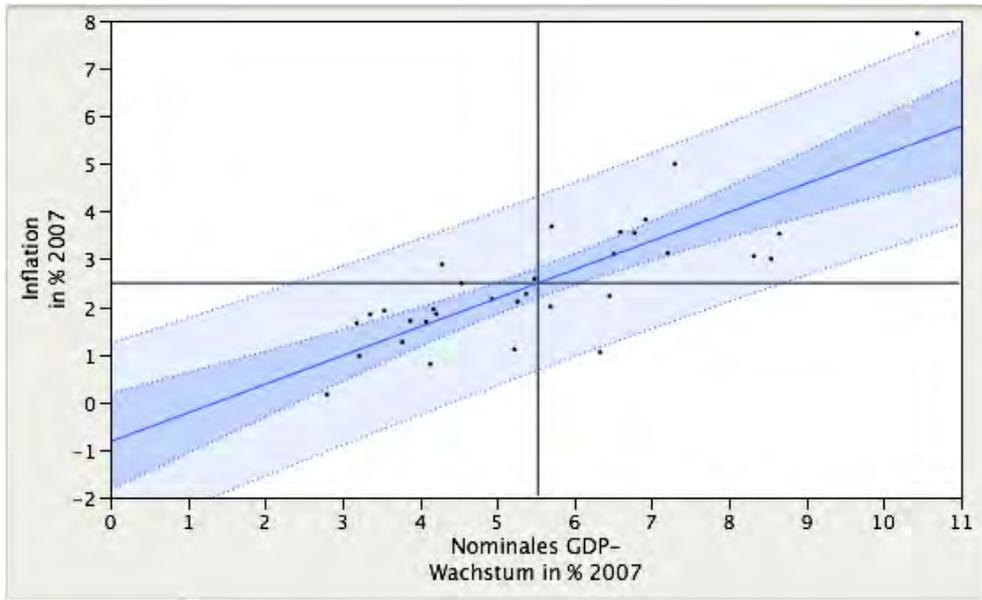


Abbildung 5.6: Geschätzte Gerade mit individuellem Prognoseband $\hat{Y}_0 \pm t(1 - \alpha/2, N - 2)\sqrt{\hat{E}[Y_0 - \hat{Y}_0]^2}$. Das Band ist wesentlich breiter als das KI für die wahre Gerade.

$$\hat{Y}_0 \pm t(1 - \alpha/2, N - 2)\hat{\sigma}\sqrt{1 + \frac{1}{N} + \frac{(X_0 - \bar{X})^2}{\sum_n (X_n - \bar{X})^2}}. \quad (5.97)$$

**Prognoseintervall
(individuelles Y_0)**

Es enthält, wie gesagt, einen zusätzlichen Term $\hat{\sigma}$ und ist somit breiter. Für große Stichproben wird das KI an der Stelle $X = \bar{X}$ immer schmaler, während das PI immer den Term $\hat{\sigma}^2(1 + \frac{1}{N}) \rightarrow \sigma^2$ enthält (Streuung des individuellen Werts).

Beispiel 5.4 (BIP 2007 und Inflationsrate (Prognose-Intervall))

Der Streuungsterm ist nun

$$\begin{aligned}\sqrt{\hat{E}[Y_0 - \hat{Y}_0]^2} &= \hat{\sigma} \sqrt{1 + \frac{1}{N} + \frac{(x_0 - \bar{x})^2}{\sum_n (x_n - \bar{x})^2}} \\ &= 0.88 \sqrt{1 + \frac{1}{32} + \frac{(x_0 - 5.5275)^2}{107.121}}.\end{aligned}$$

Das 95%-Prognose-Intervall ist damit

$$(-0.827 + 0.601 x_0) \pm 1.7972 \sqrt{1 + 1/32 + 0.00934(-5.5275 + x_0)^2}$$

mit dem Quantil $t(.975, 30) = 2.04227$. An der Stelle $x_0 = \bar{x} = 5.5275$ ergibt sich der minimale Wert

$$2.497 \pm 1.825.$$

Abb. 5.6 zeigt das Prognose-Intervall als Prognose-Band im Vergleich zum Konfidenzintervall.



5.4 Varianz-Zerlegung, Bestimmtheitsmaß und globaler F -Test

5.4.1 Varianz-Zerlegung

Die gesamte Streuung der Daten Y_n läßt sich zerlegen in

- totale Streuung: $SQT = \sum_n (Y_n - \bar{Y})^2$
- erklärte Streuung: $SQE = \sum_n (\hat{Y}_n - \bar{Y})^2$
- Rest-Streuung: $SQR = \sum_n (Y_n - \hat{Y}_n)^2$

Es gilt die Varianz-Zerlegung

Varianz-Zerlegung

$$SQT = SQE + SQR \quad (5.98)$$

Herleitung:

$$\begin{aligned}
 SQT &= \sum_n (Y_n - \bar{Y})^2 \\
 &= \sum_n (Y_n - \hat{Y}_n + \hat{Y}_n - \bar{Y})^2 \\
 &= \sum_n (Y_n - \hat{Y}_n)^2 + \sum_n (\hat{Y}_n - \bar{Y})^2 + 2 \sum_n (Y_n - \hat{Y}_n)(\hat{Y}_n - \bar{Y}) \\
 &= SQR + SQE + 2 \sum_n (Y_n - \hat{Y}_n)(\hat{Y}_n - \bar{Y})
 \end{aligned}$$

Hierbei wurde eine quadratische Ergänzung durchgeführt. Der gemischte Term ist 0 (noch zu zeigen), sodaß man die Varianz-Zerlegung erhält. Setzt man nun die prognostizierten Werte (Glg. 5.8) $\hat{Y}_n = \bar{Y} + \hat{\beta}(X_n - \bar{X})$ ein, so ergeben sich die Hilfs-Formeln

$$Y_n - \hat{Y}_n = Y_n - \bar{Y} - \hat{\beta}(X_n - \bar{X}) = \hat{\epsilon}_n \quad (5.99)$$

$$\hat{Y}_n - \bar{Y} = \hat{\beta}(X_n - \bar{X}). \quad (5.100)$$

Das Produkt dieser Terme ergibt

$$\begin{aligned}
 \sum_n (Y_n - \hat{Y}_n)(\hat{Y}_n - \bar{Y}) &= \sum_n \hat{\beta}(Y_n - \bar{Y})(X_n - \bar{X}) - \sum_n \hat{\beta}^2 (X_n - \bar{X})^2 \\
 &= \hat{\beta}^2 \sum_n (X_n - \bar{X})^2 - \hat{\beta}^2 \sum_n (X_n - \bar{X})^2 = 0,
 \end{aligned}$$

wobei $\hat{\beta}$ (vgl. 5.4) eingesetzt wurde. Somit ist der gemischte Term = 0 und die Streuungszersetzung ist gezeigt.

Mit Hilfe von 5.100 läßt sich die Streuungszersetzung in folgender Form schreiben:

$$SQE = \hat{\beta}^2 \sum_n (X_n - \bar{X})^2 \quad (5.101)$$

$$\sum_n (Y_n - \bar{Y})^2 = \hat{\beta}^2 \sum_n (X_n - \bar{X})^2 + \sum_n (Y_n - \hat{Y}_n)^2 \quad (5.102)$$

$$SQT = SQE + SQR \quad (5.103)$$

Dies zeigt, daß die erklärte Streuung proportional zum Quadrat des Regressionskoeffizienten ist. Normierung auf SQT ergibt die quadrierte Korrelation $R_{xy}^2 = SQE/SQT$.

5.4.2 Bestimmtheitsmaß

Als Maß für die Güte der linearen Prognoseregeln $\hat{\alpha} + \hat{\beta}X$ kann der Vergleich der Prognose-Fehler mit und ohne Regel herangezogen werden (ausführlich siehe KE 3):

- Der Fehler **mit** Regel (Prognose $\hat{Y}_n = \hat{\alpha} + \hat{\beta}X_n$) ist die Residualstreuung

$$F(+)=SQR=\sum_n(Y_n-\hat{Y}_n)^2$$

- Der Fehler **ohne** Regel (Prognose $\hat{Y}_n = \bar{Y} = \hat{Y}_n(\hat{\beta} = 0)$) ist die totale Streuung

$$F(-)=SQT=\sum_n(Y_n-\bar{Y})^2$$

Ein allgemeines Maß, das die proportionale Fehlerreduktion bei Verwendung einer Prognose-Regel angibt, ist das **PRE-Maß** (proportional reduction of errors)

**PRE-Maß
(Regression)**

$$PRE = \frac{F(-) - F(+)}{F(-)} \quad (5.104)$$

$$= \frac{SQT - SQR}{SQT} = \frac{SQE}{SQT}. \quad (5.105)$$

Hierbei wurde die Streuungszerlegung eingesetzt. Das Bestimmtheitsmaß ist also das Verhältnis von erklärter zu totaler Streuung der Daten Y_n . Aufgrund der Hilfsformel 5.100 ergibt sich die wichtige Darstellung:

$$PRE = \frac{SQE}{SQT} = \hat{\beta}^2 \frac{\sum_n (X_n - \bar{X})^2}{\sum_n (Y_n - \bar{Y})^2} \quad (5.106)$$

$$= R_{xy}^2. \quad (5.107)$$

**Bestimmtheitsmaß
(Regression)**

Das **Bestimmtheitsmaß**

(Verhältnis von erklärter zu totaler Streuung)

ist gleich dem quadrierten Korrelationskoeffizienten von Prädiktor X und Prädikand Y .

Es liegt zwischen 0 und 1 ($-1 \leq R_{xy} \leq 1; 0 \leq \frac{SQE}{SQT} \leq 1$).

Aufgrund der Darstellung als Verhältnis von erklärter zu totaler Streuung ist die Sprechweise

R_{xy}^2 Prozent der Streuung von Y wird durch X erklärt

üblich (etwa 62.5% im Beispiel 5.5). Im Extremfall ($R_{xy}^2 = 1$) werden 100% der Streuung erklärt und alle Datenpunkte liegen auf der Regressionsgeraden.

Beispiel 5.5 (BIP 2007 und Inflationsrate (Bestimmtheitsmaß))

Das Bestimmtheitsmaß läßt sich am einfachsten aus dem Korrelationskoeffizienten zwischen X und Y berechnen. Es war

$$\hat{\beta} = \frac{\sum_n (X_n - \bar{X})(Y_n - \bar{Y})}{\sum_n (X_n - \bar{X})^2} \quad (5.108)$$

$$= R_{xy} \frac{S_y}{S_x}. \quad (5.109)$$

Daraus ergibt sich ($N - 1$ kürzen)

$$R_{xy}^2 = \hat{\beta}^2 \frac{S_x^2}{S_y^2} = \hat{\beta}^2 \frac{\sum_n (X_n - \bar{X})^2}{\sum_n (Y_n - \bar{Y})^2} \quad (5.110)$$

Aus den bereits berechneten Werten findet man:

$$\sum_n (x_n - \bar{x})^2 = 107.121 = SQX \quad (5.111)$$

$$\sum_n (y_n - \bar{y})^2 = 61.9503 = SQT \quad (5.112)$$

$$r_{xy}^2 = 0.3616 \frac{107.121}{61.9503} = 0.625. \quad (5.113)$$

Alternativ ergibt sich

$$SQE = \hat{\beta}^2 \sum_n (x_n - \bar{x})^2 = 0.3616 \times 107.121 = 38.7396 \quad (5.114)$$

$$SQT = \sum_n (y_n - \bar{y})^2 = 61.9503 \quad (5.115)$$

$$PRE = \frac{SQE}{SQT} = \frac{38.7396}{61.9503} = 0.625, \quad (5.116)$$

also der gleiche Wert. Aus dem Computer-Output (Abb. 5.2) entnimmt man bis auf Rundungsfehler (die Daten wurden in den Beispielen nur mit 2 Kommastellen eingesetzt) die gleichen Resultate:

$$RSquare = r_{xy}^2$$

$$\text{Root Mean Square Error} = \hat{\sigma}$$

$$\text{Sum of Squares (Model)} = SQE$$

$$\text{Sum of Squares (Total)} = SQT$$

Also wird 62.5% der Streuung (Varianz) der Inflationsrate durch das BIP erklärt.

■

5.4.3 Gobaler F -Test und ANOVA-Tafel

Hierbei wird abgetestet, ob der Regressor X (unabhängige Variable) überhaupt etwas zu Erklärung von Y beiträgt.

Man prüft die Hypothese

$$H_0 : \beta = 0 \quad (5.117)$$

$$H_1 : \beta \neq 0 \quad (5.118)$$

mit Hilfe eines F -Tests. Bei nur einem Regressor ist dies äquivalent zum t -Test für β , der F -Test gilt jedoch auch simultan bei mehreren Regressoren X_1, \dots, X_q (deshalb 'global').

Man vergleicht nun die erklärte Streuung $SQE = \hat{\beta}^2 \sum_n (X_n - \bar{X})^2$ mit der Residualstreuung SQR . Bei einem Modell mit hoher Erklärungskraft ist $\hat{\beta}$ ‘groß’ (im Vergleich zu seiner Streuung) und die Residualstreuung ist ‘klein’. Daher wird der F -Test $F = SQE/SQR$ betrachtet.

Die Streuungszerlegung mit den entsprechenden Quadratsummen wird übersichtlich in einer sog. ANOVA-Tafel (Analysis of Variance – Varianzanalyse) zusammengestellt. Die gesamte Varianz wird in die Quellen

- Hypothese (lineare Prognoseregeln)
- und Fehler (Residuen)

aufgeteilt:

Varianz-Quelle	SQ		df	F -Statistik
Hypothese	SQE	$\sum_n (\hat{Y}_n - \bar{Y})^2$	1	$F = \frac{SQE}{SQR/(N-2)}$ $\sim F(1, N-2)$
Residuen	SQR	$\sum_n (Y_n - \hat{Y}_n)^2$	$N-2$	
Total	SQT	$\sum_n (Y_n - \bar{Y})^2$	$N-1$	

Da die Quadratsummen χ^2 -verteilt sind (mit den angegebenen Freiheitsgraden), ergibt der Bruch

$$F = \frac{SQE}{SQR/(N-2)} \sim F(1, N-2) \quad (5.119)$$

Globaler F-Test

einen F -Test für H_0 (Quantil $f(1-\alpha, 1, N-2)$).

Der Nenner $MQR := SQR/(N-2)$ im F -Test stimmt mit der geschätzten Varianz des Gleichungsfehlers überein, d.h.

$$\hat{\sigma}^2 = \frac{1}{N-2} \sum_n \epsilon_n^2 = \frac{1}{N-2} SQR. \quad (5.120)$$

Außerdem läßt sich die Residualstreuung auch mit Hilfe der Korrelation ausdrücken

$$\begin{aligned} SQR &= SQT - SQE = SQT(1 - SQE/SQT) \\ &= SQT(1 - R_{xy}^2). \end{aligned} \quad (5.121)$$

In der Tat geht die Residualstreuung gegen Null, wenn die Daten perfekt korreliert sind. Dies korrespondiert wiederum mit einer großen F -Statistik.

Beispiel 5.6 (BIP 2007 und Inflationsrate: Globaler F -Test)

Im Beispiel ergibt sich

Varianz-Quelle	SQ		df	F -Statistik
Hypothese	SQE	38.7396	1	$F = 50.0713$ $f(.95, 1, 30) = 4.17088$
Residuen	SQR	23.2107	30	
Total	SQT	61.9503	31	

Das F -Quantil $f(.95, 1, 30) = 4.17088$ ist wesentlich kleiner als die Teststatistik, sodaß die Nullhypothese verworfen wird. Dies stimmt mit dem Resultat des t -Tests überein (Bsp. 5.2). Die T -Statistik war 7.07611. Quadriert man diese, so ergibt sich $T^2 = 50.0713$, was mit der F -Statistik übereinstimmt. Dies ist kein Zufall, sondern folgt aus dem Zusammenhang $T(N-2)^2 = N(0, 1)^2/\chi^2(N-2) = F(1, N-2)$ der T , χ^2 und F -Statistik (vgl. Abs. 1.3.4.2, Nummer 4).

■

Der F -Bruch läßt sich auch mit Hilfe des Korrelationskoeffizienten ausdrücken, da

$$\frac{SQE}{SQR/(N-2)} = \frac{SQE}{SQT - SQE}(N-2) \quad (5.122)$$

$$= \frac{SQE/SQT}{1 - SQE/SQT}(N-2). \quad (5.123)$$

Somit gilt

$$F = \frac{R_{xy}^2}{1 - R_{xy}^2}(N-2). \quad (5.124)$$

Große (betragsmäßige) Korrelationen führen also zu großen F -Statistiken.

Im Beispiel ist $r_{xy}^2 = 0.625$ und somit $F = \frac{0.625}{1-0.625}30 = 50.07$.

5.4.4 Residualanalyse

Nach dem Schätzen der Parameter und dem Testen des Modells sollte auch eine Analyse der Residuen vorgenommen werden (Diagnose). Hiermit wird überprüft, ob die Annahmen des Modells (vgl. Abs. 5.1.2.2) zumindest approximativ erfüllt sind oder ob grobe Abweichungen vorliegen.

Diagnose

Beispielsweise sollten die Residuen $\hat{\epsilon}_n$ unsystematisch streuen und keine Abhängigkeit von den Regressoren X_n aufweisen. Dies zeigt sich im Streudiagramm Abb. 5.7. Die eingezeichnete Regressionslinie hat nur eine sehr kleine Steigung. Man hat allerdings den Eindruck, daß für große BIP-Raten eine leichte Vergrößerung der Streuung eintritt (Heteroskedastizität).

Im Streudiagramm sind auch sog. Ausreißer zu erkennen, die sich vom Rest der Daten deutlich unterscheiden (etwa der Wert für die Türkei mit BIP-Anderung = 10.43%, Inflation = 7.73%). Allerdings sind solche Extremwerte durchaus mit den Verteilungsannahmen vereinbar und sollten nicht voreilig aus dem Datensatz entfernt werden.

Das Histogramm (mit angepaßter Normalverteilung) und der Normal-Quantil-Plot (Abb. 5.8.) belegen, daß keine starken Abweichungen von der Normalverteilung auftreten. Beim Normal-Quantil-Plot werden die empirischen Quantile der Verteilung gegen diejenigen einer entsprechenden Normalverteilung aufgetragen.

Normal-Quantil-Plot

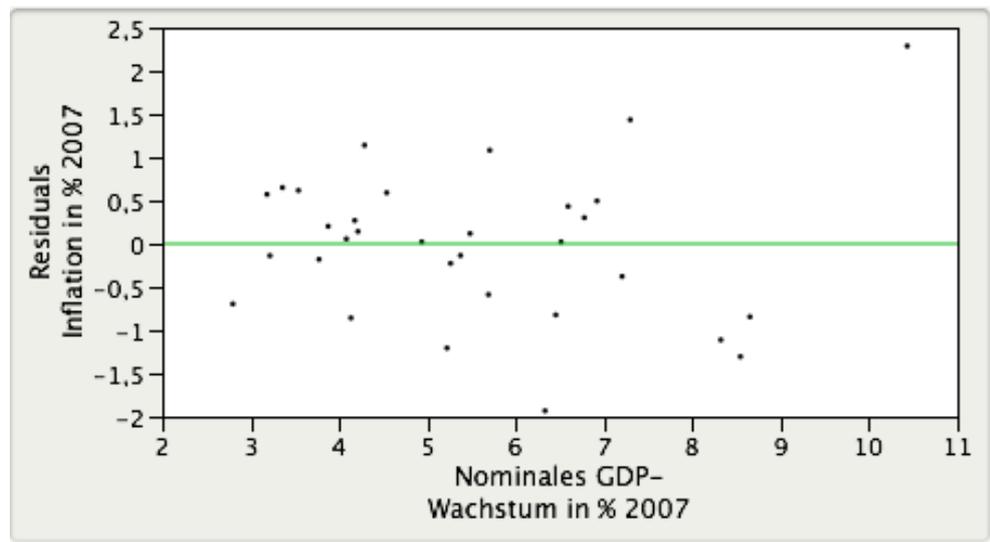


Abbildung 5.7: Streudiagramm der Residuen $\hat{\epsilon}_n$ mit den Regressoren X_n . Eingezeichnet ist auch eine geschätzte Regressionslinie.

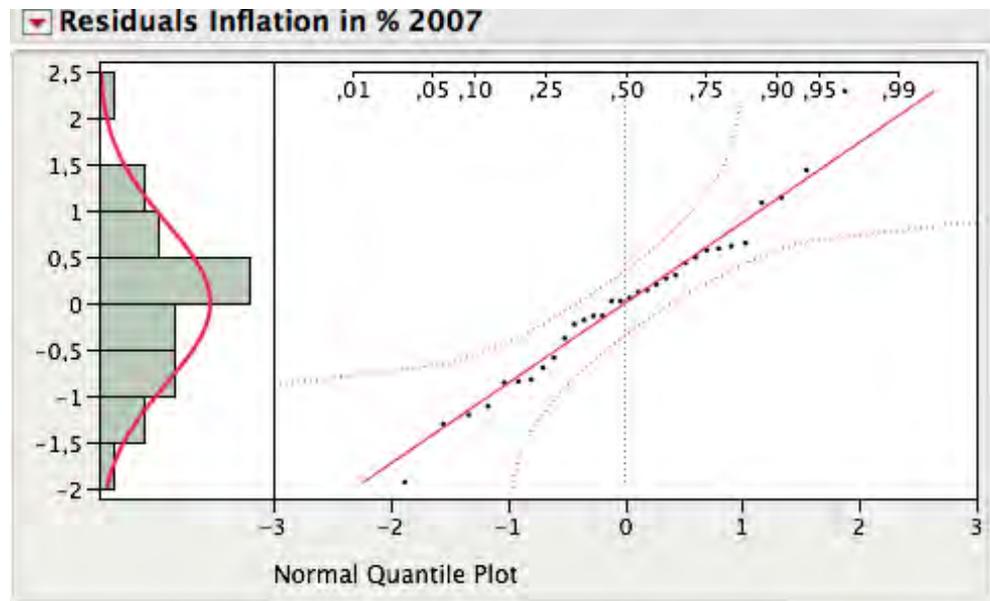


Abbildung 5.8: Histogramm und Normal-Quantil-Plot der Residuen $\hat{\epsilon}_n$. Es sind keine groben Abweichungen von der Normalverteilung zu erkennen.

Kapitel 6

Varianz-Analyse

6.1 Multiple Vergleiche

Der 2-Stichproben- t -Test vergleicht die Populations-Mittelwerte (Erwartungswerte) in 2 Gruppen, etwa West/Ost-Deutschland oder Filialen mit/ohne Werbemaßnahme. Hat man mehr als 2 Größen, so kann man alle Kombinationen gegeneinander abtesten, muß dann jedoch berücksichtigen, daß das simultane Signifikanzniveau α^* aller ausgeführten Test (6 Tests bei $I = 4$ Gruppen) nicht mehr bei α liegt, wenn die Einzeltests mit diesem Niveau ausgeführt wurden. Da mehrere Tests ausgeführt wurden, steigt die Wahrscheinlichkeit, $H_0 : \mu_i = \mu, i = 1, \dots, I$ abzulehnen, obwohl alle Mittelwerte gleich sind.

Beispiel 6.1 (Vergleich von 4 Lehrmethoden)

In einem Lehrgang wurden 4 verschiedene Unterrichtsmethoden benutzt. $N = 32$ Teilnehmer wurden auf die 4 Gruppen (Methoden) zufällig verteilt. Am Ende des Lehrgangs wurde ein Abschlußtest durchgeführt. Die entscheidende Frage lautet: Hat die Lehrmethode einen Einfluß auf den mittleren Punktwert.

Getestet wird also:

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 \text{ gegen} \quad (6.1)$$

$$H_1 : \mu_i \neq \mu_j \text{ für mindestens ein Paar } i, j. \quad (6.2)$$

Die Daten $y_{ij}, i = 1, \dots, I = 4; j = 1, \dots, J = 8$ sind in Tabelle 6.1 angegeben und in Abb. 6.1 graphisch dargestellt. Die Daten y_{ij} sind als Funktion der Lehrmethode i (x -Achse) als Punkte, Mittelwerte und Standardfehler der Mittelwerte (s_i/\sqrt{J}) aufgetragen. Um die Mittelwertunterschiede zu betonen, wurden die Mittelwerte durch Linien verbunden,

<i>Methode</i>	1	2	3	4	5	6	7	8	\bar{y}_i	s_i
1	16	18	20	15	20	15	23	19	18.25	2.82
2	16	12	10	14	18	15	12	13	13.75	2.55
3	2	10	9	10	11	9	10	9	8.75	2.82
4	5	8	8	11	1	9	5	9	7.00	3.16

Tabelle 6.1: Abschlußtest von 4 Lehrmethoden.

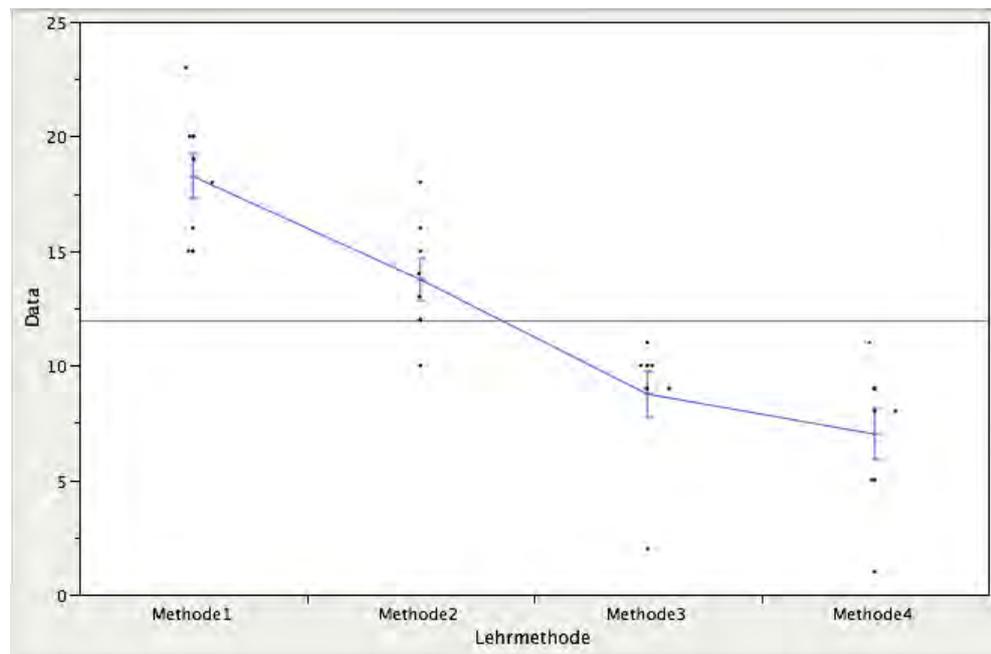


Abbildung 6.1: Abschlußtest von 4 Lehrmethoden als Funktion der Methode.

obwohl nur diskrete (nominale) Ausprägungen auf der x -Achse vorliegen. Aufgrund der Standardfehler unterscheiden sich vermutlich die Methoden 3 und 4 nicht (signifikant) voneinander.



Vergleicht man alle Gruppen mit multiplen (mehrfachen) Mittelwertsvergleichen, so sind hier (Binomialkoeffizient 4 über 2)

$$\binom{4}{2} = \frac{4!}{2!2!} = \frac{4 \times 3 \times 2 \times 1}{2 \times 2} = 6 \quad (6.3)$$

Tests auszuführen.

Wenn jeder Test mit α ausgeführt wird, kann das simultane Signifikanzniveau α^* auf maximal 6α ansteigen, etwa $6 \times 5\% = 30\%$. Man versteht dies mit Hilfe folgender Berechnung:

Die Nullhypothese der Varianzanalyse

$$H_0 : \mu_i = \mu; \quad i = 1, \dots, I \quad (6.4)$$

**Nullhypothese der
Varianzanalyse**

kann als Schnitt der Einzelhypothesen

$$H_0 = \bigcap_{i < j} H_{0,ij} = \bigcap_{i < j} \{\mu_i = \mu_j\} \quad (6.5)$$

aufgefaßt werden.

Dann gilt für die Ereignisse $E_{ij} = \{H_{0,ij} \text{ beibehalten}\}$ die Zusammensetzung $E = \bigcap_{i < j} E_{ij}$ (H_0 beibehalten). H_0 wird also abgelehnt, wenn $\bar{E} = \bigcup_{i < j} \bar{E}_{ij}$ eintritt, also mindestens eine Einzelhypothese abgelehnt wird.

Dann gilt für das exakte Signifikanzniveau mit Hilfe der Booleschen Ungleichung (Übung 6.1)

**exaktes
Signifikanzniveau**

Satz 6.1 (Bonferroni-Ungleichung)

$$\alpha^* = P_0(\bar{E}) = P_0\left(\bigcup_{i < j} \bar{E}_{ij}\right) \quad (6.6)$$

$$\leq \sum_{i < j} P_0(\bar{E}_{ij}) \quad (6.7)$$

$$= \sum_{i < j} \alpha_{ij} \quad (6.8)$$

**Bonferroni-
Ungleichung**

(P_0 = Wahrscheinlichkeit berechnet unter Gültigkeit der H_0).



Hierbei sind α_{ij} die Signifikanzniveaus der Einzelhypothesen (etwa $\alpha_{ij} = \alpha/k$; $k = I(I-1)/2 = 4 * 3/2 = 6$ im obigen Beispiel).

Durch die Benutzung von adjustierten α -Niveaus, etwa $\alpha' = \alpha/k = 0.05/6$, vergrößern sich die t -Quantile (vgl. Abb. 6.2)

Die obige Ungleichung hat zur Folge, daß das exakte multiple Signifikanzniveau deutlich kleiner sein kann als α , was zu einem konservativen Test führt.

Bei $k = 3$ Tests (3 Gruppen) hat man die exakte Formel

$$\alpha^* = P_0(\bar{E}) \quad (6.9)$$

$$= P_0(\bar{E}_{12} \cup \bar{E}_{13} \cup \bar{E}_{23}) \quad (6.10)$$

$$= P_0(\bar{E}_{12}) + P_0(\bar{E}_{13}) + P_0(\bar{E}_{23}) \quad (6.11)$$

$$- P_0(\bar{E}_{12} \cap \bar{E}_{13}) - P_0(\bar{E}_{12} \cap \bar{E}_{23}) - P_0(\bar{E}_{13} \cap \bar{E}_{23}) \quad (6.12)$$

$$+ P_0(\bar{E}_{12} \cap \bar{E}_{13} \cap \bar{E}_{23}) \quad (6.13)$$

$$= \alpha_{12} + \alpha_{13} + \alpha_{23} - \dots + . \quad (6.14)$$

Aufgrund der negativen Terme, die das simultane Ablehnen von Einzelvergleichen ausdrücken (aber schwer berechenbar sind), ist die Bonferroni-Ungleichung eine grobe, wenn auch leicht anwendbare Abschätzung.

Bei unabhängigen Einzeltests ergibt sich die Formel

$$\alpha^* = P_0(\bar{E}) = 1 - P_0\left(\bigcap_{i < j} E_{ij}\right) \quad (6.15)$$

$$= 1 - \prod_{i < j} P_0(E_{ij}) \quad (6.16)$$

$$= 1 - \prod_{i < j} (1 - \alpha_{ij}). \quad (6.17)$$

Wählt man adjustierte Niveaus $\alpha_{ij} = \alpha/k$, so hat man bei k Vergleichen

$$\alpha^* = 1 - (1 - \alpha/k)^k \rightarrow 1 - e^{-\alpha} \leq \alpha \quad (6.18)$$

$$= \begin{cases} \alpha - \frac{\alpha^2}{4} & (k = 2) \\ \alpha - \frac{\alpha^2}{3} + \frac{\alpha^3}{27} & (k = 3) \end{cases} \quad (6.19)$$

etc. Das exakte simultane Signifikanzniveau ist also kleiner als α (Abb. 6.3, oben).

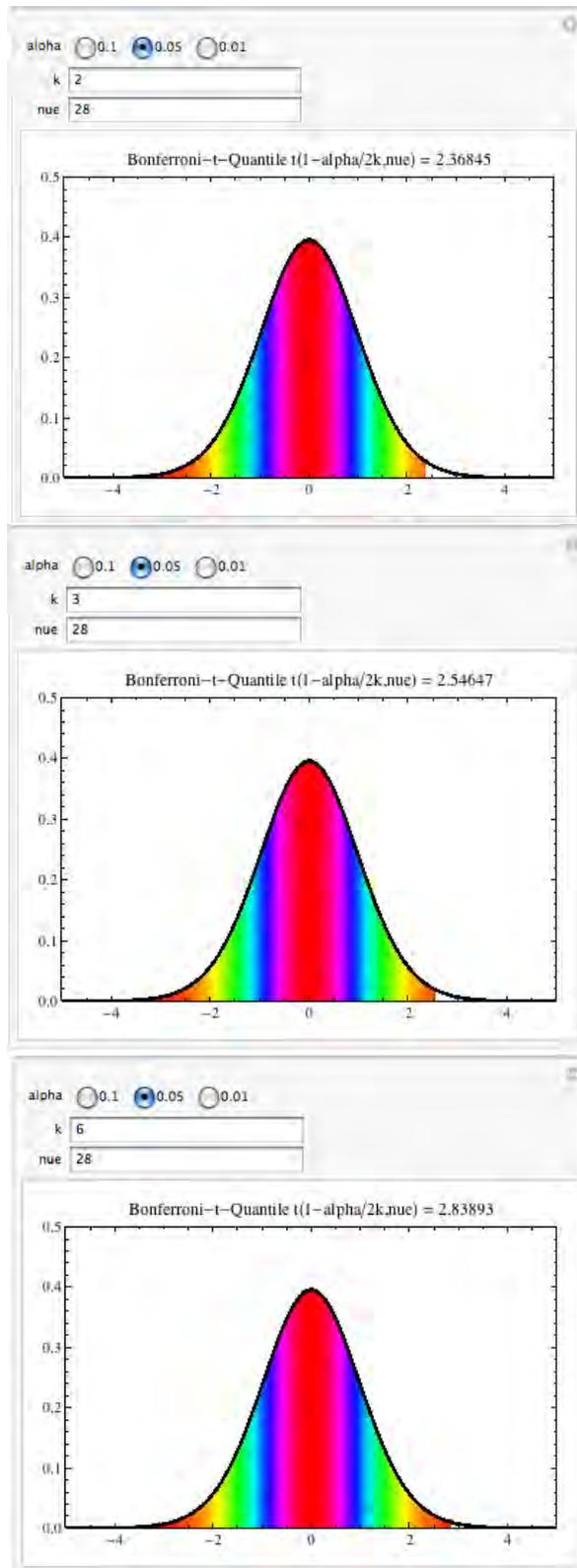


Abbildung 6.2: Bonferroni- t -Quantile $t(1 - \alpha/(2k), df = 28)$ bei $k = 2, 3, 6$ Vergleichen ($\alpha = .05$).

Ohne Adjustierung geht das simultane Niveau

$$\alpha^* = 1 - (1 - \alpha)^k \rightarrow 1 \quad (6.20)$$

gegen 1 (Abb. 6.3, unten).

Übung 6.1 (Boolesche Ungleichung)

Boolesche Ungleichung

Leiten Sie die Boolesche Ungleichung

$$P(A \cup B) \leq P(A) + P(B) \quad (6.21)$$

her.

Hinweis: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Finden Sie analoge Gleichungen für 3 Mengen $A \cup B \cup C$.

Allgemein gilt somit

$$P\left(\bigcup_i A_i\right) \leq \sum_i P(A_i). \quad (6.22)$$



6.2 Einfaktorielle Varianzanalyse

Nullhypothese der Varianzanalyse

Im vorherigen Abschnitt wurde gezeigt, daß bei simultanen Tests zur Überprüfung der Nullhypothese

$$H_0 : \mu_i = \mu; i = 1, \dots, I \quad (6.23)$$

eine Adjustierung der Einzelvergleiche durchgeführt werden muß, um eine Inflation des simultanen Signifikanzniveaus des Tests der H_0 zu vermeiden.

Daher wird zuerst ein exakter simultaner Test durchgeführt, um dann bei Ablehnung der H_0 anschließend herauszufinden, welche (Kombinationen von) Mittelwertsunterschiede die Ablehnung verursacht haben (Post-hoc-Tests).

Um die Analogie der Problemstellung zur Regression zu verdeutlichen, wird die Abhängigkeit einer quantitativen Variable Y von einer nominalen unabhängigen Variable $X = i$ betrachtet, also $Y = Y(i) = Y_i$ (vgl. ausführlich KE 3).

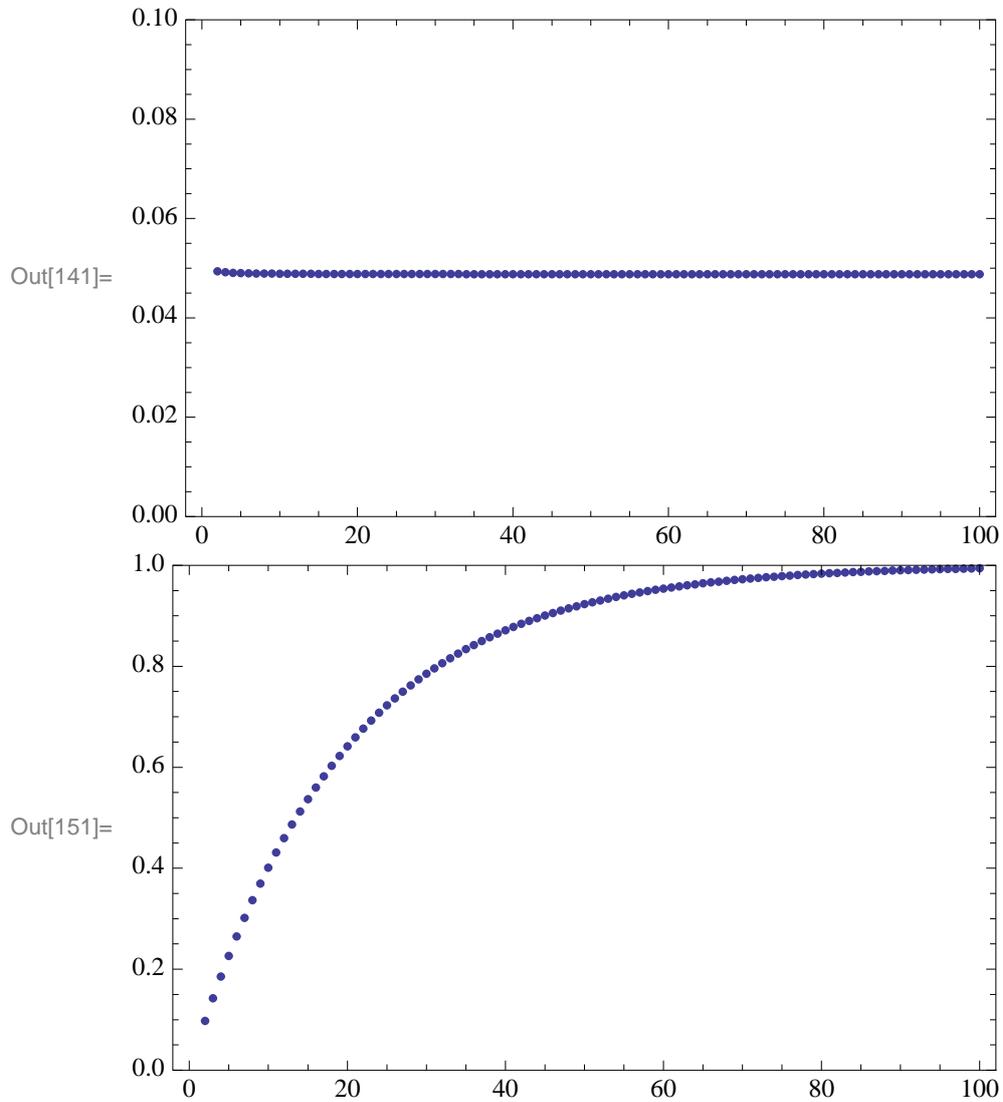


Abbildung 6.3: Simultaner α -Fehler bei unabhängigen Tests. Oben: mit Adjustierung α/k , unten: ohne Adjustierung.

**Varianzanalyse
(Grundmodell)**

Das Grundmodell der Varianzanalyse lautet somit

$$Y_{ij} = \mu_i + \epsilon_{ij} \quad (6.24)$$

wobei ein iid-Gleichungsfehler $\epsilon_{ij} \sim N(0, \sigma^2)$ unterstellt wird.¹ Hierbei ist $i = 1, \dots, I$ der Gruppen-Index und $j = 1, \dots, J$ der Personen-Index in Gruppe i .

Als Schätzung für den Gruppenmittelwert (Erwartungswert) μ_i ergibt sich

$$\hat{\mu}_i = \frac{1}{J} \sum_j Y_{ij} := \bar{Y}_i \quad (6.25)$$

Definiert man noch den Gesamt-Mittelwert

$$\bar{Y} = \frac{1}{IJ} \sum_{ij} Y_{ij}, \quad (6.26)$$

so ergibt sich die Varianz(bzw. Streuungs)-Zerlegung (die dem Verfahren den Namen gibt)

$$SQT = \sum_{ij} (Y_{ij} - \bar{Y})^2 \quad (6.27)$$

$$= \sum_{ij} (Y_{ij} - \bar{Y}_i)^2 + \sum_{ij} (\bar{Y}_i - \bar{Y})^2 \quad (6.28)$$

Streuungszerlegung

$$= SQR + SQE \quad (6.29)$$

$$= \text{residuale} + \text{erklärte Quadratsumme} \quad (6.30)$$

$$= SQ(\text{innerhalb}) + SQ(\text{zwischen}). \quad (6.31)$$

Eine Schätzung für σ^2 ist die mittlere Residual-Quadratsumme innerhalb der Gruppen

$$\hat{\sigma}^2 = \frac{1}{I} \sum_i \frac{1}{J-1} \sum_j (Y_{ij} - \bar{Y}_i)^2 \quad (6.32)$$

$$= SQR / (I(J-1)) = MQR. \quad (6.33)$$

¹iid = independent and identically distributed, d.h. alle Personen in allen Gruppen sind voneinander unabhängig und gleich verteilt.

Unter Gültigkeit der H_0 sind alle \bar{Y}_i ungefähr gleich dem Gesamtmittelwert \bar{Y} , sodaß SQE klein ist im Verhältnis zur Residualstreuung.

Unter H_0 ist somit $SQE/(I-1)/(SQR/(IJ-I)) = MQE/MQR$ F -verteilt mit Freiheitsgraden $I-1, IJ-I$.

Übersichtlich schreibt man in Form einer ANOVA-Tabelle

ANOVA-Tafel

SQ	Formel	df	F -Statistik
SQE (zwischen)	$\sum_{ij}(\bar{Y}_i - \bar{Y})^2$	$I - 1$	$\frac{SQE/(I-1)}{SQR/(I(J-1))}$ $\sim F(I-1, I(J-1))$
SQR (innerhalb)	$\sum_{ij}(Y_{ij} - \bar{Y}_i)^2$	$I(J - 1)$	
SQT (total)	$\sum_{ij}(Y_{ij} - \bar{Y})^2$	$IJ - 1$	

Für

$$F(I-1, I(J-1)) > f(1-\alpha; I-1, I(J-1)) \quad (6.34)$$

wird die H_0 auf dem Signifikanzniveau α verworfen.

Beispiel 6.2 (Vergleich von 4 Lehrmethoden)

Es gilt die vorteilhafte Umrechnung

$$SQE = \sum_{ij}(\bar{Y}_i - \bar{Y})^2 = J \sum_i \bar{Y}_i^2 - IJ\bar{Y}^2 \quad (6.35)$$

$$SQT = \sum_{ij}(Y_{ij} - \bar{Y})^2 = \sum_{ij} Y_{ij}^2 - IJ\bar{Y}^2. \quad (6.36)$$

Aus Tabelle 6.1 entnimmt man die Gruppenmittelwerte $\bar{y}_i = \{18.25, 13.75, 8.75, 7.00\}$ und den Gesamtmittelwert $\bar{y} = \frac{1}{4}(18.25 + 13.75 + 8.75 + 7.00) = 11.94$.

Daraus findet man

$$\begin{aligned} SQE &= 8(18.25^2 + 13.75^2 + 8.75^2 + 7.00^2) - 32 \times 11.94^2 = 621.3750 \\ SQT &= 5408 - 32 \times 11.94^2 = 847.8750. \end{aligned}$$

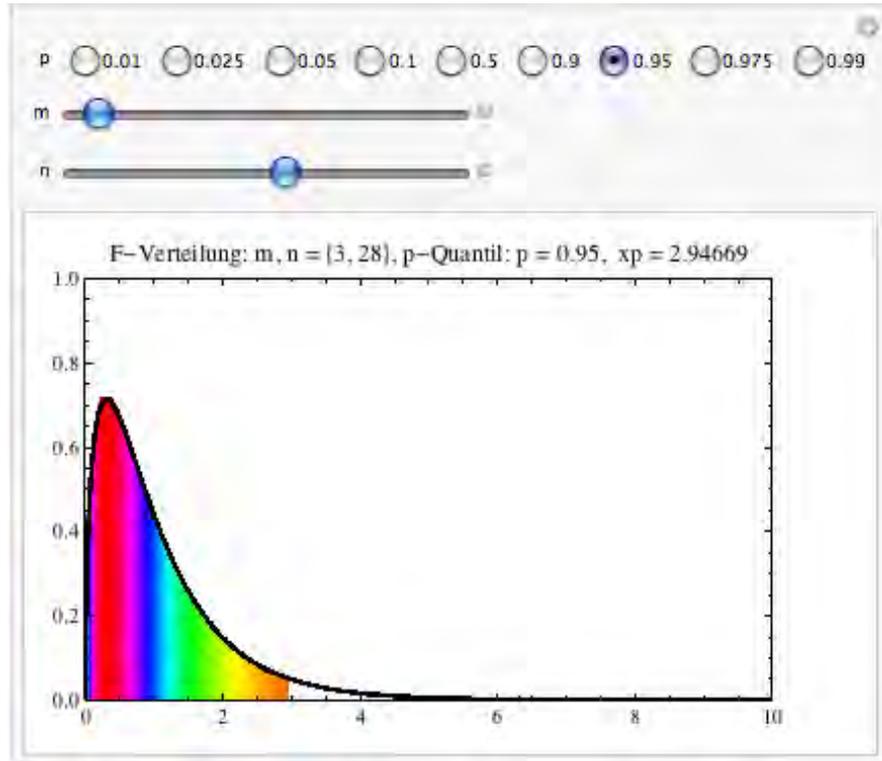
Die residuale Quadratsumme ist somit $SQR = SQT - SQE = 847.8750 - 621.3750 = 226.50$.

Für den F -Test ergibt sich

$$f(3, 28) = (621.3750/3)/(226.050/28) = 25.6049 \quad (6.37)$$

mit dem Quantil $f(.95, 3, 28) = 2.95$ (Abb. 6.4). Damit muß H_0 abgelehnt werden: die Lehrmethoden sind also nicht gleich effektiv.

■

Abbildung 6.4: $f(.95, 3, 28)$ -Quantil

Welche Mittelwertsunterschiede zu der Ablehnung der H_0 geführt haben, ist somit noch nicht geklärt.

Man führt daher im Anschluß an die signifikante Varianzanalyse T -Vergleiche (post hoc) durch, die aber mit adjustierten α -Fehlern gerechnet werden müssen, um das Signifikanzniveau des globalen Tests der H_0 einzuhalten.

Als Teststatistiken verwendet man die T -Brüche

$$T_{ii'} = \frac{\bar{Y}_i - \bar{Y}_{i'}}{S\sqrt{2/J}} \sim t(IJ - I). \quad (6.38)$$

Der Nenner ergibt sich aus der Überlegung

$$\text{Var}(\bar{Y}_i - \bar{Y}_{i'}) = \text{Var}(\bar{Y}_i) + \text{Var}(\bar{Y}_{i'}) = \frac{2}{J}\sigma^2, \quad (6.39)$$

da die Gruppen voneinander unabhängig sind und gleiche Varianzen aufweisen (Annahmen). Ersetzt man die Varianz durch die Schätzung

$$S^2 = \hat{\sigma}^2 = SQR/(I(J - 1)) = MQR, \quad (6.40)$$

so ergibt sich obige T -Statistik.

Beispiel 6.3 (Vergleich von 4 Lehrmethoden)

Im Beispiel ist $s^2 = SQR/(I(J - 1)) = 226.50/28 = 8.0839$; $s = 2.8442$.

Der Nenner der T -Statistik ist also $s\sqrt{2/J} = 2.8442\sqrt{2/8} = 1.4221$.

Die 6 T -Statistiken lauten (die signifikanten Vergleiche sind durch * gekennzeichnet)

$T_{ii'}$	18.25	13.75	8.75	7.00
18.25		3.1643*	6.6803*	7.91*
13.75			3.5159*	4.7465*
8.75				1.2306
7.00				

mit dem kritischen t -Wert $t(1 - \alpha/(2k), df) = 2.84$; $k = 6$, $df = 28$, $\alpha = 0.05$ (Bonferroni-Tabelle im Anhang oder applet http://www.fernuni-hagen.de/lis_statistik/lehre/). Das nicht adjustierte Quantil ist dagegen $t(1 - \alpha/2, df) = 2.048$, also wesentlich kleiner. Allerdings würde auch hier der Vergleich 3-4 nicht signifikant.

■

6.3 Einfaktorielle Varianzanalyse in Effektdarstellung

Das Grundmodell der Varianzanalyse lautete als Funktion der Gruppenmittelwerte

$$Y_{ij} = \mu_i + \epsilon_{ij}. \quad (6.41)$$

Parametrisiert man diese als Abweichungen von einem Gesamt-Mittelwert

$$\mu = \frac{1}{I} \sum_i \mu_i \quad (6.42)$$

und führt Abweichungen (Effekte)

$$\alpha_i = \mu_i - \mu \quad (6.43)$$

**Varianzanalyse
(Effekt-
Darstellung)**

ein, so ergibt sich das Modell in Effekt-Darstellung

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij}, \quad (6.44)$$

wobei die Effekte der **Restriktion**

$$\sum_i \alpha_i = \sum_i (\mu_i - \mu) = I\mu - I\mu = 0 \quad (6.45)$$

genügen müssen.

Die Nullhypothese ($\mu_i = \mu$) lautet nun

$$H_0 : \alpha_i = 0 \quad (6.46)$$

und Paar-Vergleiche $\mu_i = \mu_{i'}$ sind $\alpha_i = \alpha_{i'}$.

Die Schätzung der Effekte kann mit Hilfe von

$$\hat{\alpha}_i = \hat{\mu}_i - \hat{\mu} = \bar{Y}_i - \bar{Y} \quad (6.47)$$

erfolgen, was zu einer alternativen Darstellung der erklärten Streuung

$$SQE = \sum_{ij} (\bar{Y}_i - \bar{Y})^2 = J \sum_i \hat{\alpha}_i^2 \quad (6.48)$$

führt. Große Effekte (Abweichungen vom Mittelwert) spiegeln sich in einer großen erklärten Streuung wider.

Die geschätzten Effekte $\hat{\alpha}_i$ erfüllen die gleichen Restriktionen wie die Parameter α_i :

$$\sum_i \hat{\alpha}_i = \sum_i (\bar{Y}_i - \bar{Y}) = I\bar{Y} - I\bar{Y} = 0. \quad (6.49)$$

Beispiel 6.4 (Vergleich von 4 Lehrmethoden)

Im Beispiel war

$$\bar{y}_i = \{18.25, 13.75, 8.75, 7.00\} \text{ und}$$

$$\bar{y} = \frac{1}{4}(18.25 + 13.75 + 8.75 + 7.00) = 11.94.$$

Daraus ergibt sich

$$\hat{\alpha}_i = \{18.25, 13.75, 8.75, 7.00\} - 11.94 = \{6.31, 1.81, -3.19, -4.94\},$$

was sich auf 0 addiert.

Die Quadratsumme ist

$$39.85 + 3.29 + 10.16 + 24.38 = 77.67 \text{ und somit } SQE = 8 * 77.67 = 621.375.$$

■

6.4 Varianzanalyse als Prognosemethode

In Analogie zur Regressionsanalyse kann die Varianzanalyse als Modell der Abhängigkeit einer quantitativen Variable Y von einer nominalen (oder ordinalen) unabhängigen Variable $X = i$ betrachtet werden, also

$$Y_{ij} = \mu_i + \epsilon_{ij} = \mu + \alpha_i + \epsilon_{ij}. \quad (6.50)$$

Man kann nun Prognose-Regeln formulieren, die mit (+) und ohne (-) Kenntnis der unabhängigen Variablen $X = i$ arbeiten.

Prognose-Regeln:**Prognose-Regeln**

- Wenn bekannt ist, daß eine Person der Gruppe $X = i$ angehört, ist die beste Prognose der Gruppen-Mittelwert

$$\bar{Y}_i := J^{-1} \sum_j Y_{ij} = \bar{Y} + \hat{\alpha}_i.$$

- Wenn diese Information nicht vorliegt, prognostiziert man am besten den globalen Mittelwert (ohne Effekt $\hat{\alpha}_i$)

$$\bar{Y} := (IJ)^{-1} \sum_{ij} Y_{ij}.$$

Die entsprechenden Fehlermaße sind dann

$$F(+)= \sum_{ij} (Y_{ij} - \bar{Y}_i)^2 = SQR \text{ (residuale Quadratsumme)}$$

$$F(-)= \sum_{ij} (Y_{ij} - \bar{Y})^2 = SQT \text{ (totale Quadratsumme)}$$

und der Determinationskoeffizient (prozentuale Fehlerreduktion PRE) ist mit Hilfe der Streuungszerlegung

**prozentuale
Fehlerreduktion
PRE**

$$PRE = \frac{F(-) - F(+)}{F(-)} = \frac{SQT - SQR}{SQT} = \frac{SQE}{SQT}. \quad (6.51)$$

Im Beispiel ist somit $PRE = 621.357/847.875 = 0.73 = 73\%$, also wird 73% der Varianz im Testergebnis durch die Lehrmethode $X = i$ (unterschiedliche Gruppenmittelwerte) erklärt.

Alle wichtigen Resultate der 1-faktoriellen ANOVA sind nochmals im Computer-Output Abb. 6.5 zusammengefaßt.

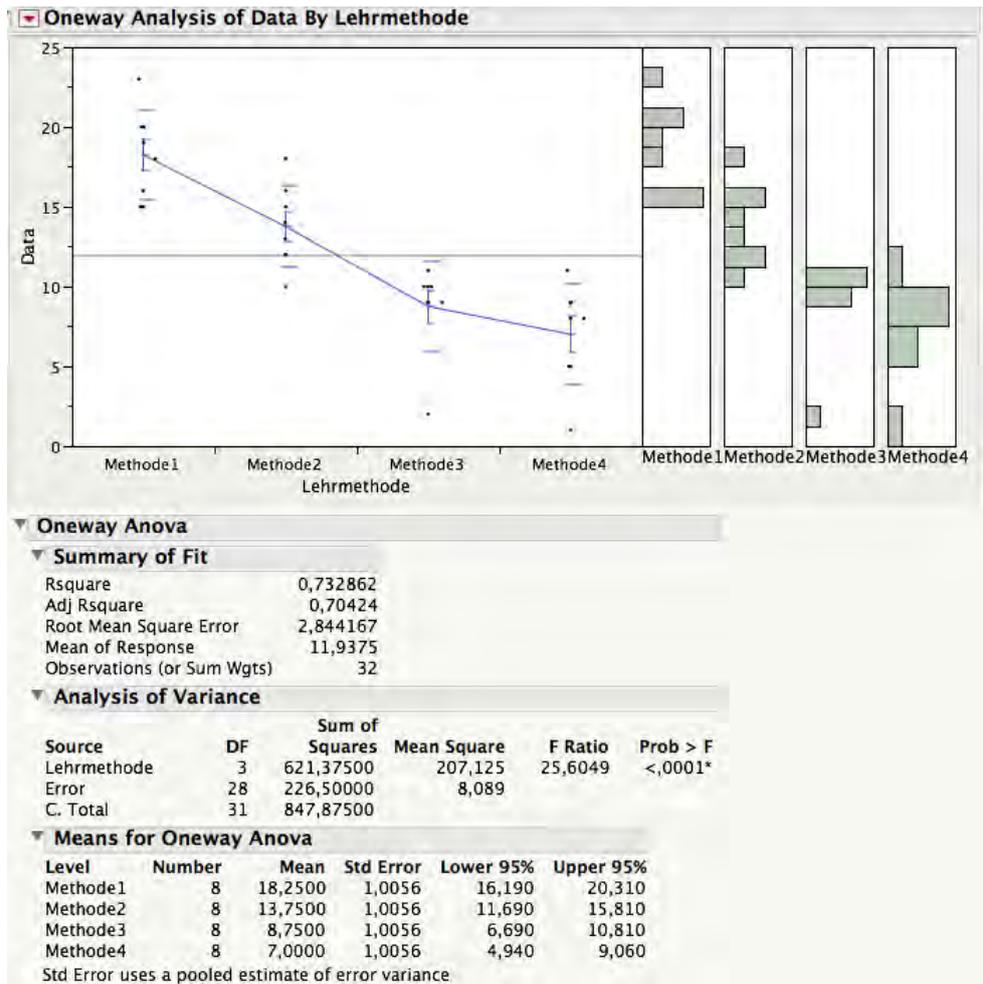


Abbildung 6.5: ANOVA (analysis of variance) für unterschiedliche Lehrmethoden.

Teil II

Empirische Sozialforschung

Kapitel 7

Statistik und Empirische Forschung

Wissenschaftsphilosophie ist für die Wissenschaftler ähnlich nützlich wie die Ornithologie für die Vögel.

Richard P. Feynman (1918 bis 1988), Physiker und Nobelpreisträger; zitiert nach Frankfurter Allgemeine Zeitung vom 22. März 2013 ¹

In den bisherigen Kapiteln (KE 1) wurden vorwiegend Methoden zur **Datenanalyse** (Auswertung) behandelt. Die Gewinnung der Daten (**Datenerhebung**) wird hierbei schon vorausgesetzt. Insofern lassen sich die statistischen Methoden weitgehend von den einzelnen Substanzwissenschaften abstrahieren, was sich in einer Vielzahl von Beispielen aus verschiedenen Disziplinen zeigt: Ein Mittelwert kann genauso an einem biologischen wie an einem betriebswirtschaftlichen Beispiel erläutert werden; der Begriff des Mittelwerts ist eine Konstruktion der statistischen Theorie.

Datenerhebung

In diesem Sinn ist die **Statistik eine übergreifende Disziplin**, die in den Substanzwissenschaften nur eine entsprechende Schwerpunktsetzung annimmt. Zwar gibt es die Teilgebiete der Biometrie, der Ökonometrie, der Psycho-, Sozio-, Techno- metrie etc.; trotzdem sind diese Bindestrich-Metrien nur **fachspezifische Anwendungen der Statistik** auf die empirische Forschung. Die schon qualitativ oder quantitativ vorliegenden

**Statistik als
übergreifende
Disziplin**

**Bindestrich-
Metrien**

¹<https://www.faz.net/aktuell/feuilleton/forschung-und-lehre/wissenschaftsphilosophie-forschung-ueber-wahrheiten-12119901.html>

Begriffe der einzelnen Fachrichtungen werden stochastisch reformuliert und dann deskriptiv oder inferentiell analysiert.

Datenerzeugung

Die Methoden der Datengewinnung, oder verschärft formuliert, der Datenerzeugung (Produktion im Meßprozeß), sind der Analyse vorgelagert und werden nicht thematisiert.

Allerdings setzt die statistische Analyse eine bestimmte Art von Daten voraus: jeder statistischen Einheit muß die **gleiche Menge von Merkmalen** zugeordnet werden können, und die Erhebung einzelner Merkmale stört nicht die Erhebung der Anderen. Die Möglichkeit, **gemeinsame Häufigkeiten**

$$f(x_1, x_2, \dots, x_p)$$

der Merkmale X_i anzugeben, wird unbefragt unterstellt. Beispielsweise werden die Merkmale Geschlecht (X_1) und Körpergröße (X_2) jeder statistischen Einheit zugeordnet und es ist möglich, daß jeder Person simultan eine Ausprägung $x_1 \in \{m, w\}$ und $x_2 \in \{0, 1, 2, 3, \dots\}$ (cm) zugeordnet werden kann. Die Reihenfolge der Messungen spielt dabei keine Rolle.

Reihenfolge der Messungen

Beispiel 7.1 (Reihenfolge von items)

Betrachtet man einen Fragebogen mit den items (Fragen, Aufgaben) A und B , so sind die Versionen $F_1 = (A, B)$ und $F_2 = (B, A)$ unterschiedlich in der Reihenfolge.

Eine saubere Notation wäre $F_1 = (A_1, B_2)$ und $F_2 = (B_1, A_2)$.

Der erste Fragebogen ergibt die Häufigkeitsverteilung

$$f(A_1, B_2) = f(B_2, A_1),$$

der zweite

$$f(B_1, A_2) = f(A_2, B_1)$$

(f ist symmetrisch in den Argumenten).

Man erhält jedoch keine Aussagen über das simultane Vorliegen von (A_1, B_1) oder (A_2, B_2) , da dies auch nicht empirisch erhoben wurde.

Um Aussagen über die simultane Verteilung von (A_1, B_1) oder (A_2, B_2) zu erhalten, müßte man die items simultan präsentieren und beantworten lassen (ist dies möglich?).

Weiterhin können sich auch die Randverteilungen

$$\begin{aligned} f(A_1) &\neq f(A_2) \\ f(B_1) &\neq f(B_2) \end{aligned}$$

unterscheiden. Etwa hat das Item A andere Mittelwerte und Streuungen, wenn es vor oder nach B plazierte wurde (Kontext-Effekt).

Zu beachten ist, daß man die Verteilung mit einem Index versehen müßte, d.h. $f_{AB}(A_1, B_2)$, $f_{BA}(B_1, A_2)$. Dies wurde oben vermieden, um die Notation nicht zu überladen.



Daß die Vernachlässigung der Reihenfolge nicht unproblematisch ist, zeigt die Quantenmechanik, bei der Orts- und Impulsmessungen nicht verträglich sind, sich gegenseitig stören und keine gemeinsame Wahrscheinlichkeitsverteilung der Variablen existiert.²

Dies ist ein Beispiel dafür, daß sich Datenerhebung und Datenanalyse nicht immer sauber voneinander trennen lassen. Der Meßprozeß ist ein **Interaktionsprozeß**, bei dem Meßobjekt und Meßapparat in eine zeitlich befristete Wechselwirkung treten, und dann anschließend im Meßapparat die **irreversible Veränderung registriert** ist.

**Meßprozeß als
Interaktion**

Beispielsweise kann bei der Ortsmessung von Elektronen Licht verwendet werden, wobei die Lichtquanten aufgrund ihrer Energie Wirkungen auf den Meßgegenstand ausüben. Ohne auf Details weiter einzugehen, kann man sagen, daß eine genaue Ortsmessung die Geschwindigkeitsmessung beeinträchtigt und umgekehrt; die Variablen sind komplementär und nicht simultan meßbar. Als Konsequenz gilt eine sog. Unschärferelation (Heisenberg) und es existiert keine simultane Wahrscheinlichkeitsverteilung der Variablen x (Ort) und Impuls p (Masse mal Geschwindigkeit). Der Einfluß des Meßprozesses läßt sich nicht beliebig verkleinern, da die Energie nur in festen Portionen (Quanten) vorliegt. Allerdings sind die Objekte der Makrowelt so riesig, daß sich diese Quantisierungseffekte im allgemeinen nicht bemerkbar machen.

**simultane
Wahrscheinlichkeitsverteilung**

Auch im Bereich der Sozialforschung sind analoge Effekte zu berücksichtigen.³ Die Gewinnung von Daten setzt einen Eingriff in das soziale Gefüge voraus. Der Forscher muß Meßinstrumente, etwa Fragebögen, Videokameras oder Beobachter in das natürliche Geschehen einbringen. Möglicherweise wird bei der Untersuchung einer Firma auf das Vorliegen bestimmter Unternehmensstrukturen eine verzerrte oder verfälschte Realität erhoben (Reaktivität; soziale Erwünschtheit), und noch so elaborierte statistische Verfahren sind nicht in der Lage, aus

Reaktivität

²Die Wahrscheinlichkeitstheorie von Kolmogoroff ist ein Spezialfall der Wahrscheinlichkeitstheorie von J. von Neumann; vgl. Nelson, 1967, Kap. 14, Satz 14.1, Gudder, 1979, Kap. 2.1-2.

³vgl. P. Jordan, 1951, Kap. 3.

schlechten Daten etwas sinnvolles zu destillieren (*garbage in, garbage out*).

**empirische
Sozialforschung
als übergreifende
Disziplin**

Insofern ist die **empirische Sozialforschung eine übergreifende Disziplin**, da sie einen methodologischen Rahmen liefert, in dem sich Statistik abspielen kann.

Die Diskussion über die Möglichkeit der Gewinnung und den Status sozialwissenschaftlicher Daten hat eine weitverzweigte Literatur hervorgebracht und war Gegenstand erbitterter Auseinandersetzungen (vgl. Adorno et al., 1979).

Kapitel 8

Qualitative vs. quantitative Methoden

Üblicherweise wird versucht, den Methodenkanon der empirischen Sozialforschung dem der Naturwissenschaft, idealerweise der (klassischen) Physik, nachzubauen.¹

Dies ist mit Begriffen wie (siehe Glossar)

- objektive Realität, Realismus,
- Quantitativ,
- Messung,
- Gesetz,
- Experiment,
- Hypothesenprüfung,
- Erklären,
- Analytisch-Nomologisch,
- Positivismus,

**Methodenkanon
der empirischen
Sozialforschung**

**quantitative
Methoden**

¹Die Ergebnisse der Quantenmechanik (etwa Unschärferelationen, Komplementarität), sowie der Relativitätstheorie (Relativitätsprinzip, Äquivalenzprinzip etc.) werden dabei nicht zur Kenntnis genommen. Aber auch Ergebnisse der Thermodynamik wie z.B. die Irreversibilität von Prozessen dienen nicht als Modell der Sozialforschung und ihrer Gesetze. Neuere Ansätze wie die Synergetik sind hier innovativer, werden allerdings meist von Physikern vorgetragen (Haken, Weidlich). Auch die Idee der Selbstorganisation ist in die Soziologie vorgedrungen (vgl. Luhmann, 1988, S. 24 ff.).

- Wertneutralität,
- Standardisierung,
- Intersubjektivität,
- Verifikation, Falsifikation

etc. verknüpft.

Auf der anderen Seite machen Vertreter der qualitativ-interpretativen Richtung geltend, daß die sozialen Gegenstände erst in einem **Prozeß der Interpretation ausgehandelt und expliziert** werden müssen.

In diesem Zusammenhang fallende Schlagworte sind etwa

**qualitative
Methoden**

- Qualitativ,
- (Subjektive) Deutung,
- Bedeutung,
- Interpretation,
- Hermeneutik,
- Verstehen,
- Offenheit,
- Alltagstheorie,
- (dichte) Beschreibung,
- (symbolische) Interaktion,
- Kommunikation,
- Konstruktivismus.

V. Pispers (2001) hat die Situation auf den Punkt gebracht:

Es gibt eine Wirklichkeit und eine gefühlte Wirklichkeit ... Und die einzigen, die darauf bislang adäquat reagiert haben sind meines Wissens die Meteorologen ... Die sagen allen Ernstes: Heute 14 Grad – gefühlt 9 Grad... Die wissen nicht nur, wie kalt es ist. Die wissen, wie kalt mir ist ... Heute weiß ich: FDP-Wähler haben einen gefühlten Steuersatz von 90 Prozent. Da kommen Sie mit keinem Taschenrechner gegen an ...

Es ist klar, daß sich aus den obengenannten Ingredienzien eine Fülle von wissenschaftstheoretischen und methodologischen Positionen und Spielarten mischen lassen. Im Rahmen eines einführenden Statistik- und Methodenkurses erscheint es mir sinnvoll, nur auf den ersten Ansatz näher einzugehen und für das qualitative Paradigma auf die Literatur zu verweisen.

Es muß aber darauf hingewiesen werden, daß auch Bestrebungen zu einer Integration oder Synthese der Standpunkte zu verzeichnen sind. Etwa werden bei Methodenkonferenzen immer wieder Sessions zu integrativen Methoden angeboten.²

Eine kurze Gegenüberstellung der Standpunkte ist in Kromrey, 2002, Kap. 1, 9.2; auch Kurs 03607) zu finden. Ausführliche Darstellungen finden sich in den Kursmaterialien der FernUniversität: Heinze/Kleining (1999; Kurs 3655, 3656), Bergmann (1988; Kurs 3772), Strauss (1984; Kurs 3758).

²vgl. etwa RC33 conference on social science methodology; Session: Mixing Qualitative and Quantitative Methods oder MMIRA (Mixed Methods International Research Association); Abb. 8.1.

MMIRA Mixed Methods International Research Association

Home About Membership Events Discussion Forum Mixed Methods Resources News

Qualitative Quantitative

MMIRA supports the development of a mixed methods research community.

Join the MMIRA now!

We cordially invite you to join the Mixed Methods International Research Association (MMIRA), an exciting, new professional association created to promote the development of an international and interdisciplinary mixed methods research community. Among the many benefits of membership will be:

- Online subscription to the *Journal of Mixed Methods Research*
- Networking and training opportunities
- [Access to online resources](#)

To join the MMIRA, check out the [Membership](#) page.

Here is the link to our new **Facebook** page: <https://www.facebook.com/groups/mmira>

[We are pleased to announce that the MMIRA Conference Schedule and Preliminary Program can be viewed by clicking here!](#)

[Visit our Conference Sponsor Page!](#)

[MMIRA Election 2014 Results](#)

[Information for Conference Participants Staying on Campus](#)

Upcoming Conference

[MMIRA Annual Conference 2014](#)

News

[Welcome to MMIRA.org](#)

[MMIRA Election Results](#)

[Symposium at Aalborg Psychiatric Hospital in Aalborg, Denmark, August 25-29, 2014](#)

[Mixed Methods in Mental Health Services Research at Johns Hopkins June 23 - 24, 2014: M/Tu 8:30 am - 5:00 pm](#)

© MMIRA Powered by [Wild Apricot](#) Membership Software

Abbildung 8.1: Methodenkonferenz zu integrativen Methoden

Kapitel 9

Der Forschungsprozeß

Im Rahmen des quantifizierenden Paradigmas lassen sich Phasen des Forschungsprozesses unterscheiden, die möglicherweise mehrmals durchlaufen werden. Ein graphisches Schema ist Abb. 9.1 zu entnehmen (vgl. Schnell et al. 1999, S. 8, Bortz, 2005, S. 3). Nach Bortz (2005, S. 3 ff.) werden die Stadien

**Phasen des
Forschungs-
prozesses**

1. Erkundungsphase,
2. Theoretische Phase
3. Planungsphase
4. Untersuchungsphase
5. Auswertungsphase
6. Entscheidungsphase

unterschieden.

9.1 Erkundungsphase

In der Erkundungsphase muß das Problemfeld exploriert werden (Literatur-Recherche, Kontakt zu einschlägigen Forschern und Praktikern bzw. den entsprechenden Institutionen oder Firmen). Dabei soll die eigene Studie in einen theoretischen Kontext eingeordnet werden. Je nach Gegenstand gibt es elaborierte Theorien oder man betritt wissenschaftliches Neuland. Aus Theorien können dann Folgerungen und Hypothesen abgeleitet werden. In der explorativen Phase ist ein besonders

Exploration

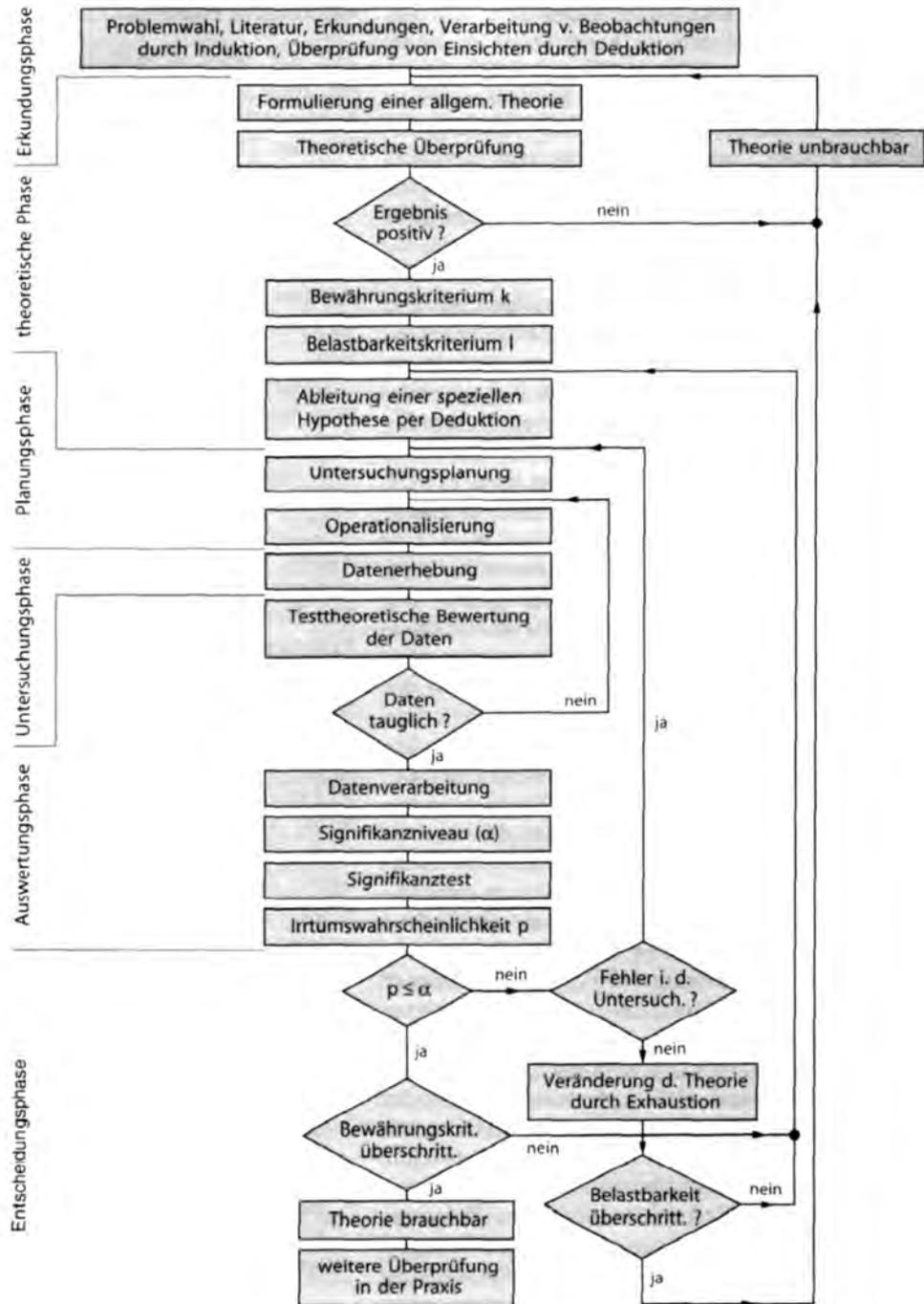


Abbildung 9.1: Phasen der empirischen Forschung (Bortz, 2005).

starkes Wechselspiel zwischen Theorie und Empirie zu beobachten, das die größte Nähe zu den qualitativen Methoden aufweist. Auch sind hier erste Voruntersuchungen (explorative Studien) einzuordnen.

**explorative
Studien**

9.2 Theoretische Phase

Empirische Überprüfungen einer Theorie sind nur sinnvoll, wenn zumindest ihre theoretische Struktur bestimmte Gütekriterien erfüllt. Man muß prüfen, ob

**theoretische
Struktur**

1. die Theorie präzise formuliert ist,
2. ob sie einen Informationsgehalt besitzt,
3. logisch konsistent ist,
4. mit anderen Theorien vereinbar
5. und empirisch überhaupt überprüfbar ist.

9.2.1 Erläuterungen zu den einzelnen Punkten:

1. Grundlegend ist die möglichst *präzise Definition der Begriffe*, die in einer Theorie vorkommen (Konzeptspezifikation, Operationalisierung).

Beispielsweise ist der Begriff *ethnische Identität* zunächst unklar, da zuerst die Teilbegriffe *Ethnisch* und *Identität* definiert und abgegrenzt werden müssen. Damit in Zusammenhang steht die Frage, ob es beobachtbare Sachverhalte (Indikatoren) gibt, die mit den theoretischen Begriffen möglichst übereinstimmen.

Indikatoren

Die Frage, wie den Begriffen die Indikatoren zugeordnet werden, wird unter dem Titel *Operationalisierung* geklärt. Dies beinhaltet Anweisungen, wie Messungen vorgenommen werden sollen.

Operationalisierung

Etwa wird *Intelligenz* durch Ausfüllen eines Intelligenz-Tests und einer bestimmten Aggregationsmethode der Teilaufgaben (items) operationalisiert. Meistens ist dies die Summe (der Rohwerte; Likert-Skala, vgl. 14.10.1) und darauf folgende Standardisierungen.

Likert-Skala

Falsifikatoren**Konditionalsatz**

2. Der *Informationsgehalt* (*empirische Gehalt*) der Aussagen einer Theorie bezieht sich auf ihre Falsifikatoren. Betrachtet man sogenannte Konditionalsätze (*wenn-dann-Satz* oder *je-desto-Satz*), so steigt der Informationsgehalt mit der Zahl der Ereignisse, die mit dem *dann* (bzw. *desto*)-Teil in Widerspruch stehen. Beispielsweise ist für den Satz

- *A: Wenn der Blutalkoholspiegel 0.5 Promille übersteigt, sinkt die Reaktionsfähigkeit*

der Nachweis einer verbesserten Reaktionsfähigkeit ein Falsifikator. Dagegen sind Sätze wie

- *B: Wenn der Hahn kräht auf dem Mist, ändert sich das Wetter oder es bleibt wie es ist*

nicht falsifizierbar, da der Dann-Teil immer wahr ist. Auch verringern vage, unpräzise Begriffe den Informationsgehalt eines Satzes, etwa kann die Reaktionsfähigkeit durch präzise Reaktionszeitmessungen oder lediglich durch Beobachtung ermittelt werden. Im ersteren Fall gibt es mehr Ereignisse, die dem Satz widersprechen.

Tautologien,**Kontradiktionen**

3. Theoretische Aussagen sollten keine Tautologien oder Kontradiktionen sein, die immer wahr oder falsch sind. Etwa ist Satz *B* tautologisch, da er immer wahr ist (und daher auch nicht empirisch überprüft werden muß). Versteckte Tautologien stecken in *Kann*-Sätzen, etwa

- *C: Rauchen kann Krebs verursachen.*

In diesem Fall ist sowohl das Auftreten als auch das Nicht-Auftreten von Krebs mit der Aussage vereinbar. Überprüfbar wird der Satz erst durch eine Häufigkeits- oder Wahrscheinlichkeitsaussage, etwa

- *D: Bei Rauchern ist die Wahrscheinlichkeit für Krebs höher als bei Nichtrauchern.*

Entsprechend müssen die bedingten Häufigkeiten für Krebs in den Gruppen der Raucher/Nichtraucher ermittelt und statistisch getestet werden, jedoch ist im Einzelfall keine empirische Überprüfung möglich.

Wichtig ist jedoch, daß alle Kombinationsmöglichkeiten (Krebs (ja/nein) \times Raucher/Nichtraucher) untersucht werden.



Abbildung 9.2: Behauptungen auf einer Zigarettenspackung.

Übung 9.1

- (a) Lassen sich die in Abb. 9.2 gezeigten Behauptungen empirisch überprüfen, und wenn ja, in welcher Form?
- (b) Wie ist die zweite Behauptung vermutlich gemeint?
- (c) Formulieren Sie die Behauptungen in Form einer statistischen Hypothese.



4. Liegen mehrere Theorien vor, die sich auf den gleichen Gegenstandsbereich beziehen, so muß untersucht werden, ob logische Widersprüche zwischen den Theorien bestehen. Sind keine logischen Widersprüche auffindbar, so bedeutet dies nicht, daß die

**logische
Widersprüche**

**Verifikation,
Falsifikation**

Theorien wahr sind. Dies kann, wie gesagt, nur durch empirische Überprüfung herausgefunden werden.

5. Schließlich muß die *empirische Überprüfbarkeit* (bzw. Falsifizierbarkeit) der Theorie analysiert werden. Es ist möglich, daß eine Theorie im Prinzip falsifizierbar ist, jedoch beim gegenwärtigen Stand der Forschung die Begriffe noch nicht genau oder weit genug meßbar sind. Dann müssen erst geeignete Meßinstrumente entwickelt werden.

Beispiel 9.1

In der physikalischen Forschung sind bestimmte Theorien erst dann überprüfbar, wenn neue Beschleuniger gebaut werden, die Prozesse mit hoher Energie zum Nachweis bestimmter Elementarteilchen erlauben. ■

Im allgemeinen kann eine Theorie nicht vollständig überprüft werden, sondern nur bestimmte Folgerungen und deduzierte Teilaspekte.

Beispiel 9.2

Die allgemeine Relativitätstheorie sagt vorher, daß Licht durch Massen um einen bestimmten Winkel abgelenkt wird. Dieser (im Falle der Sonne 0.00049 Grad) wurde von Einstein (1915) berechnet und bei einer Sonnenfinsternis im Jahre 1919 gemessen. Da Messung und Vorhersage 'gut' übereinstimmten, wurde dies als empirische Bestätigung der Theorie gefeiert. Auch heute wird die Theorie noch überprüft (Abb. 9.3, 10.1). Andere Vorhersagen wie z.B. die Gravitonen wurden bis heute nicht entdeckt, was an der Grobheit der Meßtechnik liegen mag. ■



Abbildung 9.3: Zeitungsausschnitt zur allgemeinen Relativitätstheorie (27. 4. 2005).

Beispiel 9.3 (Autoritärer Unterricht)

(nach Bortz, a.a.o.)

Die 'Theorie'

- *T: Autoritärer Unterricht hat negative Auswirkungen auf das Sozialverhalten von Kindern.*

ist zunächst mit vagen Begriffen definiert. Zunächst muß eine Definition der Begriffe *Autoritärer Unterricht* sowie *Sozialverhalten* der Kinder gefunden werden. Dann können speziellere Hypothesen aus *T* abgeleitet werden. Beispielsweise könnte Kooperationsbereitschaft als Komponente von Sozialverhalten identifiziert werden. Außerdem könnte man nur Mädchen in der 8. Klasse untersuchen, da in der Theorie kein Alter und kein Geschlecht spezifiziert wurde.

**Kooperations-
bereitschaft**

Daraus folgt die abgeleitete Hypothese

- H_1 : *Autoritärer Unterricht hat negative Auswirkungen auf die Kooperationsbereitschaft von Mädchen in der 8. Klasse.*

Unschwer lassen sich andere Hypothesen ableiten, wobei sich die Frage stellt, wieviele untersucht werden müssen, um die Theorie zu prüfen. Da die Theorie sehr allgemein formuliert wurde, stellt sich die Frage nach der Gültigkeit für Jungen, andere Altersklassen, in verschiedenen Schultypen, in anderen Ländern etc. Allerdings würde eine negative Evidenz (Falsifikation) die Theorie belasten (Exhaustion), was jedoch durch zusätzliche Antezedenz-Bedingungen (etwa Persönlichkeitseigenschaften der Schüler etc.) wieder behoben werden kann:

Antezedenz-Bedingungen

- T : *Autoritärer Unterricht (A_1) und eine autoritäre Persönlichkeits-Struktur (A_2) hat negative Auswirkungen auf das Sozialverhalten von Kindern.*

Auch diese kompliziertere Hypothese kann überprüft werden, nachdem der Begriff *autoritäre Persönlichkeits-Struktur* expliziert worden ist. ■

Exhaustion

Exhaustionen sind laut Holzkamp nur insoweit zulässig, als die Behauptung, daß Abweichungen zwischen Theorie und Daten durch störende Bedingungen verursacht werden, auch begründet werden kann (Holzkamp, 1972, S. 95).

9.3 Planungsphase

Ist die theoretische Überprüfung positiv verlaufen, so muß Aufbau und Ablauf der Untersuchung geplant werden. Damit wird vermieden, daß während der Datenerhebung Pannen auftreten, die sich nicht mehr korrigieren lassen.

9.3.1 Variablen-Auswahl

Variablen

Zunächst werden *Variablen* ausgewählt, die für die Untersuchung relevant sein können. Darunter sind Merkmale zu verstehen, die mindestens 2 Ausprägungen annehmen können (etwa Geschlecht = {weiblich,männlich}, Unterrichtsstil = {autoritär,demokratisch}).

Außerdem wird festgelegt, auf welchem *Meßniveau* die Daten erhoben werden sollen (Kap. 1.3).

Weiterhin ist zu klären, welche Variablen als unabhängig (X -Variablen; UV), und welche als abhängig (Y -Variablen; AV) entsprechend der Funktion $Y = f(X)$ zu betrachten sind.

Dies ist insbesondere wichtig, wenn ein experimentelles Design geplant ist, bei dem die Werte der X -Variablen vom Versuchsleiter gesetzt werden.

Bei reinen Beobachtungs(Survey)-Studien liegen allerdings die Werte der Variablen in ihrem natürlichen Umfeld faktisch vor, so daß die Unterscheidung in unabhängig/abhängig entweder durch theoretische Überlegungen oder statistische Modelle (etwa Strukturgleichungsmodelle) erfolgen muß.

Sehr wichtig ist, bei Analysen von Beobachtungs-Studien zu beachten, daß auch die UV (X) zufällig sind. Dann lassen sich die y -Werte nicht bei festen x -Werten wiederholen, sondern es gibt nur zufällige Wertepaare $(x_n, y_n), n = 1, \dots, N$. Die üblichen Resultate sind dann nur **bedingt** auf die gemessenen $X_n = x_n$ -Werte zu interpretieren (vgl. Kap. 5.1.2.2, 14.8.4).

Strukturgleichungsmodelle wie LISREL modellieren die Regressoren ξ_n von vorneherein stochastisch, d.h.

$$\eta_n = \mathbf{B}\eta_n + \mathbf{\Gamma}\xi_n + \zeta_n. \quad (9.1)$$

Die Kovarianzmatrix $\Phi = \text{Var}(\xi_n)$ der Regressoren enthält freie Parameter des Modells, die mitgeschätzt werden.

Weitere Variablen, die weder unabhängig noch abhängig sind, werden als Kontrollvariablen (Moderatorvariablen) bezeichnet. Sie können entweder miterhoben, konstant gehalten oder gar nicht berücksichtigt werden.

Meßniveau**(un)abhängige Variablen****Survey-Studie****stochastische Regressoren****Kontrollvariablen, Moderatorvariablen**

Beispiel 9.4

Das Geschlecht der Probanden kann als Kontrollvariable Z aufgefaßt werden, die den Zusammenhang zwischen X und Y moderiert. Mißt man auch Z , so kann der Zusammenhang geschlechtsspezifisch untersucht werden, hält man Z konstant, so ergibt sich nur der Zusammenhang in einer Gruppe (etwa der Mädchen). Wird Z nicht erhoben, so ist der gefundene

Zusammenhang eine Mixtur aus den gruppenspezifischen Effekten, die sich möglicherweise aufheben.

Faßt man die Variablen (X, Y) als Vektor auf, so wirkt Z simultan auf beide Variablen. Der Zusammenhang zwischen X und Y läßt sich dann als mittlerer Zusammenhang bei kontrolliertem Z plus Zusammenhang der Prognosen (erklärte Werte) interpretieren

(Totale Kovarianz = Residuale Kovarianz + Erklärte Kovarianz).

Scheinkorrelation

Das Problem der Scheinkorrelation wird ausführlich in Kap. 14.8 diskutiert. ■

Weiterhin ist zu überlegen, welche Variablen die untersuchte Gesetzmäßigkeit beeinflussen können, jedoch nicht berücksichtigt wurden (weder theoretisch noch empirisch).

Beispiel 9.5

Wichtige zusätzliche Variablen könnten sein: Alter des Lehrers, Intelligenz und Motivation der Kinder, Lärmbelästigung, soziale Schicht. ■

Störvariablen

Die potentiell wichtigen, jedoch nicht miterhobenen Variablen werden oft als Störvariablen bezeichnet. Dies umfaßt auch Größen, deren Einfluß überhaupt nicht bekannt ist. Störvariablen werden häufig als latente stochastische Größen modelliert (etwa Personeneffekte), da sie nicht erhoben wurden, jedoch auch nicht vernachlässigt werden können.

9.3.2 Untersuchungs-Designs

Labor vs. Feld Von großer Relevanz ist auch der Typ der Untersuchung (Design), etwa ob im Labor oder im natürlichen 'Feld' untersucht wird. Einerseits kann unter Laborbedingungen eine Fülle von Variablen kontrolliert oder ausgeschaltet werden, andererseits besteht der Nachteil einer geringen Generalisierbarkeit auf komplexe reale Bedingungen.

Labor vs. Feld

Beispiel 9.6

Kann das Erlernen von sinnlosen Silben unter verschiedenen Geräusch-Bedingungen (im Rahmen der Gedächtnisforschung) verallgemeinerbare Resultate für Alltags-Situationen erbringen? ■

Da die Zahl von Störvariablen im Labor niedriger ist, spricht man ihr eine hohe *interne Validität (Gültigkeit)* zu, da die Zahl von plausiblen Alternativ-Erklärungen durch die Kontrolle von Störvariablen geringer ausfällt.

interne Validität

Auf der anderen Seite hat die Labor-Untersuchung nur eine geringe *externe Validität*, da sie nur bedingt auf natürliche Situationen übertragen werden kann.

externe Validität

Experiment vs. Beobachtung Eine andere Unterscheidung von Untersuchungs-Designs erfolgt entlang der Gesichtspunkte Experiment vs. Beobachtung. Während im Labor die Zahl der *situativen* Störvariablen kontrolliert wird, ist bei einem Experiment die Zahl der *personellen* Störvariablen unter Kontrolle. Etwa können den Probanden die verschiedenen Ausprägungen der unabhängigen Variablen zufällig zugeweiht werden (Randomisierung). Die unabhängige Variable $X = \text{Unterrichtsstil}$ wird also vom Experimentator vorgegeben und liegt nicht nur faktisch im Feld vor.

Experiment vs. Beobachtung

Randomisierung

Beispiel 9.7

Von 20 Schülern werden je 10 zufällig auf die Gruppen Unterrichtsstil = {autoritär, demokratisch} verteilt. Damit sind (im Mittel) die personenbezogenen Merkmale (etwa Geschlecht, Intelligenz, Alter etc.) auf alle Ausprägungen der UV gleichmäßig verteilt. ■

Alternativ könnte man Schüler gleichen Geschlechts, gleicher Intelligenz und gleichen Alters auf die Gruppen verteilen.

Problematisch ist auf jeden Fall die externe Validität des Experiments, die durch eine repräsentative Stichprobe sicherzustellen wäre.

Beobachtungs-Studie

Bei einer Beobachtungsstudie ist die Zuordnung der UV zu den Probanden nicht willkürlich, sondern faktisch (natürlich) vorgegeben.

Beispiel 9.8

Etwa hat der Forscher keinen Einfluß darauf, welche Lehrer welchen Klassen zugeordnet werden und welche Gesichtspunkte dabei eine Rolle spielen.

Etwa könnten bestimmte ‘schwierige’ Klassen durch autoritäre Lehrer unterrichtet werden (Zuordnung durch den Rektor). ■

Beispiel 9.9**Selektionsprozesse**

Vergleicht man die Auswirkungen von Führungsstilen in verschiedenen Unternehmen, so sind die Mitarbeiter schon fest vorgegeben und können nicht zufällig verteilt werden. Hierbei spielen auch Selektionsprozesse (wer bewirbt sich bei welcher Firma) eine Rolle. ■

Konstanthaltung von Störvariablen

Insofern wird Beobachtungs-Studien eine geringe interne Validität zugeschrieben. Diese kann erhöht werden, wenn die zu vergleichenden Gruppen nach relevanten Störvariablen parallelisiert werden.

Beispiel 9.10

Etwa können Schulklassen so selektiert werden, daß Lehrer mit autoritärem oder demokratischem Unterrichtsstil zumindest gleichaltrig sind oder daß der Lärmpegel in den Klassen etwa gleich ist. ■

Beispiel 9.11

Eine ähnliche Paarbildung findet in der Zwillingforschung statt (Kontrolle der genetischen Variabilität). ■

Beispiel 9.12

Wenn eine Klasse im Zeitablauf von verschiedenen Lehrern unterrichtet wird, so werden die Schülervariablen im Prinzip konstant gehalten (bis auf das Lebensalter und andere fluktuierende Größen).

Insofern können Unterschiede leichter nachgewiesen werden, da eine geringere Streuung vorliegt.

Meßwiederholungen

Dieser Effekt wird häufig in Designs mit Meßwiederholungen ausgenutzt (Kap. 3.3.2). Hierbei wird der Effekt einer Intervention an einer abhängigen Stichprobe untersucht (gleiche Probanden vor und nach der Intervention). Die Varianz der Differenz

$$\text{Var}(Y - X) = \text{Var}(X) + \text{Var}(Y) - 2\text{Cov}(X, Y) \quad (9.2)$$

$$< \text{Var}(X) + \text{Var}(Y) \quad (9.3)$$

ist kleiner als bei unabhängigen Stichproben (**Stabilität** der Probanden vor und nach der Intervention: $\text{Cov}(X, Y) > 0$).

Mittelwertsunterschiede lassen sich somit leichter nachweisen. ■

9.3.3 Operationalisierung

In Rahmen der Theoretischen Phase wurde schon erwähnt, daß die Definition der Begriffe eine entscheidende Rolle für die Untersuchung spielt. Es muß nun festgelegt werden, mit welchen Indikatoren und Operationen die UV, AV sowie Moderatorvariablen erfaßt und quantifiziert werden.

Indikatoren

Während manche Variablen relativ leicht zu erfassen sind (Alter, Zahl der Geschwister, Geschlecht etc.), sind komplexe Variablen wie autoritärer Unterrichtsstil, Intelligenz etc. über mehrere Indikatoren (beobachtbare Variablen; vgl. Friedrichs, Kap. 2.3, S. 79 ff.) operational zu ermitteln.

Oft liegen hierzu schon Fragebögen, Beobachtungsbögen oder Tests vor, deren psychometrische Qualitäten (etwa Reliabilität) schon untersucht sind. Bei der Entwicklung neuer Meßinstrumente müssen in Voruntersuchungen die Eigenschaften der Fragebögen, Testbatterien etc. ermittelt werden.

Beispielsweise ist es bei einer Likert-Skala erforderlich, daß die Teilfragen (items) mit dem Summenwert korreliert sind, da ansonsten dieses item nicht zum Skalenwert beiträgt (sog. Item-Analyse; vgl. Kap. 14.10.2).

Item-Analyse

9.3.4 Stichprobe

Die Stichprobe ist ein Ausschnitt aus der Grundgesamtheit (GG), die die Eigenschaften der GG möglichst genau abbilden soll.

Grundgesamtheit

In diesem Zusammenhang fällt häufig der Begriff *Repräsentativität*. Beispielsweise wäre bei einer Befragung in der Fußgängerzone, bei der viele Personen die Teilnahme verweigern, die Stichprobe nicht repräsentativ bezüglich der Variable *Bereitschaft, an dieser Untersuchung teilzunehmen*.

Repräsentativität

Ist über die Verteilung der untersuchungsrelevanten Merkmale nichts bekannt, so sollte eine einfache Zufallsstichprobe untersucht werden, bei der jedes Element der GG die gleiche Wahrscheinlichkeit hat, ausgewählt zu werden.

**einfache
Zufallsstichprobe**

Oft sind jedoch nur sog. Klumpenstichproben verfügbar (Schulklassen, Städte, Landkreise etc.), die bereits vorgruppiert sind.

**Klumpen-
stichproben**

Kennt man Variablen, die Einfluß auf die Verteilung der relevanten

geschichtete Stichproben	Merkmale haben, so kann die Stichprobe so erhoben werden, daß sie repräsentativ bezüglich dieser Variablen ist. Man spricht dann von <i>geschichteten Stichproben</i> . Etwa sollte bei einer Untersuchung, die <i>repräsentativ</i> für die Bundesrepublik ist, Personen aus allen Bundesländern untersucht werden.
Quoten-Auswahl	Bei Umfragen des Instituts für Demoskopie Allensbach werden <i>Quoten</i> für die Variablen Geschlecht, Alter, Berufsgruppen, Wohnortgrößen und Bundesland benutzt (Quelle: Bamberg et al., 2008, S. 9f). Derartige Stichproben sind aber nicht mehr zufällig.
Stichprobengröße	Auch die Größe N der Stichprobe ist eine wichtige Frage, die meistens in der statistischen Beratung gestellt wird:

- *wieviele Versuchspersonen brauche ich, damit ich ein signifikantes Resultat erhalte?*

Dies läßt sich allerdings nur mit Hilfe eines Genauigkeitskriteriums (Effektstärke) beantworten, da die Schätzfehler mit der Stichprobengröße abnehmen. Will man etwa einen Mittelwertsunterschied prüfen, so muß zur Abschätzung der nötigen Stichprobengröße angegeben werden, welchen Unterschied man mindestens nachweisen will.

Dies kann der Statistiker nicht beantworten.

Signifikanz und Relevanz	Umgekehrt kann man auch geringste Unterschiede nachweisen, wenn nur N genügend groß ist (Verwechslung von Signifikanz und Relevanz).
---------------------------------	--

Beispiel 9.13 (Gauß-Test)

Will man einen Unterschied von $D = 10$ Intelligenzpunkten nachweisen können, so muß man in der Signifikanzbedingung für die Teststatistik

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{(1/N)(\sigma_X^2 + \sigma_Y^2)}} > z = z(1 - \alpha/2) \quad (9.4)$$

nach der Stichprobengröße N auflösen

$$N > \frac{\sigma_X^2 + \sigma_Y^2}{(\bar{X} - \bar{Y})^2} z^2. \quad (9.5)$$

und die (zufällige) Mittelwertsdifferenz durch den gewünschten Effekt D ersetzen.

Somit muß N um so größer gewählt werden, je kleiner die Irrtumswahrscheinlichkeit α , die nachzuweisende Differenz und je größer die Streuung der Daten ist.

Da sich in einer Untersuchung ein zufälliger Wert $\bar{X} - \bar{Y}$ ergibt, hat man keine Gewähr, daß man den gesuchten Effekt auch nachweisen kann.

Numerisch ergibt sich [$D = 10, \sigma_X = \sigma_Y = 15, z = z(.975) = 1.96$]

$$N > \frac{450}{100} 1.96^2 = 17.3. \quad (9.6)$$



9.3.5 Planung der statistischen Auswertung

In der Praxis werden häufig Daten erhoben, ohne an die spätere Auswertung zu denken. Dabei kann es passieren, daß sich bestimmte Verfahren nicht anwenden lassen, weil das notwendige Skalenniveau nicht erreicht wurde.

Will man beispielsweise eine Varianzanalyse anwenden, so muß die abhängige Variable mindestens **intervallskaliert** sein; bei Vorliegen einer Ordinal-Skala muß auf andere Verfahren zurückgegriffen werden, die auf Rängen basieren.

Varianzanalyse

Dies ist jedoch eine Vergrößerung der Messung, die man bei besserer Planung hätte vermeiden können.

Weiterhin sollte die Untersuchung nicht völlig auf eine einzige Auswertungsmethode zugeschnitten sein, da die Erfassung der Realität komplett 'durch die Brille' dieser einen Methode gesehen wird.

Idealerweise sollte eine Theorie bzw. ihre empirische Realisation *durch den Gegenstand* und **nicht** durch die zur Verfügung stehenden (bzw. dem Forscher bekannten) Auswertungsmethoden bestimmt sein.

Beispielsweise wurde die Faktorenanalyse von Thurstone erfunden, um ein adäquates statistisches Instrument zur Intelligenzforschung zur Verfügung zu haben.

Faktorenanalyse

Verkürzung der Realität

Als etwas moderatere Forderung kann man sagen, daß möglichst gute Kenntnisse der verschiedenen Methoden erforderlich sind. Allerdings kann man davon ausgehen, daß Praktiker nur Verfahren anwenden, die auch in den gängigen Statistik-Paketen (SPSS, SAS, STATA etc.) implementiert sind und die **Realität** somit auf diese Verfahren **verkürzen**.

9.4 Untersuchungsphase

Datenerhebung

Nach einer sorgfältigen Planung kann die eigentliche Datenerhebung beginnen. Je nach Art der Meßinstrumente und der Untersuchungsteilnehmer sowie der räumlichen, zeitlichen, personellen sowie finanziellen Rahmenbedingungen lassen sich keine generellen Leitlinien erstellen.

Grundsätzlich ist zu sagen, daß der Erhebungsvorgang den natürlichen Ablauf **möglichst wenig stören** soll und daß die Datenerhebung möglichst gleichartig in den verschiedenen Subgruppen erfolgt.

Versuchsleiter-Effekte

In diesem Zusammenhang ist das Verhalten von Versuchsleitern (VL), Interviewern, Test-Instruktoren etc. von großer Wichtigkeit (Versuchsleiter-Effekte). Da das Verhalten der VL die Resultate wesentlich beeinflussen kann, muß eine möglichst hohe Standardisierung und Homogenisierung erreicht werden. Etwa lassen sich Instruktionen von Tonband oder per Video in identischer Weise bei allen Untersuchungsgruppen präsentieren. Alternativ kann versucht werden, in einer Schulung die Untersucher auf vergleichbare Verhaltensmuster zu trainieren. Inter(und intra)individuelle Unterschiede werden sich jedoch kaum vermeiden lassen.

Es ist also immer mit unkontrollierbaren Einflüssen des **Messinstruments Mensch** zu rechnen, die zu Unschärfen führen (vgl. Fußnote 3).

9.5 Auswertungsphase

Kodierung

Im Anschluß an die Erhebung werden die Daten statistisch analysiert. Dazu muß für jede statistische Einheit und für jedes Merkmal (Variable) eine Kodierung der Ausprägungen vorgenommen werden. Meist werden

	treatment	April AM	April PM	May AM	May PM	June AM	June PM
1	A	278	280	284	208	171	175
2	A	278	281	195	199	185	189
3	A	276	280	215	219	179	181
4	A	276	281	201	211	183	188
5	A	279	285	188	192	170	174
6	B	266	270	220	224	180	184
7	B	280	284	228	232	200	204
8	B	284	288	233	237	175	179
9	B	273	277	215	219	202	207
10	B	281	285	237	241	199	205
11	Control	278	282	273	277	281	285
12	Control	273	277	274	278	280	282
13	Control	282	285	276	281	279	282
14	Control	274	277	285	289	274	278
15	Control	277	281	279	284	284	285
16	Placebo	279	283	278	284	288	272
17	Placebo	277	279	291	291	280	285
18	Placebo	275	279	280	283	281	283
19	Placebo	276	282	277	282	274	279
20	Placebo	282	286	281	285	282	285

Abbildung 9.4: Datenmatrix: Zeilen = Personen, Spalten = Variablen. Die Variable *treatment* wurde verbal kodiert.

hierfür Zahlen benutzt (z.B. männlich = 1, weiblich = 2; oder trifft zu = 1; trifft nicht zu = 0), jedoch können auch andere Symbole (etwa m, w) benutzt werden. Dies setzt allerdings voraus, daß das benutzte Programm auch solche Kodierungen verarbeiten kann (vgl. Abb. 9.4).

Liegen die Daten nur als Ton- und Video-Protokoll oder in schriftlicher Form vor, so müssen daraus mit Hilfe von Beurteilern und Operationalisierung-Vorschriften den einzelnen Variablen Merkmalsausprägungen (zumindest in nominaler Form) zugewiesen werden.

Beurteiler

Beispiel 9.14**Rating-Skala**

Zur Beurteilung einer Video-Aufzeichnung wird dem Beurteiler ein Beurteilungsbogen zur Erzeugung einer Rating-Skala (engl.: rating = Einschätzung) vorgelegt:

Die Gruppenatmosphäre in der ersten Video-Aufzeichnung empfinde ich als

gespannt 1 2 3 4 5 gelöst

oder

Die Gruppenatmosphäre in der ersten Video-Aufzeichnung empfinde ich als gespannt

trifft zu trifft nicht zu

(Bortz/Döring, 2006, Kap. 4.2.4) ■

**statistische
Einheiten
Datenmatrix**

Nach der Übertragung der Erhebungsprotokolle (Fragebögen, Beobachtungsbögen, Video-Aufzeichnungen etc.) in eine elektronische Form (Datei; engl.: file), die normalerweise die Meßwerte der Personen (statistische Einheiten) als Zeilen sowie die Variablen in den Spalten enthält, wird diese sogenannte Datenmatrix (Abb. 9.4) in ein statistisches Auswertungsprogramm eingelesen.

Fehlende Werte

Fehlende Werte (missing data) werden hierbei mit speziellen Symbolen codiert (z.B. ein Punkt . in SAS). Dann können die Spalten (Variablen) untersucht werden.

In einem ersten Schritt muß die psychometrische Qualität der Daten überprüft werden. Hierbei ist zunächst das Gütekriterium der *Objektivität* zu nennen. Es muß untersucht werden, ob mehrere Beobachter oder Auswerter den Vpn aufgrund der Untersuchungsergebnisse die gleiche Merkmalsausprägung (bzw. Zahlenwert im numerischen Relativ) zuordnen (Intersubjektivität). Bei Fragebogen, die von den Probanden ausgefüllt werden (z.B. Intelligenztest) ist dies i.a. unproblematisch, jedoch kann es vorkommen, daß bei unscharfen Operationalisierungen die Beurteiler (sog. *Rater*; engl.: rating = Einschätzung) zu unterschiedlichen Einschätzungen, etwa der Kooperationsbereitschaft der Kinder, kommen. Damit zusammenhängend ist die Inter-Rater-Reliabilität, welche die Übereinstimmung verschiedener Beurteiler ausdrückt. Die Beobachterübereinstimmung kann z.B. mit Hilfe des Kappa(κ)-Koeffizienten von Cohen ausgedrückt werden, der sich als Quotient

$$\kappa = \frac{F(-) - F(+)}{F(-)} = \frac{G(+) - G(-)}{1 - G(-)} \quad (9.7)$$

berechnen läßt, wobei $F(+) = 1 - G(+)$ der Prozentsatz der Nichtübereinstimmung zweier Beurteiler und $F(-)$ die Nichtübereinstimmung bei rein zufälliger Beurteilung ist (proportionale Fehlerreduktion PRE; vgl. Nagl, 1992, Kap. 3.2).

Kappa ist also die Differenz der Übereinstimmungen (tatsächliche minus zufällige: $G(+) - G(-)$) bezogen auf die perfekte Übereinstimmung minus der zufälligen Übereinstimmung $1 - G(-)$. Somit kann κ maximal 1 werden (perfekte Beobachterübereinstimmung).

Weiterhin müssen Reliabilitätsuntersuchungen der betrachteten Skalen durchgeführt werden (Kap. 14.9). Das Gütekriterium der *Reliabilität* ist ein Maß für die Genauigkeit bzw. Zuverlässigkeit der erhobenen Daten.

Im Rahmen der klassischen Testtheorie setzt sich der beobachtbare Wert X aus wahren Wert ($T = \text{true score}$) und Meßfehler ϵ zusammen, d.h.

$$X = T + \epsilon \quad (9.8)$$

$$\text{Var}(X) = \text{Var}(T) + \text{Var}(\epsilon), \quad (9.9)$$

wenn T und ϵ unkorreliert sind. Je kleiner die Meßfehler-Varianz, um so

**psychometrische
Qualität**

Objektivität

**numerisches
Relativ**

Rater

Cohen κ

PRE

Reliabilität

**klassische
Testtheorie**

größer ist die Reliabilität (erklärte durch gesamte Streuung)

$$rel = \frac{\text{Var}(T)}{\text{Var}(X)} = \frac{\text{Var}(T)}{\text{Var}(T) + \text{Var}(\epsilon)} \quad (9.10)$$

(maximal = 1, wenn kein Meßfehler vorliegt). Bei geringen Reliabilitäten kann versucht werden, durch bessere Operationalisierungen oder mehr items eine Verringerung des Meßfehlers zu erreichen.

Eine geringe Reliabilität verringert die Güte von Tests. Soll eine Hypothese über T geprüft werden, etwa

$$H_0 : \mu := E[T] = \mu_0 \quad (9.11)$$

so hat die Teststatistik \bar{X} die Varianz

$$(1/N)(\text{Var}(T) + \text{Var}(\epsilon)), \quad (9.12)$$

d.h. der Meßfehler ϵ vergrößert die Streuung der Testgröße.

Somit muß die Differenz $\bar{X} - \mu_0$ größer werden, um ein signifikantes Resultat zu erhalten (bei gleicher Stichprobengröße und gleichem Signifikanzniveau).

explorative Statistik

Bevor inferenzstatistische Untersuchungen (Tests; Modell-Anpassung) vorgenommen werden, muß man sich einen deskriptiven (explorativen) Überblick der Daten verschaffen. Dies erfolgt mit Hilfe von graphischen Darstellungen (stem-leaf-Diagramme, Histogramme, Tortendiagramme, Streudiagramme etc.). Dabei kann eine Unterteilung der Gesamtstichprobe in Sub-Gruppen (z.B Frauen/Männer; Gruppierungsvariable Geschlecht) vorgenommen werden. Auch sollten Zusammenhänge zwischen Variablen graphisch dargestellt werden (etwa kann in einem Streudiagramm ein nichtlinearer Zusammenhang entdeckt werden, der bei einer schematischen Korrelationsberechnung leicht untergeht).

graphische Datenanalyse

Die graphische Datenanalyse ist nicht unwissenschaftlich, sondern bildet die Voraussetzung zur Entdeckung von Zusammenhängen, die mit entsprechenden Modellen geprüft werden können.

Statistiken

Auch ist die Berechnung von statistischen Kennwerten (Statistiken) für alle Variablen erforderlich (je nach Skalenniveau: Mittelwert, Median, Modus, Standardabweichung, Quantile, Schiefe, Kurtosis).

symmetrische, asymmetrische Fragestellung

Anschließend werden dann Modelle über den Zusammenhang (symmetrisch) oder die Abhängigkeit (asymmetrisch) von Variablen geschätzt.



Abbildung 9.5: JMP-Output: Vierfeldertafel nach Skalenniveau (nominal n, ordinal o, continuous c) der UV und AV

Je nach Skalenniveau der UV und AV kommen Regressionsmodelle, kategoriale Regression, Varianzanalyse, etc. in Frage. Hilfreich ist hierbei das Vierfelder-Schema in Abb. (9.5) (SAS/JMP erkennt anhand des Skalenniveaus automatisch das richtige Verfahren).

Da die UV nominal skaliert (Behandlung A, B, Kontrolle, Placebo) und die AV metrisch ist (Cholesterol-Wert), wird eine Varianzanalyse gerechnet, d.h. der Einfluß der UV *Behandlung* auf den Cholesterol-Wert wird analysiert. Anders formuliert: gibt es Mittelwertsunterschiede zwischen den Behandlungs-Gruppen? Die graphische Darstellung orientiert sich dabei an den üblichen Funktionsdiagrammen $y = f(x)$, oder $AV = f(UV)$. In Abb. (9.6) werden auf der x-Achse die Behandlungsarten und auf der y-Achse die Cholesterol-Werte, sowie Mittelwerte und Box-Plots mit den Statistiken Median, Quartile und Maximum, Minimum dargestellt.

Auch können Signifikanztests durchgeführt werden:

Beispielsweise wird die Hypothese gleicher Mittelwerte $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ mit Hilfe eines F -Tests überprüft (Abb. 9.6, unten). Da der p -Wert 0.9512 sehr groß ist, muß H_0 beibehalten werden. Dies ist aufgrund der Daten auch plausibel.

Hierbei ist p die Wahrscheinlichkeit, mit der die gemessene Teststatistik ($F = .1131$) überschritten wird (Überschreitungswahrscheinlichkeit). Dieser Wert sollte jedoch bei Ablehnung von H_0 kleiner als das gewählte

Varianzanalyse

Überschreitungswahrscheinlichkeit (p -Wert)

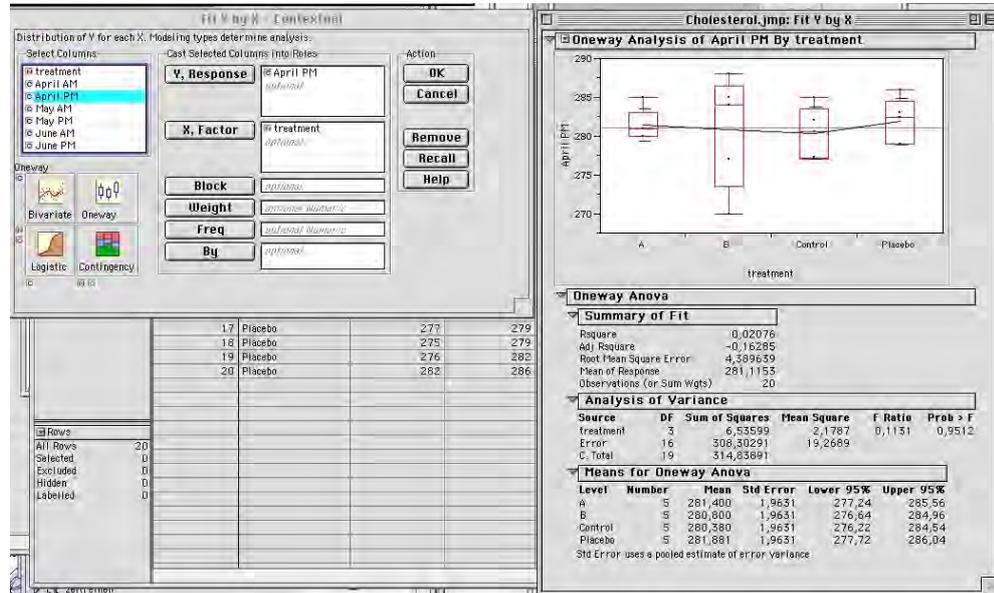


Abbildung 9.6: SAS/JMP-Output: Vierfeldertafel (links) nach Skalenniveau (nominal n, ordinal o, continuous c) der UV und AV sowie Graphik und Auswertung (rechts). Da UV=treatment=Nominal und AV=Cholesterol-Wert=metrisch, wird eine Varianzanalyse (ANOVA) gerechnet.

Signifikanzniveau (etwa $\alpha = 0.05, 0.01$ etc.) sein.

Signifikanzniveau

Fehler 1. Art (α -Fehler)

Wichtig ist, daß α vor der Untersuchung (bzw. Auswertung) gewählt und nicht nach Berechnung von F bzw. p passend ausgesucht wird. Es handelt sich nämlich um den Fehler 1. Art des Tests, d.h die Wahrscheinlichkeit H_0 abzulehnen, obwohl H_0 richtig ist. Diese Fehlentscheidung sollte eine möglichst kleine Wahrscheinlichkeit besitzen, jedoch a priori gewählt werden.

9.6 Entscheidungsphase

Falls nun $p < \alpha$ (signifikantes Ergebnis; H_0 ablehnen), so wird man H_1 annehmen, und der dabei begangene Fehler ist kleiner als $\alpha = 0.05, 0.01, \dots$. Damit gilt H_1 als (ziemlich sicher) bestätigt. Würde also autoritärer Unterrichtstil negatives Sozialverhalten fördern, so müßte ein signifikanter Mittelwertsunterschied zwischen den Gruppen (= UV) *autoritärer/nichtautoritärer Unterrichtstil* auf der AV *negatives Sozialverhalten* vorliegen.

Anders ist das Problem, wenn $p > \alpha$, also H_0 beibehalten werden muß. Da bei üblichen Hypothesen ($H_0 : \mu_1 = \mu_2$ gegen $H_1 : \mu_1 \neq \mu_2$) der Fehler 2. Art (β -Fehler) nicht kontrolliert (und maximal $1 - \alpha$) werden kann, spricht man vorsichtig von 'H₀ beibehalten', also nicht etwa von 'H₀ bestätigt' etc. Hierbei ist β die Wahrscheinlichkeit, H_0 beizubehalten, obwohl H_1 richtig ist. Man kann vielleicht nur aufgrund der Datenlage (kleine Stichprobe, Test mit geringer *Macht*) die Nullhypothese nicht ablehnen. Weiterhin könnten zusätzliche Antezedenzbedingungen notwendig sein:

- T' : *Autoritärer Unterricht (A₁) und eine autoritäre Persönlichkeits-Struktur (A₂) hat negative Auswirkungen auf das Sozialverhalten von Kindern (Y).*

Dann müßte der Mittelwertsvergleich auf dem 2-faktoriellen Design $Y = f(A_1, A_2)$ erfolgen, wobei A_1, A_2 jeweils 2 Ausprägungen hat (also 4 Kombinationen). Es könnte z.B. ein Mittelwertsunterschied in der Subgruppe: *Kinder mit autoritärer Persönlichkeits-Struktur* vorliegen (Interaktionseffekt). Derartige Modifikationen einer Theorie aufgrund einer Falsifikation werden als *Exhaustion* bezeichnet (Holzkamp, 1972, S. 75 ff, S. 147 ff).

Falls nun auch die erweiterte Theorie T' zu keinem signifikanten Ergebnis führt, könnte man weitere Antezedenzbedingungen hinzunehmen (etwa $A_3 =$ Schüler gehört der Oberschicht an etc.) und diese Theorie überprüfen. Andererseits wird der Erklärungswert immer unbedeutender, da im Wenn-Teil der Aussage immer mehr einschränkende Bedingungen genannt werden. In der Praxis wird das Interesse an solchen immer komplizierter werdenden Theorien nachlassen und neuen Ansätzen Platz machen. P. Meehl (1978, S. 807) schreibt:

There is period of enthusiasm about a new theory, a period of attempted application of several fact domains, a period

**Fehler 2. Art
(β -Fehler)**

Interaktionseffekt

Exhaustion

of disillusionment as negative data come in, a growing bafflement about inconsistent and unreplicable results, multiple resort to ad hoc excuses and then finally people just sort of lose interest in the thing and pursue other endeavors.

Falsifikation

Im einfachsten Fall gilt für die Falsifikation (modus tollens, Kontrapositionsgesetz ¹; T = Theorie, O = empirisch überprüfbare Resultate, \bar{O} = nicht O)

$$T \Rightarrow O \quad (9.13)$$

$$\bar{O} \Rightarrow \bar{T} \quad (9.14)$$

d.h. wenn O nicht gilt, ist auch T falsch.

**Zusatzhypothesen
Randbedingungen**

Andererseits benötigt man zusätzlich zu einer Theorie auch noch Zusatzhypothesen A und Randbedingungen R , bis man empirisch überprüfbare Resultate erhält (\wedge = logisches *und*, \vee = logisches *oder*). Dies ist die Aussage der sog. Duhem-Quine-These (vgl. Duhem, 1998 (1906)), daß sich Beobachtungssätze nur aus Satzsystemen ableiten lassen:

**Duhem-Quine-
These**

$$T \wedge A \wedge R \Rightarrow O \quad (9.15)$$

$$\bar{O} \Rightarrow \overline{T \wedge A \wedge R} = \bar{T} \vee \bar{A} \vee \bar{R}. \quad (9.16)$$

Kerntheorie

Falls also O nicht gilt, ist entweder T , A oder R falsch. Der Satz von Zusatzhypothesen und Randbedingungen schützt also den harten Kern der Theorie wie eine Speckschicht. ² Man kann immer neue Antezedenzbedingungen dazufügen und somit die Falsifikation der Theorie verhindern. Wenn jedoch ein signifikantes Resultat vorliegt, so wird der postulierte Zusammenhang

¹ $A \Rightarrow B \iff \bar{B} \Rightarrow \bar{A}$;

Popper (1984, S. 15f) schreibt hierzu: ...*Asymmetrie zwischen Verifizierbarkeit und Falsifizierbarkeit, die mit der logischen Form der allgemeinen Sätze zusammenhängt; diese sind nämlich nie aus besonderen Sätzen ableitbar, können aber mit besonderen Sätzen in Widerspruch stehen. Durch rein deduktive Schlüsse (mit Hilfe des sogenannten 'modus tollens' der klassischen Logik) kann man daher von besonderen Sätzen auf die 'Falschheit' allgemeiner Sätze schließen...*

Es sind ja immer gewisse Auswege möglich, um einer Falsifikation zu entgehen, – etwa ad hoc eingeführte Hilfhypothesen oder ad hoc abgeänderte Definitionen...(S. 16)

²vgl. die Diskussion der Argumente von Lakatos, Feyerabend und Grünbaum in Meehl (1978). Siehe auch Grünbaum, 1976, Harding, 1976

- *O*: *Autoritärer Unterricht (X) hat negative Auswirkungen auf das Sozialverhalten von Kindern (Y).*

als bestätigt angesehen, wobei mit der Wahrscheinlichkeit $\alpha = .05, .01, ..$ eine Fehlentscheidung (H_1 angenommen, obwohl H_0 richtig) vorliegen kann. Die absolute Richtigkeit der Theorie ist also durch ein signifikantes Resultat nicht bewiesen. Der Rückschluß

$$O \Rightarrow T \quad (9.17)$$

bzw.

$$O \Rightarrow T \wedge A \wedge R \quad (9.18)$$

ist falsch.

Eine probabilistische Theorie wird in Abs. 14.4.2 diskutiert (vgl. Glg. 14.28).

Kapitel 10

Falsifikation und statistische Hypothesenprüfung

K. Popper ist der Auffassung, daß aus Theorien (T) abgeleitete Hypothesen (O), wenn sie sich nicht bestätigen, zu einer Falsifikation der Theorie führen ($\bar{O} \Rightarrow \bar{T}$). Bei einer Bestätigung von O hat das System seine Prüfung vorläufig bestanden, jedoch ist (durch den fehlenden Induktionsschluß) keine endgültige Bestätigung der Theorie möglich. Dagegen ist die Falsifikation schon durch ein Gegenbeispiel möglich. Dies ist bei All-Sätzen (alle Schwäne sind weiß) oder bei mathematischen Sätzen einleuchtend, wo durch ein widersprechendes Beispiel der Satz widerlegt ist (etwa *Alle stetigen Funktionen sind differenzierbar* wird durch das Gegenbeispiel $f(x) = \max(x, 0)$ widerlegt).

Gegenbeispiel

In der empirisch-statistischen Forschung sind jedoch immer Kollektive (Stichproben) beteiligt, so daß die Gültigkeit einer Hypothese 'im Durchschnitt' untersucht werden muß (etwa: 99% der Schwäne sind weiß).

Beispiel 10.1

- O : *Männer sind größer als Frauen.*

Betrachtet man etwa Frau Merkel und Herrn Lafontaine, so wäre die Behauptung falsifiziert.

Gemeint ist jedoch vermutlich etwas anderes:

- O : *Männer sind im Durchschnitt größer als Frauen.*

Hypothesen	Sachverhalt	Fehler
Falsifikation		
H_0	O	$\beta = P\{H_0 H_1\} = P\{O \bar{O}\} = P\{O \bar{T}\} = P\{V \bar{T}\}$
H_1	\bar{O}	$\alpha = P\{H_1 H_0\} = P\{\bar{O} O\} = P\{\bar{T} O\} = P\{F O\}$
Verifikation		
H_0	\bar{O}	$\beta = P\{H_0 H_1\} = P\{\bar{O} O\} = P\{\bar{T} O\} = P\{F O\}$
H_1	O	$\alpha = P\{H_1 H_0\} = P\{O \bar{O}\} = P\{O \bar{T}\} = P\{V \bar{T}\}$

Tabelle 10.1: Hypothesen, Fehler 1. und 2. Art bei Verifikation und Falsifikation (siehe Haupttext).



In der Sprache des Signifikanztests formuliert:

- $O = H_1: \mu_y > \mu_x;$
- $\bar{O} = H_0: \mu_y < \mu_x.$

Signifikanztest

Die zu verifizierende Hypothese (O) wird mit der H_1 identifiziert, damit bei Annahme von H_1 der Fehler 1. Art α klein gehalten werden kann (etwa 5%, 1%,...).

Bei Beibehaltung von $H_0 = \bar{O}$ wäre O falsifiziert, jedoch ist der Fehler 2. Art β i.a. nicht klein und kann bei Hypothesenmengen, die fließend ineinander übergehen, maximal $1 - \alpha = .95, .99, ..$ werden (vgl. etwa Bosch, 1992, S. 364). Dies ist im Beispiel der Fall:

$$H_0 : \mu := \mu_y - \mu_x < 0, \quad H_1 : \mu := \mu_y - \mu_x > 0.$$

Eine derartige Falsifikation ist nicht glaubwürdig, da die Wahrscheinlichkeit, eine Falsifikation zu folgern, obwohl $H_1 = O$ richtig ist, groß ist.

Ein Anhänger von Popper müßte die Hypothese O als H_0 formulieren, um sie (mit kleinem α) verwerfen zu können. Die unterschiedlichen Ziele der Falsifizierer und der Verifizierer sind in Tabelle 10.1 dargestellt.

Hierbei bedeutet

- $\alpha = P\{H_1|H_0\} = P\{H_0 \text{ verwerfen}|H_0 \text{ richtig}\},$
- $\beta = P\{H_0|H_1\} = P\{H_0 \text{ beibehalten}|H_1 \text{ richtig}\},$

Fehler 1. Art

(α -Fehler)

Fehler 2. Art

(β -Fehler)

da die **Hypothesen keine Ereignisse** (denen man Wahrscheinlichkeiten zuordnen kann), sondern **Untermengen des Parameterraums** sind, etwa $H_0 : \mu_y < \mu_x$.

Entsprechend ist $P\{O|\bar{T}\} = P\{V|\bar{T}\}$ eine Kurzschreibweise für die Wahrscheinlichkeit, O beizubehalten, obwohl T falsch ist, also O fälschlicherweise zu verifizieren (V).

In der Forschungspraxis wird man vorwiegend den Fall der Verifikation vorfinden, da O möglichst bestätigt werden soll. Negative Resultate werden oft nicht berichtet, da ja *nichts herausgekommen ist*.

Verifikation

Einerseits kann also eine Theorie nicht wirklich falsifiziert werden (Duhem-Quine-These; Glg. 9.15), andererseits führen auch mehrere Verifikationen nicht zu einem induktiven Beweis der Theorie im logischen Sinne (\rightarrow Verifikation, Falsifikation).

Es bleibt die vage Forderung, daß die Gesamt-Theorie mit den Daten konsistent gehalten werden muß.

Die Tatsache, daß eine Theorie im logischen Sinne nicht verifiziert werden kann, wird durch eine Äußerung des bekannten Astro-Physikers Harald Lesch treffend erläutert (Abb. 10.1).



Abbildung 10.1: Interview mit dem Astrophysiker Harald Lesch (Quelle: Lüdenscheider Nachrichten, 29.5.2009).

Kapitel 11

Hypothesenprüfung mit Prädiktionsregeln

11.1 Wahrheitstafeln

Umgangssprachlich formulierte Hypothesen wie

- *O: Männer sind größer als Frauen.*

sind unscharf, daher ist es besser, die Aussage möglichst präzise zu formulieren. Eine Methode besteht darin, die verschiedenen Ausprägungen des Wenn-(Dann)-Teils in Form einer Kreuztabelle darzustellen (der Einfachheit halber wird zunächst nur nominales und ordinales Skalenniveau betrachtet). In Tabelle 11.1 ist der Zusammenhang zwischen Geschlecht und Körpergröße symbolisch angegeben. Im Sinne des Prädiktions-Konzepts wird jeder Ausprägung der UV (x) eine oder mehrere Ausprägungen der AV (y) zugeordnet. Diese Zuordnung ist in Tabelle 11.1 mit 1 codiert. Männern werden im Beispiel die Größen mittel oder groß zugeordnet. Die Zellen mit Eintrag 0 werden auch als Fehlerzellen bezeichnet, da sie bei Gültigkeit der Regel nicht besetzt sein dürfen. Die Aussage

Fehlerzellen

- *O: Wenn der Blutalkoholspiegel 0.5 Promille übersteigt (A), sinkt die Reaktionsfähigkeit (B),*

kann wie in Tabelle 11.2 als Wahrheitstafel dargestellt werden. Dabei

Wahrheitstafel

Männer sind größer als Frauen			
	klein	mittel	groß
männlich	0	1	1
weiblich	1	1	0

Tabelle 11.1: Kreuztabelle des Zusammenhangs der Variablen Geschlecht und Körpergröße.

Wenn A , dann B		
$A \Rightarrow B$	B	\bar{B}
A	1	0
\bar{A}	1	1

Tabelle 11.2: Kreuztabelle der Variablen $A =$ 'Blutalkoholspiegel > 0.5 Promille', $B =$ 'Verminderte Reaktion' (Wahrheitstafel der Implikation $A \Rightarrow B$; 1 = wahr, 0 = falsch).

Implikation

Äquivalenz

wurde die Aussage als Implikation interpretiert. Dies bedeutet, daß im Fall geringen Blutalkoholspiegels sowohl $B =$ 'Verminderte Reaktion' als auch $\bar{B} =$ 'Normale Reaktion' in Frage kommt. Sie könnte aber auch im Sinne einer Äquivalenz aufgefaßt werden (Tabelle 11.3). Dies bedeutet, daß im Fall hohen Alkoholspiegels $A =$ 'Blutalkoholspiegel > 0.5 Promille' nur $B =$ 'Verminderte Reaktion', bei geringem Blutalkoholspiegel \bar{A} nur $\bar{B} =$ 'Normale Reaktion' in Frage kommt. Eine weitere Möglichkeit wäre die logische Und-Verknüpfung $A \wedge B$ (Tab. 11.4).

ordinale Variablen

Bei ordinalen Variablen kann auch eine asymmetrische Relation zwischen A und B modelliert werden (vgl. Nagl, 1992, S.80).

Übereinstimmung von Beobachtern

Auch die Übereinstimmung von Beobachtern (Rater) etc. kann in Form einer Kreuztabelle beschrieben werden.

Liegen etwa 5 verschiedene Urteilmöglichkeiten vor:

Die Gruppenatmosphäre in der ersten Video-Aufzeichnung empfinde ich als

gespannt 1 2 3 4 5 gelöst

Genau dann, wenn A , dann B		
$A \Leftrightarrow B$	B	\bar{B}
A	1	0
\bar{A}	0	1

Tabelle 11.3: Kreuztabelle der Variablen $A =$ 'Blutalkoholspiegel > 0.5 Promille', $B =$ 'Verminderte Reaktion' (Wahrheitstafel der Äquivalenz $A \Leftrightarrow B$).

A und B		
$A \wedge B$	B	\bar{B}
A	1	0
\bar{A}	0	0

Tabelle 11.4: Kreuztabelle der Variablen $A =$ 'Blutalkoholspiegel > 0.5 Promille', $B =$ 'Verminderte Reaktion' (Wahrheitstafel der Und-Verknüpfung $A \wedge B$).

Frauen heiraten Männer mit höherer Ausbildung			
$X < Y$	y_1	y_2	y_3
x_1	0	1	1
x_2	0	0	1
x_3	0	0	0

Tabelle 11.5: Asymmetrische Relation zwischen $X =$ Ausbildung der Frau und $Y =$ Ausbildung des Manns. $\{x_1, x_2, x_3\} =$ Ausprägungen des Bildungsniveaus = {gering, mittel, hoch}.

ungefähre Übereinstimmung von Ratern					
$X \approx Y$	y_1	y_2	y_3	y_4	y_5
x_1	1	1	0	0	0
x_2	1	1	1	0	0
x_3	0	1	1	1	0
x_4	0	0	1	1	1
x_5	0	0	0	1	1

Tabelle 11.6: Übereinstimmung von Ratern mit den Urteilsausprägungen $\{x_1, x_2, x_3, x_4, x_5\}$ und $\{y_1, y_2, y_3, y_4, y_5\}$. Es wurde keine komplette Übereinstimmung zugrundegelegt.

so kann die Übereinstimmung zweier Rater bei verschiedenen Video-Sequenzen durch Tabelle (11.6) beschrieben werden. In Abweichung von der üblichen völligen Übereinstimmung wurde hier eine ungefähre Übereinstimmung (\pm eine Rating-Klasse) zugrundegelegt.

11.2 A priori-Regeln

Die so aufgestellten Hypothesen müssen natürlich mit empirischen Daten konfrontiert werden. Dies erfolgt bei a priori (aus inhaltlichen Theorien) abgeleiteten Hypothesen mit dem (schon erwähnten) Kappa-Koeffizienten, der sich als Quotient

apriori-Fehlermaße

Cohen κ

$$\kappa = \frac{F(-) - F(+)}{F(-)} = \frac{G(+) - G(-)}{1 - G(-)} \quad (11.1)$$

berechnen läßt, wobei $F(+)$ der Prozentsatz der Fehler (Kosten, Fehldiagnosen; Zellen mit 0) und $F(-)$ der Prozentsatz der Fehler bei Unabhängigkeit der Merkmale ist (vgl. Nagl, 1992, Kap. 3.2).

Kappa ist also die Differenz aus hypothesenkonformen Zellen $G(+)$ in den empirischen Daten und den hypothesenkonformen Zellen $G(-)$ bei Unabhängigkeit (Produkt der Randhäufigkeiten) bezogen auf die perfekte Hypothesenkonformität minus der zufälligen Hypothesenkonformität $1 - G(-)$.

Beispielsweise gibt es einen bestimmten Prozentsatz von zufälliger Übereinstimmung, auch wenn beide Rater völlig unabhängig voneinander urteilen. Dieser Prozentsatz wird bei Berechnung von Kappa bereits in Rechnung gestellt, so daß die Verbesserung der Prädiktion durch die Hypothese (Regel) gegen den Hintergrund zufälliger Regelkonformität gemessen wird:

$$f(y|x) = f(y) \text{ bei Unabhängigkeit} \iff f(y, x) = f(x) * f(y).$$

(gemeinsame relative Häufigkeiten sind das Produkt der Randhäufigkeiten).

Beispiel 11.1

*Wenn der Blutalkoholspiegel 0.5 Promille übersteigt (A),
sinkt die Reaktionsfähigkeit (B)*

Analyse der Implikations-Regel: $A \Rightarrow B$
mit fiktiven empirischen Daten:

absolute Häufigkeiten

	1	2	3
1	40	10	50
2	25	25	50
3	65	35	100

relative Häufigkeiten

	1	2	3
1	0.4	0.1	0.5
2	0.25	0.25	0.5
3	0.65	0.35	1.

Unabhaengigkeitstafel $f_{i.} * f_{.j}$

	1	2
1	0.325	0.175
2	0.325	0.175

bedingte relative Haeufigkeiten

	1	2		1	2
1	0.8	0.2	1	0.615385	0.285714
2	0.5	0.5	2	0.384615	0.714286

ChiQuadrat, Kontingenzkoeff., K^+
 9.89011 0.3 0.424264

Hypothese, Cohen Kappa
 1 2 0.428571
 1 1 0
 2 1 1

Mit den gleichen Daten gilt für die Äquivalenz $A \Leftrightarrow B$:

ChiQuadrat, Kontingenzkoeff., K^+
 9.89011 0.3 0.424264

Hypothese, Cohen Kappa
 1 2 0.3
 1 1 0
 2 0 1

Im 2. Fall ist κ kleiner, da die Zelle $f_{21} = .25$ bei der Äquivalenzregel eine Fehlerzelle ist, bei der Implikation dagegen nicht. Die Koeffizienten χ^2 und der Kontingenzkoeffizient sind in beiden Fällen gleich, da sie die Unabhängigkeit der Merkmale prüfen ■

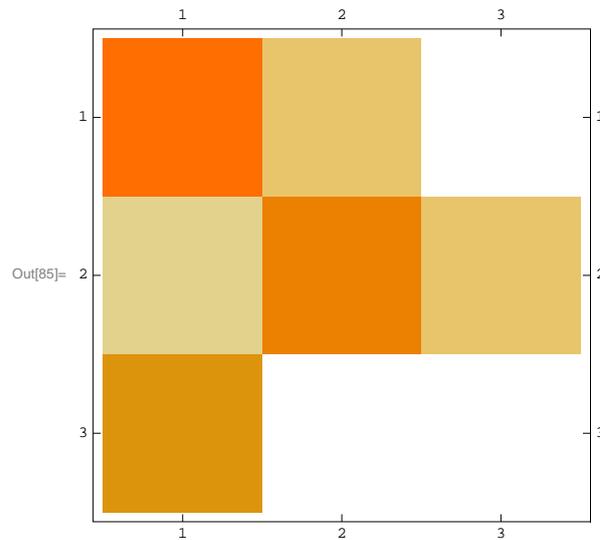


Abbildung 11.1: u-förmiger Zusammenhang der Variablen *Angst-Niveau* und *Leistung*. Graphische Darstellung der relativen Häufigkeiten f_{ij} als Farbmatrix (Temperaturen: hohe Werte = rot, geringe Werte gelb, negative Werte blau)

Beispiel 11.2

u-förmiger Zusammenhang zwischen Angst-Niveau (niedrig, mittel, hoch) und Leistung (niedrig, mittel, hoch) (Abb. 11.1).

relative Häufigkeiten

	1	2	3	4
1	0.3	0.1	0.	0.4
2	0.05	0.25	0.1	0.4
3	0.2	0.	0.	0.2
4	0.55	0.35	0.1	1.

Unabhängigkeitstafel $f_{i.} * f_{.j}$

	1	2	3
1	0.22	0.14	0.04
2	0.22	0.14	0.04
3	0.11	0.07	0.02

bedingte relative Haeufigkeiten

	1	2	3		1	2	3
1	0.75	0.25	0.	1	0.545455	0.285714	0.
2	0.125	0.625	0.25	2	0.0909091	0.714286	1.
3	1.	0.	0.	3	0.363636	0.	0.

ChiQuadrat, Kontingenzkoeff., K+
 55.1948 0.596363 0.730392

Hypothese, Cohen Kappa
 1 2 3 0.365079
 1 1 0 0
 2 0 0 1
 3 1 0 0

Hätte man die Hypothese in der Form

	1	2	3
1	1	0	0
2	0	1	1
3	1	0	0

gewählt, so würde sich ein höheres Cohen-Kappa = 0.694 ergeben, da mittlere Angst-Werte sowohl mittlere als auch hohe Leistungen bewirken. Eine solche Zuordnungsregel paßt also besser zu den Daten und wäre vorzuziehen (geringere Prädiktions-Fehleranzahl) ■

11.3 A posteriori-Regeln

Im letzten Abschnitt wurden a priori-Hypothesen anhand von Daten überprüft. Andererseits kann anhand von Daten auch eine Prädiktionsregel entwickelt werden (a posteriori).

a posteriori-

Fehlermaße

Betrachten wir etwa folgende Kreuztabelle

gemeinsame Häufigkeiten			
f_{ij}	y_1	y_2	$f_{i.}$
x_1	0.4	0.1	0.5
x_2	0.25	0.25	0.5
$f_{.j}$	0.65	0.35	1.0

der gemeinsamen Häufigkeiten von X und Y (Häufigkeiten der Modalwerte fett). Bei Vorliegen von x_1 (1. Zeile) ist y_1 die Ausprägung mit der größten Häufigkeit = 0.4 (Modalwert). Bei Vorliegen von x_2 (2. Zeile) ist y_1 oder y_2 die Ausprägung mit größten Häufigkeit = 0.25 (Modalwert). Ordnet man bei Vorliegen von x_i den bedingten Modalwert $y(x_i) = \text{mod}(Y|X = x_i)$ zu, so sind die **nicht** ausgewählten Ausprägungen **Fehlerzellen** (hier y_2 in der 1. Zeile sowie keine Zelle in der 2. Zeile). Bei Unkenntnis von X wird der Modalwert der Randhäufigkeiten $f_{.j}$, d.h. $y_1 = \text{mod}(Y)$ prädiziert und die anderen Ausprägungen y_2 sind Fehlerzellen.

Fehlerzellen

Die a posteriori-Regel (Regel = 1; Fehlerzellen = 0) lautet also

Modalregel		
	y_1	y_2
x_1	1	0
x_2	1	1

und die Fehler sind

- $F(+)$ = 0.1 = $\sum_{\text{Fehlerzellen}} f_{ij}$ (Prädiktions-Regel mit x)
- $F(-)$ = 0.35 = $\sum_{\text{Fehlerzellen}} f_{.j}$ (Prädiktions-Regel ohne x).

Das zugehörige Prädiktionsmaß $\lambda_{x \rightarrow y}$ ist dann

$$\lambda_{x \rightarrow y} = \frac{F(-) - F(+)}{F(-)} = \frac{0.35 - 0.1}{0.35} = 0.714. \quad (11.2)$$

Analog kann $\lambda_{y \rightarrow x}$ sowie $\lambda_{x \leftrightarrow y}$ gebildet werden. Man erhält also auf empirischem Weg die Implikations-Regel: $X \Rightarrow Y$.

Der Datensatz ist identisch mit obigem Beispiel für die Regel

- *Wenn der Blutalkoholspiegel 0.5 Promille übersteigt (A), sinkt die Reaktionsfähigkeit (B).*

Dort ergab sich ein Kappa-Koeffizient von 0.429 für die a priori-Implikationsregel. Vergleicht man die dortigen Fehlersummen

- $F(+)$ = 0.1 = $\sum_{\text{Fehlerzellen}} f_{ij}$ (Prädiktions-Regel mit x)

- $F(-) = 0.175 = \sum \text{Fehlerzellen } f_{i \cdot} * f_{\cdot j}$ (Prädiktions-Regel ohne x; Unabhängigkeit).
- $\kappa = (F(-) - F(+))/F(-) = (0.175 - 0.1)/0.175 = 0.429$

so erkennt man, daß in diesem Beispiel bei Kappa eine geringere proportionale Fehlerreduktion stattfindet als bei der Lambda-Regel.

Glossar

Sammlung von Techniken und Methoden zur korrekten Durchführung der wissenschaftlichen Untersuchung menschlichen Verhaltens und gesellschaftlicher Phänomene.	empirische Sozialforschung
Interaktion zwischen Meßapparat und Untersuchungsobjekt, wobei die Veränderung in der Apparatur als Aussage über das Meßobjekt interpretiert wird.	Messung Meßprozeß
Form der empirischen Sozialforschung, bei der durch Vorstrukturierung des Untersuchungsgegenstands in Form von Hypothesen sowie einer Standardisierung der Erhebungssituation eine quantitative Auswertung ermöglicht wird. Hierbei können auch qualitative Variablen quantitativ ausgewertet werden (etwa: $X = \text{Unterricht ist autoritär: trifft zu} = 1/\text{trifft nicht zu} = 0$; Der Mittelwert von X ist der Prozent-Anteil autoritären Unterrichts). Quantitative Analyse setzt daher nicht unbedingt quantitative (metrische) oder komparative (ordinale) Variablen voraus. Entscheidend ist, daß jeder Untersuchungseinheit die gleichen Merkmale/Variablen zugeordnet werden können, damit eine Häufigkeitsverteilung berechnet werden kann. Auf jedem der Skalenniveaus Nominal, Ordinal, Metrisch (Intervall, Verhältnis) kann eine quantitative Analyse der Ausprägungen vorgenommen werden. Davon ist die Diskussion um die standardisierte vs. qualitativ-offene Perspektive bei der Datenerhebung und Interpretation zu unterscheiden (→ qualitative Methoden).	quantitative Methoden/ quantitatives Paradigma
Erkenntnistheoretischer Realismus, bei dem eine reale, tatsächliche Welt (Gegenstände, Ereignisse und ihre Beziehungen) unabhängig von ihrer Wahrnehmung/Messung durch einen Beobachter postuliert wird.	objektive Realität/ Realismus

Gesetz

Oft wird nicht klar zwischen Hypothesen, Gesetzen und Theorien unterschieden, jedoch handelt es sich im Unterschied zu Definitionen um allgemeine Aussagen über Zusammenhänge zwischen empirischen Sachverhalten. Hypothesen können als Wenn-Dann oder Je-Desto-Aussagen formuliert werden. Quantitative Hypothesen nehmen die Form einer mathematischen Funktion oder Relation an. Der Begriff Hypothese wird oft für noch unbestätigte Zusammenhänge benutzt, während Gesetze schon als bewährt angesehen werden. Mehrere Hypothesen und Gesetze werden zu einer Theorie zusammengefaßt.

Beispiel 11.3

Quanten-Hypothese (QH): Annahme von Max Planck, daß Energie nur in Portionen vorkommt)

Quanten-Theorie: ausgearbeitetes mathematisches System, das auch die QH als Spezialfall beinhaltet. ■

Experiment

Systematische Variation unabhängiger (X , UV) und abhängiger Variablen (Y , AV). Untersuchungsanordnung zur Kausalanalyse (Aufdeckung und Quantifizierung von Ursache-Wirkungs-Zusammenhängen). **Die Ausprägungen der UV werden vom Experimentator gesetzt** und liegen nicht nur faktisch vor. Entsprechend ist bei einem (klassischen) linearen Regressionsmodell

$$Y = a + bX + \epsilon$$

die UV X **deterministisch**. Dagegen liegen bei einer Erhebung (Survey-Studie; quasi-experimentelle (ex-post-facto-) Anordnung) die Ausprägungen und Verteilungen der Variablen faktisch vor und eine Kausalanalyse kann nur mit Hilfe von strukturellen Regressionsmodellen (Strukturgleichungsmodelle; Pfadmodelle) vorgenommen werden. **Die UV ist hier stochastisch**, ihre Verteilung ist vorgegeben, was bei der Analyse berücksichtigt werden muß (vgl. KE III).

Weiterhin ist jedoch zu beachten, daß eine mindestens zweimalige Messung (2-Wellen-Panel) vorgenommen werden muß, da eine reine Querschnittsanalyse nur eine korrelative (symmetrische) oder strukturelle Analyse (mit strukturellen Regressionsmodellen, etwa LISREL) erlaubt.

Echte Kausalanalysen setzen jedoch voraus, daß Vorgang Y auf Vorgang X zeitlich folgt.

Bei Feldexperimenten wird versucht, die relevanten Variablen auch im Feld (in der natürlichen Umweltsituation) systematisch zu manipulieren. Zwar wird die externe Validität erhöht, da die Untersuchung vor Ort stattfindet, jedoch sind die Möglichkeiten, Störeffekte zu kontrollieren oder zu randomisieren, geringer als im Labor (höhere interne Validität).

Bei der *deduktiv-nomologischen Erklärung* (Hempel-Oppenheim) wird mit Hilfe von Antezedenzbedingungen (A_1, \dots, A_n) und Gesetzen (G_1, \dots, G_m) ein Sachverhalt E erklärt, d.h. das Explanans (A, G) erklärt das Explanandum E . Die deterministische Gesetzesaussage hat also eine Wenn-Komponente und eine Dann-Komponente (bzw. Je-Desto-Komponenten).

Erklärung**Beispiel 11.4**

A : Antezedenzbedingung: In der BRD sind die Scheidungsgesetze freizügig.

G : Gesetz: Wenn in einem Land die Scheidungsgesetze freizügig sind, dann ist die Scheidungsquote hoch.

E : Die Scheidungsquote in der BRD ist hoch.



Eine *induktiv-statistische Erklärung* legt statt dessen probabilistische Gesetze zugrunde:

Beispiel 11.5

A : Antezedenzbedingung: X ist in der BRD geboren.

G : Gesetz: 5.5% aller lebend Geborenen in der BRD sind uneheliche Kinder.

E : X ist mit 5.5% Wahrscheinlichkeit unehelich geboren.



Analytisch- Nomologische Wissenschaft	Annahme der Existenz grundlegender Gesetzmäßigkeiten auch im Bereich des Sozialen. Diese Regelhaftigkeiten haben nomologischen Charakter, sind also in ihrem Geltungsanspruch weder räumlich noch zeitlich relativiert. Die Gegenstände der Welt stehen in geordneter Weise in Beziehung (Struktur), Ereignisfolgen laufen nach gleichbleibenden Regeln ab (Gesetze; griech.: <i>nomos</i>), für jedes Ereignis (Wirkung) gibt es eine Ursache (Kausalitätsprinzip). Die Tatsachen liegen bereits vor und müssen nur noch entdeckt werden. Die Erfahrungswissenschaften unterscheiden sich lediglich in ihrem Gegenstand, jedoch nicht in ihrem Vorgehen (Postulat der Einheitswissenschaft).
Positivismus	Philosophische Tradition, die eine Abkehr von der Metaphysik und ein Primat des Realen fordert.
Positivismusstreit	Im Positivismusstreit (Adorno, Popper, Dahrendorf, Habermas, Albert, u.a.) wurde von Vertretern der Frankfurter Schule (Adorno, Habermas) Kritik an der herrschaftsstabilisierenden Funktion der analytisch-nomologischen Methodologie geübt. Popper betont die strikte Anerkennung des Wertfreiheits-Postulats für den → Begründungszusammenhang.
Wertneutralität, Wertfreiheit	Subjektive Werte, Urteile und Vorlieben der Forscher dürfen im Forschungsprozeß (bei der Datenerhebung und Auswertung) keine Rolle spielen (Wertfreiheit im Begründungszusammenhang).
Begründungs- zusammenhang	Wissenschaftliche Aussagen über die Realität sollen objektiv sein und nicht durch subjektive Werthaltungen beeinflusst werden (Wertfreiheit im Begründungszusammenhang). Im Gegensatz dazu ist der Entdeckungszusammenhang (Auswahl von Forschungsfragen) und der Verwertungszusammenhang zwangsläufig Werturteilen unterworfen.
Standardisierung	Die Bedingungen der Datenerhebung müssen von Fall zu Fall gleich sein (Standardisierung der Meßsituation).
Intersubjektivität	Der Forschungsprozeß soll so dokumentiert werden, daß er von an-

deren Personen (Forschern etc.) nachvollzogen oder sogar wiederholt (repliziert) werden kann.

Wenn empirische Befunde (Beobachtungsaussagen) mit vorher formulierten Hypothesen in Einklang sind, spricht man von empirischer Bewährung (Verifikation); bei einem Widerspruch liegt eine Falsifikation (Widerlegung) der Hypothese vor. Sie muß dann ganz verworfen oder umformuliert werden (\rightarrow Exhaustion). Allerdings ist nach Popper die Induktionslogik, d.h. der Schluß von (auch noch so vielen) besonderen Sätzen (Beobachtungen, Experimenten) auf allgemeine Sätze (Hypothesen, Theorien) nicht zulässig:

Bekanntlich berechtigen uns auch noch so viele Beobachtungen von weißen Schwänen nicht zu dem Satz, daß alle Schwäne weiß sind.

(Popper, 1984, S. 3)

Im Gegensatz dazu wird eine deduktive Überprüfung von Theorien vertreten:

Aus dem System werden...empirisch möglichst leicht nachprüfbar...Folgerungen (Prognosen) deduziert...Fällt die Entscheidung positiv aus, werden die singulären Folgerungen anerkannt, verifiziert, so hat das System die Prüfung vorläufig bestanden; wir haben keinen Anlaß, es zu verwerfen. Fällt eine Entscheidung negativ aus, werden Folgerungen falsifiziert, so trifft ihre Falsifikation auch das System, aus dem sie deduziert wurden... Induktionslogische Elemente treten...nicht auf... Auch durch ihre verifizierten Folgerungen können Theorien niemals als wahr oder auch nur wahrscheinlich erwiesen werden. (S.8) ...Ein empirisch-wissenschaftliches System muß an der Erfahrung scheitern können (S. 15).

Die logisch unhaltbare Induktion wird durch vage Erläuterungen zur Entdeckung neuer Gesetze ersetzt, etwa wenn Einstein zitiert wird:

Höchste Aufgabe des Physikers ist also das Aufsuchen jener allgemeinsten...Gesetze, aus denen durch reine Deduktion das Weltbild zu gewinnen ist. Zu diesen...Gesetzen führt kein logischer Weg, sondern nur die auf Einfühlung in die Erfahrung sich stützende Intuition (Einstein, 1934).

**Verifikation,
Falsifikation**

Duhem-Quine-These

Beobachtungssätze lassen sich nur aus Satzsystemen ableiten, nicht aus einzelnen Hypothesen. Experimentelle Resultate müssen immer mit einer Theorie als ganzer konfrontiert werden, nicht nur mit einzelnen isolierten Hypothesen. Kerntheorie und Instrumententheorie (Meßtheorie, Beobachtungstheorie etc.) lassen sich nur als Einheit testen.

qualitative Methoden

Qualitative Sozialforschung schreibt bei der Datenerhebung den subjektiven Perspektiven der Beforschten einen entscheidenden Stellenwert zu. Dabei wird die Rollentrennung zwischen Forscher und Forschungsgegenstand nicht fallengelassen, jedoch sollen die Begriffe aus der Sicht der Betroffenen heraus entwickelt werden. Es handelt sich also um ein interpretatives Paradigma. Die erforschte Realität soll zutreffend gedeutet werden, ohne dem Forschungsobjekt ein Erhebungsschema, sowie einen Interpretations- und Bezugsrahmen überzustülpen. Datengewinnung ist als kommunikative Leistung anzusehen, bei der umgangssprachliche Äußerungen und Handlungen nur aus dem Gesamtkontext der Erhebungssituation heraus zu begreifen sind.

Beispielsweise gilt für konversationsanalytische Untersuchungen: *Interpretationshypothesen gewinnen ihre Gültigkeit nicht über statistische Operationen, sondern über den sorgfältig geführten Nachweis, daß die Deutung einen Sinnzusammenhang erfaßt, auf den die Interagierenden selbst hinorientiert sind und der von ihnen im Interaktionsfortgang reproduziert wird*¹

Exhaustion

Modifikation einer Theorie aufgrund einer Falsifikation aus ihr abgeleiteter Hypothesen. *Die Beibehaltung einer Theorie trotz abweichender Daten unter Rückgriff auf 'störende Bedingungen' ist das Exhaustions-Verfahren* (Holzkamp, 1972, S. 90).

¹Bergmann, 1987, S. 58

Teil III

**Computergestützte
Datenanalyse**

Vorbemerkung

Der immer stärker werdende Einsatz von Computern hat die Statistik und ihre Einsatzmöglichkeiten völlig verändert. Während bis ca. 1990 nur Großrechner zur Verfügung standen, laufen heute auf Desktop- und Notebook-Rechnern Programme, deren Speicherhunger ungeahnte Ausmaße angenommen hat. Auf der anderen Seite kann jeder Anwender, mit oder ohne Statistik-Kenntnisse, sofort ausgefeilte Analysen, Tabellen und Graphiken erzeugen.

Statistische Methoden sind heutzutage ein unverzichtbarer Teil empirisch-wissenschaftlicher Arbeit. Auch Unternehmensdaten werden zunehmend in statistisch aufbereiteter Form dargestellt. Daher sind statistische Kenntnisse ein wichtiges Qualifikationsmerkmal von Mitarbeitern.

Auf der anderen Seite hat die leichte Bedienbarkeit von Statistik-Programmen zu der Tendenz geführt, daß komplexe Verfahren und Algorithmen angewandt werden, ohne den statistischen Hintergrund genau zu verstehen. Das Computerprogramm fungiert dabei als eine 'black box', deren Resultate unkritisch übernommen und präsentiert werden.

Ziel dieser Kurseinheit ist, exemplarisch den mathematisch-statistischen Hintergrund der Verfahren zu erarbeiten und auf empirische Daten anzuwenden.

Kapitel 12

Fallstudien

12.1 Einführung

Statistik hat einen dubiosen Ruf. Sätze wie: *Traue nur der Statistik, die du selbst gefälscht hast* weisen daraufhin, daß bei der Datenaufbereitung ein erheblicher Interpretationsspielraum besteht. Schon die Wahl der Achse bei einem Aktienkurs-Verlauf (Chart) kann eine geringe oder eine sehr starke Kursschwankung suggerieren (Abb. 12.1). Die statistische Aufbereitung von Daten ist inzwischen in fast alle Lebensbereiche vorgedrungen, man denke nur an Sportübertragungen, wo die Zahl der gewonnenen Zweikämpfe etc. tabellarisch aufbereitet wird. Über den Sinn solcher Übungen soll hier nicht diskutiert werden, sie tragen jedenfalls zum Ruf der Statistiker als Fliegenbein- bzw. Erbsenzähler bei.

Ernsthafter ist Benutzung von Statistiken als Argument der politischen

Statistiken

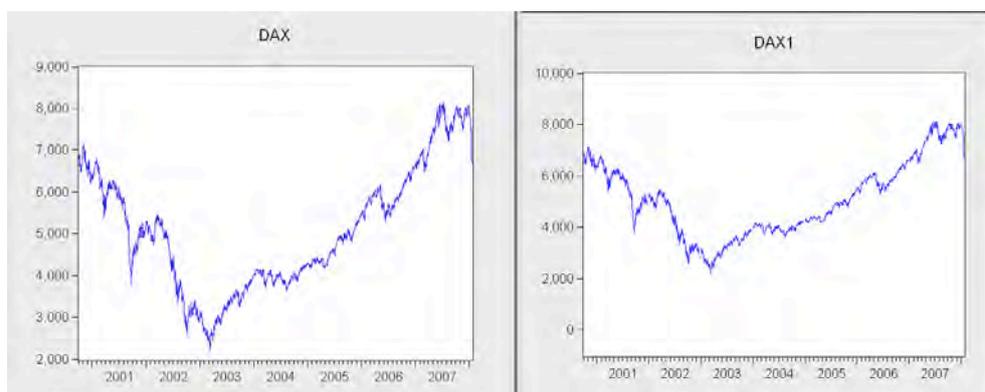


Abbildung 12.1: Dax-Verlauf mit unterschiedlicher Skalierung der y-Achse.

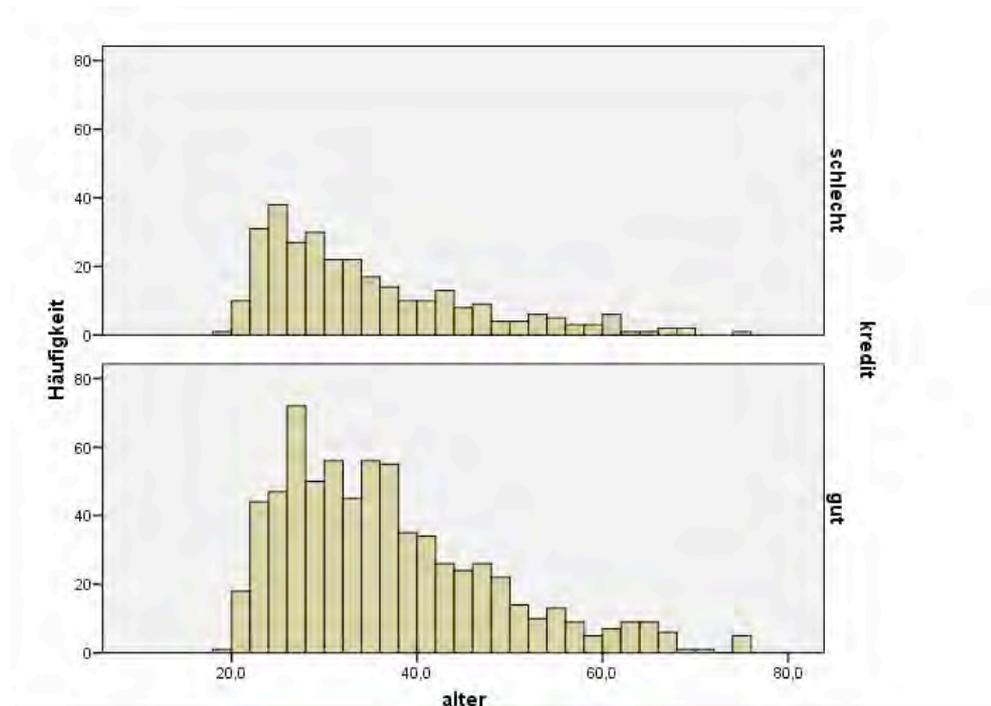


Abbildung 12.2: Altersverteilung in den Gruppen: schlechte und gute Kredite.

Auseinandersetzung, etwa in der Wahlberichterstattung (Politbarometer und ähnliche Sendungen). Auch Unternehmensdaten werden zunehmend in statistisch aufbereiteter Form präsentiert. Statistik kann jedoch auch direkt zum Erfolg eines Unternehmens beitragen.

Beispiel 12.1 (Credit-Scoring)

Credit-Scoring

Z.B. wird im Rahmen des sog. Credit-Scorings nach Merkmalen (etwa Alter, Bürge, Laufzeit etc.) gesucht, die sich in den Gruppen der guten bzw. schlechten Schuldner im Mittelwert bzw. ganz allgemein in ihrer Verteilung unterscheiden. Dann kann die Vergabe neuer Kredite an diesen Kriterien orientiert werden, um die Wahrscheinlichkeit zu minimieren, daß der Kredit nicht ordnungsgemäß zurückgezahlt wird (Abb. 12.2; Datensatz Kreditrisiko.LMU.sav).



	land	gdp	inflatio	arbkost	arbeitsl	arbprod	brutanl	nachfr	langfris	gdp2006	var	var
1	Australia	6,59	3,57	2,26	4,96	1,77	7,11	3,90	5,76	7,24		
2	Austria	4,08	1,68	,80	5,47	1,88	3,60	2,11	4,04	4,29		
3	Belgium	4,21	1,85	,76	8,27	1,19	4,23	2,52	4,04	4,29		
4	Canada	5,22	1,11	2,98	6,57	1,45	3,68	3,45	4,22	6,22		
5	Czech Republic	8,32	3,06	2,21	6,80	3,70	5,22	3,61	4,58	7,15		
6	Denmark	5,36	2,27	2,27	3,27	1,61	5,76	2,90	3,94	5,08		
7	Finland	3,77	1,26	1,47	7,57	2,33	2,80	2,28	4,06	3,88		
8	France	3,88	1,71	1,23	8,51	1,19	3,26	2,30	3,97	3,72		
9	Germany	3,18	1,66	-,26	7,65	1,15	3,40	1,04	3,95	2,24		
10	Greece	7,20	3,13	3,69	9,24	2,34	4,56	3,44	4,13	7,25		
11	Hungary	7,30	4,99	1,37	7,74	2,40	7,03	4,03	6,85	6,16		
12	Iceland	5,71	3,69	6,11	2,53	-,78	-14,67	-2,72	12,53	8,44		
13	Ireland	8,54	3,00	2,53	4,40	2,39	5,03	5,54	4,04	8,41		
14	Italy	3,54	1,92	1,71	6,80	,80	3,10	1,40	4,34	2,79		
15	Japan	2,80	-,16	-,29	3,86	1,93	1,34	1,43	2,13	2,14		
16	Korea	6,33	1,05	1,27	3,38	3,56	3,80	4,09	5,27	6,26		
17	Luxembourg	6,51	3,11	1,69	4,74	,71	3,05	2,62	3,54	8,80		
18	Mexico	6,92	3,83	2,71	3,40	1,04	6,25	4,18	7,15	9,71		
19	Netherlands	4,13	,80	,52	4,19	1,49	3,13	1,69	4,01	4,18		
20	New Zealand	4,29	2,89	1,94	4,37	1,70	-2,69	,88	5,56	3,27		
21	Norway	5,69	2,01	3,65	3,28	2,19	2,64	2,68	5,48	10,90		
22	Poland	6,45	2,23	2,39	12,61	2,99	8,50	4,59	-9,00	5,96		
23	Portugal	3,36	1,84	2,17	7,43	,57	2,84	1,66	4,14	3,59		
24	Slovak Republic	8,65	3,53	1,48	12,23	5,86	8,02	5,10	4,86	8,52		
25	Spain	6,78	3,55	2,95	7,83	,65	4,98	3,92	3,95	7,17		
26	Sweden	5,26	2,11	1,73	5,31	1,97	4,24	2,47	4,53	5,57		
27	Switzerland	3,22	,97	1,45	3,55	1,34	3,11	1,74	2,88	3,19		
28	Turkey	10,43	7,73	1,52	9,77	3,30	13,50	6,88	18,85	11,56		
29	United Kingdom	4,54	2,49	2,54	5,70	1,48	5,19	2,80	4,67	4,55		
30	United States	5,48	2,59	2,61	4,79	1,69	3,90	3,19	4,78	6,71		
31	Euro area	4,18	1,95	1,06	7,44	1,00	3,67	2,05	4,05	3,79		
32	Total OECD	4,94	2,16	1,82	5,83	1,61	3,97	2,79	-9,00	5,35		

Abbildung 12.3: SPSS-Datensatz (Regression.sav) der OECD-Makrovariablen (2007): Datenansicht.

12.2 Makroökonomische Daten

Beispiel 12.2 (Makroökonomische Daten)

Anhand eines makroökonomischen Datensatzes (Regression.sav, Regression.jmp) soll die Verwendung statistischer Methoden zunächst illustriert werden. Es handelt sich um Daten aus dem Jahr 2007, z.B. Arbeitslosenraten in verschiedenen Ländern (Abb. 12.3). Die einzelnen Variablen werden in Abb. (12.4) erläutert. Es handelt sich um die sog. Variablenansicht, in der die Charakteristika der Variablen, z.B. das Skalenniveau (Spalte rechts aussen) eingetragen sind. Bei der Variable *Land* handelt es sich um eine sog. nominale Variable, bei der nur die Bezeichnung unterschieden werden kann. Dagegen sind die anderen Größen metrischer Natur, d.h. sie können quantitativ interpretiert und ausgewertet werden (etwa Durchschnittswerte etc.). Eine Betrachtung der Daten ist zwar interessant, jedoch nicht sehr aussagekräftig. Einen besseren Eindruck vermitteln Graphiken, welche die Verteilung der einzelnen Variablen (Spalten), etwa GDP-Änderung, Arbeitslosenrate etc. darstellen (Abb. 12.5–12.6). Diese werden als Histogramme bezeichnet. Dabei wird die relative Häufigkeit (Prozentwerte), daß die Daten in bestimm-

Variablenansicht

nominale Variable

metrische Variable

Histogramme

	Name	Typ	Spaltenformat	Dezimalstellen	Variablenlabel	Werteformat	Fehlende Werte	Spalten	Ausrichtung	Maßniveau
1	land	String	15	2	Land	Kein	Kein	15	Links	Nominal
2	gdp	Numerisch	11	2	Nominales GDP-Wachstum in % 2007	Kein	Kein	8	Rechts	Metrisch
3	inflato	Numerisch	11	2	Inflation in % 2007	Kein	Kein	8	Rechts	Metrisch
4	arbkost	Numerisch	11	2	Kostenänderung einer Arbeitseinheit in % 2007	Kein	Kein	8	Rechts	Metrisch
5	arbeitsl	Numerisch	11	2	Arbeitslosenrate in % der Arbeitsbevölkerung 2007	Kein	Kein	8	Rechts	Metrisch
6	arbprod	Numerisch	11	2	Änderung der Arbeitsproduktivität in % 2007	Kein	Kein	8	Rechts	Metrisch
7	brutand	Numerisch	11	2	Änderung des realen Brutto Anlagevermögens in % 2007	Kein	Kein	8	Rechts	Metrisch
8	nachf	Numerisch	11	2	Änderung der realen aggregierten Nachfrage in % 2007	Kein	Kein	8	Rechts	Metrisch
9	langfrs	Numerisch	11	2	Langfristiger Zinssatz in % 2007	Kein	9,00	8	Rechts	Metrisch
10	gdp2006	Numerisch	8	2	Nominales GDP-Wachstum in % 2006	Kein	Kein	8	Rechts	Metrisch

Abbildung 12.4: SPSS-Datensatz (Regression.sav) der OECD-Makrovariablen (2007): Variablenansicht.

te Klassen fallen, als Säulendiagramm gezeichnet. In den Graphiken ist auch der Mittelwert, die Streuung (Standardabweichung) sowie die Stichprobengröße ($N = 32$) angegeben. Weiterhin ist in den Histogrammen eine theoretische Verteilungskurve, die sog. Gaußsche Glockenkurve oder Normalverteilung, eingezeichnet.

Normalverteilung

Box-Plot

Eine gute Übersicht der Verteilung mehrerer Variablen liefert der sog. Box-Plot (Abb. 12.7), bei dem der Median, das 25% sowie das 75% - Quantil und der minimal- bzw. der maximale Wert eingezeichnet sind.

bivariate Analyse

In der bisherigen Darstellung wurden die Variablen einzeln betrachtet (univariate Analyse). Es ist jedoch auch von großem Interesse, wie zwei und mehrere Variablen miteinander zusammenhängen. Die einfachste Möglichkeit ist, die Datenpaare von 2 Variablen als Punkte in einem Streudiagramm zu zeichnen (bivariate Analyse). In Abb. 12.8 ist das $x - y$ -Streudiagramm zu sehen, wobei jeder Datenpunkt zu einem Land gehört. Offenbar ist ein positiver Zusammenhang zwischen den Variablen vorhanden, da hohe (niedrige) GDP-Werte tendenziell mit hohen (niedrigen) Inflationswerten einhergehen. Man spricht in diesem Zusammenhang von einer positiven Korrelation.

Korrelation

Dieser Zusammenhang kann jedoch nicht als Kausalwirkung interpretiert werden (etwa daß ein hohes GDP-Wachstum eine hohe Inflation erzeugt). Dazu wären weitere Untersuchungen notwendig, etwa die Iden-

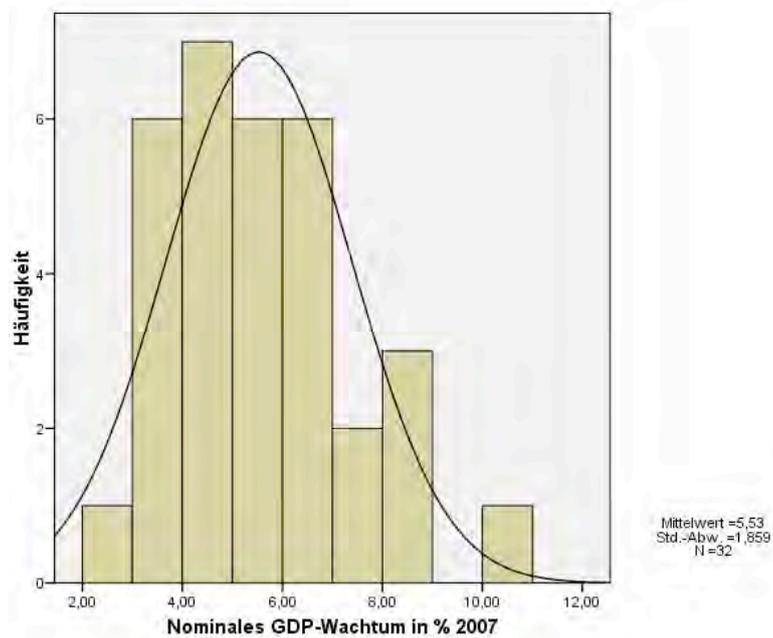


Abbildung 12.5: Histogramm des Nettoinlandsprodukts der OECD-Länder.

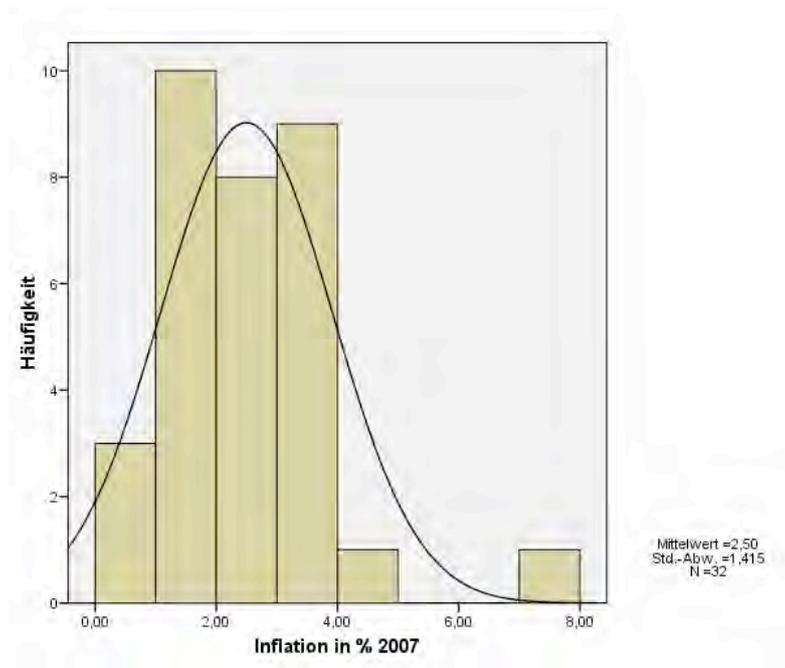


Abbildung 12.6: Histogramm der Inflationsrate der OECD-Länder.

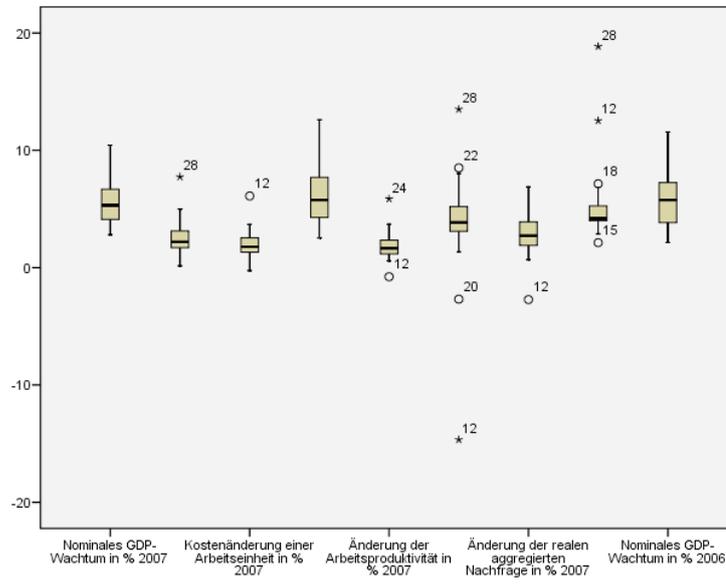


Abbildung 12.7: Box-Plots der OECD-Daten.

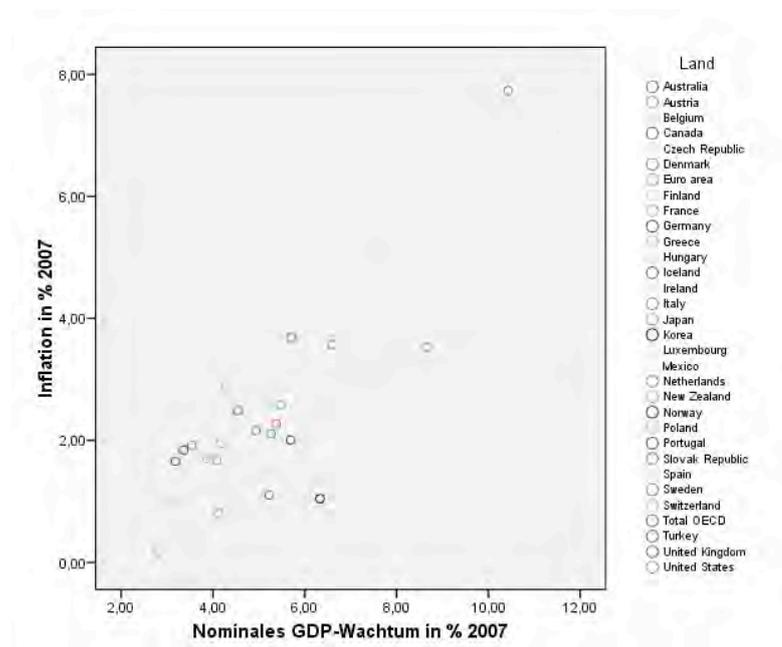


Abbildung 12.8: Streudiagramm der Variablen GDP und Inflationsrate.

tifikation weiterer Variablen, die den Zusammenhang bewirken (Problem der Scheinkorrelation). Etwa weisen Variablen, die einen gleichzeitigen positiven Trend aufweisen, eine hohe Korrelation auf, die jedoch nur durch die simultane Bewegung bedingt ist. Dieser Sachverhalt wird in der Ökonometrie durch den Begriff der Kointegration thematisiert.

Scheinkorrelation

Ein wichtiger Gesichtspunkt bei der Einteilung der statistischen Verfahren ist die Symmetrie/Asymmetrie des Modells. Etwa kann zwischen abhängigen und unabhängigen Variablen (AV/UV) unterschieden werden. In Anlehnung an die übliche mathematische Konvention werden diese als Y -Variablen (links) und X -Variablen (rechts) bezeichnet. Ein asymmetrisches Modell versucht, die Y -Werte aus den X -Werten vorherzusagen, etwa als Funktion

Kointegration**Symmetrie,
Asymmetrie**

$$Y_n = f(X_n) \quad (12.1)$$

während ein symmetrisches Modell nur die Kovariation (Korrelation) der Werte (X_n, Y_n) untersucht. Beispielsweise ist ein Streudiagramm (Abb. 12.9) eine symmetrische Darstellung, während eine geschätzte Kurve (Abb. 12.10) eine asymmetrische Darstellung ist.

Meistens werden in der Ökonometrie (asymmetrische) Regressionsmodelle der Form

$$Y_n = f(X_n) + \epsilon_n, \quad (12.2)$$

Regressionsmodelle

etwa das lineare Modell

lineares Modell

$$Y_n = \alpha + \beta X_n + \epsilon_n, \quad (12.3)$$

spezifiziert und geschätzt (d.h. Schätzungen $\hat{\alpha}$ und $\hat{\beta}$ für α und β berechnet). Hierbei ist ϵ_n ein Gleichungsfehler, der die Abweichung der Punkte von der Kurve beschreibt. Ein rein deterministisches Modell (etwa eine Gerade $Y_n = \alpha + \beta X_n$) ist meist nicht ausreichend.

Die Rolle von X ist die eines Prädiktors für die Variable Y (d.h. es wird eine Kausalwirkung $X \rightarrow Y$ unterstellt). Allerdings ist hierbei zu beachten, daß auch eine inverse Kurve $X = g(Y)$ aus den selben Daten geschätzt werden kann. Insbesondere lassen sich aus Daten, die nur korrelativ zusammenhängen (kovariieren), asymmetrische Modelle schätzen (etwa $X = \text{Zahl der Störche}$, $Y = \text{Zahl der Kinder}$).

Prädiktor

Abb. 12.10 zeigt auch, daß eine beliebige Zahl verschiedener Kurven geschätzt werden kann. Das lineare Modell ist nur eine einfache Näherung (Approximation) für gekrümmte Kurven, d.h.

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) \quad (12.4)$$

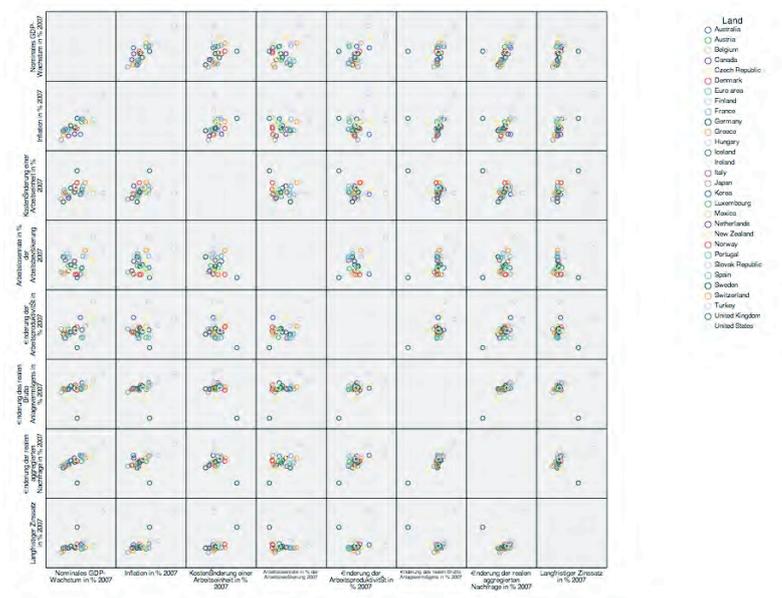


Abbildung 12.9: Streudiagramm verschiedener makroökonomischer Variablen aus dem OECD-Bereich.

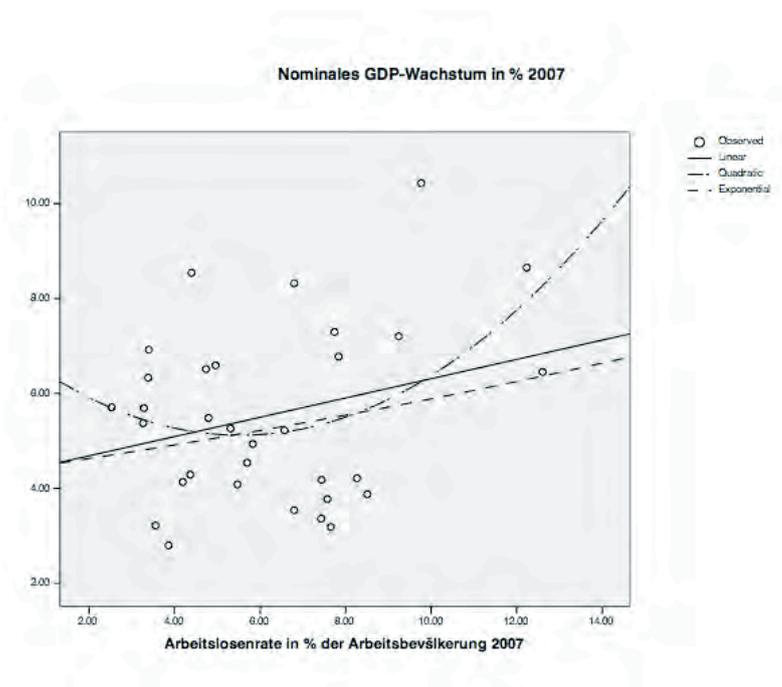


Abbildung 12.10: Geschätzte Kurven (linear, quadratisch, exponentiell) für Länder aus dem OECD-Bereich.

($f'(x_0)$ ist die Ableitung bei x_0). Im allgemeinen hat man mehrere erklärende Variablen X_1, X_2, \dots, X_q , sodaß eine mehrfaches (multiples) Regressionsmodell in Frage kommt:

$$Y_n = \alpha + \beta_1 X_{n1} + \beta_2 X_{n2} + \epsilon_n, \quad (12.5)$$

hier in linearer Form für $q = 2$ Variablen.

Von großem Interesse ist auch, ob sich die Variablen im Zeitablauf ändern, etwa GDP_t , die prozentualen GDP-Änderungen (nominales GDP-Wachstum in %) im Jahr t . Sie sind ähnlich, wie Abb. 12.11 zeigt (Streudiagramm GDP_{2006}, GDP_{2007}).

Ob sich eine signifikante Mittelwerts-Veränderung eingestellt hat, kann durch einen t -Test geprüft werden (Abb. 12.12). Es ist sehr wichtig, daß hier die Abhängigkeit der Daten (die Werte aus den beiden Jahren sind autokorreliert) berücksichtigt wird. Dies wird durch einen t -Test für gepaarte Stichproben erreicht. In diesem Fall wird die Differenz der Werte

$$D = GDP_{2007} - GDP_{2006} = -.36 \quad (12.6)$$

auf Null geprüft, wodurch die Abhängigkeit der Daten erfaßt wird. Bei einem Standardfehler der Mittelwertsdifferenz von 1.32 ergibt sich ein t -Wert von -1.535 , sodaß erst bei einem alpha-Niveau oberhalb von $p = .135$ (p -Wert) ein signifikanter Unterschied bestünde.

Abhängige Stichproben treten bei sog. Paneldaten auf, wo die gleiche Variable im Zeitablauf untersucht wird (d.h. $X_{nt}, n = 1, \dots, N, t = 1, \dots, T$, etwa GDP_{nt} für Land n zum Zeitpunkt t). Ähnlich wie bei Zeitreihen X_t muß hier die temporale Abhängigkeit der Daten berücksichtigt werden.

■

12.3 Aktienkurs-Renditen

Beispiel 12.3 (Rendite-Korrelation mehrerer Aktien)

Kursbewegungen von Aktien sind häufig gleichsinnig, da makroökonomische Einflüsse auf alle Papiere ähnlich wirken, etwa die US-Wirtschaft, Zinsänderungen, neue Konjunkturdaten, politische Veränderungen etc. Andererseits sind auch spezifische Einflüsse vorhanden, etwa Führungsprobleme in einer Firma, Absatzprobleme einer Branche etc.

**multiples
Regressionsmodell**

**t -Test für gepaarte
Stichproben**

Paneldaten

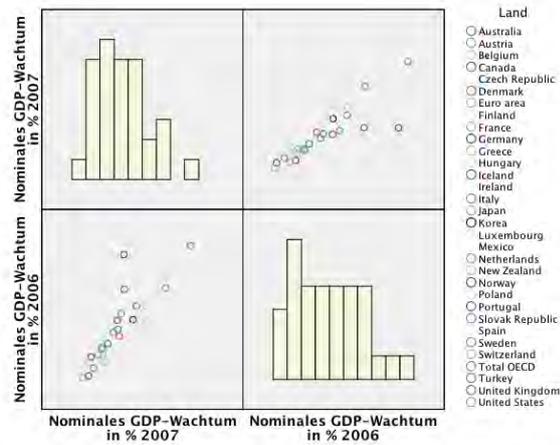


Abbildung 12.11: GDP-Änderungen der Jahre 2006 und 2007 für Länder aus dem OECD-Bereich: Histogramme und Streudiagramme.

Statistik bei gepaarten Stichproben

		Mittelwert	N	Standardabweichung	Standardfehler des Mittelwertes
Paaren 1	Nominales GDP-Wachstum in % 2007	5,5283	32	1,85928	,32868
	Nominales GDP-Wachstum in % 2006	5,8877	32	2,46622	,43597

Korrelationen bei gepaarten Stichproben

	N	Korrelation	Signifikanz
Paaren 1	32	,849	,000

Test bei gepaarten Stichproben

		Gepaarte Differenzen				T	df	Sig. (2-seitig)	
		Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz				
					Untere				Obere
Paaren 1	Nominales GDP-Wachstum in % 2007 - Nominales GDP-Wachstum in % 2006	-,35944	1,32463	,23416	-,83702	,11814	-1,535	31	,135

Abbildung 12.12: GDP-Änderungen der Jahre 2006 und 2007 für Länder aus dem OECD-Bereich. t-Test für gepaarte (abhängige) Stichproben.

Die Zusammenhänge zwischen den Aktienkurs-Renditen, d.h. den prozentualen Änderungen der Kurse werden im folgenden dargestellt. Formal ist die (einfache) Rendite als

$$X_t = \frac{S_t - S_{t-1}}{S_{t-1}} \approx \log \frac{S_t}{S_{t-1}} \quad (\text{log-Rendite}) \quad (12.7)$$

definiert ¹, wobei S_t der Kurs zur Zeit $t = 1, \dots, T$ ist. Im folgenden wird als Index für die Beobachtungen $t \rightarrow n$ und als Stichprobengröße $T \rightarrow N$ benutzt (dies ist die übliche Konvention in der multivariaten Statistik).



12.4 Graphische Zusammenhänge zwischen Variablen

Im Gegensatz zur *univariaten* Statistik, bei der gleichzeitig nur *eine* Variable betrachtet wird (etwa ihre Verteilung, Mittelwert, Standardabweichung etc.), werden in der *multivariaten* Statistik mehrere Variablen *simultan* untersucht. Diese Betrachtungsweise führt über die bloße Wiederholung univariater Analysen hinaus, da auch die Zusammenhänge *zwischen* den Variablen analysiert werden. Im allgemeinen werden zuerst die Zusammenhänge graphisch dargestellt, wobei das sog. Streudiagramm (vgl. Abb. 12.13) besonders einfach zu interpretieren ist. Man trägt die einzelnen Datenpunkte ($N =$ Stichprobengröße)

Streudiagramm

$$(x_n, y_n), n = 1, \dots, N \quad (12.8)$$

in ein $x - y$ -Koordinatensystem ein, wobei dann, ähnlich wie auf einer Schießscheibe, ein Muster zu sehen ist. Wenn die Variablen einen deterministischen Zusammenhang aufweisen, etwa $y = a + bx$ oder $y = x^2$, würde man Muster sehen, die den üblichen Funktionsdarstellungen in der Mathematik ähneln. Bei empirischen Daten sind jedoch immer zufällige Einflußgrößen überlagert, die den Zusammenhang maskieren (etwa Rauschen bei Fernsehbildern, Meßfehler bei der Datenerhebung) oder es fehlt eine systematischer Zusammenhang.

Im allgemeinen hat man nicht nur 2, sondern p Variablen, sodaß die Zusammenhänge zwischen

$$(x_{n1}, x_{n2}, \dots, x_{np}), n = 1, \dots, N \quad (12.9)$$

¹Für $\log(x)$ gilt die Taylor-Entwicklung $(x - 1) + O((x - 1)^2)$

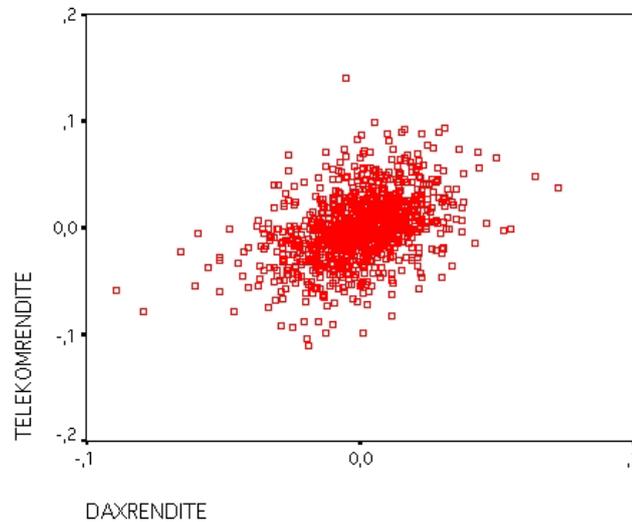


Abbildung 12.13: Streudiagramm der Renditen von Dax und Telekom. Datensatz **renditen.sav**

Matrix- Streudiagramm

dargestellt werden müssen, am besten in Form einer Tabelle von Streudiagrammen, wobei in der i -ten Zeile und der j -ten Spalte der Graphik das Streudiagramm zwischen den Variablen X_i und X_j mit den Meßwerten (Realisationen)

$$(x_{ni}, x_{nj}), n = 1, \dots, N \quad (12.10)$$

aufgetragen ist (Abb. 12.14). Man kann auch versuchen, den 4-dimensionalen Datensatz anders zu visualisieren, etwa in Form einer 3-dimensionalen Punktwolke, die perspektivisch in 2 Dimensionen abgebildet ist (Abb. 12.15). Inwieweit dies zu einem verbesserten Verständnis für den Datensatz führt, mag dahingestellt sein. Programme wie SAS/JMP bieten die Möglichkeit, eine interaktive Rotation der Daten-Wolke durchzuführen (Abb. 12.16).

Es ist sinnvoll, die Daten in Form einer Tabelle aufzulisten, wobei üblicherweise jede Zeile einer Untersuchungseinheit (Person, Firma, Zeitpunkt etc.) entspricht. Die Daten $x_{ni}, n = 1, \dots, N, i = 1, \dots, p$ werden

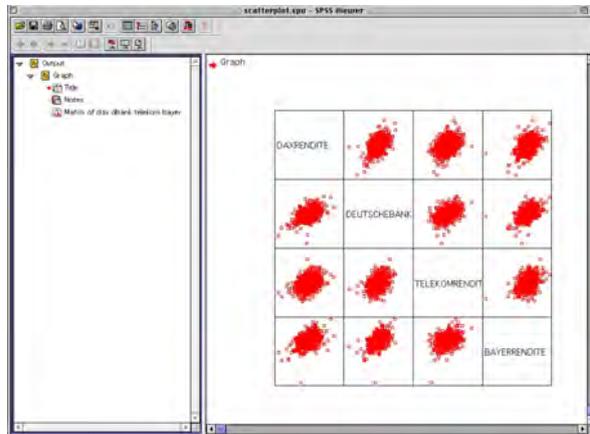


Abbildung 12.14: Matrix der Streudiagramme der Renditen von: Dax, Telekom, Deutsche Bank und Bayer.

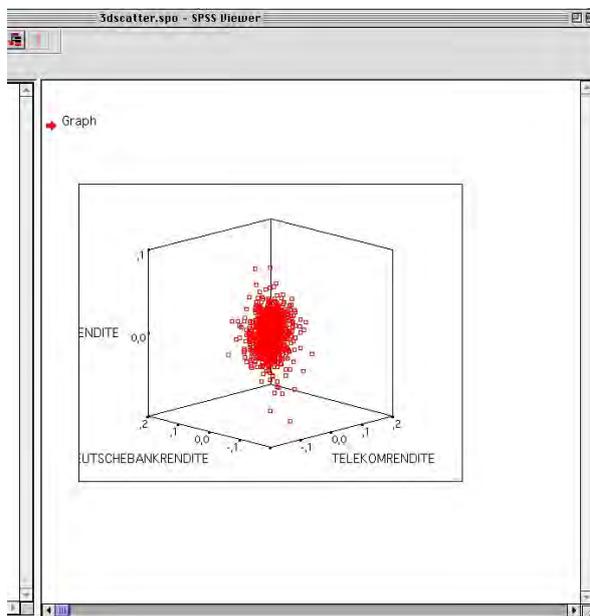


Abbildung 12.15: 3-D Streudiagramm der Renditen von: Dax, Telekom und Deutsche Bank (SPSS).

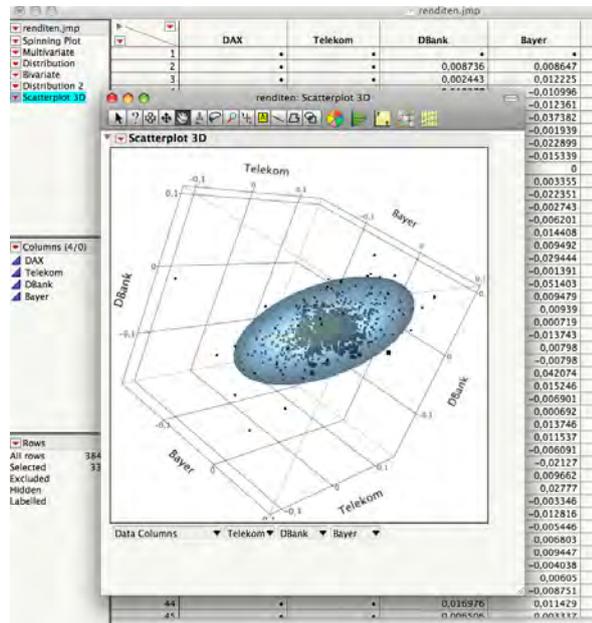


Abbildung 12.16: Spinning plot: 3-D Streudiagramm der Renditen von: Dax, Telekom und Deutsche Bank (SAS/JMP).

dann als Daten-Matrix (Tabelle)

Daten-Matrix

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \dots & \dots & \dots & \dots \\ x_{N1} & x_{N2} & \dots & x_{Np} \end{bmatrix} : N \times p \quad (12.11)$$

dargestellt. Dies ist auch die Form der Speicherung in Dateien (englisch: files) für Statistikprogramme und Tabellenkalkulationen.

Häufig werden den Spalten noch Variablennamen ('Preis', 'Kosten für Werbung', ..., 'Zahl der Vertreterbesuche') oder ('Dax', 'Telekom', 'Deutsche Bank', 'Bayer') und den Zeilen Namen für die statistischen Einheiten (Filiale 1, ..., Filiale N) oder (Zeitpunkt 1, ..., Zeitpunkt N) zugewiesen. Ein SPSS-Datensatz ist in Abb. 12.17 zu sehen.

Weiterhin können die Daten in Form von univariaten Histogrammen, Quantil-Graphiken etc. dargestellt werden (Abb. 12.18–12.19).

Dichteschätzungen

Der bivariate Zusammenhang läßt sich mit Hilfe von 2-dimensionalen Dichteschätzungen $\hat{f}(x, y)$ abbilden. Diese können als Histogramm oder als Kerndichte-Schätzung berechnet werden. Abb. (12.20, oben) zeigt die Daten und die Dichteschätzung in Form von Höhenlinien (Landkarte) oder als räumliche Fläche (unten).

	dax	telekom	dbank	bayer	MBF	MBF	MBF
3822	,0035	-,0011	,0120	,0173			
3823	-,0182	-,0287	-,0299	,0156			
3824	-,0007	,0153	,0026	,0084			
3825	-,0272	-,0490	-,0248	-,0098			
3826	,0050	,0035	,0026	-,0085			
3827	-,0183	-,0305	-,0159	,0286			
3828	,0221	-,0006	,0159	,0132			
3829	-,0312	-,0365	-,0267	-,0193			
3830	-,0348	-,0197	-,0435	-,0168			
3831	-,0301	-,0453	-,0604	-,0114			
3832	-,0129	,0213	-,0258	-,0057			
3833	-,0887	-,0582	-,1235	-,0531			
3834	,0143	,0523	,0215	-,0091			
3835	,0131	-,0099	,0393	-,0374			
3836	-,0650	-,0218	-,0608	-,0405			
3837	,0284	,0458	,0198	,0281			
3838	-,0094	,0209	-,0463	-,0195			
3839	-,0372	,0075	-,1005	-,0388			
3840	-,0591	-,0050	-,0834	-,0872			
3841	-,0059	-,0008	,0665	-,0075			
3842	,0643	,0488	,0967	,0929			
3843	-,0073	-,0096	,0172	-,0207			
3844	,0213	,0161	-,0036	,0138			
3845	,0215	-,0185	,0161	,0271			
3846	,0291	,0291	,0585	,0393			
3847	-,0160	-,0351	-,0462	-,0277			

Abbildung 12.17: Ausschnitt aus der Datenmatrix (SPSS-Datei renditen.sav) von: Dax, Telekom, Deutsche Bank und Bayer.

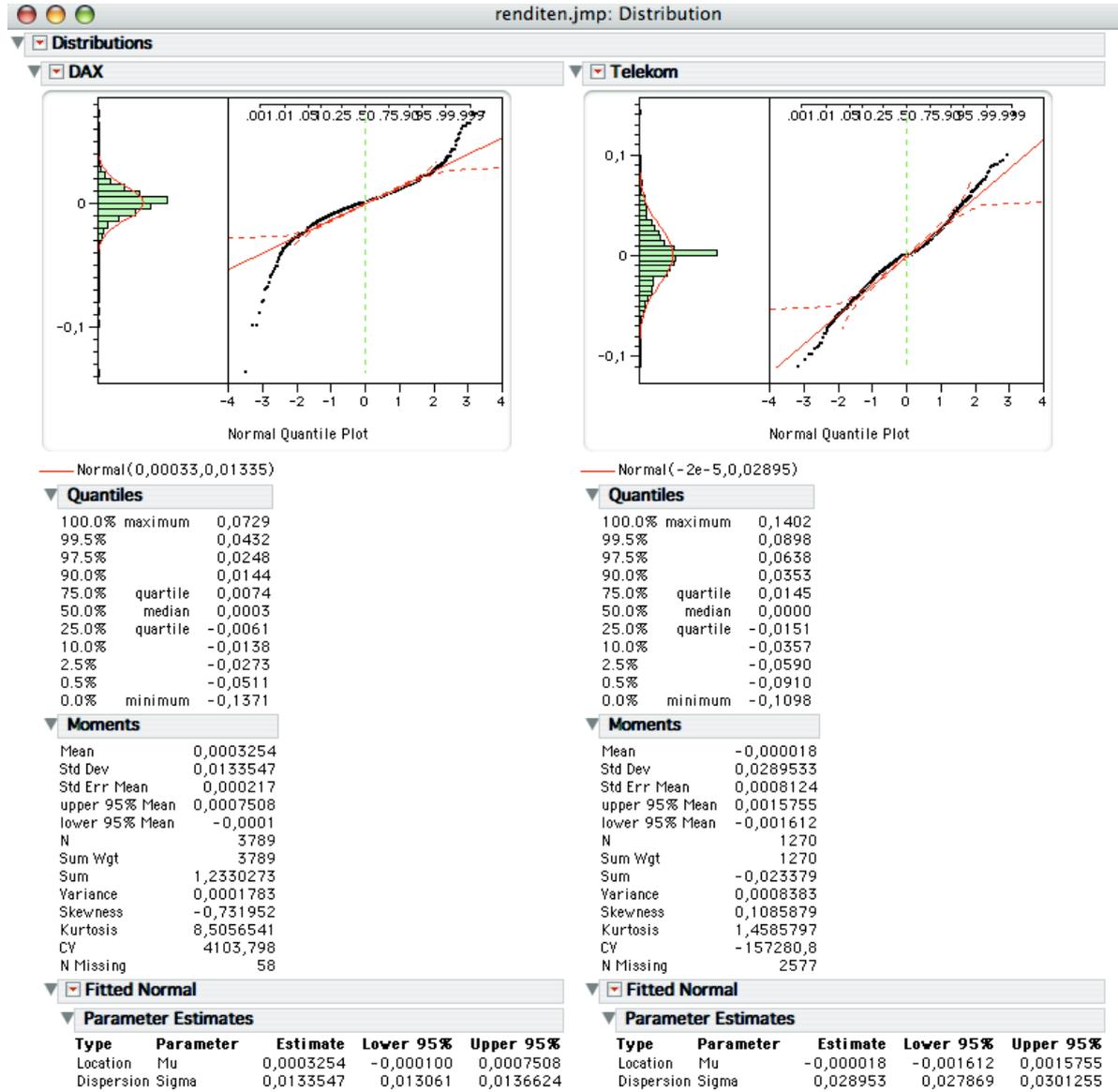


Abbildung 12.18: Histogramme, angepaßte Normalverteilung und univariate Statistiken von Dax und Telekom (SAS/JMP).

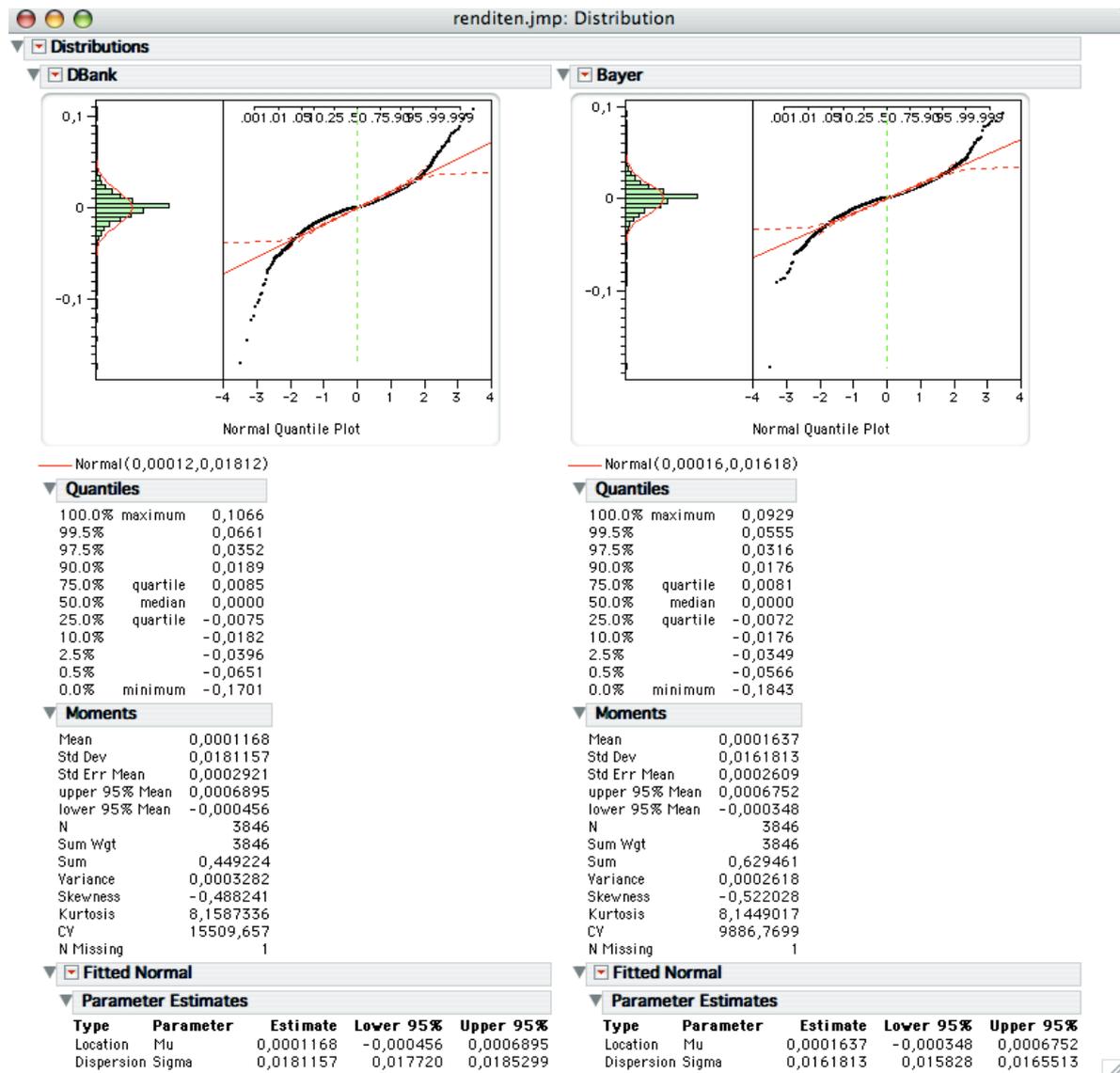


Abbildung 12.19: Histogramme, angepaßte Normalverteilung und univariate Statistiken von Deutscher Bank und Bayer (SAS/JMP).

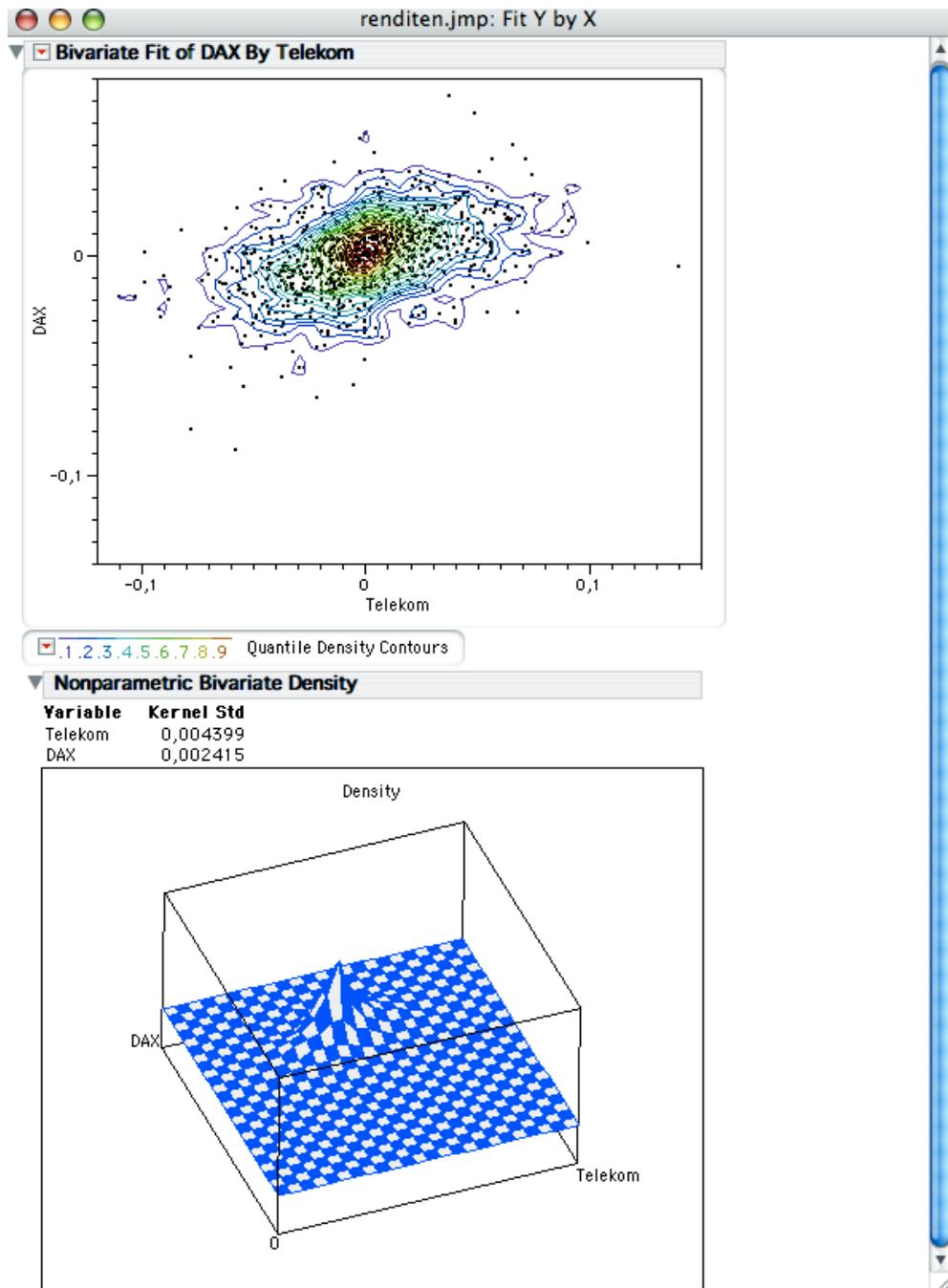


Abbildung 12.20: Streudiagramm, Dichte-Graphik (Höhenlinien) (oben) und 3-D-Darstellung der bivariaten empirischen Dichte (unten) von Dax und Telekom-Rendite (SAS/JMP).

12.5 Deskriptive Statistik

Geht man über graphische Analysen hinaus, so berechnet man nicht nur die empirischen Stichprobenvarianzen s_1^2, \dots, s_p^2 für die p Variablen (Realisationen) x_1, \dots, x_p ,² sondern auch die empirischen Kovarianzen s_{ij} zwischen den Variablen x_i und x_j

$$s_{ij} = (N - 1)^{-1} \sum_n (x_{ni} - \bar{x}_i)(x_{nj} - \bar{x}_j) \quad (12.12)$$

$$s_{ii} = s_i^2 \quad (12.13)$$

Diese geben den Zusammenhang (Kovarianz) s_{ij} , in standardisierter Form auch die sog. Korrelation

$$r_{ij} = \frac{s_{ij}}{s_i s_j} \quad (12.14)$$

zwischen x_i und x_j an. In übersichtlicher Form läßt sich dieses Zusammenhangsmuster in Form einer Tabelle darstellen, die als Korrelationsmatrix \mathbf{R} (für $p = 4$ Variablen)

$$\mathbf{R} = \begin{bmatrix} 1 & r_{12} & r_{13} & r_{14} \\ r_{21} & 1 & r_{23} & r_{24} \\ r_{31} & r_{32} & 1 & r_{34} \\ r_{41} & r_{42} & r_{43} & 1 \end{bmatrix} \quad (12.15)$$

bezeichnet wird. In der Diagonale stehen die Werte 1, da $s_{ii} = s_i^2, i = 1, \dots, p$ für die Einzelvarianzen gilt.

Beispiel 12.4 (Fortsetzung: Rendite-Korrelation von Aktien)

Wie gesagt sind Kursbewegungen von Aktien häufig gleichsinnig, da makroökonomische Einflüsse auf alle Papiere ähnlich wirken, etwa die US-Wirtschaft, Zinsänderungen, neue Konjunkturdaten, politische Veränderungen etc. Andererseits sind auch spezifische Einflüsse vorhanden, etwa Führungsprobleme in einer Firma, Absatzprobleme einer Branche etc. Daher wird die Korrelation der Aktienkurs-Renditen positiv, aber nicht gleich 1 sein. In Abb. 12.21 sind die paarweisen Korrelationen

$$\mathbf{R} = \begin{bmatrix} 1.000 & .420 & .596 & .567 \\ .420 & 1.000 & .392 & .351 \\ .596 & .392 & 1.000 & .566 \\ .567 & .351 & .566 & 1.000 \end{bmatrix} \quad (12.16)$$

²vgl. Abb. 12.18–12.19 für andere Statistiken, z.B. die Kurtosis, die bei Renditen häufig größer als 3 ist. Für den DAX ergibt sich $m_4 = 8.51$. Ausnahme ist die Telekom-Aktie mit $m_4 = 1.46$.

**Zusammenhang
Korrelation**

**Korrelations-
Matrix**

Descriptive Statistics

	Mean	Std. Deviation	N
DAXRENDITE	,000325	,0133547	3789
TELEKOMRENDITE	-,00002	,0289533	1270
DEUTSCHEBANKRENDITE	,000117	,0181157	3846
BAYERRENDITE	,000164	,0161813	3846

Correlations

		DAXRENDITE	TELEKOMRENDITE	DEUTSCHEBANKRENDITE	BAYERRENDITE
DAXRENDITE	Pearson Correlation	1	,420**	,596**	,567**
	Sig. (2-tailed)		,000	,000	,000
	N	3789	1270	3789	3789
TELEKOMRENDITE	Pearson Correlation	,420**	1	,392**	,351**
	Sig. (2-tailed)	,000		,000	,000
	N	1270	1270	1270	1270
DEUTSCHEBANKRENDITE	Pearson Correlation	,596**	,392**	1	,566**
	Sig. (2-tailed)	,000	,000		,000
	N	3789	1270	3846	3846
BAYERRENDITE	Pearson Correlation	,567**	,351**	,566**	1
	Sig. (2-tailed)	,000	,000	,000	
	N	3789	1270	3846	3846

**. Correlation is significant at the 0.01 level (2-tailed).

Abbildung 12.21: Korrelationsmatrix von Dax, Telekom, Deutsche Bank und Bayer mit univariaten Signifikanztests (casewise deletion, siehe Text).

der Aktienkurs-Renditen von Dax, Telekom, Deutsche Bank und Bayer als Tabelle dargestellt, zusammen mit univariaten Tests (in der Tabelle stehen p -Werte, die fälschlicherweise von SPSS als Signifikanz bezeichnet werden): wenn p kleiner als ein vorgegebenes α (z.B. 5%) ist, wird die Nullhypothese $\rho_{ij} = 0$ abgelehnt. Die angegebenen p -Werte sind sehr klein (fast 0), sodaß die Nullhypothese praktisch auf jedem α -Niveau abgelehnt wird. Die Stichprobengrößen der Datensätze sind unterschiedlich, da etwa der Telekom-Kurs erst ab Börseneinführung verfügbar ist; die nicht vorhandenen Werte werden als fehlende Daten bezeichnet. Sie sind in SPSS oder JMP standardmäßig als Punkt (.) codiert. Datenpaare (x_{in}, x_{jn}) mit fehlenden Werten auf einer Variable werden weggelassen (casewise deletion).

Man kann die Korrelation jedoch auch multivariat berechnen, wobei nur

p -Wert

fehlende Daten

casewise deletion

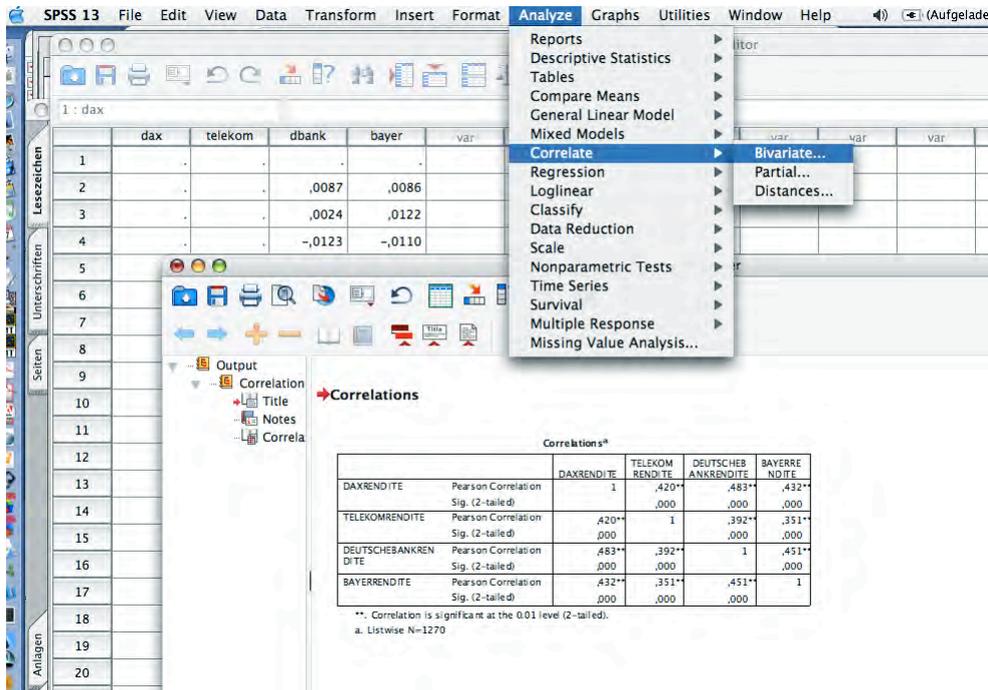


Abbildung 12.22: Ausklappmenü SPSS/Analyze/Correlate/Bivariate...: Korrelationsmatrix von Dax, Telekom, Deutsche Bank und Bayer mit univariaten Signifikanztests; listwise deletion, siehe Text.

Beobachtungen mit vorhandenen Werten auf allen 4 Variablen benutzt werden (listwise deletion; Abb. 12.23). Diese Option kann auch in

listwise deletion

SPSS/Analyze/Correlate/Bivariate.../Options

angeklickt werden (Abb. 12.22).



12.6 Statistische Tests

Bei der berechneten Korrelationsmatrix R handelt es sich um eine empirische Größe (Statistik), die statistischen Schwankungen unterworfen ist (etwa bei Berechnung in verschiedenen Zeiträumen). Will man über eine rein deskriptive Analyse hinaus gehen, können statistische Tests oder Vertrauensintervalle für die Korrelationen berechnet werden (KE 1).

Statistik

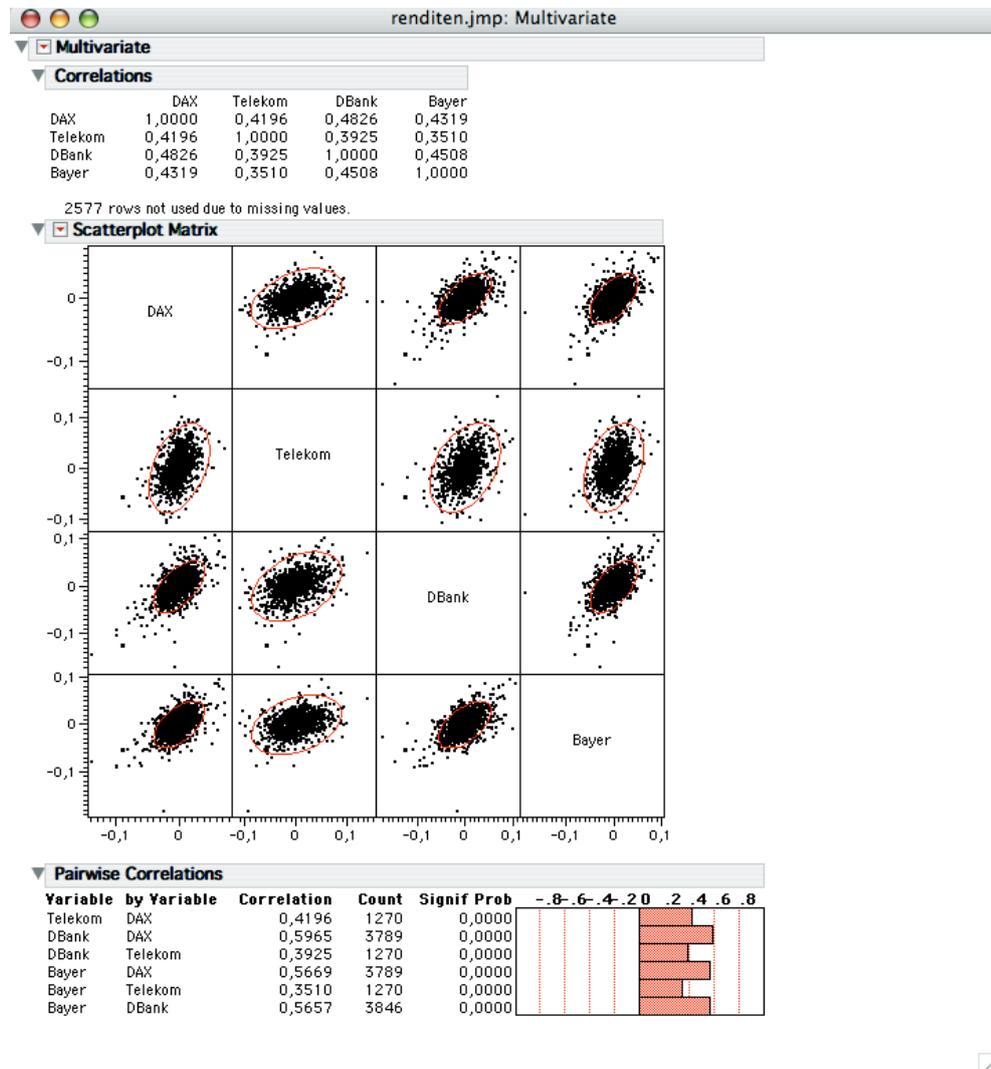


Abbildung 12.23: Korrelationsmatrix, Scatterplots und Streuungs-Ellipsen von Dax, Telekom, Deutsche Bank und Bayer mit univariaten Signifikanztests; listwise deletion, siehe Text (SAS/JMP).

Die einfachste Vorgehensweise besteht darin, die einzelnen Korrelationen r_{ij} zu prüfen. Man führt dann $p(p-1)/2$ univariate Signifikanztests aus, etwa für die Einzel-Hypothesen

$$H_{0,ij} : \rho_{ij} = 0 \quad (12.17)$$

$i = 1, \dots, p; j = i + 1, \dots, p$. Die Diagonale von R muß nicht getestet werden, da sie auf jeden Fall 1 ist. Außerdem ist die Korrelationsmatrix symmetrisch, sodaß man nur die obere (oder untere) Hälfte testen muß. Für $p = 4$ erhält man somit $4 * 3/2 = 6$ Tests.

Getestet wird (wie in jedem Test) nicht die berechnete Korrelation r_{ij} , sondern der hypothetische Korrelations-Parameter ρ_{ij} in der Grundgesamtheit. Die empirische Korrelation r_{ij} ist hierfür eine Schätzung, die in großen Stichproben nur noch wenig vom (unbekannten) ρ_{ij} abweicht. Die Abweichung ist umgekehrt proportional zur Wurzel aus der Stichprobengröße N , also $N^{-1/2}$.

Man kann sich nun fragen, warum der multivariate Test nicht als Wiederholung univariater Vorgehensweisen aufgefaßt werden kann (etwa die $p(p-1)/2$ -fache Wiederholung aller interessierenden Korrelationstests). Ein Problem liegt darin, daß die einzelnen Tests mit einem Signifikanzniveau α , z.B. 5% ausgeführt werden. Wiederholt man mehrere Tests auf diesem Niveau, so ist das simultane Signifikanzniveau α^* für alle Tests wesentlich höher, etwa $\alpha^* \leq 6\alpha = 30\%$ für $k = 6$ Tests. Dies bedeutet, daß die Wahrscheinlichkeit, die Nullhypothese abzulehnen, obwohl H_0 richtig ist

$$\alpha^* = P(H_0 \text{ ablehnen} | H_0 \text{ richtig}) \quad (12.18)$$

(Fehler 1. Art), sehr groß werden kann.

Man kann sagen, daß bei der Ausführung vieler Tests signifikante Resultate zufällig entstehen, obwohl gar kein Effekt vorliegt (also Korrelation gleich Null).

Im Spezialfall voneinander unabhängiger Einzel-Tests kann man sogar zeigen, daß

$$\alpha^* = 1 - (1 - \alpha)^k; k = \text{Anzahl der Tests} \quad (12.19)$$

gilt. Daher geht die simultane Irrtumswahrscheinlichkeit gegen 1, wenn man nur genügend viele Tests ausführt. In der Praxis sind $k = 100$ Tests

**univariate
Signifikanztests**

Grundgesamtheit

Signifikanzniveau

keine Seltenheit. Dann findet man auf jeden Fall Effekte, obwohl keine vorliegen.

Übung 12.1 (Simultanes Signifikanzniveau)

Berechnen Sie die simultanen Irrtumswahrscheinlichkeiten für $\alpha = 5\%$ und $k = 2, \dots, 10$ durchgeführte unabhängige Tests.



Übung 12.2 (Simultanes Signifikanzniveau: Herleitung)

Leiten Sie Gleichung 12.19 her.

Hinweis:

Für Ereignisse A gilt: $P(\bar{A}) = 1 - P(A)$ (Komplement) sowie $P(A \cap B) = P(A)P(B)$ bei Unabhängigkeit.

Definieren Sie die Ereignisse $E_i : H_{0i}$ beibehalten und $E = \cap_i E_i : H_0$ beibehalten und berechnen Sie $P(\bar{E})$.



Man behilft sich in der Praxis damit, die Signifikanzniveaus für die Einzelprüfungen abzusenken, etwa $\alpha \rightarrow \alpha/k$, sodaß $\alpha^* \leq \alpha = 5\%$ ist (Bonferroni-Methode). Allerdings werden dann die Einzeltests mit dem kleinen Wert α/k ausgeführt, was insgesamt zu einem **konservativen** Vorgehen führt. Dies bedeutet, daß das tatsächliche Signifikanzniveau α^* deutlich kleiner als das vorgegebene $\alpha = 5\%$ sein kann. Man findet also nichts, obwohl in den Daten Effekte vorliegen.

Bonferroni-Methode

Daher sind exakte multivariate Tests notwendig, deren Signifikanzniveau genau gleich α ist, etwa der Korrelationstest (Sphärizitäts-Test)³

Sphärizitäts-Test

$$H_0 = \bigcap_{i < j} H_{0,ij} : \mathbf{P} = \mathbf{I}_p \quad (12.20)$$

wobei hierbei die gesamte Matrix

$$\mathbf{P} = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{21} & 1 & \rho_{23} & \rho_{24} \\ \rho_{31} & \rho_{32} & 1 & \rho_{34} \\ \rho_{41} & \rho_{42} & \rho_{43} & 1 \end{bmatrix} \quad (12.21)$$

³Der Name ergibt sich aus der Tatsache, daß die quadratische Form $\sum_{i,j} x_i \rho_{ij} x_j = r^2$ eine p -dimensionale Kugel mit Radius r beschreibt, wenn ρ gleich der Einheitsmatrix ist.

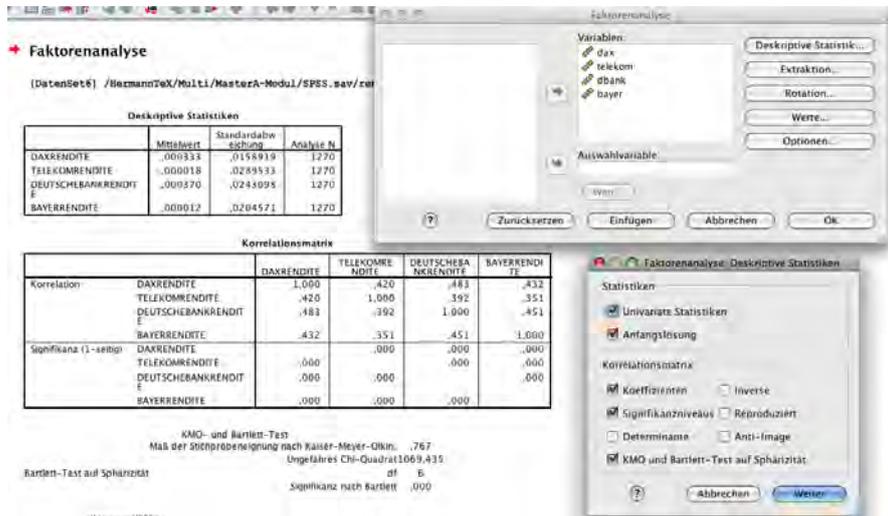


Abbildung 12.24: Bartlett-Test auf Sphärizität für die Korrelationsmatrix der DAX-Renditen.

simultan auf Übereinstimmung mit der Einheitsmatrix

$$I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{12.22}$$

getestet wird. Die Nullhypothese $H_0 = \bigcap_{i < j} H_{0,ij}$ läßt sich hierbei als Schnittmenge der univariaten Hypothesen auffassen (das sog. *Union-Intersection-Prinzip*).

In SPSS läßt sich dieser Test sehr einfach im Rahmen der Faktorenanalyse durchführen (Abb. 12.24). Da die p -Werte der univariaten Tests praktisch 0 sind, ist es nicht verwunderlich, daß auch der simultane Test hochsignifikant ausfällt.

Union-Intersection-Prinzip

12.7 Übungen mit SPSS und SAS/JMP

Damit die dargestellte Theorie für den Leser auch praktisch umgesetzt werden kann, wurden im Text eine Vielzahl von Beispielen integriert, die mit Hilfe der leicht bedienbaren Softwarepakete SPSS (Statistical Package for the Social Sciences) und SAS/JMP nachvollzogen werden können. Alle Verfahren sind mit Hilfe einer Menü-Steuerung zugänglich und können ohne textbezogene Kommandos durch Maus-Klicks bedient

**SPSS
SAS/JMP**

werden. JMP unterstützt auch eine interaktive Bearbeitung bereits bestehender Outputs (in diesen können zusätzliche Analysen durchgeführt und graphische Elemente, z.B. Konfidenzellipsen, hinzugefügt werden).

Die selbsttätige Durchführung der Computerübungen ist für das Begreifen des Lehrstoffs unbedingt erforderlich. Es reicht nicht aus, den Text nur zu lesen, sondern die in den Graphiken und Tabellen dargestellten Resultate auch selbst zu erzeugen. Die Datensätze und Programme sind in den Verzeichnissen SPSS und JMP der beigefügten Kurs-CD enthalten.

SPSS kann für die Studierenden der FernUniversität über ein Softwareportal des ZMI (Zentrum für Medien und IT) bezogen werden.

Bitte beachten Sie die Hinweise auf der Homepage des Lehrstuhls:

http://www.fernuni-hagen.de/ls_statistik/kurse/k42221.shtml

Für die Nutzung von JMP verweise ich auf die Homepage von SAS/JMP:

<http://www.jmp.com/>

Kapitel 13

Erste Auswertungsschritte

13.1 Daten-Sätze und Daten-Matrizen

Der wichtigste Schritt bei der computergestützten Analyse ist die Erstellung eines Datensatzes, der vom Statistik-Programm verarbeitet werden kann. Im Prinzip handelt es sich um eine rechteckige Daten-Tabelle, die ähnlich wie in Tabellenkalkulationen aufgebaut ist.

Ganz unabhängig von Software-Programmen definiert man zunächst die Daten-Matrix:

Definition 13.1 (Daten-Matrix)

Daten-Matrix

Die **Daten-Matrix** \mathbf{X} ist ein rechteckiges Schema, bei dem die Daten $x_{ni}, n = 1, \dots, N, i = 1, \dots, p$ in N Zeilen (Personen) und p Spalten (Variablen) angeordnet werden.

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \dots & \dots & \dots & \dots \\ x_{N1} & x_{N2} & \dots & x_{Np} \end{bmatrix} : N \times p \quad (13.1)$$

Etwa lassen sich mit Hilfe der Daten-Matrix sehr leicht Statistiken wie

der Mittelwert $\bar{x}_i = (1/N) \sum_n x_{ni}$, $i = 1, \dots, p$, aller p Spalten berechnen:

Mittelwert

$$\bar{\mathbf{x}} = [\bar{x}_1, \dots, \bar{x}_p]' : p \times 1 \quad (13.2)$$

$$= \frac{1}{N} \mathbf{X}' \mathbf{1} \quad (13.3)$$

$$\mathbf{1} = [1, \dots, 1]' : N \times 1 \text{ Spaltenvektor von Einsen} \quad (13.4)$$

**transponierte
Matrix**

Hierbei ist $\mathbf{X}' : p \times N$ ist die transponierte Daten-Matrix.

Die empirische Kovarianzmatrix S ergibt sich als (vgl. Glg. 12.12)

empirische

Kovarianzmatrix

$$\mathbf{S} = \frac{1}{N-1} \sum_n (x_n - \bar{x})(x_n - \bar{x})' = \frac{N}{N-1} \mathbf{S}_* \quad (13.5)$$

$$\mathbf{S}_* = \frac{1}{N} \mathbf{X}' \mathbf{X} - \bar{\mathbf{x}} \bar{\mathbf{x}}', \quad (13.6)$$

wobei $\mathbf{S}_* = \frac{1}{N} \sum_n (\mathbf{x}_n - \bar{\mathbf{x}})(\mathbf{x}_n - \bar{\mathbf{x}})'$ der Maximum Likelihood-Schätzer der Kovarianz-Matrix $\Sigma : p \times p$ ist.

13.2 Erzeugen einer SPSS-System-Datei

13.2.1 Öffnen einer SPSS-System-Datei

Am einfachsten ist die Analyse, wenn bereits ein Datensatz vorliegt. Starten Sie zunächst das Programm (SPSS oder JMP) durch einen Doppelklick. Nach Anklicken des Menüs

SPSS/Datei/Öffnen/...

wird der Datensatz geladen (Fig. 13.1) Die einzelnen Spalten tragen Überschriften (Variablenamen). In den Zeilen sind die Daten der einzelnen statistischen Einheiten (hier Länder) eingetragen.

13.2.2 Öffnen einer Tabelle anderer Programme

Der Datensatz wurde ursprünglich durch Öffnen einer Tabelle aus einem Tabellenkalkulationsprogramm erzeugt (auch Copy-Paste wäre möglich). Nach Anklicken des Menüs

SPSS/Datei/Öffnen/Daten/...

wird der Datensatz geladen (Fig. 13.2). Zu beachten ist, daß als

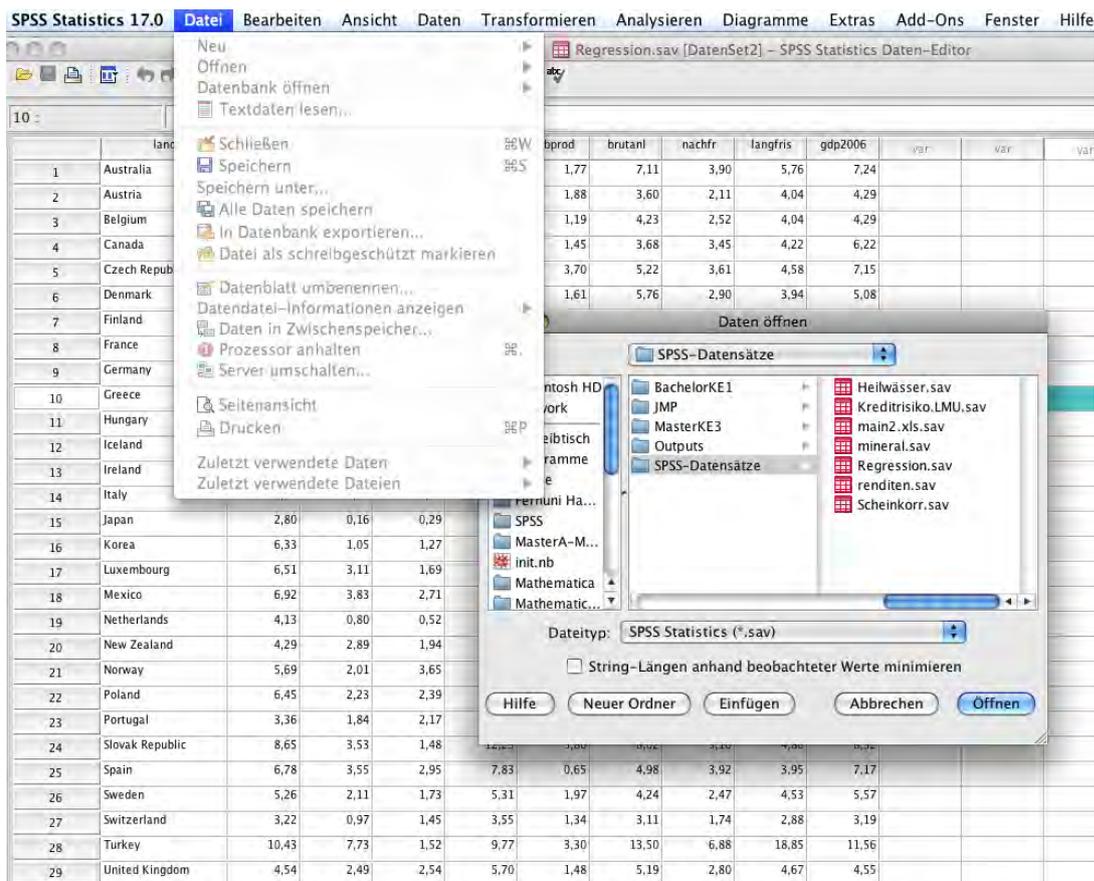


Abbildung 13.1: Laden eines bereits bestehenden Datensatzes.

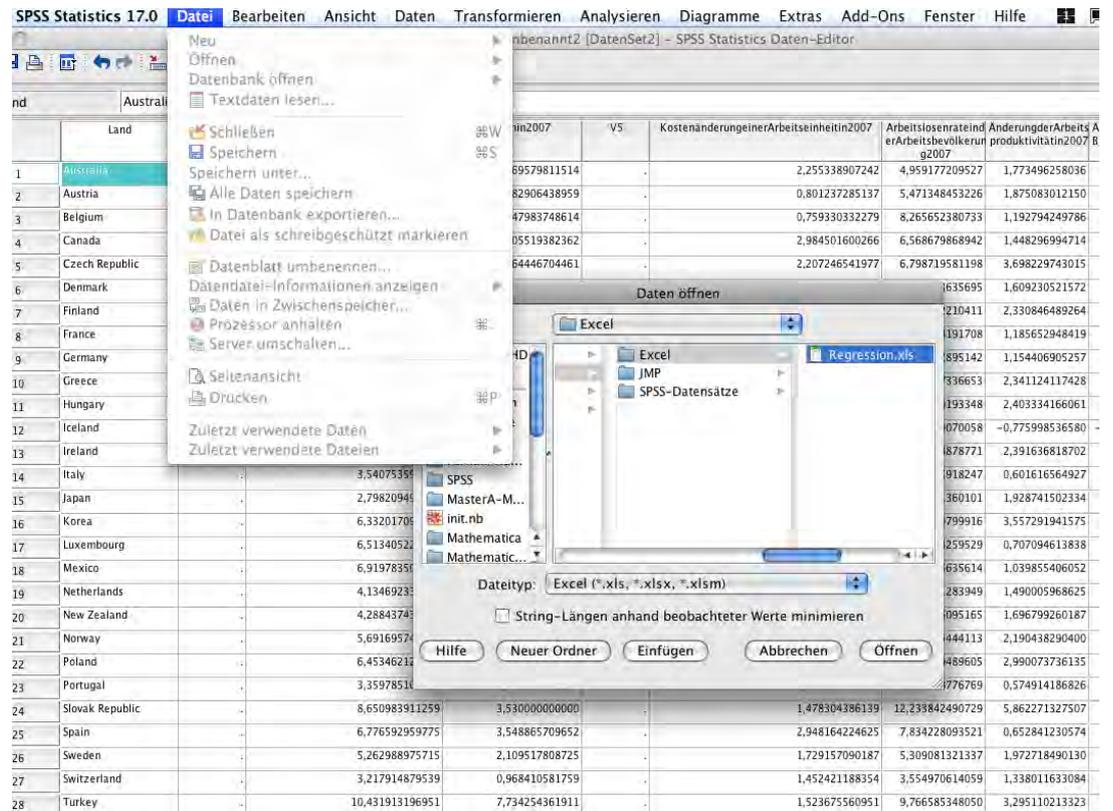


Abbildung 13.2: Laden eines Excel-Spreadsheets in SPSS.

	Name	Typ	Spaltenf...	Dezimal...	Variablenlabel	Wertelabels	Fehlende Werte	Spalten	Ausrichtung	Messniveau
1	land	String	15	0	Land	Keine	Keine	15	Linksbünd...	Nominal
2	gdp2007	Numerisch	11	2	Nominales GDP-Wachstum in % 2007	Keine	Keine	8	Rechtsbü...	Metrisch
3	inflatio	Numerisch	11	2	Inflation in % 2007	Keine	Keine	8	Rechtsbü...	Metrisch
4	arbkost	Numerisch	11	2	Kostenänderung einer Arbeitseinheit in % 2007	Keine	Keine	8	Rechtsbü...	Metrisch
5	arbeitsl	Numerisch	11	2	Arbeitslosenrate in % der Arbeitsbevölkerung 2007	Keine	Keine	8	Rechtsbü...	Metrisch
6	arbprod	Numerisch	11	2	Änderung der Arbeitsproduktivität in % 2007	Keine	Keine	8	Rechtsbü...	Metrisch
7	brutanl	Numerisch	11	2	Änderung des realen Brutto Anlagevermögens in % 2007	Keine	Keine	8	Rechtsbü...	Metrisch
8	nachfr	Numerisch	11	2	Änderung der realen aggregierten Nachfrage in % 2007	Keine	Keine	8	Rechtsbü...	Metrisch
9	langfris	Numerisch	11	2	Langfristiger Zinssatz in % 2007	Keine	-9,00	8	Rechtsbü...	Metrisch
10	gdp2006	Numerisch	8	2	Nominales GDP-Wachstum in % 2006	Keine	Keine	8	Rechtsbü...	Metrisch
11										
12										
13										
14										

Abbildung 13.3: Variablen-Ansicht des OECD-Datensatzes.

Dateityp Excel (*.xls,...) ausgewählt werden muß. Auch System-Dateien (also Daten-Files) anderer Programme, etwa SAS, STATA etc. können geladen werden. Die Spalten-Überschriften aus Excel werden übernommen. In der sog. Variablen-Ansicht (Abb. 13.3) können diese Namen verändert werden. Bei Bedarf können für die Ausprägungen der Variablen Werte-Labels vergeben werden (z.B. 0 = männlich, 1 = weiblich etc.).

Variablen-Ansicht

Weiterhin sind in der letzten Spalte die Meßniveaus der Variablen angegeben (Land = nominal, etc.). In JMP werden die statistischen Verfahren konsequent nach den angegebenen Skalenniveaus ausgewählt. Beispielsweise wird bei metrischem Skalenniveau von X und Y automatisch eine Regressionsanalyse, bei nominalem X und metrischem Y eine Varianzanalyse gerechnet (siehe die Tabelle in Abb. 13.4). Der Benutzer kann also ohne Kenntnisse der Modell-Typen korrekte Analysen durchführen.

Übung 13.1 (Automatische Modell-Wahl)

Denken Sie bitte über Vor- und Nachteile einer derartigen Menü-Steuerung nach ■

13.2.3 Manuelle Eingabe im SPSS-Dateneditor

Wenn die Daten nur als Fragebogen oder in gedruckter Form vorliegen, müssen sie in eine leere Excel-Tabelle oder ein leeres SPSS-Dateneditor-Fenster eingegeben werden. Bei sehr großen Datenmengen kann man auch

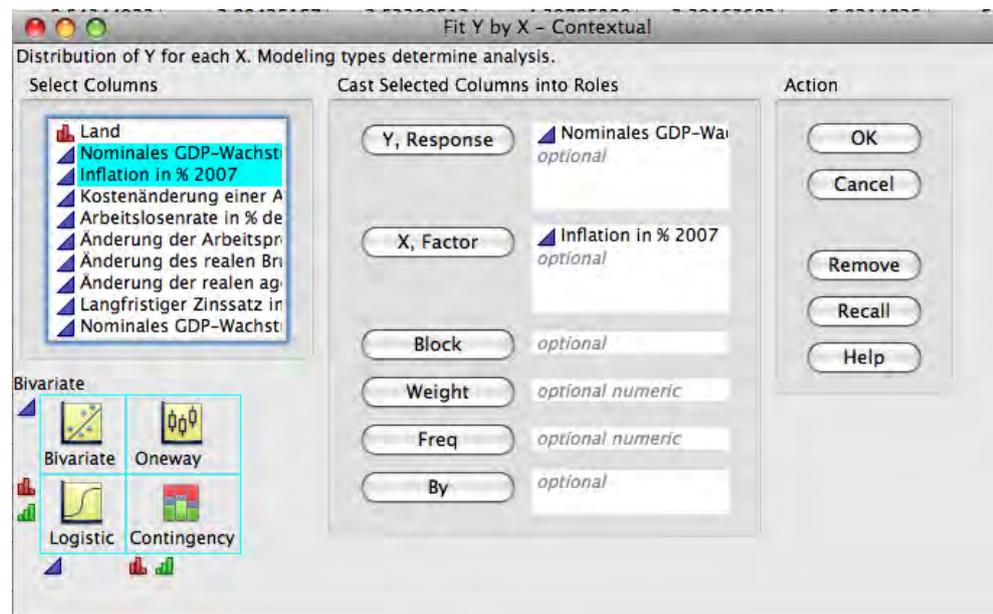


Abbildung 13.4: JMP: Variablen-Ansicht des OECD-Datensatzes. Die Skalenniveaus der abhängigen (Y , Response) und unabhängigen Variablen (X , Factor) determinieren die Auswertungsmethode (Bivariate = Regression, Oneway = Varianzanalyse, Logistic = logistische Regression, Contingency = Kreuztabellen).

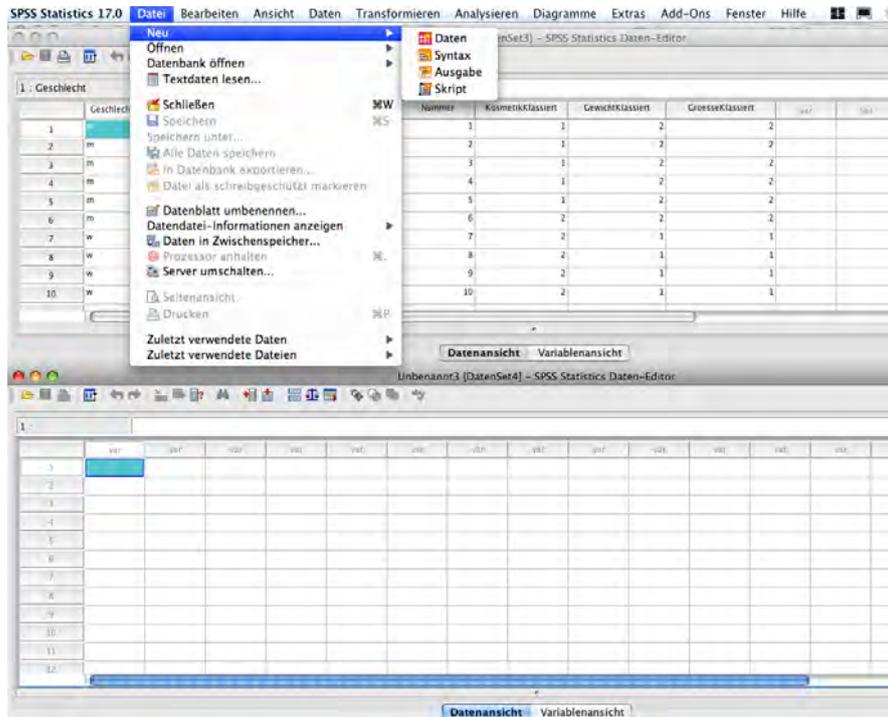


Abbildung 13.5: Erstellen eines kleinen SPSS-Datensatzes.

über die Möglichkeit nachdenken, die Daten zu scannen und dann eine Texterkennung durchzuführen.

Übung 13.2 (Erstellen eines Datensatzes)

Es ist instruktiv, selbst einen kleinen Datensatz einzugeben. Laden Sie etwa den Datensatz **Scheinkorr.sav** und öffnen Sie eine leeres Datenfenster (Abb. 13.5). Tragen Sie dann die Daten manuell in das leere Fenster ein.

Vielleicht haben Sie auch eigene Daten, die für Sie interessant sind ■

13.3 Explorative Datenanalyse

13.3.1 Univariate Analyse: Histogramme und summarische Statistiken

Tabellen sind meist sehr unübersichtlich. Daher ist es nötig, mit Hilfe graphischer Darstellungen und summarischer Werte (Statistiken) einen Überblick zu erhalten. Moderne Programme enthalten eine Vielzahl von

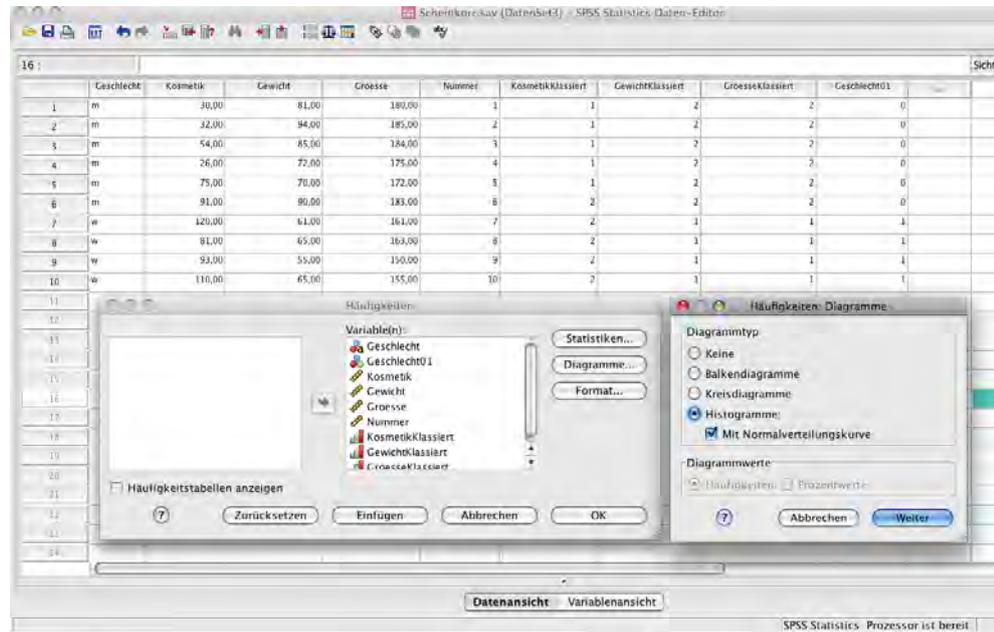


Abbildung 13.6: Menü-Befehle für Häufigkeiten.

graphischen Darstellungen und statistischen Kennwerten, die Ihnen teilweise sicherlich vertraut sind (etwa Säulendiagramm, Tortendiagramm, Mittelwert, Streuung etc.). In Kap. 12 wurde bereits ein Überblick gegeben.

Skalenniveau

Auch hier ist wichtig, daß die **Skalenniveaus** der Variablen beachtet werden (vgl. Kap. 14). Beispielsweise macht es bei einer nominalen Variablen wie Geschlecht (Codierung m/w) nur wenig Sinn, ein Histogramm mit überlagerter Normalverteilung zu zeichnen.¹ Abb. 13.6 zeigt die Menüauswahl

SPSS/Analysieren/Deskriptive Statistiken/Häufigkeiten

für die Variablen des Datensatzes Scheinkorr.sav. Die nominale Variable **Geschlecht** wurde auf 2 Arten codiert:

1. als String (Text): m/w bzw. als
2. 0-1-Variable (**Geschlecht01**) mit den Wertelabels 0=männlich, 1=weiblich.

¹plotten auf Computerdeutsch

Warnungen
Geschlecht ist ein String. Deshalb kann kein Histogramm erstellt werden.

Statistiken

	Geschlecht	Geschlecht01	Kosmetik	Gewicht	Groesse	Nummer	Kosmetik (Klassiert)	Gewicht (Klassiert)	Groesse (Klassiert)
N	10	10	10	10	10	10	10	10	10
Fehlend	0	0	0	0	0	0	0	0	0
Mittelwert		,40	71,2000	73,8000	170,8000	5,50	1,50	1,60	1,60
Standardfehler des Mittelwertes		,163	10,76909	4,13333	4,03815	,957	,167	,163	,163
Median		,00	78,0000	71,0000	173,5000	5,50	1,50	2,00	2,00
Modus		0	26,00 ^a	65,00	150,00 ^a	1 ^a	1 ^a	2	2
Standardabweichung		,516	34,05486	13,07075	12,76976	3,028	,527	,516	,516
Varianz		,267	1159,733	170,844	163,067	9,167	,278	,267	,267
Schiefe		,484	-,119	,255	-,435	,000	,000	-,484	-,484
Standardfehler der Schiefe		,687	,687	,687	,687	,687	,687	,687	,687
Kurtosis		-2,277	-1,446	-1,215	-1,359	-1,200	-2,571	-2,277	-2,277
Standardfehler der Kurtosis		1,334	1,334	1,334	1,334	1,334	1,334	1,334	1,334
Spannweite		1	94,00	39,00	35,00	9	1	1	1
Minimum		0	26,00	55,00	150,00	1	1	1	1
Maximum		1	120,00	94,00	185,00	10	2	2	2
Summe		4	712,00	738,00	1708,00	55	15	16	16

a. Mehrere Modi vorhanden. Der kleinste Wert wird angezeigt.

Histogramm

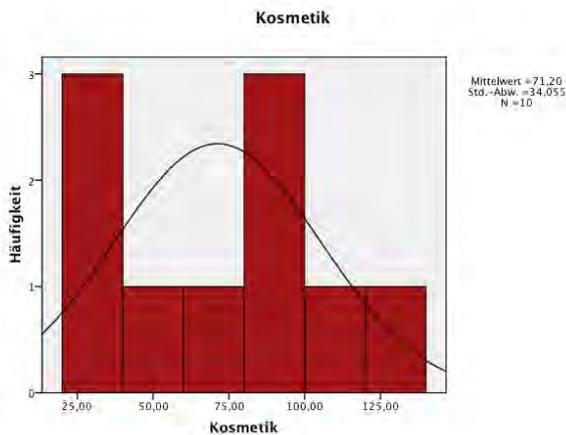
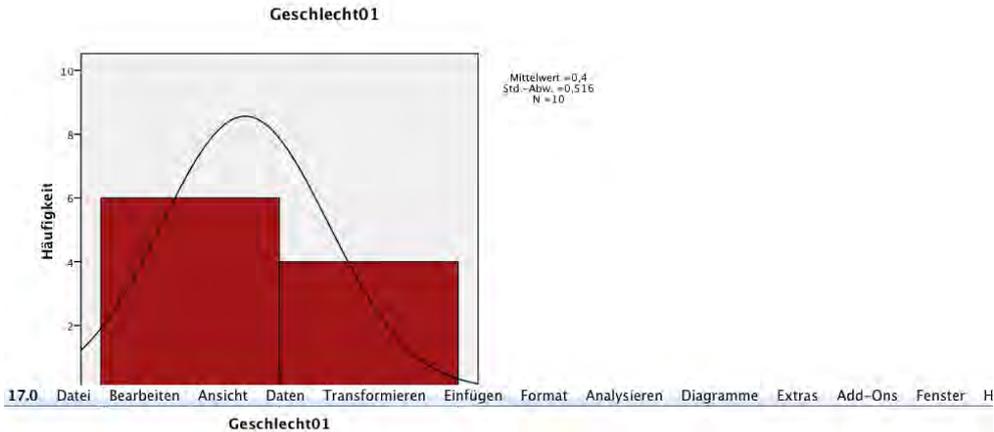


Abbildung 13.7: Statistiken und Histogramme für die Indikatorvariable *Geschlecht01* und die metrische Variable *Ausgaben für Kosmetik-Produkte*.

Im ersten Fall werden keine Statistiken und auch kein Histogramm erstellt², während für die 0-1 (Null-Eins)-Variable Mittelwerte etc. berechnet werden.

Indikator- Variablen

Derartig codierte dichotome (binäre) Variablen werden auch als Indikator-Variablen bezeichnet, da sie das Vorliegen einer bestimmten Qualität anzeigen (etwa Arbeitslos = 0 bedeutet *nicht arbeitslos*).

Abb. 13.7 zeigt die Statistiken sowie ein Histogramm für `Geschlecht01`. Der Mittelwert von `Geschlecht01` stimmt mit der relativen Häufigkeit p für die Ausprägung 1 = weiblich überein. Es macht daher Sinn, den Mittelwert von Indikatoren zu berechnen, auch wenn nur nominales Skalenniveau vorliegt.

Auch die Stichprobenvarianz S^2 läßt sich interpretieren, nämlich als Streuung der Bernoulli-verteilten Variable $X \sim B(1, \pi)$ mit $E[X] = \pi$, $\text{Var}(X) = E[X^2] - \pi^2 = E[X] - \pi^2 = \pi(1 - \pi)$.

Bernoulli- Verteilung

Es gilt

$$\bar{X} = N^{-1} \sum_{n=1}^N X_n = P \quad (13.7)$$

$$E[\bar{X}] = \pi \quad (13.8)$$

und für die Stichprobenvarianz

$$S^2 = (N - 1)^{-1} \sum_{n=1}^N (X_n - P)^2 \quad (13.9)$$

$$= \frac{N}{N - 1} P(1 - P) \quad (13.10)$$

$$E[S^2] = \frac{N}{N - 1} E[P(1 - P)] = \pi(1 - \pi). \quad (13.11)$$

Die Stichprobenvarianz S^2 ist also ein unverzerrter Schätzer für $\text{Var}(X) =$

²SPSS schreibt: Warnungen: Geschlecht ist ein String. Deshalb kann kein Histogramm erstellt werden.

$\pi(1 - \pi)$, die Varianz der Indikatorvariablen X .³

13.3.2 Bivariate Analyse: Kreuztabellen und Streudiagramme

Will man den Zusammenhang von Variablen untersuchen, bieten sich je nach Skalenniveau Kreuztabellen (beide Variablen Nominal- oder Ordinalskaliert) oder Streudiagramme an (beide Variablen metrisch bzw. eine binär). In Kap. 12 wurde bereits eine Vielzahl von Streudiagrammen präsentiert.

Nominale Variablen lassen sich als Tabellen der gemeinsamen relativen Häufigkeiten

$$f_{ij} = f(x_i, y_j); \quad i = 1, \dots, I, j = 1, \dots, J \quad (13.22)$$

**gemeinsame
Häufigkeiten**

³Es gilt

$$S_*^2 = N^{-1} \sum_{n=1}^N X_n^2 - P^2 \quad (13.12)$$

$$= N^{-1} \sum_{n=1}^N X_n - P^2 = P(1 - P) \quad (13.13)$$

$$E[S_*^2] = E[P(1 - P)] = \frac{N-1}{N} \pi(1 - \pi) \quad (13.14)$$

da $X_n^2 = X_n$ und wegen der Unabhängigkeit der Stichprobenvariablen X_n

$$E[P^2] = N^{-2} \sum_{nm} E[X_n^2 X_m^2] \quad (13.15)$$

$$= N^{-2} \left(\sum_n E[X_n^4] + \sum_{n \neq m} E[X_n^2] E[X_m^2] \right) \quad (13.16)$$

$$= N^{-2} (N\pi + N(N-1)\pi^2) = N^{-1} (\pi + (N-1)\pi^2). \quad (13.17)$$

Daher findet man

$$E[S_*^2] = E[P - P^2] \quad (13.18)$$

$$= \pi - N^{-1} (\pi + (N-1)\pi^2) \quad (13.19)$$

$$= \frac{N-1}{N} \pi(1 - \pi) \quad (13.20)$$

$$E[S^2] = \frac{N}{N-1} E[S_*^2] = \pi(1 - \pi). \quad (13.21)$$

S^2 ist also ein unverzerrter Schätzer für $\text{Var}(X) = \pi(1 - \pi)$.

(x_i, y_j sind die Ausprägungen) oder der bedingten relativen Häufigkeiten

bedingte Häufigkeiten(Spalte)

$$f(x_i|y_j) = \frac{f_{ij}}{f_{.j}} \quad (13.23)$$

(bedingt auf Spalten) bzw.

bedingte Häufigkeiten(Zeile)

$$f(y_j|x_i) = \frac{f_{ij}}{f_{i.}} \quad (13.24)$$

(bedingt auf Zeilen) aufbereiten. Hierbei sind

$$f_{i.} = \sum_j f_{ij} \quad (13.25)$$

Randhäufigkeiten

$$f_{.j} = \sum_i f_{ij} \quad (13.26)$$

die Randhäufigkeiten der Kreuztabelle (Randverteilung, marginale Verteilung).

Beispiel 13.1 (Filialgestaltung von Bankfilialen)

Der Datensatz `main2.xls.sav` enthält die Daten einer empirischen Untersuchung zur Auswirkung unterschiedlicher Filialgestaltung von Bankfilialen auf das Kundenverhalten. Er wird in Kap. 15 ausführlich besprochen.

Abb. 13.8 zeigt die Menüsteuerung zur Erzeugung einer Kreuztabelle (gemeinsame und relative Häufigkeiten) sowie verschiedener Zusammenhangs-Koeffizienten. Die Kreuztabelle (Abb. 13.9) ist in den Zeilen nach der Variable Geschlecht und in den Spalten nach der Berufsgruppe geordnet und enthält die Häufigkeiten $f_{ij} = f(x_i, y_j)$ sowie die bedingten Häufigkeiten.

Etwa unterscheidet sich die Verteilung in der Gruppe der Männer stark von der Verteilung in der Gruppe der Frauen

$$f(\text{Hausfrau/mann} \mid \text{männlich}) = 0.012 \quad (13.27)$$

$$f(\text{Hausfrau/mann} \mid \text{weiblich}) = 0.199 \quad (13.28)$$

was kaum überraschen dürfte.

Sehr übersichtlich ist das **gruppierte Balkendiagramm**, das die Verteilung in den Gruppen (männlich/weiblich) darstellt.



Abbildung 13.8: Menü zur Erzeugung einer Kreuztabelle der Variablen Geschlecht*Berufsgruppe.

Abb. 13.10 zeigt den Mosaik-Plot, der die relativen Häufigkeiten als Flächen innerhalb der Gruppen darstellt. Die Zusammenhänge lassen sich auch zahlenmäßig ausdrücken (Abb. 13.11).

Das χ^2 -Maß = 33.9675 mit 5 Freiheitsgraden (siehe Kap. 4.3) ist hochgradig signifikant (p -Wert fast 0), ebenso der Phi-Koeffizient

$$\phi = \sqrt{\frac{\chi^2}{N}} = \sqrt{33.968/377} = 0.300166, \quad (13.29)$$

Cramérs v

$$v = \sqrt{\frac{\chi^2}{N(m-1)}} \quad (13.30)$$

$$= \sqrt{33.968/377} = 0.300166 \quad (13.31)$$

$$m = \min(I, J) = \min(2, 6) = 2 \quad (13.32)$$

und der Kontingenzkoeffizient (normiertes χ^2 -Maß)

$$\sqrt{\frac{\chi^2}{\chi^2 + N}} = \sqrt{33.968/(33.968 + 377)} = 0.287494. \quad (13.33)$$

Chi-Quadrat

Phi-Koeffizient

Cramérs v

Kontingenzkoeffizient

GESCHLECHT * BERUF Kreuztabelle

			BERUF					Gesamt	
			Schüler	Arbeiter	Hausfrau	Selbständig	Rentner		ohne Beschäftigung
GESCHLECHT	männlich	Anzahl	15	71	2	13	55	15	171
		% innerhalb von GESCHLECHT	8,8%	41,5%	1,2%	7,6%	32,2%	8,8%	100,0%
		% innerhalb von BERUF % der Gesamtzahl	46,9%	50,0%	4,7%	56,5%	49,1%	60,0%	45,4%
	weiblich	Anzahl	17	71	41	10	57	10	206
		% innerhalb von GESCHLECHT	8,3%	34,5%	19,9%	4,9%	27,7%	4,9%	100,0%
		% innerhalb von BERUF % der Gesamtzahl	53,1%	50,0%	95,3%	43,5%	50,9%	40,0%	54,6%
Gesamt	Anzahl	32	142	43	23	112	25	377	
	% innerhalb von GESCHLECHT	8,5%	37,7%	11,4%	6,1%	29,7%	6,6%	100,0%	
	% innerhalb von BERUF	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	
	% der Gesamtzahl	8,5%	37,7%	11,4%	6,1%	29,7%	6,6%	100,0%	

Balkendiagramm

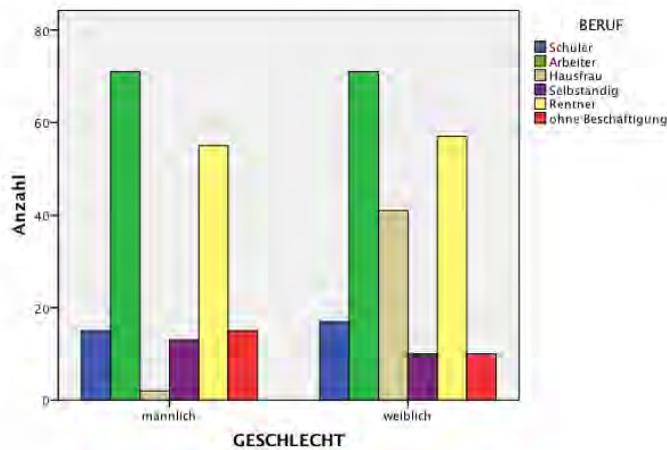


Abbildung 13.9: Kreuztabelle und gruppiertes Balkendiagramm der Variablen Geschlecht*Berufsgruppe.

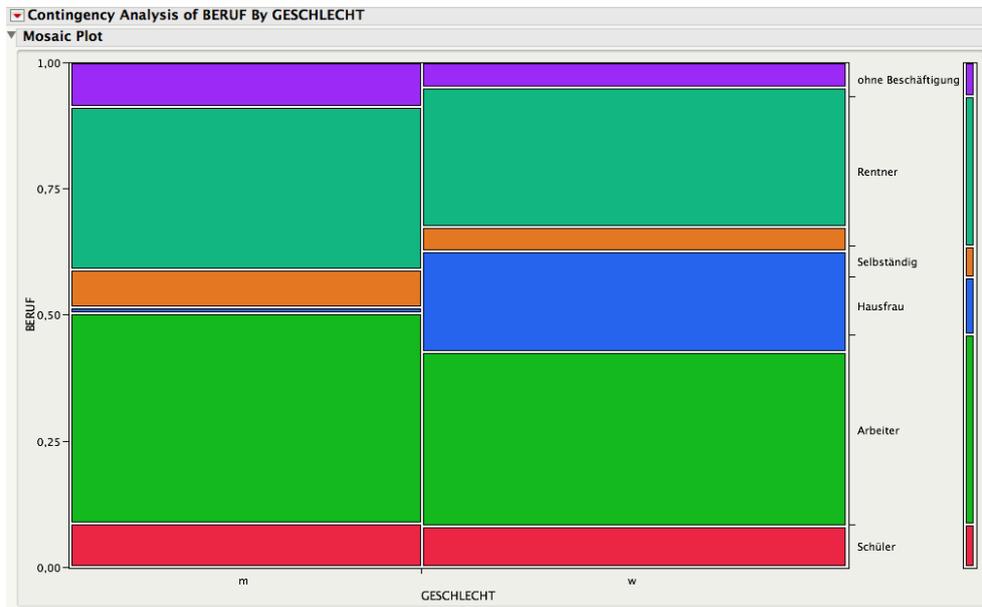


Abbildung 13.10: Mosaik-Plot der Variablen Geschlecht*Berufsgruppe.

Dagegen ist das Lambda-Maß $\lambda(y \rightarrow x)$ (Goodman-Kruskal) ⁴ nicht signifikant. Dies hat den Grund, daß bei der spaltenweisen Prädiktion (in der Spalte $Y = \text{Arbeiter}$) eine gleiche Anzahl von Fällen ($h_{12} = h_{22} = 71$) vorliegt, also der bedingte Modalwert in dieser Spalte nicht eindeutig ist. Daher ist der Fehler in dieser Spalte sehr groß und die Modalregel in den Spalten (**mit** Prädiktor Y) prognostiziert nicht viel besser als die Modalregel der zeilenweisen Randverteilung $f_{i.} = [0.454, 0.546]'$ (Geschlecht). Ohne den Prädiktor **Beruf** würde man einfach die Modal-Kategorie **weib-**

Goodman-Kruskal
 λ

⁴Beim Lambda-Maß $\lambda(y \rightarrow x)$ wird als Prognosewert der Modus $\hat{i}(j)$ für jeden Wert des Prädiktors $Y = j$ genommen. Der Anteilswert der korrekten Prädiktionen **mit** Y -Modalregel $G(+)$ wird mit der Modalregel der Randverteilung verglichen (korrekte Prädiktionen **ohne** Y : $G(-)$) und bezogen auf den maximalen Wert $G(+)=1$. Daher gilt

$$\lambda(y \rightarrow x) = \frac{G(+)-G(-)}{1-G(-)} \leq 1.$$

Umgekehrt für $\lambda(x \rightarrow y)$. Siehe KE 2.

Chi-Quadrat-Tests

	Wert	df	Asymptotische Signifikanz (2-seitig)
Chi-Quadrat nach Pearson	33,968 ^a	5	,000
Likelihood-Quotient	41,739	5	,000
Zusammenhang linear-mit-linear	,895	1	,344
Anzahl der gültigen Fälle	377		

a. 0 Zellen (.0%) haben eine erwartete Häufigkeit kleiner 5. Die minimale erwartete Häufigkeit ist 10,43.

Richtungsmaße

			Wert	Asymptotischer Standardfehler ^a	Näherungsweise ^b	Näherungsweise Signifikanz
Nominal- bzgl. Nominalmaß	Lambda	Symmetrisch	,020	,034	,581	,561
		GESCHLECHT abhängig	,047	,079	,581	,561
	Goodman-und-Kruskal-Tau	GESCHLECHT abhängig	,090	,017		,000 ^d
		BERUF abhängig	,015	,005		,000 ^d
	Unsicherheitskoeffizient	Symmetrisch	,050	,012	4,007	,000 ^e
		GESCHLECHT abhängig	,080	,020	4,007	,000 ^e
	BERUF abhängig	,036	,009	4,007	,000 ^e	

a. Die Null-Hypothese wird nicht angenommen.

b. Unter Annahme der Null-Hypothese wird der asymptotische Standardfehler verwendet.

c. Kann nicht berechnet werden, weil der asymptotische Standardfehler gleich Null ist.

d. Basierend auf Chi-Quadrat-Näherung

e. Chi-Quadrat-Wahrscheinlichkeit für Likelihood-Quotienten.

Symmetrische Maße

		Wert	Näherungsweise Signifikanz
Nominal- bzgl. Nominalmaß	Phi	,300	,000
	Cramer-V	,300	,000
	Kontingenzkoeffizient	,287	,000
Maß der Übereinstimmung	Kappa	, ^a	
Anzahl der gültigen Fälle		377	

a. Die Kappa-Statistik kann nicht berechnet werden. Sie benötigt eine symmetrische 2-Wege-Tabelle, in der die Werte der ersten Variable mit den Werten der zweiten Variable übereinstimmen.

Abbildung 13.11: Zusammenhangsmaße der Variablen *Geschlecht* und *Berufsgruppe*.

lich prognostizieren (Fehler $F(-) = .454$). Es gilt also für $\lambda(y \rightarrow x)$:

$$\lambda(y \rightarrow x) = \frac{F(-) - F(+)}{F(-)} = \frac{G(+)-G(-)}{1-G(-)} \quad (13.34)$$

$$G(+)=\sum_j f_{\hat{i}(j)j} \quad (13.35)$$

$$= \text{Anteilswert der Modalprädiktionen} \quad (13.36)$$

$$= 0.045 + 0.18 + 0.108 + 0.034 + 0.15 + 0.039$$

$$= 0.5676 \quad (13.37)$$

$$G(-) = f_{\hat{i}} = 0.5464$$

$$= \text{Maximum der Randverteilung (Zeilen)} \quad (13.38)$$

$$\hat{i}(j) = \text{Modalwert in Spalte } j. \quad (13.39)$$

$$\lambda(y \rightarrow x) = (0.5676 - 0.5464)/(1 - 0.5464) = 0.0467372. \quad (13.40)$$

(bitte nachrechnen!).

Daher ist die Fehlerreduktion bei der Vorhersage des Geschlechts aus der Berufsgruppe nur sehr gering (im Verhältnis zur einfachen Modalregel: Proband ist weiblich).

Umgekehrt gilt: Bei der Prädiktion **Geschlecht** \rightarrow **Beruf** ergibt sich in beiden Zeilen die Modal-Kategorie **Arbeiter**, was auch ohne den Prädiktor **Geschlecht** die Modalkategorie der Randverteilung $f_{.j}$ ist. Daher ist in diesem Fall sogar $F(+)=F(-)$ und $\lambda(x \rightarrow y)=0$.

Falls Sie Koeffizienten begegnen, die Ihnen nicht bekannt sind, kann auch die Online-Hilfe von SPSS konsultiert werden (Abb. 13.12).

Fundiertere Erläuterungen erhält man in der Literatur (vgl. etwa Nagl, 1991, Kap. 3; insbesondere sei auf Tabelle 3.1 (Zusammenhangsmaße) verwiesen).



[Base System](#) > [Analysieren von Daten](#) > [Kreuztabellen](#) >

[Zurück](#) [Weiter](#)

Kreuztabellen: Statistik

Chi-Quadrat. Für Tabellen mit zwei Zeilen und zwei Spalten wählen Sie **Chi-Quadrat** aus, um das Pearson-Chi-Quadrat, das Likelihood-Quotienten-Chi-Quadrat, den exakten Test nach Fisher und das korrigierte Chi-Quadrat nach Yates (Kontinuitätskorrektur) zu berechnen. Für 2×2 -Tabellen wird der exakte Test nach Fisher berechnet, wenn eine Tabelle, die nicht aus fehlenden Zeilen oder Spalten einer größeren Tabelle entstanden ist, eine Zelle mit einer erwarteten Häufigkeit von weniger als 5 enthält. Für alle anderen 2×2 -Tabellen wird das korrigierte Chi-Quadrat nach Yates berechnet. Für Tabellen mit einer beliebigen Anzahl von Zeilen und Spalten wählen Sie **Chi-Quadrat** aus, um das Pearson-Chi-Quadrat und das Likelihood-Quotienten-Chi-Quadrat zu berechnen. Wenn beide Tabellenvariablen quantitativ sind, ergibt **Chi-Quadrat** den Zusammenhangstest linear-mit-linear.

Korrelationen. Für Tabellen, in denen sowohl Zeilen als auch Spalten geordnete Werte enthalten, ergeben die **Korrelationen** den Korrelationskoeffizienten nach Spearman, also Rho (nur numerische Daten). Der Korrelationskoeffizient nach Spearman ist ein Zusammenhangsmaß zwischen den Rangordnungen. Wenn beide Tabellenvariablen (Faktoren) quantitativ sind, ergibt sich unter **Korrelationen** der Korrelationskoeffizient nach Pearson, r , der ein Maß für den linearen Zusammenhang zwischen den Variablen darstellt.

Nominal. Für nominale Daten (ohne implizierte Reihenfolge, wie beispielsweise katholisch, protestantisch, jüdisch) können Sie **Kontingenzkoeffizient**, **Phi** (Koeffizient) und **Cramér-V**, **Lambda** (symmetrische und asymmetrische Lambdas sowie Goodman-und-Kruskal-**Tau**) und **Unsicherheitskoeffizient** auswählen.

[Einzelheiten ausblenden](#)

- **Kontingenzkoeffizient.** Ein auf der Chi-Quadrat-Statistik basierendes Zusammenhangsmaß. Dieser Koeffizient liegt immer zwischen 0 und 1, wobei 0 angibt, dass kein Zusammenhang zwischen Zeilen- und Spaltenvariable besteht und Werte nahe 1 auf einen starken Zusammenhang zwischen den Variablen hindeuten. Der maximale Wert hängt von der Anzahl der Zeilen und Spalten in der Tabelle ab.
- **Phi und Cramer-V.** Phi ist ein auf der Chi-Quadrat-Statistik basierendes Zusammenhangsmaß. Es ergibt sich als Wurzel aus dem Quotienten aus Chi-Quadrat und dem Stichprobenumfang. Cramer-V ist ebenfalls ein Zusammenhangsmaß auf der Basis der Chi-Quadrat-Statistik.
- **Lambda.** Ein Zusammenhangsmaß für die proportionale Fehlerreduktion, wenn Werte der unabhängigen Variablen zur Vorhersage von Werten der abhängigen Variablen verwendet werden. Der Wert 1 bedeutet, dass die abhängige Variable durch die unabhängige Variable vollständig vorhergesagt werden kann. Der Wert 0 bedeutet, dass die Vorhersage der abhängigen Variablen durch die unabhängige Variable nicht unterstützt wird.
- **Unsicherheitskoeffizient.** Ein Zusammenhangsmaß, das die proportionale Fehlerreduktion angibt, wenn Werte einer Variablen zur Vorhersage von Werten der anderen Variablen verwendet werden. Ein Wert von 0,83 gibt z. B. an, dass die Kenntnis einer Variablen den Fehler bei der Vorhersage der Werte der anderen Variablen um 83 % reduziert. Das Programm berechnet beide Versionen des Unsicherheitskoeffizienten, die symmetrische und die asymmetrische.

Abbildung 13.12: SPSS-Online-Hilfe: Zusammenhangsmaße der nominalen Variablen *Geschlecht* und *Berufsgruppe*.

13.4 Einfache Unterschieds-Hypothesen

0-1-codierte Variablen können auch als Regressorvariablen in Regressionsgleichungen der Form (Kap. 5)

$$Y_n = \alpha + \beta X_n + \epsilon_n \quad (13.41)$$

eingesetzt werden. Man hat dann die beiden Formen

$$Y_n = \alpha + \epsilon_n; X_n = 0 \quad (13.42)$$

$$Y_n = \alpha + \beta + \epsilon_n; X_n = 1. \quad (13.43)$$

Daher kann β als Mittelwertsunterschied von 2 Gruppen interpretiert werden.

Schätzt man den Regressionsparameter mit der KQ-Methode (kleinste Quadrate), so ergibt sich

KQ-Schätzwerte

$$\hat{\alpha} = \bar{y} - \hat{\beta} \bar{x} \quad (13.44)$$

$$\hat{\beta} = \frac{\frac{1}{N} \sum_n y_n x_n - \bar{y} \bar{x}}{\frac{1}{N} \sum_n (x_n - \bar{x})^2} \quad (13.45)$$

$$= \frac{s_{xy}}{s_x^2} = r_{xy} \frac{s_y}{s_x}. \quad (13.46)$$

In der letzten Gleichung wurde der Produkt-Moment-Korrelationskoeffizient

$$r_{xy} = \frac{\sum_n (x_n - \bar{x})(y_n - \bar{y})}{\sqrt{\sum_n (x_n - \bar{x})^2} \sqrt{\sum_n (y_n - \bar{y})^2}} \quad (13.47)$$

eingesetzt.

Im Fall von 0-1-Variablen ergibt sich folgende Vereinfachung ($\sum x_n := N_1$)

Produkt-Moment-Korrelation

$$\bar{x} = N_1/N = p \quad (13.48)$$

$$\bar{y} = \frac{1}{N} \sum_n y_n = \frac{1}{N} (N_0 \bar{y}_0 + N_1 \bar{y}_1) \quad (13.49)$$

$$= (1-p) \bar{y}_0 + p \bar{y}_1 \quad (13.50)$$

$$\hat{\beta} = \frac{p \bar{y}_1 - p \bar{y}}{p(1-p)} = \bar{y}_1 - \bar{y}_0 \quad (13.51)$$

$$\hat{\alpha} = \bar{y} - (\bar{y}_1 - \bar{y}_0)p = \bar{y}_0 \quad (13.52)$$

(bitte nachrechnen!).

Daher ist der geschätzte Regressionsparameter der Mittelwertsunterschied der Gruppen $X = 0$ und $X = 1$ und der Achsenabschnitt ist der Mittelwert in Gruppe $X = 0$.

Der **Produkt-Moment-Korrelationskoeffizient zwischen stetigen und 0-1-Variablen** ist somit

$$r_{xy} = \frac{\hat{\beta} s_{*x}}{s_{*y}} = \frac{\bar{y}_1 - \bar{y}_0}{s_{*y}} \sqrt{p(1-p)} \quad (13.53)$$

d.h. die standardisierte Mittelwertsdifferenz der Y -Variable multipliziert mit der Standardabweichung der 0-1-Variable X . Bei gleicher Gruppengröße $N_0 = N_1 = N/2$ ist $p = 1/2$ und somit $r_{xy} = \frac{\bar{y}_1 - \bar{y}_0}{2s_{*y}}$.

Die Formel (13.13) wurde hier eingesetzt, um Faktoren $\sqrt{(N-1)/N}$ zu vermeiden.

Beispiel 13.2 (Filialgestaltung von Bankfilialen: Fortsetzung)

Die Variable ALTNEU unterscheidet das alte bzw. neue Filialkonzept (Schalter vs. offene Gestaltung), während ZufriedU die Kundenzufriedenheit erfaßt. Abb. 13.13 zeigt die Parameterschätzungen der Regression

$$\text{ZufriedU}_n = \alpha + \beta * \text{ALTNEU}_n + \epsilon_n \quad (13.54)$$

Die Konstante ist $\hat{\alpha} = .207$ und der Gruppenunterschied $\hat{\beta} = -.356$. Der F -Test für die Hypothese $H_0 : \beta = 0$ ergibt einen p -Wert von .001 [mit dem Quantil $F(1, n - 2) = 11.226$]. Vergleichen man dies mit dem t -Test für die Hypothese

$$H_0 : \mu_0 = \mu_1 \quad (13.55)$$

(Gleichheit der Erwartungswerte der Variablen ZufriedU für die Gruppen ALTNEU = 0, 1), so ergibt sich das Resultat aus Abb. 13.14.

Das Gruppenmittel $\bar{y}_0 = 0.2071$ entspricht der Konstanten $\hat{\alpha}$ aus der Regression, während die Differenz $\bar{y}_1 - \bar{y}_0 = -0.355592$ betragsmäßig der Schätzung $\hat{\beta} = -.356$ entspricht (Glg. 13.51–13.52).

Weiterhin kann man aus der t -Test-Tabelle folgende Größen entnehmen:

$$N_0 = 148; N_1 = 206; N = N_0 + N_1 = 354;$$

$$p = N_1/N = 0.581921;$$

$$s_{*x} = \sqrt{p(1-p)} = 0.493243;$$

$$\bar{y}_0 = \hat{\alpha} = .2071; \bar{y}_1 = -.1488;$$

$$\hat{\beta} = \bar{y}_1 - \bar{y}_0 = -.1488 - .2071 = -0.3559;$$

$$s_0 = .81898; s_1 = 1.08985; s_{*y} \approx \sqrt{s_0^2(1-p) + s_1^2p} = 0.985701;$$

$$r = \hat{\beta}s_{*x}/s_{*y} = -0.178092;$$

$$r^2 = 0.0317167$$

Auch gilt $t^2(352) = 3.351^2 = 11.226 = F(1, 352)$, da $N = 354$ die gesamte Stichprobengröße ist (vgl. Kap. 1.3.4.2).

Die Vorgehensweisen sind also äquivalent.

SPSS-Menübefehle:

Analysieren/Regression/Kurvenanpassung/...

Analysieren/Mittelwerte vergleichen/T-Test bei unabhängigen Stichproben/...



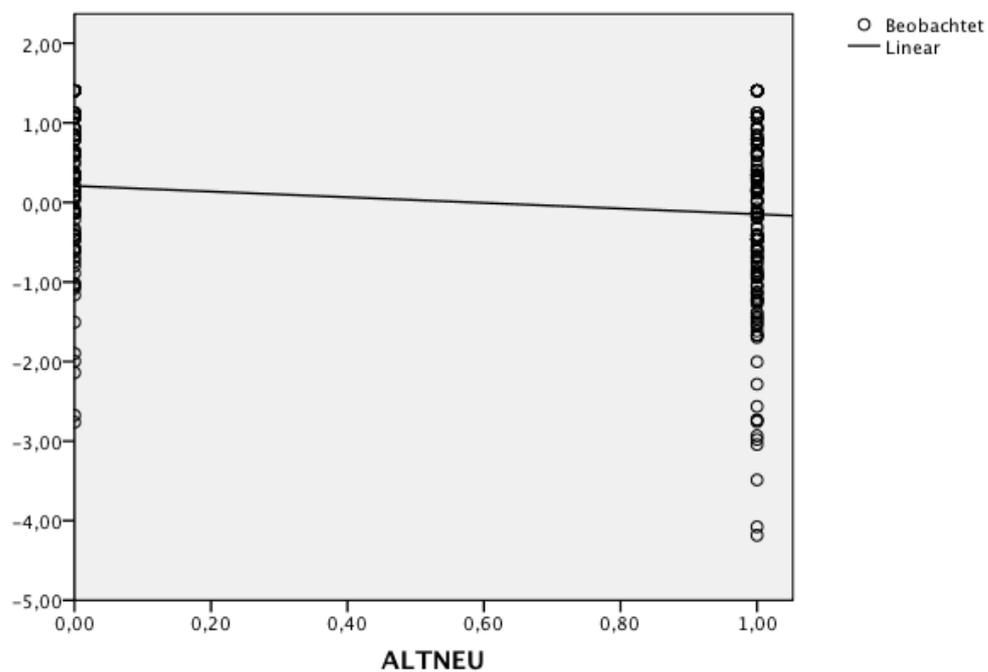
Modellzusammenfassung und Parameterschätzer

Abhängige Variable: ZufriedU

Gleichung	Modellzusammenfassung					Parameterschätzer	
	R-Quadrat	F	Freiheitsgrad e 1	Freiheitsgrad e 2	Sig.	Konstante	b1
Linear	,031	11,226	1	352	,001	,207	-,356

Die unabhängige Variable ist ALTNEU.

ZufriedU

Abbildung 13.13: Regressionsanalyse mit der Indikatorvariablen *altneu* als unabhängiger Variable.

```
T-TEST GROUPS=ALTNEU (0 1)
/MISSING=ANALYSIS
/VARIABLES=ZufriedU
/CRITERIA=CI (.95).
```

T-Test

[DatenSet5] /HermannTex/Multi/MasterA-Modul/Daten/SPSS-Datensätze/main2.xls.sav

Gruppenstatistiken

	ALTNEU	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
ZufriedU	,00	148	,2071	,81898	,06732
	,100	206	-,1488	1,08985	,07593

Test bei unabhängigen Stichproben

	Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					95% Konfidenzintervall der Differenz	
	F	Signifikanz	T	df	Std. 12-seitig	Mittlere Differenz	Standardfehler der Differenz	Untere	Oberer
ZufriedU	8,691	,003	3,351	352	,001	,35592	,10623	,14700	,56484
			3,507	351,271	,001	,35592	,10148	,15634	,55550

Abbildung 13.14: t-Test mit der Indikatorvariablen *altneu* als Gruppenvariable.

Zusammenfassend man sagen:

Dichotome nominalskalierte Variablen X sollten als **Indikator-Variablen** (0-1, dummy, binäre Variablen) codiert werden, da ...

Indikator-Variablen

1. Mittelwerte und Streuungen interpretiert werden können.
2. In Regressionsgleichungen der Regressions-Koeffizient $\hat{\beta}$ als Mittelwertsunterschied $= \bar{Y}_1 - \bar{Y}_0$ der durch X definierten Gruppen interpretiert werden kann.
Außerdem ist die Korrelation ein Maß für den standardisierten Mittelwertsunterschied.
3. In Regressionsgleichungen Interaktionseffekte der Form $\beta_{12}X_1X_2$ definiert werden können
(simultanes Vorliegen von $X_1 = 1, X_2 = 1$).
4. Sich gruppenspezifische Steigungen $(\beta + \beta_1X_1)X$
für die stetige Variable X definieren lassen .
5. Produkt-Moment-Korrelationen von X mit anderen Indikatoren Y als ϕ -Koeffizient oder als biseriale Korrelation ($Y =$ Intervallskaliert) gelesen werden können (Bsp. 14.2).

Kapitel 14

Gesichtspunkte bei statistischen Analysen

14.1 Übersicht

In diesem Abschnitt sollen einige Gesichtspunkte tabellarisch aufgelistet werden, die bei der Datenanalyse nützlich sind. Sie bilden sozusagen ein **Koordinatensystem**, das hilft, die einzelnen Vorgehensweisen und Verfahren systematisch einzuordnen. Bei der Vielzahl statistischer Konzepte und Methoden ist dies unabdingbar. Anstatt die einzelnen Methoden aufzulisten, werden Sie im folgenden anhand dieser Systematik klassifiziert. Die meisten Begriffe sollten Ihnen aus früheren Kapiteln bekannt sein. Gewisse Gesichtspunkte sind ähnlich oder sogar identisch, andere sehr unterschiedlich:

Systematik

- **Skalen-Niveau**

Nominal, Ordinal, Metrisch (Intervall, Verhältnis)

Beispiele und zulässige Relationen:

- Nominal:
Geschlecht (männlich/weiblich): $x = y, x \neq y$
- Ordinal (Rang):
Schulnoten (1, ..., 6): $x < y, x = y$
Ratings von Firmen: $AAA > AA > \dots > B \dots$
- Intervall:
Temperaturen (Celsius): $\Delta x / \Delta y$

- Verhältnis:
Temperaturen (Kelvin): x/y
Alter (60.9 Jahre): x/y
- **qualitative/quantitative Merkmale**
 - entspricht der Unterscheidung nominal/metrisch (oder ordinal).
 - ordinal = qualitativ geordnet (A besser als B ; $A > B$)
- **diskrete/metrische Merkmale**
 - diskret = Nominal- und Ordinal-Skalen
 - metrisch = Intervall- und Verhältnis-Skala
- **kategoriale/kontinuierliche Merkmale**
entspricht diskret/metrisch
- **Statistiken und Skalenniveau**
 - Meßniveau berücksichtigen!
 - *Beispiel:*
Mittelwert nur sinnvoll bei metrischen Variablen.
Schulnoten sind nur ordinal.
- **Symmetrische Zusammenhänge: $Y \longleftrightarrow X$**
 - Korrelationen $\text{Corr}(X, Y)$
 - Es gibt nur Zusammenhänge (Assoziationen),
keine Ursache-Wirkungs-Relationen.
- **Scheinkorrelationen $Y \longleftrightarrow Z \longleftrightarrow X$**
 - Die Korrelation zwischen X und Y ist durch die Wirkung einer Drittvariable Z vermittelt.
 - Korrigierte Variablen (Prognose-Fehler, Innovationen)
 $\tilde{X} = X - \hat{X}(Z)$
 $\tilde{Y} = Y - \hat{Y}(Z)$
durch Berücksichtigung der Information in Z
 $\hat{X}(Z) = \text{Prognose von } X \text{ durch } Z$

- Partielle Korrelation: $\text{Corr}(\tilde{X}, \tilde{Y})$
- **Asymmetrische Zusammenhänge: $Y \longleftarrow X$**
 - Abhängige Variablen (AV, Y) \longleftarrow Unabhängige Variablen (UV, X)
 - Bedingte Verteilung $P(Y = y|X = x) = f(y|x)$
 - entspricht der logischen Implikation $B \longleftarrow A$
- **Optimale Prognose**
 - $\hat{Y}(X)$: Prognose von Y durch X
 - $\hat{Y} = E[Y|X]$:
bedingter Erwartungswert von Y , gegeben X
(bekannte Information)
 - $\tilde{Y} = Y - \hat{Y}$: Prognosefehler
 - $E[\tilde{Y}^2] = E[\text{Var}(Y|X)]$
Prognosefehler-Varianz (Residual-Streuung)
 - $\text{Var}(Y) = E[\text{Var}(Y|X)] + \text{Var}(E[Y|X])$
Zerlegung:
Totale Streuung = Residual-Streuung + Erklärte Streuung
- **Optimale lineare Prognose**
 - $E[Y|X] = E[Y] + \text{Cov}(Y, X)\text{Cov}(X, X)^{-1}(X - E[X])$
 $:= \alpha + \beta X$
lineare Prognose von Y durch X (lineare Regression)
 - $\text{Var}(Y|X) = \text{Var}(Y) - \text{Cov}(Y, X)\text{Cov}(X, X)^{-1}\text{Cov}(X, Y)$
bedingte Varianz ist unabhängig von X
- **univariate, bivariate, multivariate Analyse**
 - $P(X = x)^1, P(X = x, Y = y), P(X = x, Y = y, Z = z)$:
univariate, bivariate und multivariate Verteilung.

¹Bei stetigen Variablen nimmt man die Dichtefunktion $f(x)dx = P(x \leq X \leq x + dx)$ etc.

- $Y \leftarrow X$:
 $P(Y = y|X = x)$:
bedingte Verteilung von Y , gegeben X .
- $Y \leftarrow (X, Z)$:
 $P(Y = y|X = x, Z = z)$:
bedingte Verteilung von Y , gegeben X und Z .
- **Wahrscheinlichkeit/relative Häufigkeit**
 - $P(A)$: Wahrscheinlichkeit, daß das Ereignis A eintritt (Würfel zeigt 6).
 - $f_n(A)$: relative Häufigkeit, daß das Ereignis A bei n Würfeln eintritt.
In Stichproben kann nur $f_n(A)$ berechnet werden.
 - In großen Stichproben strebt $f_n(A)$ gegen $P(A)$.
- **Latente/manifeste Variablen**
 - latente Variablen = Faktoren
(Konstrukte, verborgene Variablen)
 - manifeste Variablen = Indikatoren = meßbare Größen
- **Konstrukte (Faktoren)/Indikatoren**
siehe Latente/manifeste Variablen.
- **Meßfehler, wahrer Wert, Reliabilität, Abschwächung**
 - manifeste Variablen = latente Variablen + Meßfehler
 - Reliabilität = wahre Streuung/gesamte Streuung
 - Abschwächung: Meßfehler reduzieren Korrelationen und Mittelwertsunterschiede (t -Statistiken).

Die einzelnen Gesichtspunkte werden im folgenden genauer erläutert.

	Symbol	nominal	ordinal	metrisch
Häufigkeiten				
relative Häufigkeit	$f(x)$	x	x	x
kumulierte relative Häufigkeit	$F(x)$		x	x
Lage-Parameter				
Modus	x_{mod}	x	x	x
Median	x_{med}		x	x
Mittelwert	\bar{x}			x
Streuungs-Parameter				
Spannweite	SP			x
Quartilsabstand	QA			x
Standard-Abweichung	s			x

Tabelle 14.1: Univariate Statistiken nach Skalenniveau.

14.2 Univariate Statistiken nach Skalen-Niveaus

Sind die Variablenausprägungen numerisch codiert (etwa männlich=1, weiblich = 2), so können problemlos Statistiken wie der Mittelwert, Standardabweichung etc. berechnet werden, allerdings ist die Frage, ob sich diese Werte sinnvoll interpretieren lassen.

Wie Tabelle 14.1 zeigt, lassen sich bei nominalem Skalenniveau nur die relative Häufigkeit und der Modalwert sinnvoll berechnen. Für die kumulierte Häufigkeitsverteilung $F(x) = \sum_{y \leq x} f(y)$ ist eine Ordnungsrelation $x < y$, also mindestens ordinales Skalenniveau erforderlich.

Beispiel 14.1 (Filialgestaltung von Bankfilialen: Fortsetzung)

Die nominale Variable Beruf hat die Ausprägungen 1–6 mit den Wertelabels ‘Schüler’, ‘Arbeiter’, ‘Hausfrau’, ‘Selbständig’, ‘Rentner’, ‘ohne Beschäftigung’. Ruft man in SPSS die Prozedur

**Interpretation von
Statistiken**

Analysieren/Deskriptive Statistiken/Häufigkeiten

Modus

auf, so ergibt sich folgendes Resultat (Abb. 14.1). SPSS berechnet klaglos einen Mittelwert von 3.31, was sich durch die stark vertretenen Ausprägungen 2 = Arbeiter und 5 = Rentner erklären läßt. Trotzdem ist dieser Wert nicht interpretierbar. Sinnvoll ist nur der Modalwert (Modus) = 2, also die Kategorie mit der größten Häufigkeit (37.5%). Die kumulierte Häufigkeit ist problematisch, da die Ausprägungen keine natürliche Reihenfolge aufweisen (etwa relative Häufigkeit der Personen bis zur Ausprägung Hausfrau = 57.4%).

Indikatoren

Wie schon ausführlich erläutert (Abschnitt 13.3), sollte man dichotome (zweiwertige) Variablen mit 0-1 codieren (Indikatoren, binäre Variablen), was zu einer stark verbesserten Interpretierbarkeit führt.



14.3 Bivariate Korrelationen nach Skalen-Niveaus

Auch bei Zusammenhangsmaßen sind die Skalenniveaus der beteiligten Variablen entscheidend für die Auswahl:

Symmetrische Verfahren $Y \longleftrightarrow X$			
	Y		
X	nominal	ordinal	metrisch
nominal	χ^2, K, ϕ, v	χ^2, K	χ^2, K
ordinal	χ^2, K	r_S, τ, γ	r_S, τ, γ
metrisch	χ^2, K	r_S, τ, γ	r_{PM}

In der Tabelle sind folgende Maße eingetragen:

- χ^2 = Chi-Quadrat-Koeffizient, K = Kontingenzkoeffizient (auch der ϕ -Koeffizient und Cramér's v).
- τ = Kendall-Tau
- γ = Goodman-Kruskal-Gamma

Statistiken

BERUF		
N	Gültig	378
	Fehlend	1
Mittelwert		3,31
Standardfehler des Mittelwertes		,081
Median		3,00
Modus		2
Standardabweichung		1,573
Varianz		2,475
Schiefe		,218
Standardfehler der Schiefe		,125
Kurtosis		-1,457
Standardfehler der Kurtosis		,250
Spannweite		5
Minimum		1
Maximum		6
Summe		1252

BERUF

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Schüler	32	8,4	8,5	8,5
	Arbeiter	142	37,5	37,6	46,0
	Hausfrau	43	11,3	11,4	57,4
	Selbständig	23	6,1	6,1	63,5
	Rentner	113	29,8	29,9	93,4
	ohne Beschäftigung	25	6,6	6,6	100,0
	Gesamt	378	99,7	100,0	
Fehlend	System	1	,3		
	Gesamt	379	100,0		

Statistics 17.0 Datei Bearbeiten Ansicht Daten Transformieren Einfügen Format A

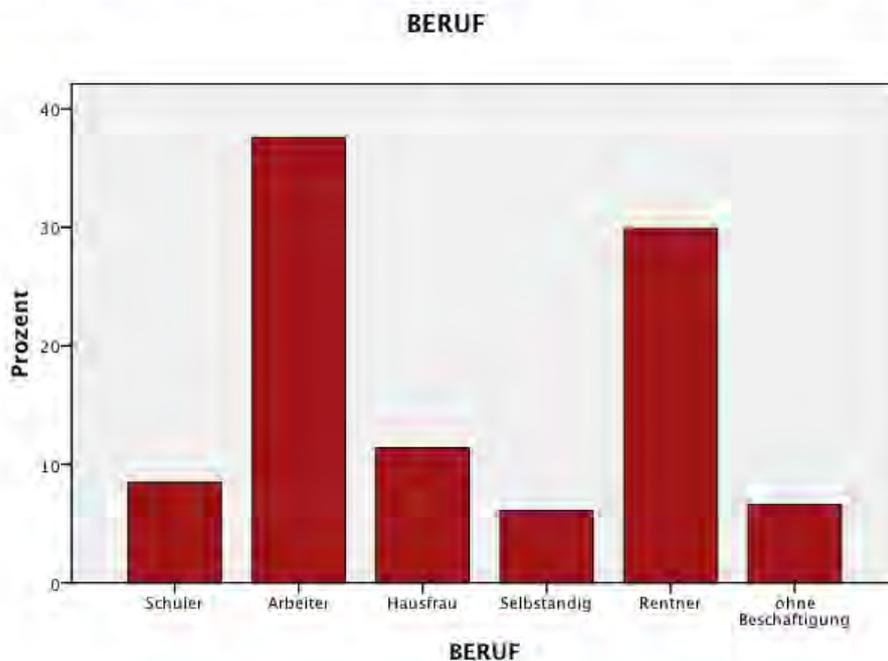


Abbildung 14.1: Statistiken und Häufigkeiten der nominalen Variablen *Berufsgruppe*.

- r_S = Spearman-Rang-Korrelation
- r_{PM} = Produkt-Moment-Korrelation (Bravais-Pearson)

Bei 0-1-Codierung der nominalen Variablen kann der Produkt-Moment-Korrelationskoeffizient (Bravais-Pearson) in Kombination mit stetigen Variablen berechnet werden (biserial Korrelation) (vgl. ausführlich Abs. 13.4, sowie Bortz, Kap. 6).

Sind beide Variablen 0-1-codiert, so ergibt sich

$$r_{PM} = \phi = \sqrt{\chi^2/N}. \quad (14.1)$$

Beispiel 14.2 (Korrelation bei 0-1-Variablen)

Für die beiden 0-1-Variablen X und Y ist folgende Kreuztabelle der relativen Häufigkeiten a, b, c, d gegeben:

gemeinsame Häufigkeiten f_{ij}			
X/Y	0	1	$f_{i.}$
0	a	b	$a + b$
1	c	d	$c + d = \bar{x}$
$f_{.j}$	$a + c$	$b + d = \bar{y}$	$a + b + c + d = 1$

Korrelation

Der Korrelationskoeffizient

$$r_{xy} = \frac{s_{*xy}}{s_{*x}s_{*y}} \quad (14.2)$$

$$s_{*xy} = N^{-1} \sum_n (x_n - \bar{x})(y_n - \bar{y}) \quad (14.3)$$

$$s_{*x} = \sqrt{N^{-1} \sum_n (x_n - \bar{x})^2} \quad (14.4)$$

$$s_{*y} = \sqrt{N^{-1} \sum_n (y_n - \bar{y})^2} \quad (14.5)$$

läßt sich mit Hilfe der obigen Kreuztabelle folgendermaßen umschreiben:

$$r_{xy} = \frac{d - (c + d)(b + d)}{\sqrt{(c + d)(a + b)(b + d)(a + c)}}, \quad (14.6)$$

da $c + d = \bar{x} = N^{-1} \sum_n x_n$, $1 - c - d = a + b$ und $d = N^{-1} \sum_n x_n y_n$.

Für den Zähler gilt

$$d - (c + d)(b + d) = d - cb - cd - db - d^2 \quad (14.7)$$

$$d - d^2 = d(1 - d) = d(a + b + c) \quad (14.8)$$

$$d - cb - cd - db - d^2 = ad - bc. \quad (14.9)$$

Damit findet man den ϕ -Koeffizienten

$$r_{xy} = \frac{ad - bc}{\sqrt{(c + d)(a + b)(b + d)(a + c)}} = \phi. \quad (14.10)$$

Setzt man in

$$\chi^2 = N \sum_{ij} \frac{(f_{ij} - f_{i.} f_{.j})^2}{f_{i.} f_{.j}} \quad (14.11)$$

die Werte aus der Kreuztabelle ein (etwa $f_{00} = a$; $f_{01} = b$; $f_{10} = c$; $f_{11} = d$; $f_{1.} = c + d$ etc.), so ergibt sich

$$\chi^2/N = \phi^2. \quad (14.12)$$



Die ordinalen Zusammenhangsmaße Kendall- $\tau_{a,b}$ und Goodman-Kruskal- γ sind über Konkordanzen bzw. Diskordanzen definiert (ausführlich siehe Nagl, 1992, Kap. 3). Hierbei werden Paare (i, j) von Untersuchungseinheiten genommen und untersucht, ob für die ordinale Relation

ordinale Zusammenhangsmaße

$$x_i < x_j \quad (14.13)$$

$$y_i < y_j \text{ (Konkordanz)} \quad (14.14)$$

$$y_i > y_j \text{ (Diskordanz)} \quad (14.15)$$

auftritt (entsprechend für $x_i > x_j$). Die entsprechenden Anzahlen n_{konk} , n_{disk} bei allen Paarvergleichen lassen sich aus den Häufigkeiten

	<	=	>
<	$n_{<<}$	$n_{<=}$	$n_{<>}$
=	$n_{=<}$	$n_{==}$	$n_{=>}$
>	$n_{><}$	$n_{>=}$	$n_{>>}$

leicht berechnen. Aufgrund der Radialsymmetrie der obigen Kreuztabelle gilt

$$n_{konk} = n_{<<} = n_{>>} \quad (14.16)$$

$$n_{disk} = n_{><} = n_{<>} \quad (14.17)$$

(Übung).

Der Gamma-Koeffizient ist dann

**Gamma-
Koeffizient
(Goodman-
Kruskal)**

$$\gamma = \frac{F(-) - F(+)}{F(-)} = \frac{(n_{konk} + n_{disk}) - 2n_{disk}}{n_{konk} + n_{disk}} \quad (14.18)$$

$$= \frac{n_{konk} - n_{disk}}{n_{konk} + n_{disk}}, \quad (14.19)$$

wobei

- das Fehlermaß $F(+)$ (mit Prädiktor X) die diskordanten Vergleiche (zeilenweise) summiert ($n_{<>} + n_{><} = 2n_{disk}$) und
- das Fehlermaß $F(-)$ (ohne Prädiktor X) die Anzahlen der kleiner-Prädiktion $y_i < y_j$ aufsummiert ($n_{<<} + n_{><} = n_{konk} + n_{disk}$).

Vertauscht man die Rolle von X und Y , ergibt sich das gleiche Resultat, sodaß ein symmetrischer Koeffizient resultiert. Das γ -Maß unterstellt, daß nur Paarvergleiche gezählt werden, bei denen keine Bindungen (ties) vorkommen, also Vergleiche der Form $(x_i = x_j, y_i \leq y_j)$.

**Tau-Koeffizient
(Kendall)**

Das gleiche gilt für Kendall- τ_a , während bei τ_b auch Bindungen gezählt werden (etwa $n_{>=} = \text{Anzahl}(x_i > x_j, y_i = y_j)$).

Die Tau-Koeffizienten sind folgendermaßen definiert:

$$\tau_a = \frac{n_{konk} - n_{disk}}{N(N-1)/2} \quad (14.20)$$

$$\tau_b = \frac{n_{konk} - n_{disk}}{\sqrt{(n_{konk} + n_{disk} + n_{>=})(n_{konk} + n_{disk} + n_{>})}}. \quad (14.21)$$

14.4 Asymmetrische Verfahren nach Skalen-Niveaus

Symmetrische Maße geben einen Zusammenhang (Assoziation), lassen sich jedoch **nicht** als Wirkungsweisen interpretieren. Dazu sind asymmetrische Ansätze in der Lage, wo zwischen abhängigen ($AV = Y$) und unabhängigen Variablen ($UV = X$) unterschieden wird:

Asymmetrische Verfahren $Y = f(X)$		
	AV = Y	
UV = X	diskret	metrisch
diskret	Kreuztabellen	Varianz-Analyse
metrisch	kategoriale Regression	Regression

14.4.1 Implikation

Die aus der Aussagenlogik bekannte Implikation

$$B \Leftarrow A \quad (14.22)$$

mit der Wahrheitstafel (Tabelle 14.2) nimmt für die Ereignisse (Mengen) $A = \{\omega | X(\omega) = 1\}$, $B = \{\omega | Y(\omega) = 1\}$ folgende probabilistische Formulierung an:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \quad (14.23)$$

Implikation

Ereignisse

Bayes-Formel

Wenn A , dann B		
$A \Rightarrow B$	$B = 1$	$B = 0$
$A = 1$	1	0
$A = 0$	1	1

Tabelle 14.2: Kreuztabelle der Aussagen $A =$ 'Blutalkoholspiegel > 0.5 Promille', $B =$ 'Verminderte Reaktion'. *Wahrheitstafel der Implikation $A \Rightarrow B$* ; 1 = wahr, 0 = falsch.

$P(B A)$		
	B	\bar{B}
A	1	0
\bar{A}	1/2	1/2

Tabelle 14.3: Kreuztabelle der Ereignisse $A = \{X = 1\}$ ('Blutalkoholspiegel > 0.5 Promille') und $B = \{Y = 1\}$ ('Verminderte Reaktion'). Zeilenweise bedingte Wahrscheinlichkeiten $P(Y = y|X = x)$.

$P(A B)$		
	B	\bar{B}
A	2/3	0
\bar{A}	1/3	1

Tabelle 14.4: Kreuztabelle der Ereignisse $A = \{X = 1\}$ ('Blutalkoholspiegel > 0.5 Promille', $B = \{Y = 1\}$ ('Verminderte Reaktion'). Spaltenweise bedingte Wahrscheinlichkeiten $P(X = x|Y = y)$ bei Randverteilung $P(X = 1) = P(A) = 0.5$. Prognose der Prämisse (A) aus der Conclusio (B).

$P(A \cap B)$		
	B	\bar{B}
A	$a = 1/2$	$b = 0$
\bar{A}	$c = 1/4$	$d = 1/4$

Tabelle 14.5: Kreuztabelle der Variablen X und Y . Gemeinsame Wahrscheinlichkeiten $P(A \cap B)$.

(bedingte Wahrscheinlichkeit für das Ereignis B , wenn A bereits eingetreten ist).

Insbesondere gilt für

$$P(B|A) = 1, \tag{14.24}$$

daß B auf jeden Fall eintritt, wenn A vorliegt.

14.4.2 Kontraposition

In umgekehrter Schlußrichtung gilt dann

$$P(\bar{A}|\bar{B}) = 1, \tag{14.25}$$

da

$$\begin{aligned} P(\bar{A}|\bar{B}) &= 1 - P(A|\bar{B}) \\ &= 1 - P(\bar{B}|A)P(A)/P(\bar{B}) = 1, \end{aligned}$$

wegen $P(\bar{B}|A) = 1 - P(B|A) = 1 - 1 = 0$. Dies läßt sich auch direkt aus Tabelle 14.5 ablesen, da $P(A \cap \bar{B}) = b = 0$ ist.

Dies entspricht dem sogenannten *Kontrapositionsgesetz (modus tollens)*

$$\bar{A} \iff \bar{B}. \tag{14.26}$$

**modus tollens
 Kontrapositions-
 gesetz**

(\bar{A} = Komplement von A , vgl. Tab. 14.4).

Angewandt auf die Tabellen 14.3–14.4 ergibt sich:

Wenn $\bar{B} = \{Y = 0\}$ (keine verminderte Reaktion),
 dann muß $\bar{A} = \{X = 0\}$ vorliegen ('Blutalkoholspiegel ≤ 0.5 Promille').

$P(X = x Y = y)$		
	$Y = 1$	$Y = 0$
$X = 1$	$\frac{a}{a+c}$	$\frac{b}{b+d} = 0$
$X = 0$	$\frac{c}{a+c}$	$\frac{d}{b+d} = 1$

Tabelle 14.6: Kreuztabelle der Variablen X und Y . Spaltenweise bedingte Wahrscheinlichkeiten $P(X = x|Y = y)$.

Der modus tollens ergibt sich aus der Tabelle der zeilenweisen bedingten Wahrscheinlichkeiten, obwohl die Wahrscheinlichkeitstabelle nur unvollständig spezifiziert ist.

Man weiß zunächst nur, daß $P(X = 1, Y = 0) = P(A \cap \bar{B}) = b = 0$.

Will man weitere Aussagen treffen, müssen die anderen Einträge $a = P(A \cap B), c = P(\bar{A} \cap B), d = P(\bar{A} \cap \bar{B})$ der gemeinsamen Wahrscheinlichkeits-Tabelle 14.5 spezifiziert werden.

Im Gegensatz zur Aussagenlogik, wo aus dem Vorliegen der Conclusio ($B = 1 = \text{wahr}$) nichts über die Gültigkeit der Prämisse A gesagt werden kann (die Implikation ist wahr sowohl für $A = 1$ als auch $A = 0$; vgl. Tab. 14.2), lassen sich im probabilistischen Fall auch Aussagen der Form $P(X = x|Y = y)$ machen, die man aus der spaltenweisen Tabelle der bedingten Wahrscheinlichkeiten ablesen kann (Tabellen 14.4, 14.6).

Bei einer a priori-Wahrscheinlichkeit $P(X = 1) = P(A) = 0.5$ der Prämisse $A = \text{'Blutalkoholspiegel} > 0.5 \text{ Promille}'$ ist die bedingte Wahrscheinlichkeit für A , gegeben das Ereignis $B = \text{'Verminderte Reaktion}'$

$$P(A|B) = P(B|A)P(A)/P(B) = 1 \cdot (1/2)/(3/4) = 2/3 \quad (14.27)$$

(siehe Tabelle 14.4). Diese plausible Tatsache wird im folgenden Abschnitt weiter verdeutlicht.

Zum Rückschluß von Y auf X wird die Bayes-Formel

**Wahrscheinlichkeit
der Ursache**

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)} \quad (14.28)$$

benutzt (vgl. McGrayne, 2014). Der Rückschluß erfolgt also mit einer Wahrscheinlichkeit, die i.a. kleiner als 1 ist (da $P(A \cap B) \leq P(B)$).

Insbesondere ergibt sich für $P(B|A) = 1$ (A impliziert B) die Form

$$P(A|B) = \frac{P(A)}{P(B)} \geq P(A). \quad (14.29)$$

Für $P(A|B) = 1$ muß gelten

$$1 = \frac{P(A)}{P(B)} = \frac{a}{a+c} \quad (14.30)$$

und somit $c = P(\bar{A} \cap B) = 0$.

Die Kreuztabelle der gemeinsamen Wahrscheinlichkeiten (14.5) ist daher diagonal, was der Äquivalenz $A \Leftrightarrow B$ ($A \Rightarrow B$ und $B \Rightarrow A$) entspricht.

Äquivalenz

Probabilistisch ausgedrückt gilt also für $P(A \cap \bar{B}) = b = 0$ und $P(\bar{A} \cap B) = c = 0$ (diagonale Kreuztabelle) in beiden Richtungen

$$P(A|B) = 1 \quad (14.31)$$

$$P(B|A) = 1. \quad (14.32)$$

Der Rückschluß von \bar{B} auf \bar{A} läßt sich als

$$P(\bar{A}|\bar{B}) = P(\bar{B}|\bar{A}) \frac{P(\bar{A})}{P(\bar{B})} = d/(b+d) \quad (14.33)$$

ausdrücken, wenn die gemeinsamen Wahrscheinlichkeiten $d = P(\bar{A} \cap \bar{B})$ und $b = [1 - P(B|A)]P(A)$ bekannt sind.

Nur wenn $b = 0$, also $P(B|A) = 1$ ist, liegt eine perfekte Kontraposition vor, d.h. $P(\bar{A}|\bar{B}) = 1$.

Zusammenfassend kann man schreiben (vgl. Jaynes, 2003, S. 4 ff.):

1. $P(B|A) = 1$

Wenn A , dann B .

Implikation

2. $P(\bar{A}|\bar{B}) = 1$

Wenn \bar{B} , dann \bar{A} .

Kontraposition

3. $P(A|B) \geq P(A)$

Wenn B , dann ist A plausibler.

Die Ursache ist plausibler, wenn die Wirkung eintritt.

4. $P(B|\bar{A}) \leq P(B)$

Wenn A nicht eintritt, dann ist B weniger plausibel.

Wenn die Ursache nicht eintritt,

ist die Wirkung weniger plausibel.

Noch schwächer ist die Aussage

5. $P(B|A) \geq P(B) \Rightarrow P(A|B) \geq P(A)$

Wenn A , dann ist B plausibler \Rightarrow

Wenn B , dann ist A plausibler.

Herleitung: Als Übung

Hinweis:

$$P(B|\bar{A}) = P(\bar{A}|B)P(B)/P(\bar{A}) =$$

$$[1 - P(A|B)]P(B)/P(\bar{A}) =$$

$$[1 - P(B|A)P(A)/P(B)]P(B)/P(\bar{A}) =$$

$$[P(B) - P(A)]/[1 - P(A)] =$$

$$c/(c + d) = (1 - a - d)/(1 - a) =$$

$$(1 - d)(1 - d - a)/(1 - d - a + ad) \leq 1 - d = P(B).$$

Beispiel 14.3 (Kundenzufriedenheit und Einstellung)

In diesem hypothetischen Beispiel soll die Interpretation von Kreuztabellen als asymmetrischem Ursache-Wirkungszusammenhang noch einmal verdeutlicht werden.

In der Gruppe der unzufriedenen Kunden weisen alle Personen eine generell negative Einstellung zu Banken auf, während bei den zufriedenen Kunden fast alle positiv eingestellt sind (Tab. 14.7).

Daraus läßt sich aber *nicht* folgern, daß die negativ eingestellten Kunden überwiegend unzufrieden sind.

Wie die Tabelle 14.8 der gemeinsamen Wahrscheinlichkeiten zeigt, ist die Gruppe der Unzufriedenen klein [$P(\text{Unzufrieden}) = 0.01$]. Daher sind die auf die Spalte **Einstellung** = Negativ bedingten Wahrscheinlichkeiten fast gleich, obwohl alle Unzufriedenen negativ eingestellt sind (Tab. 14.9). Dieses Phänomen hängt mit dem starken Einfluß der a priori-Wahrscheinlichkeiten zusammen (fast alle sind Zufrieden, daher hat die Kenntnis der negativen Einstellung kaum einen prädiktiven Effekt).

Die Rolle der a priori-Wahrscheinlichkeiten bei der inversen Prädiktion zeigt sich sehr transparent in folgender Schreibweise der Bayes-Formel

$$P(X = x|Y = y) = P(Y = y|X = x) \cdot \frac{P(X = x)}{P(Y = y)}. \quad (14.34)$$

Bayes-Formel



bedingte Wahrscheinlichkeiten (Zeile i): $p_{j i} = P(Y = j X = i)$			
	Negativ	Positiv	$p_{\cdot i}$
Unzufrieden	1	0	1
Zufrieden	0.01	0.99	1

Tabelle 14.7: Kreuztabelle der Variablen $X =$ Kundenzufriedenheit und $Y =$ generelle Einstellung zu Banken. Bedingte Wahrscheinlichkeiten $P(Y = y|X = x)$.

gemeinsame Wahrscheinlichkeiten: $p_{ij} = P(X = i, Y = j)$			
	Negativ	Positiv	$p_{i\cdot}$
Unzufrieden	0.01	0	0.01
Zufrieden	0.0099	0.9801	0.99
$p_{\cdot j}$	0.0199	0.9801	1

Tabelle 14.8: Gemeinsame Wahrscheinlichkeiten $P(X = x, Y = y)$.

bedingte Wahrscheinlichkeiten (Spalte j): $p_{i j} = P(X = i Y = j)$			
	Negativ	Positiv	
Unzufrieden	0.5025	0	
Zufrieden	0.4975	1	
$p_{\cdot j}$	1	1	

Tabelle 14.9: Bedingte Wahrscheinlichkeiten $P(X = x|Y = y)$.

Zusammenfassend kann man sagen, daß die

Implikationsregel der Aussagenlogik (Aussagen A, B)

$$B \Leftarrow A \quad (14.35)$$

bei Ereignissen A, B durch die *Wahrscheinlichkeitsaussage*

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = 1 \quad (14.36)$$

ausgedrückt werden kann.

Im Rückschluß gilt das

Kontrapositionsgesetz (modus tollens)

$$P(\bar{A}|\bar{B}) = 1. \quad (14.37)$$

Die *probabilistische Version* ist allgemeiner als die Form der Aussagenlogik, da sie auch für

$$P(B|A) < 1 \quad (14.38)$$

gültig ist.

Es gilt dann die **Bayes-Formel**

$$P(\bar{A}|\bar{B}) = 1 - [1 - P(B|A)]P(A)/P(\bar{B}) \leq 1 \quad (14.39)$$

bzw. ganz allgemein

$$P(X = x|Y = y) = P(Y = y|X = x) \frac{P(X = x)}{P(Y = y)} \quad (14.40)$$

14.5 Zusammenhänge in Kreuztabellen I: symmetrische Maße

Die Zusammenhänge zwischen den Variablen X und Y lassen sich mit Hilfe der symmetrischen Koeffizienten quantifizieren. Es wird folgende empirische Kreuztabelle (vgl. die theoretische Kreuztabelle aus Bsp. 14.3) verwendet:

gemeinsame relative Häufigkeiten: f_{ij}			
	Negativ	Positiv	$f_{i.}$
Unzufrieden	.01	0	.01
Zufrieden	.01	.98	.99
$f_{.j}$.02	.98	1

Damit lassen sich die Koeffizienten

- **Chi-Quadrat-Koeffizient**

$$\chi^2 = N \sum_{ij} \frac{(f_{ij} - f_{i.}f_{.j})^2}{f_{i.}f_{.j}} = 49.7487$$

$f_{i.} f_{.j}$ = relative Häufigkeit bei Unabhängigkeit von X und Y

$f_{i.} = \sum_j f_{ij}$ (Randhäufigkeit)

- **Kontingenz-Koeffizient**

$$K = \sqrt{\chi^2 / (\chi^2 + N)} = 0.5764$$

- **normierter Kontingenz-Koeffizient**

$$K^+ = K / K_{max} = 0.8151 < 1$$

$$K_{max} = \sqrt{(m-1)/m}; m = \min(I, J) = 2$$

- **Phi-Koeffizient**

$$\phi = \sqrt{\chi^2 / N} = 0.7053$$

numerisch ausdrücken. Zur Berechnung von χ^2 wurde angenommen, daß eine Stichprobe von $N = 100$ zur Verfügung stand und die Wahrscheinlichkeiten durch relative Häufigkeiten ersetzt wurden (siehe Übung 14.1).

14.6 Zusammenhänge in Kreuztabellen II: Prädiktionsregeln

In diesem Abschnitt werden Prädiktionsregeln eingeführt, die bestimmte Zuordnungsregeln zwischen X und Y an den empirischen Daten testen (vgl. die Kapitel: Regression, Varianzanalyse und Hypothesenprüfung mit Prädiktionsregeln).

Grundlegend ist ein

- Fehlermaß $F(+)$ (mit Prädiktor) und ein
- Fehlermaß $F(-)$ (ohne Prädiktor).

Entsprechend definiert man Gütemaße $G = 1 - F$.

Die relative (prozentuale) Differenz wird als proportionale Fehlerreduktion bezeichnet. Ähnlich wie quadrierte Korrelationskoeffizienten lassen sich PRE-Maße als Determinationskoeffizienten (Bestimmtheitsmaße) interpretieren (Kap. 5.4.2).

Prädiktionsregeln

**proportionale
Fehlerreduktion
PRE-Maß**

PRE-Maße:

$$PRE = \frac{F(-) - F(+)}{F(-)} = \frac{G(+)-G(-)}{1-G(-)} \quad (14.41)$$

(proportional reduction of errors)

$$\frac{G(+)-G(-)}{1-G(-)} \leq 1, \quad (14.42)$$

da $G(+)\leq 1$ (tatsächliche Güte der Prädiktion bezogen auf die maximale Güte der Prädiktion).

Spezifische PRE-Maße:

- Cohen κ

$$G(+)=\sum_{ij,H} f_{ij}$$

Cohen κ

$$G(-) = \sum_{ij,H} f_{i.} f_{.j}$$

In diesem Fall wird eine **a-priori-Hypothese** H formuliert, welche Ausprägungen von X und Y zusammenhängen (etwa Diagonale: Beobachterübereinstimmung). In diesem Fall ist $H = \text{diag}(1, \dots, 1)$. Die Summation erfolgt nur über die Zellen der Kreuztabelle, wo $H_{ij} = 1$ ist.

Die Güte dieser Regel wird in der empirischen Tabelle und unter Unabhängigkeit berechnet. Dies stellt sicher, daß die Randhäufigkeiten korrekt berücksichtigt wurden. Dies ist notwendig, da auch bei unabhängigen Variablen hypothesengerechte Zellbesetzungen auftreten.

Im Bsp. 14.3 ergibt sich:

Cohen $\kappa = 0.66443$, wenn die Hypothese H : *Unzufriedene Kunden haben eine negative Einstellung, zufriedene Kunden haben eine positive Einstellung* zugrundegelegt wurde.

Goodman-Kruskal

$\lambda(x \rightarrow y)$

• **Goodman-Kruskal** $\lambda(x \rightarrow y)$

Es handelt sich um ein **a-posteriori-Maß**, da die Modalwerte aus den Daten (empirische Kreuztabellen) gewonnen werden.

$G(+)$ = $\sum_i f_{i\hat{j}(i)}$ = Summe der Modalhäufigkeiten der Zeilen i

$G(-)$ = $f_{. \hat{j}}$ = Maximum der Randverteilung (Spalten j)

\hat{j} = Modalwert der Randverteilung (Spalten j)

$\hat{j}(i)$ = Modalwert in Zeile i =

Modalregel in den Zeilen (Ausprägungen von X)

Im Bsp. 14.3 ergibt sich:

Goodman-Kruskal $\lambda(x \rightarrow y) = 0.502513$

Goodman-Kruskal

$\lambda(y \rightarrow x)$

• **Goodman-Kruskal** $\lambda(y \rightarrow x)$

$G(+)$ = $\sum_j f_{\hat{i}(j)j}$ = Summe der Modalhäufigkeiten der Spalten j

$G(-)$ = $f_{\hat{i} .}$ = Maximum der Randverteilung (Zeilen i)

\hat{i} = Modalwert der Randverteilung (Zeilen i)

$\hat{i}(j)$ = Modalwert in Spalte j =

Modalregel in den Spalten (Ausprägungen von Y)

Im Bsp. 14.3 ergibt sich:

Goodman-Kruskal $\lambda(y \rightarrow x) = 0.01$

Übung 14.1 (Kundenzufriedenheit und Einstellung)

Erzeugen Sie bitte einen SPSS-Datensatz, der die Kreuztabellen aus Bsp. 14.3 ergibt und berechnen Sie die symmetrischen sowie asymmetrischen Koeffizienten von Hand und mit SPSS.

Hinweis: codieren Sie die Ausprägungen mit 0 und 1 und verwenden Sie die Tabelle

gemeinsame absolute Häufigkeiten: h_{ij}			
	Negativ	Positiv	$h_{i.}$
Unzufrieden	1	0	1
Zufrieden	1	98	99
$h_{.j}$	2	98	100

da sich die in Bsp. 14.3 geforderten Wahrscheinlichkeiten für $N = 100$ numerisch nicht genau als relative Häufigkeiten ausdrücken lassen.

Die Lösung ist in Abb. 14.2–14.3 angegeben.

**14.7 Regressions-Modelle (Prognose)**

Die bisherige Diskussion war für nominale Prädiktoren und Prädikanden gültig. Die Idee, die bedingte Wahrscheinlichkeit $P(Y = y|X = x)$ zur Grundlage der Vorhersage zu machen, gilt jedoch auch für stetige oder gemischte Variablen. Die folgende Aufstellung listet die Kombinationen auf:

- lineare **Regressions-Analyse**

(Y metrisch, X metrisch oder Indikator)

$$E(Y|X) = \alpha + \beta X \quad (14.43)$$

- **Varianz-Analyse**

(Y metrisch, $X = i$ nominal oder ordinal)

$$E(Y|X = i) = \mu_i \quad (14.44)$$

Verarbeitete Fälle

	Fälle					
	Gültig		Fehlend		Gesamt	
	N	Prozent	N	Prozent	N	Prozent
Zufrieden * Einstellung	100	100,0%	0	,0%	100	100,0%

Zufrieden * Einstellung Kreuztabelle

Anzahl	Einstellung			Gesamt
	,00	1,00		
Zufrieden ,00	1	0		1
1,00	1	98		99
Gesamt	2	98		100

Chi-Quadrat-Tests

	Wert	df	Asymptotische Signifikanz (2-seitig)	Exakte Signifikanz (2-seitig)	Exakte Signifikanz (1-seitig)
Chi-Quadrat nach Pearson	49,495 ^a	1	,000		
Kontinuitätskorrektur ^b	11,874	1	,001		
Likelihood-Quotient	8,428	1	,004		
Exakter Test nach Fisher				,020	,020
Zusammenhang linear-mit-linear	49,000	1	,000		
Anzahl der gültigen Fälle	100				

a. 3 Zellen (75,0%) haben eine erwartete Häufigkeit kleiner 5. Die minimale erwartete Häufigkeit ist ,02.
b. Wird nur für eine 2x2-Tabelle berechnet

Richtungsmaße

			Wert	Asymptotischer Standardfehler ^a	Näherungsweise T ^b	Näherungsweise Signifikanz
Nominal- bzgl. Nominalmaß	Lambda	Symmetrisch	,333	,629	,448	,654
		Zufrieden abhängig	,000	1,414	,000	1,000
		Einstellung abhängig	,500	,354	1,005	,315
	Goodman-und-Kruskal-Tau	Zufrieden abhängig	,495	,354		,000 ^c
		Einstellung abhängig	,495	,251		,000 ^c
	Unsicherheitskoeffizient	Symmetrisch	,547	,307	1,067	,004 ^d
Zufrieden abhängig		,752	,147	1,067	,004 ^d	
Einstellung abhängig		,430	,332	1,067	,004 ^d	

- a. Die Null-Hypothese wird nicht angenommen.
- b. Unter Annahme der Null-Hypothese wird der asymptotische Standardfehler verwendet.
- c. Basierend auf Chi-Quadrat-Näherung
- d. Chi-Quadrat-Wahrscheinlichkeit für Likelihood-Quotienten.

Abbildung 14.2: SPSS-Output zur Übung 14.1.

Symmetrische Maße

		Wert	Asymptotischer Standardfehler ^a	Näherungsweise T ^b	Näherungsweise Signifikanz
Nominal- bzgl. Nominalmaß	Phi	,704			,000
	Cramer-V	,704			,000
	Kontingenzkoeffizient	,575			,000
Intervall- bzgl. Intervallmaß	Pearson-R	,704	,251	9,800	,000 ^c
Ordinal- bzgl. Ordinalmaß	Korrelation nach Spearman	,704	,251	9,800	,000 ^c
Maß der Übereinstimmung	Kappa	,662	,316	7,035	,000
Anzahl der gültigen Fälle		100			

a. Die Null-Hyphothese wird nicht angenommen.

b. Unter Annahme der Null-Hyphothese wird der asymptotische Standardfehler verwendet.

c. Basierend auf normaler Näherung

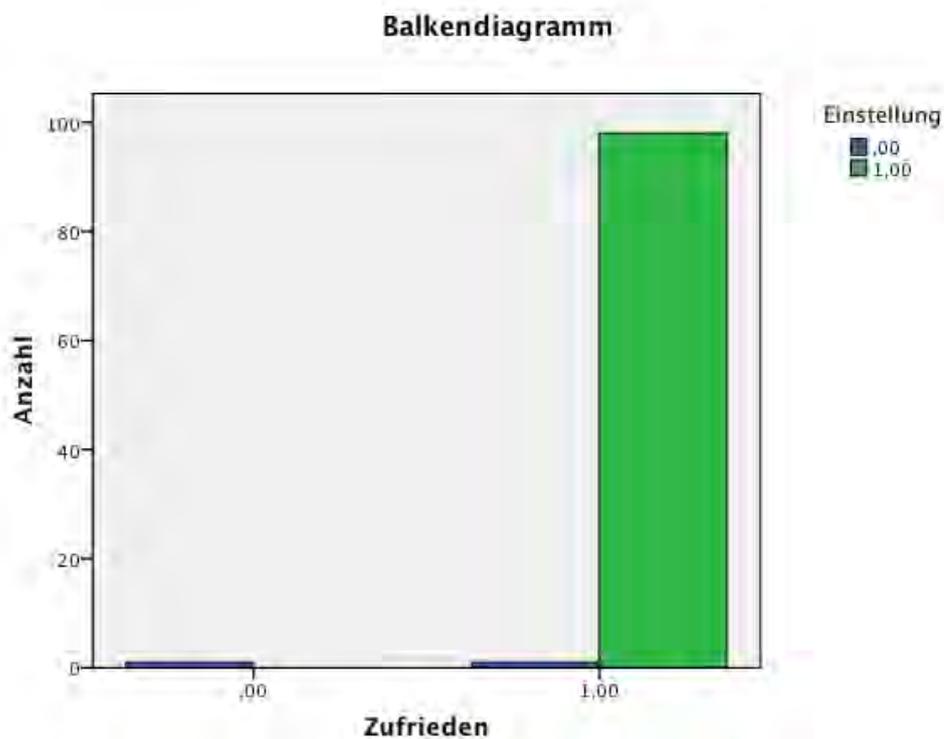


Abbildung 14.3: SPSS-Output zur Übung 14.1.

- **kategoriale Regression**

(Y nominal oder ordinal, X metrisch oder Indikator)

$$P(Y = 1|X) = F(\alpha + \beta X) \leq 1 \quad (14.45)$$

- **Modal-Regel in Kreuztabellen**

$$\hat{j}(i) = \arg \max_j P(Y = j|X = i), \quad (14.46)$$

d.h. der Index j , bei dem $P(Y = j|X = i)$ maximal ist.

Aufgabe der statistischen Analyse ist es, die unbekannt Parameter $\alpha, \beta, \mu_i, P(Y = j|X = i) = p_{j|i}$ aus Daten $(X_n, Y_n), n = 1, \dots, N$ zu schätzen.

14.7.1 Einfach-Regression (X und Y univariat)

linearer Prädiktor

Beispielsweise ist die lineare Regression eine Prädiktionsregel (Geradenregel), die es erlaubt, den linearen Prädiktor

$$\hat{Y}(X) = E(Y|X) = \alpha + \beta X \quad (14.47)$$

zu berechnen. Der Prognose-Fehler ist dann $(\hat{Y}_n = \alpha + \beta X_n)$

$$F(+) = \sum_n (Y_n - \hat{Y}_n)^2 := SQR. \quad (14.48)$$

KQ-Schätzer

Dies ist die bekannte Quadratsumme der Abweichungen (residuale Quadratsumme SQR), deren Minimierung bzgl. α und β zum kleinsten-Quadrat-Schätzer (KQ) $\hat{\alpha}$ und $\hat{\beta}$ führt:

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} \quad (14.49)$$

$$\hat{\beta} = \frac{\sum_n (Y_n - \bar{Y})(X_n - \bar{X})}{\sum_n (X_n - \bar{X})^2} \quad (14.50)$$

$$= \frac{S_{xy}}{S_x^2} \quad (14.51)$$

$$= R_{xy} \frac{S_y}{S_x} \quad (14.52)$$

In der letzten Gleichung wurde der Produkt-Moment-Korrelationskoeffizient

$$R_{xy} = \frac{\sum_n (X_n - \bar{X})(Y_n - \bar{Y})}{\sqrt{\sum_n (X_n - \bar{X})^2} \sqrt{\sum_n (Y_n - \bar{Y})^2}} \quad (14.53)$$

Produkt-Moment-Korrelation

eingesetzt.

Ohne Prädiktor hat man den Fehler (totale Quadratsumme SQT)

$$F(-) = \sum_n (Y_n - \bar{Y})^2 \quad (14.54)$$

$$:= SQT = (N - 1)S_y^2. \quad (14.55)$$

Aufgrund der bekannten Streuungszerlegung

$$SQT = SQE + SQR \quad (14.56)$$

$$SQE = \sum_n (\hat{Y}_n - \bar{Y})^2 \quad (14.57)$$

$$= \sum_n (\hat{\alpha} + \hat{\beta}X_n - \bar{Y})^2 \quad (14.58)$$

$$= \sum_n [\hat{\beta}(X_n - \bar{X})]^2 \quad (14.59)$$

$$= \hat{\beta}^2(N - 1)S_x^2 \quad (14.60)$$

$$= R_{xy}^2(N - 1)S_y^2 \quad (14.61)$$

Streuungszerlegung

ergibt sich also

$$PRE = \frac{F(-) - F(+)}{F(-)} \quad (14.62)$$

$$= \frac{SQT - SQR}{SQT} = \frac{SQE}{SQT} \quad (14.63)$$

$$= R_{xy}^2. \quad (14.64)$$

Daher gilt die Identität:

prozentuale Fehlerreduktion =
erklärte Streuung/totale Streuung =
quadrierte Korrelation =
Determinationskoeffizient.

Die quadrierte Korrelation R_{xy}^2 ist der prozentuale Anteil der Streuung von Y , die durch X erklärt wird.

14.7.2 Varianz-Analyse (X und Y univariat)

experimentelles Design

In vielen Fällen hat die unabhängige Variable nur nominales oder ordinale Skalenniveau, insbesondere bei experimentellen Designs. Wenn nur 2 Abstufungen vorliegen, kann man eine 0-1-Codierung vornehmen und mit Regressionsanalysen arbeiten.

Varianz-Analyse (ANOVA)

Im allgemeinen verwendet man jedoch die Methode der Varianz-Analyse (ANOVA), bei der die bedingten Erwartungswerte

$$E(Y|X = i) = \mu_i$$

für die Gruppen $X = i$ geschätzt werden (ausführlich siehe Kap. 6).

Bei I Gruppen und J Personen pro Gruppe gilt für die $N = I * J$ Daten Y_{ij} das bedingte Modell (feste Effekte)

$$Y_{ij} = \mu_i + \epsilon_{ij} \quad (14.65)$$

wobei ein iid-Gleichungsfehler $\epsilon_{ij} \sim N(0, \sigma^2)$ unterstellt wird.²

Wenn sich die Gruppen (etwa verschiedene Gestaltungen von Bankfilialen, Berufsgruppen, Lehrmethoden) auf der Y -Variable nicht unterscheiden

²iid = independent and identically distributed, d.h. alle Personen in allen Gruppen sind voneinander unabhängig. Dies ist nicht selbstverständlich und wird bei Designs mit Meßwiederholungen benutzt, um bei Gruppenvergleichen die Streuung innerhalb der Gruppen zu verkleinern (im einfachsten Fall ein t -Test mit abhängigen Stichproben; vgl. Kap. 3.3.2).

den (also der Regressor $X = i$ keine Wirkung hat), müssen die Erwartungswerte in allen Gruppen gleich sein. Die zu prüfende Nullhypothese ist also

**Nullhypothese
(ANOVA)**

$$H_0 : \quad \mu_i = \mu; i = 1, \dots, I \quad (14.66)$$

gegen die Alternative

$$H_1 : \quad \mu_i \neq \mu_{i'} \text{ für mindestens ein Paar } i, i' = 1, \dots, I. \quad (14.67)$$

Prognose-Regeln:

Prognose-Regeln

- Wenn bekannt ist, daß eine Person der Gruppe $X = i$ angehört, ist die beste Prognose der Gruppen-Mittelwert

$$\bar{Y}_i := J^{-1} \sum_j Y_{ij}.$$

- Wenn diese Information nicht vorliegt, prognostiziert man am besten den globalen Mittelwert

$$\bar{Y} := (IJ)^{-1} \sum_{ij} Y_{ij}.$$

Die entsprechenden Fehlermaße sind dann

$$F(-) = \sum_{ij} (Y_{ij} - \bar{Y})^2 = SQT \text{ (totale Quadratsumme)}$$

$$F(+) = \sum_{ij} (Y_{ij} - \bar{Y}_i)^2 = SQR \text{ (residuale Quadratsumme).}$$

Aufgrund der Streuungszerlegung gilt ³

$$SQT = \sum_{ij} (Y_{ij} - \bar{Y})^2 \quad (14.68)$$

$$= \sum_{ij} (Y_{ij} - \bar{Y}_i)^2 + \sum_{ij} (\bar{Y}_i - \bar{Y})^2 \quad (14.69)$$

$$= SQR + SQE \quad (14.70)$$

$$= \text{residuale} + \text{erklärte Quadratsumme} \quad (14.71)$$

$$= SQ(\text{innerhalb}) + SQ(\text{zwischen}). \quad (14.72)$$

Determinationskoeffizient

Das PRE-Maß (Determinationskoeffizient 1. Art) ist somit

$$PRE = \frac{F(-) - F(+)}{F(-)} = \frac{SQT - SQR}{SQT} = \frac{SQE}{SQT}. \quad (14.73)$$

F-Test

Der übliche F -Test für die Nullhypothese der Gleichheit der Gruppenmittel

$$F = \frac{SQE/(I - 1)}{SQR/(I(J - 1))} = \frac{SQE/(I - 1)}{(SQT - SQE)/(I(J - 1))} \quad (14.74)$$

$$= \frac{PRE/(I - 1)}{(1 - PRE)/(I(J - 1))} \sim F(I - 1, I(J - 1)) \quad (14.75)$$

läßt sich somit als Funktion des PRE-Maßes ausdrücken. Geht das PRE-Maß gegen 1, so liegt ein starker Gruppen-Effekt vor (Mittelwertsunterschiede zwischen den Gruppen), und die F -Statistik wird groß, während für einen geringen Einfluß der Gruppe (kleines PRE-Maß, ähnliche Gruppenmittelwerte) die F -Statistik gegen 0 geht.

1 - α-Quantil

Die genauen Verhältnisse erhellt das $1 - \alpha$ -Quantil der F -Verteilung, d.h. für

$$F(I - 1, I(J - 1)) > f(1 - \alpha; I - 1, I(J - 1)) \quad (14.76)$$

wird die H_0 auf dem Signifikanzniveau α verworfen.

ANOVA-Tafel

Die verschiedenen Quadratsummen lassen sich übersichtlich in einer ANOVA-Tafel (Varianztabelle) zusammenstellen:

SQ	Formel	df	F -Statistik
SQE (zwischen)	$\sum_{ij}(\bar{Y}_i - \bar{Y})^2$	$I - 1$	$\frac{SQE/(I-1)}{SQR/(I(J-1))}$ $\sim F(I - 1, I(J - 1))$
SQR (innerhalb)	$\sum_{ij}(Y_{ij} - \bar{Y}_i)^2$	$I(J - 1)$	
SQT (total)	$\sum_{ij}(Y_{ij} - \bar{Y})^2$	$IJ - 1$	

Freiheitsgrade

Die Freiheitsgrade (degrees of freedom df) der Varianzquellen (Quadratsummen) sind analog zur Streuungszerlegung additiv.

Übung 14.2 (Freiheitsgrade der Quadratsummen)

³wobei die Produkt-Terme verschwinden:
 $\sum_{ij}(Y_{ij} - \bar{Y}_i)(\bar{Y}_i - \bar{Y}) = J \sum_i(\bar{Y}_i - \bar{Y})(\bar{Y}_i - \bar{Y}) = 0$

- Warum sind die Quadratsummen χ^2 -verteilt?
- Zeigen Sie, daß sich die Freiheitsgrade der Quadratsummen addieren.
- Warum hat die erklärte Quadratsumme $I - 1$ Freiheitsgrade?
- Warum hat die totale Quadratsumme $IJ - 1 = N - 1$ Freiheitsgrade?

Konsultieren Sie gegebenenfalls Kap. 1.3.2.



Beispiel 14.4 (Kundenzufriedenheit und Berufsgruppe)

Die Mittelwertsunterschiede auf der abhängigen Variable (AV) Kundenzufriedenheit als Funktion der UV Berufsgruppe lassen sich mit Hilfe einer Varianzanalyse untersuchen.

Geben Sie dazu in SPSS den Befehl

Analysieren/Mittelwerte vergleichen/Einfaktorielle ANOVA/

und wählen Sie

Optionen/Diagramm der Mittelwerte sowie

Post Hoc/Bonferroni

Es ergeben sich Mittelwertsunterschiede der Kundenzufriedenheit in den Berufsgruppen, deren Signifikanz durch den globalen F -Test überprüft wird ($p = 0.003$). Daher wird die H_1 : *mindestens ein signifikanter Unterschied* angenommen.

Wo genau die Unterschiede liegen, kann zunächst nicht gesagt werden. Daher führt man i.a. **Nach(Post Hoc)-Tests** durch, z.B. nach dem Bonferroni-Verfahren. Dabei werden alle k interessierenden Mittelwertsvergleiche durchgeführt (Kap. 3.3.1.1, 6), aber auf dem adjustierten Signifikanzniveau $\alpha' = \alpha/k$ (etwa $0.05/15=0.0033$, wenn alle $k = I(I - 1)/2 = 15$ Vergleiche gerechnet werden). Dann ist das simultane Signifikanzniveau aller Tests $\alpha^* \leq k\alpha' = \alpha$.

Es ist instruktiv, den Bonferroni-Test mit einem einzelnen t -Test zu vergleichen (Abb. 14.6, oben; Schüler vs. Arbeiter) Hier ist der p -Wert 0.116 (0.085 bei ungleichen Varianzen), während für den Bonferroni- t -Test $p = 1$ herauskommt. Anders ausgedrückt liegen durch die α -Adjustierung die kritischen Werte $t(1 - \alpha/(2k), df)$ der t -Statistik sehr weit außen (Abb.

**Bonferroni-
Verfahren**

**adjustiertes
Signifikanzniveau**

ONEWAY deskriptive Statistiken

	N	Mittelwert	Standardabweichung	Standardfehler	95%-Konfidenzintervall für den Mittelwert		Minimum	Maximum
					Untergrenze	Obergrenze		
Schüler	29	-,3523	,87388	,16228	-,6847	-,0199	-2,28	1,12
Arbeiter	132	-,0255	1,03472	,09006	-,2036	,1527	-3,04	1,40
Hausfrau	39	,2289	1,06988	,17132	-,1179	,5757	-4,08	1,40
Selbständig	22	-,4445	1,12840	,24058	-,9448	,0558	-4,18	1,12
Rentner	108	,2138	,79213	,07622	,0627	,3649	-1,70	1,40
ohne Beschäftigung	23	-,3393	1,27195	,26522	-,8893	,2107	-3,49	1,40
Gesamt	353	,0024	1,00037	,05324	-,1023	,1071	-4,18	1,40

ONEWAY ANOVA

ZufriedU	Quadratsumme	df	Mittel der Quadrate	F	Signifikanz
Zwischen den Gruppen	17,657	5	3,531	3,662	,003
Innerhalb der Gruppen	334,605	347	,964		
Gesamt	352,262	352			

Mittelwert-Diagramme

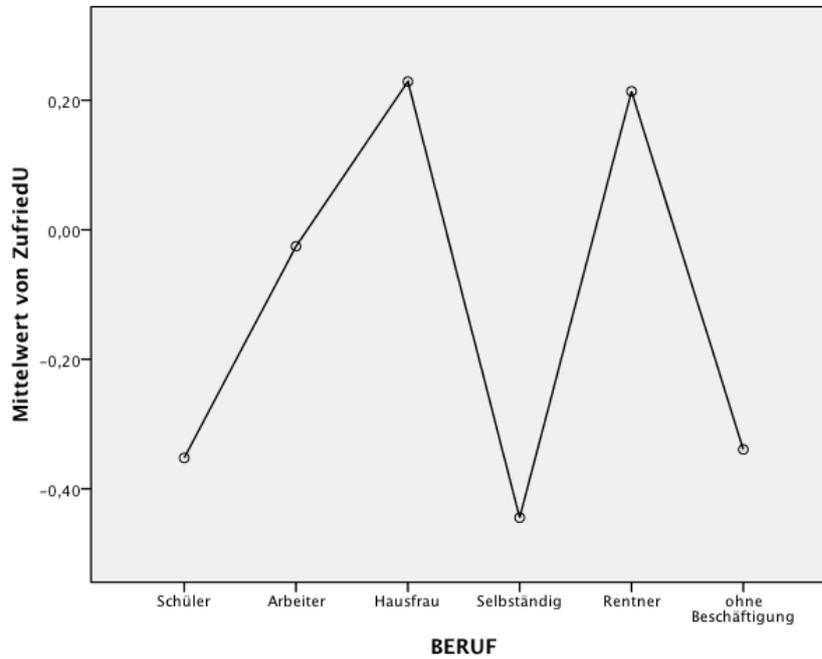


Abbildung 14.4: Mittelwertsunterschiede der Variable *Kundenzufriedenheit* in den Berufsgruppen.

Post-Hoc-Tests

Mehrfachvergleiche

ZufriedU
Bonferroni

(I) BERUF	(J) BERUF	Mittlere Differenz (I-J)	Standardfehler	Signifikanz	95%-Konfidenzintervall	
					Untergrenze	Obergrenze
Schüler	Arbeiter	-,32681	,20139	1,000	-,9220	,2684
	Hausfrau	-,58120	,24078	,245	-1,2929	,1305
	Selbständig	,09227	,27764	1,000	-,7283	,9129
	Rentner	-,56606	,20538	,092	-1,1731	,0410
	ohne Beschäftigung	-,01298	,27418	1,000	-,8234	,7974
Arbeiter	Schüler	,32681	,20139	1,000	-,2684	,9220
	Hausfrau	-,25439	,17897	1,000	-,7834	,2746
	Selbständig	,41908	,22613	,970	-,2493	1,0874
	Rentner	-,23925	,12741	,919	-,6158	,1373
	ohne Beschäftigung	,31382	,22188	1,000	-,3420	,9696
Hausfrau	Schüler	,58120	,24078	,245	-,1305	1,2929
	Arbeiter	,25439	,17897	1,000	-,2746	,7834
	Selbständig	,67347	,26183	,158	-,1004	1,4474
	Rentner	,01514	,18345	1,000	-,5271	,5574
	ohne Beschäftigung	,56821	,25817	,426	-,1948	1,3313
Selbständig	Schüler	-,09227	,27764	1,000	-,9129	,7283
	Arbeiter	-,41908	,22613	,970	-1,0874	,2493
	Hausfrau	-,67347	,26183	,158	-1,4474	,1004
	Rentner	-,65833	,22969	,066	-1,3372	,0206
	ohne Beschäftigung	-,10525	,29284	1,000	-,9708	,7603
Rentner	Schüler	,56606	,20538	,092	-,0410	1,1731
	Arbeiter	,23925	,12741	,919	-,1373	,6158
	Hausfrau	-,01514	,18345	1,000	-,5574	,5271
	Selbständig	,65833	,22969	,066	-,0206	1,3372
	ohne Beschäftigung	,55307	,22551	,220	-,1135	1,2196
ohne Beschäftigung	Schüler	,01298	,27418	1,000	-,7974	,8234
	Arbeiter	-,31382	,22188	1,000	-,9696	,3420
	Hausfrau	-,56821	,25817	,426	-1,3313	,1948
	Selbständig	,10525	,29284	1,000	-,7603	,9708
	Rentner	-,55307	,22551	,220	-1,2196	,1135

Abbildung 14.5: Paarweise Mittelwertsunterschiede der Variable *Kundenzufriedenheit* in den Berufsgruppen: Bonferroni-*t*-Test.

Gruppenstatistiken

	BERUF	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
ZufriedU	Schüler	29	-,3523	,87388	,16228
	Arbeiter	132	-,0255	1,03472	,09006

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					95% Konfidenzintervall der Differenz	
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	Untere	Obere
ZufriedU	Varianzen sind gleich	,644	,424	-1,580	159	,116	-,32681	,20678	-,73519	,08157
	Varianzen sind nicht gleich			-1,761	46,953	,085	-,32681	,18559	-,70018	,04657

Gruppenstatistiken

	BERUF	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
ZufriedU	Selbständig	22	-,4445	1,12840	,24058
	Rentner	108	,2138	,79213	,07622

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					95% Konfidenzintervall der Differenz	
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	Untere	Obere
ZufriedU	Varianzen sind gleich	,567	,453	-3,286	128	,001	-,65833	,20032	-1,05470	-,26196
	Varianzen sind nicht gleich			-2,609	25,378	,015	-,65833	,25236	-1,17769	-,13897

Gruppenstatistiken

	BERUF	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
ZufriedU	Selbständig	22	-,4445	1,12840	,24058
	Hausfrau	39	,2289	1,06988	,17132

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					95% Konfidenzintervall der Differenz	
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	Untere	Obere
ZufriedU	Varianzen sind gleich	,026	,872	-2,315	59	,024	-,67347	,29092	-1,25560	-,09134
	Varianzen sind nicht gleich			-2,280	41,763	,028	-,67347	,29534	-1,26959	-,07734

Abbildung 14.6: Mittelwertsunterschied der Variable *Kundenzufriedenheit* in den Berufsgruppen Schüler vs. Arbeiter, Selbständig vs. Rentner und Selbständig vs. Hausfrau: Einzelne-*t*-Tests.

6.2), sodaß große Werte der t -Statistik nötig sind, um einen signifikanten Unterschied nachzuweisen.

Die einzigen relevanten Ergebnisse, die bei einer simultanen Betrachtung in Frage kommen, sind der Paarvergleich **Selbständig-Rentner** ($p = .066$) und der Paarvergleich **Selbständig-Hausfrau** ($p = .158$).

Bei einer einzelnen Betrachtung (Abb. 14.6, mitte, unten) ergibt sich der Vergleich **Selbständig-Rentner** ($p = .001, 0.015$ (ungleiche Varianzen)) und **Selbständig-Hausfrau** ($p = .024, 0.028$ (ungleiche Varianzen)).

Alternativ könnte man $k = 6 * 5/2 = 15$ einzelne t -Tests rechnen, dürfte dann aber nur p -Werte unterhalb $\alpha' = \alpha/k$ (etwa $p < 0.05/15 = 0.0033$) als signifikante Tests (Mittelwertvergleiche) berücksichtigen.



14.8 Scheinkorrelationen, optimale Prognose, Residualstreuung

Häufig werden Korrelationen zwischen Variablen X und Y berichtet, die zwar zweifelsfrei bestehen, jedoch nicht sinnvoll interpretierbar sind, da sie durch die Wirkung einer dritten Variable Z erzeugt werden.

Klassisch ist hier die Korrelation zwischen der **Zahl der Störche** (X) und der **Geburtenrate** (Y). Daraus kann man folgern, daß weniger Störche mit weniger Geburten korrespondieren (oder umgekehrt), jedoch ist dies keine Bestätigung der Theorie, daß Störche die Kinder bringen. Es ist klar, daß Variablen wie $Z =$ **Industrialisierungsgrad** etc. sowohl indirekt die Zahl der Störche als auch die Geburtenrate beeinflussen.

Betrachtet man die Variablen $X =$ **Konsum an Kochsalz**, $Y =$ **Blutdruck**, so wird die Sache schon problematischer. Hier hat man das Empfinden, daß der Genuß von Salz direkt den Blutdruck beeinflusst. *Alternative Erklärung:* Möglicherweise verwenden ältere Personen mehr Salz, sind jedoch auch gleichzeitig mit höherem Blutdruck belastet (Kontrollvariable $Z =$ **Alter**).

Kontrollvariablen

Ein drittes Beispiel wäre der negative Zusammenhang zwischen **Gewicht** und **Ausgaben für Kosmetik** (etwa pro Monat). Vermehrter Gebrauch von kosmetischen Artikeln kann also eine Gewichtsreduktion herbeiführen. Die folgenden Graphiken und Tabellen erhellen die genauen Verhältnisse. Abb. 14.7 zeigt das Streudiagramm der Variablen, gesamt (oben) und gruppiert nach der Variablen $Z =$ **Geschlecht** (unten). Betrachtet man das gruppierte Streudiagramm, so ist innerhalb der Gruppen $Z = 0, 1$ ein

geringer Zusammenhang der Variablen, insgesamt hat jedoch die Position der Gruppen im Streudiagramm einen negativen Effekt.

Abb. 14.8 (oben) zeigt die Korrelationsmatrizen der Variablen **Gewicht**, **Ausgaben für Kosmetik** und **Geschlecht**. Die Variable **Geschlecht**, die als Indikator codiert wurde, hat eine positive Korrelation mit **Ausgaben für Kosmetik** und eine negative Korrelation mit **Gewicht**, was die negative Korrelation zwischen **Gewicht** und **Ausgaben für Kosmetik** zur Folge hat. In Abb. 14.8 (unten) sind die Korrelationen bei Aufteilung der Stichprobe in Männer und Frauen dargestellt (d.h. bedingt auf die Kontrollvariable $Z = \text{Geschlecht}$).

Schließlich ist in Abb. 14.9 (oben) die partielle Korrelation zwischen **Gewicht** und **Ausgaben für Kosmetik** zu sehen. Sie ist sehr gering und spiegelt die Zusammenhänge wider, wenn die Dritt(Kontroll)-Variable **Geschlecht** korrigiert wurde.

14.8.1 diskrete Kontrollvariablen

Man kann den Einfluß der Kontrollvariablen Z durch eine MANOVA-Spezifikation⁴ mit

$$\mathbf{y} = \begin{bmatrix} \text{Gewicht} \\ \text{Ausgaben für Kosmetik} \end{bmatrix} \quad (14.77)$$

als abhängigen Variablen und $Z = i = \text{Geschlecht}$ als unabhängiger Variable modellieren.

Das Modell lautet für die Gruppen i und Personen (Einheiten) j (vgl. 14.65)

$$\mathbf{y}_{ij} = \boldsymbol{\mu}_i + \boldsymbol{\epsilon}_{ij} \quad (14.78)$$

wobei ein iid-Gleichungsfehler $\boldsymbol{\epsilon}_{ij} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ unterstellt wird.

Die **partielle Korrelation** ist dann die residuale Korrelationsmatrix in der Streuungszersetzung $SQT = SQE + SQR$ (Totale Streuung = Erklärte Streuung + Residuale Streuung).

MANOVA

Prognose-Regeln

Die Streuungszersetzung ergibt sich aus den Prognose-Regeln:

- Wenn bekannt ist, daß eine Person der Gruppe $Z = i$ mit Gruppengröße J_i angehört, ist die beste Prognose der Gruppen-Mittelwert für den Vektor $\mathbf{y} = [X, Y]'$:

$$\bar{\mathbf{y}}_i := J_i^{-1} \sum_j \mathbf{y}_{ij} := \hat{\mathbf{y}}(Z = i).$$

⁴Multivariate Varianzanalyse; vgl. auch Abs. 14.7

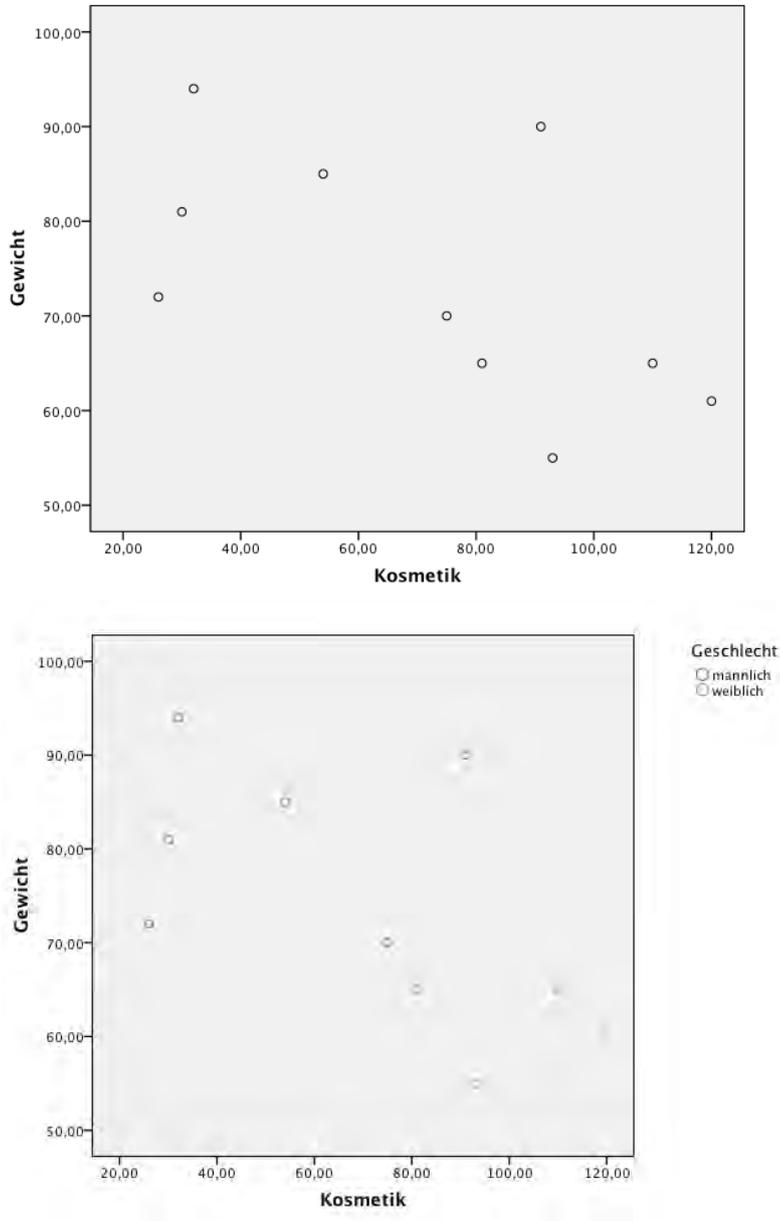


Abbildung 14.7: Oben: Streudiagramm (scatterplot) der Variablen *Gewicht* und *Ausgaben für Kosmetik*. Negative Korrelation. Unten: Streudiagramm (scatterplot) der Variablen *Gewicht* und *Ausgaben für Kosmetik*. Aufteilung nach Gruppen männlich/weiblich.

Korrelationen

		Kosmetik	Gewicht	Geschlecht01
Kosmetik	Korrelation nach Pearson	1	-,588	,753*
	Signifikanz (2-seitig)		,074	,012
	N	10	10	10
Gewicht	Korrelation nach Pearson	-,588	1	-,810**
	Signifikanz (2-seitig)	,074		,005
	N	10	10	10
Geschlecht01	Korrelation nach Pearson	,753*	-,810**	1
	Signifikanz (2-seitig)	,012	,005	
	N	10	10	10

*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

** . Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Korrelationen

Geschlecht01			Kosmetik	Gewicht
männlich	Kosmetik	Korrelation nach Pearson	1	,065
		Signifikanz (2-seitig)		,903
		N	6	6
	Gewicht	Korrelation nach Pearson	,065	1
		Signifikanz (2-seitig)	,903	
		N	6	6
weiblich	Kosmetik	Korrelation nach Pearson	1	,016
		Signifikanz (2-seitig)		,984
		N	4	4
	Gewicht	Korrelation nach Pearson	,016	1
		Signifikanz (2-seitig)	,984	
		N	4	4

Abbildung 14.8: Oben: Korrelationsmatrix der Variablen *Gewicht*, *Ausgaben für Kosmetik* und *Geschlecht*. Unten: Korrelationsmatrix der Variablen *Gewicht* und *Ausgaben für Kosmetik*. Aufteilung nach Gruppen männlich/weiblich.

Korrelationen

Kontrollvariablen			Gewicht	Kosmetik
Geschlecht01	Gewicht	Korrelation	1,000	,057
		Signifikanz (zweiseitig)	.	,884
		Freiheitsgrade	0	7
	Kosmetik	Korrelation	,057	1,000
		Signifikanz (zweiseitig)	,884	.
		Freiheitsgrade	7	0

Matrix der Quadratsummen und Kreuzprodukte für Residuen

		Kosmetik	Gewicht
Quadratsumme und Kreuzprodukte	Kosmetik	4517,333	88,000
	Gewicht	88,000	529,000
Kovarianz	Kosmetik	564,667	11,000
	Gewicht	11,000	66,125
Korrelation	Kosmetik	1,000	,057
	Gewicht	,057	1,000

Basiert auf Typ III Quadratsumme

Abbildung 14.9: Oben: Partielle Korrelation der Variablen *Gewicht*, *Ausgaben für Kosmetik*. Kontrollvariable: *Geschlecht*. Unten: MANOVA: Residuale Korrelationsmatrix der Variablen *Gewicht* und *Ausgaben für Kosmetik*. Unabhängige Variable *Geschlecht*. Partielle Korrelation und Residualkorrelation ergeben das gleiche Resultat (.057).

- Wenn diese Information nicht vorliegt, prognostiziert man am besten den globalen Mittelwert

$$\bar{\mathbf{y}} := J_+^{-1} \sum_{ij} \mathbf{y}_{ij}.$$

mit $J_+ = \sum_i J_i = N$ (Gesamtstichprobengröße).

Die Streuungszerlegung lautet

Streuungszerlegung

$$SQT = \sum_{ij} (\mathbf{y}_{ij} - \bar{\mathbf{y}})(\mathbf{y}_{ij} - \bar{\mathbf{y}})' \quad (14.79)$$

$$= \sum_{ij} (\mathbf{y}_{ij} - \bar{\mathbf{y}}_i)(\mathbf{y}_{ij} - \bar{\mathbf{y}}_i)' + \sum_{ij} (\bar{\mathbf{y}}_i - \bar{\mathbf{y}})(\bar{\mathbf{y}}_i - \bar{\mathbf{y}})' \quad (14.80)$$

$$= SQR + SQE \quad (14.81)$$

$$= \text{residuale} + \text{erklärte Quadratsumme} \quad (14.82)$$

$$= SQ(\text{innerhalb}) + SQ(\text{zwischen}). \quad (14.83)$$

Die gemischten Terme verschwinden, was bedeutet, daß Prognose und Prognose-Fehler unkorreliert (orthogonal) sind.

Es handelt sich bei der Streuungszerlegung um 2×2 (i.a. $p \times p$)-Matrizen. Das Element $SQR_{12} = 88.000$ (Abb. 14.9, unten) ist die residuale Quadratsumme der Variablen X und Y . Normiert man auf die Wurzeln der Diagonale (4517.333, 529.000) ergibt sich die partielle Korrelation (Residualkorrelation)

**partielle
Korrelation**

$$SQR_{12} / \sqrt{SQR_{11} SQR_{22}} = 88 / \sqrt{4517.333 * 529} = .057.$$

In der Matrix der residualen Streuung

$$SQR = \sum_{ij} (\mathbf{y}_{ij} - \bar{\mathbf{y}}_i)(\mathbf{y}_{ij} - \bar{\mathbf{y}}_i)' \quad (14.84)$$

werden die korrigierten Variablen (Residuen, Innovationen)

$$\tilde{\mathbf{y}}_{ij} = \mathbf{y}_{ij} - \bar{\mathbf{y}}_i \quad (14.85)$$

benutzt, welche schon die Information $\bar{\mathbf{y}}_i = \hat{\mathbf{y}}(i)$ der Kontrollvariable $Z = i$ enthalten. Daher ist $[J_+ - I = \sum_i (J_i - 1)]$

$$MQR = \frac{1}{J_+ - I} \sum_{ij} (\mathbf{y}_{ij} - \bar{\mathbf{y}}_i)(\mathbf{y}_{ij} - \bar{\mathbf{y}}_i)' \quad (14.86)$$

$$= \sum_i \left(\frac{J_i - 1}{\sum_i (J_i - 1)} \right) \frac{1}{J_i - 1} \sum_j (\mathbf{y}_{ij} - \bar{\mathbf{y}}_i)(\mathbf{y}_{ij} - \bar{\mathbf{y}}_i)'$$

$$:= \sum_i p_i MQR(i) \quad (14.87)$$

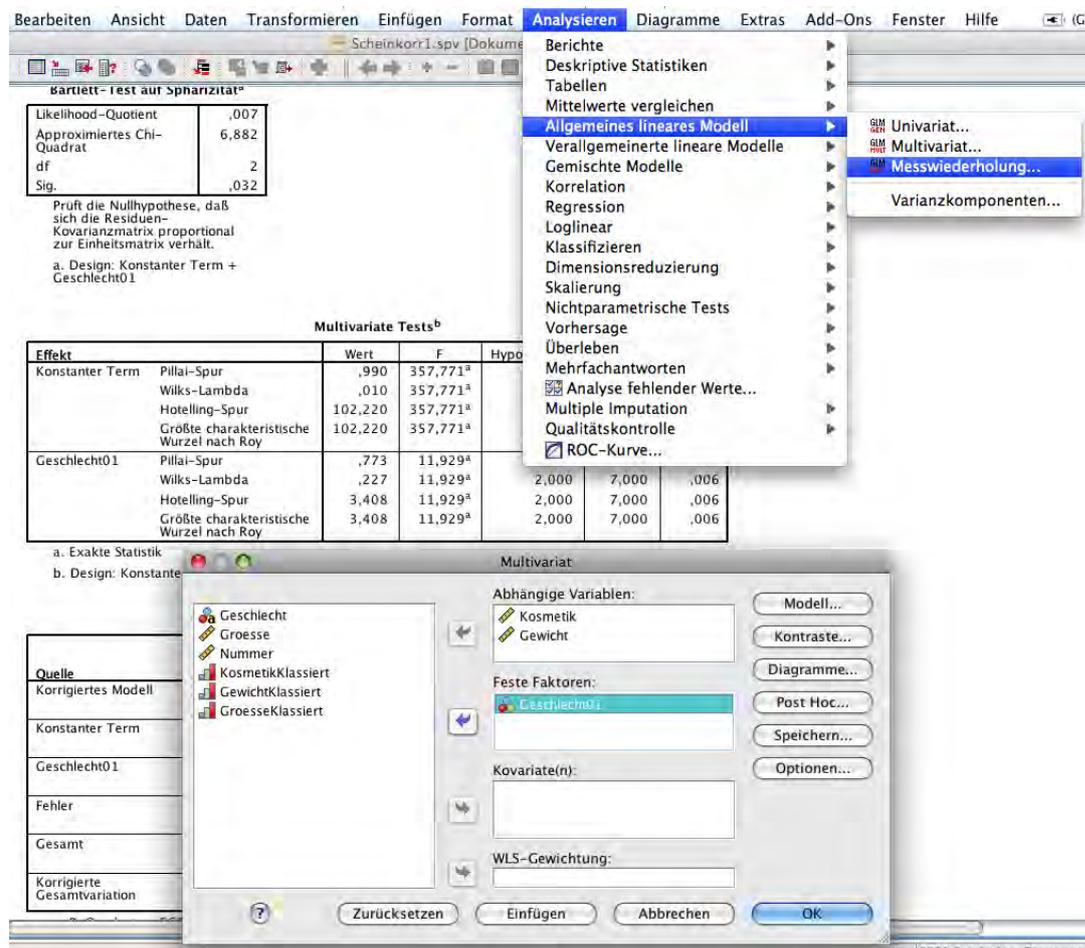


Abbildung 14.10: MANOVA: SPSS-Menü.

die über alle Gruppen gemittelte bedingte Kovarianzmatrix, aus der die partielle Korrelationsmatrix berechnet wird (siehe Abb. 14.9, unten; $MQR_{12} := MQR(X, Y)$).

Man kann natürlich auch direkt die Residuen \tilde{y}_{ij} und deren Korrelation berechnen.

SPSS-Menü-Befehle für MANOVA siehe Abb. 14.10.

Zusammenfassend gilt:

Streuungszerlegung

$$SQT = SQR + SQE \quad (14.88)$$

$$= \text{residuale} + \text{erklärte Quadratsumme} \quad (14.89)$$

$$= SQ(\text{innerhalb}) + SQ(\text{zwischen}) \quad (14.90)$$

$$SQT = \sum_{ij} (\mathbf{y}_{ij} - \bar{\mathbf{y}})(\mathbf{y}_{ij} - \bar{\mathbf{y}})' \quad (14.91)$$

$$SQR = \sum_{ij} (\mathbf{y}_{ij} - \bar{\mathbf{y}}_i)(\mathbf{y}_{ij} - \bar{\mathbf{y}}_i)' \quad (14.92)$$

$$SQE = \sum_{ij} (\bar{\mathbf{y}}_i - \bar{\mathbf{y}})(\bar{\mathbf{y}}_i - \bar{\mathbf{y}})' \quad (14.93)$$

Die gesamte Streuungs-Matrix SQT (bzw. Kovarianz-Matrix MQT) setzt sich also aus der Streuung **innerhalb** der Gruppen und der Streuung **zwischen** den Gruppen zusammen.

Die Residual-Streuung MQR ergibt nach Normierung die Matrix der **partiellen Korrelationen** zwischen den Komponenten von \mathbf{y} , da Z kontrolliert wurde.

Mittelwertsunterschiede zwischen den Gruppen (MQE) können Korrelationen der abhängigen Variablen induzieren, obwohl innerhalb der Gruppen keine oder umgekehrte Korrelationen vorliegen.

Im allgemeinen sind die relevanten Kontrollvariablen unbekannt.

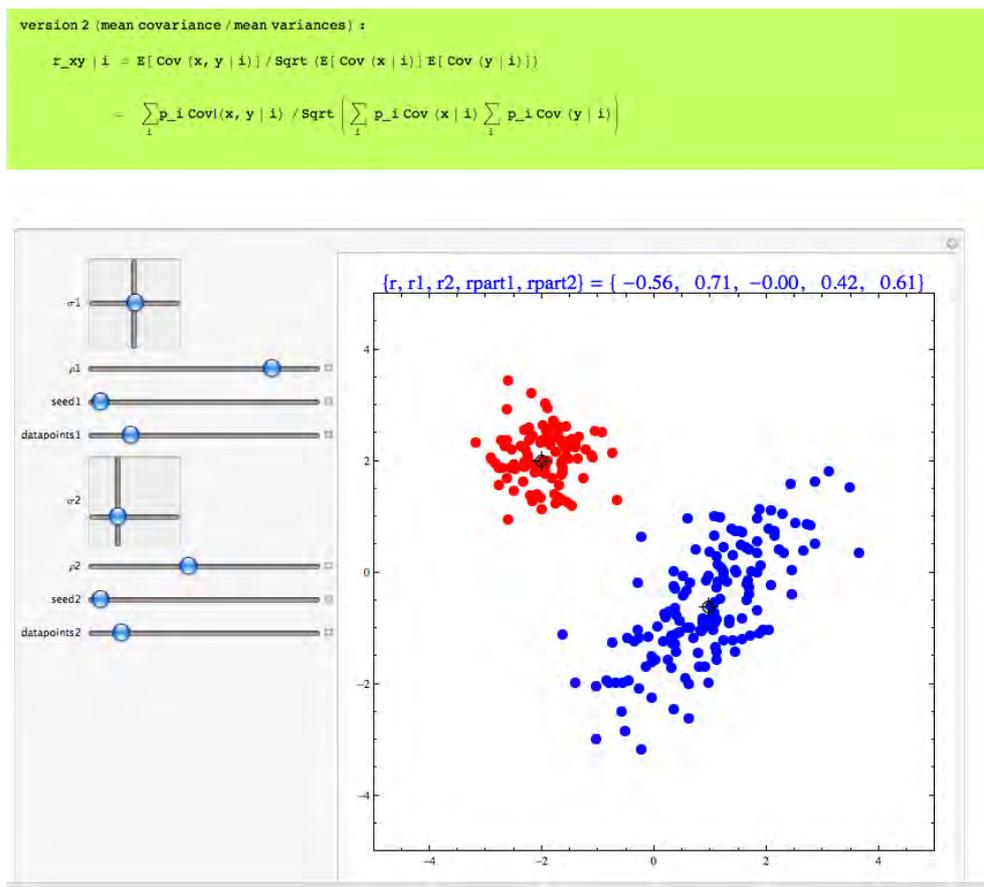


Abbildung 14.11: Partielle Korrelation der Variablen $y = [X, Y]'$ in 2 Gruppen $Z = (blau, rot)$. Obwohl die gesamte Korrelation $r_{xy} = -0.56$ ist, findet man innerhalb der Gruppen positive ($r_1 = 0.71$) oder Null-Korrelationen ($r_2 = 0$). Die partielle Korrelation $MQR(X, Y) / \sqrt{MQR(X, X)MQR(Y, Y)} = .61$ kontrolliert den Einfluß der Variable Gruppe (Applet *SpuriousCorrelation.nbp*).

http://www.fernuni-hagen.de/ls_statistik/lehre/

14.8.2 Stetige Kontrollvariablen: Allgemeiner Fall

Bisher wurden nur nominale oder ordinale Kontrollvariablen (Gruppen) diskutiert. Häufig hat man auch stetige Kontrollvariablen $\mathbf{z} = [Z_1, \dots, Z_q]'$, deren Einfluß auf die abhängigen Variablen $\mathbf{y} = [X, Y]'$ kontrolliert werden muß.

Wie im Fall der diskreten Kontrollvariablen $Z = i$ muß zunächst ein Prognosewert für die abhängigen Variablen berechnet werden, also

$$\hat{\mathbf{y}} = \hat{\mathbf{y}}(\mathbf{z}). \quad (14.94)$$

Ganz allgemein gilt:

Die **optimale Prognose** (im quadratischen Mittel) ist der bedingte Erwartungswert von \mathbf{y} , gegeben die Kontrollvariablen \mathbf{z}

optimale Prognose

$$\hat{\mathbf{y}} = E[\mathbf{y}|\mathbf{z}]. \quad (14.95)$$

Dies bedeutet, daß die optimale Prognose im allgemeinen eine nichtlineare Funktion der Kontrollvariablen \mathbf{z} ist, also

$$E[\mathbf{y}|\mathbf{z}] = \int \mathbf{y}f(\mathbf{y}|\mathbf{z})d\mathbf{y} = g(\mathbf{z}). \quad (14.96)$$

**nichtparametrische
Regressions-
modelle**

Die genaue Form von $g(\mathbf{z})$ ergibt sich aus der bedingten Dichtefunktion $f(\mathbf{y}|\mathbf{z}) = f(\mathbf{y}, \mathbf{z})/f(\mathbf{z})$, die i.a. unbekannt ist. In der Praxis kann $f(\mathbf{y}|\mathbf{z})$ aus Daten geschätzt werden, etwa als Kernschätzer $\hat{f}(\mathbf{y}|\mathbf{z})$. Dann erhält man sogenannte nichtparametrische Regressionsmodelle (vgl. Fahrmeir et al., 2007, Kap. 12).⁵

Die Prognose (14.95) ist **erwartungstreu** (unverzerrt) und es gilt

$$E[\hat{\mathbf{y}}] = E[E[\mathbf{y}|\mathbf{z}]] = E[\mathbf{y}] = \boldsymbol{\mu}. \quad (14.97)$$

Prognose-Fehler

Der Prognose-Fehler ist definiert als Abweichung

$$\tilde{\mathbf{y}} := \mathbf{y} - \hat{\mathbf{y}} \quad (14.98)$$

$$E[\tilde{\mathbf{y}}] = E[\mathbf{y} - E[\mathbf{y}|\mathbf{z}]] = E[\mathbf{y}] - E[\mathbf{y}] = \mathbf{0} \quad (14.99)$$

und die Prognose-Fehler-Varianz (mittlere Unsicherheit der Prognose) lautet

⁵ $\hat{E}[\mathbf{y}|\mathbf{z}] = \int \mathbf{y}\hat{f}(\mathbf{y}|\mathbf{z})d\mathbf{y} = \hat{g}(\mathbf{z})$.

$$E[\tilde{\mathbf{y}}\tilde{\mathbf{y}}'] = E\{E[\tilde{\mathbf{y}}\tilde{\mathbf{y}}'|\mathbf{z}]\} \quad (14.100)$$

$$= E[\text{Var}(\mathbf{y}|\mathbf{z})], \quad (14.101)$$

**Prognose-Fehler-
Varianz**

da $\hat{\mathbf{y}} = E[\mathbf{y}|\mathbf{z}]$ und $E[\tilde{\mathbf{y}}\tilde{\mathbf{y}}'|\mathbf{z}] = E[(\mathbf{y} - \hat{\mathbf{y}})(\mathbf{y} - \hat{\mathbf{y}})'|\mathbf{z}] = \text{Var}(\mathbf{y}|\mathbf{z})$. Aus der Prognose $E[\mathbf{y}|\mathbf{z}]$ und der bedingten Streuungsmatrix $\text{Var}(\mathbf{y}|\mathbf{z})$ lassen sich Sicherheitsbereiche (Prognose-Intervalle) konstruieren, etwa

$$E[\mathbf{y}|\mathbf{z}] \pm 1.96 \text{diag}(\text{Var}(\mathbf{y}|\mathbf{z}))^{1/2} \quad (14.102)$$

(approximatives 95% -Prognose-Intervall). Hierbei ist $\text{diag}(\mathbf{A})$ die Diagonale der Matrix \mathbf{A} .

Aus dem bisher Gesagten ergibt sich die allgemeine Streuungszzerlegung für die abhängige Variable \mathbf{y}

$$\text{Var}(\mathbf{y}) = E[(\mathbf{y} - \boldsymbol{\mu})(\mathbf{y} - \boldsymbol{\mu})'] \quad (14.103)$$

$$= E[(\mathbf{y} - \hat{\mathbf{y}} + \hat{\mathbf{y}} - \boldsymbol{\mu})(\mathbf{y} - \hat{\mathbf{y}} + \hat{\mathbf{y}} - \boldsymbol{\mu})'] \quad (14.104)$$

$$= E[\tilde{\mathbf{y}}\tilde{\mathbf{y}}'] + E[(\hat{\mathbf{y}} - \boldsymbol{\mu})(\hat{\mathbf{y}} - \boldsymbol{\mu})'] \quad (14.105)$$

$$= E[\text{Var}(\mathbf{y}|\mathbf{z})] + \text{Var}(E[\mathbf{y}|\mathbf{z}]). \quad (14.106)$$

Die gemischten Terme

$$E[\tilde{\mathbf{y}}(\hat{\mathbf{y}} - \boldsymbol{\mu})'] = E[E[\tilde{\mathbf{y}}(\hat{\mathbf{y}} - \boldsymbol{\mu})'|\mathbf{z}]] = E[E[\tilde{\mathbf{y}}|\mathbf{z}](E[\mathbf{y}|\mathbf{z}] - \boldsymbol{\mu})'] = \mathbf{O}$$

verschwinden, da $E[\tilde{\mathbf{y}}|\mathbf{z}] = \hat{\mathbf{y}} - E[\hat{\mathbf{y}}|\mathbf{z}] = \mathbf{0}$ ($\mathbf{O} : p \times p$ -Nullmatrix).

Dies bedeutet, daß die Prognose $E[\mathbf{y}|\mathbf{z}]$ und der Prognosefehler $\tilde{\mathbf{y}} = \mathbf{y} - E[\mathbf{y}|\mathbf{z}]$ unkorreliert sind, d.h.

$$E[(\mathbf{y} - E[\mathbf{y}|\mathbf{z}])(E[\mathbf{y}|\mathbf{z}] - \boldsymbol{\mu})'] = \text{Cov}(\tilde{\mathbf{y}}, \hat{\mathbf{y}}) = \mathbf{O}. \quad (14.107)$$

Eine geometrische Interpretation besagt, daß Prognose und Prognosefehler orthogonal sind (Abb. 14.12). Da Orthogonalität gilt, ergibt sich die allgemeine Streuungszzerlegung

**geometrische
Interpretation**

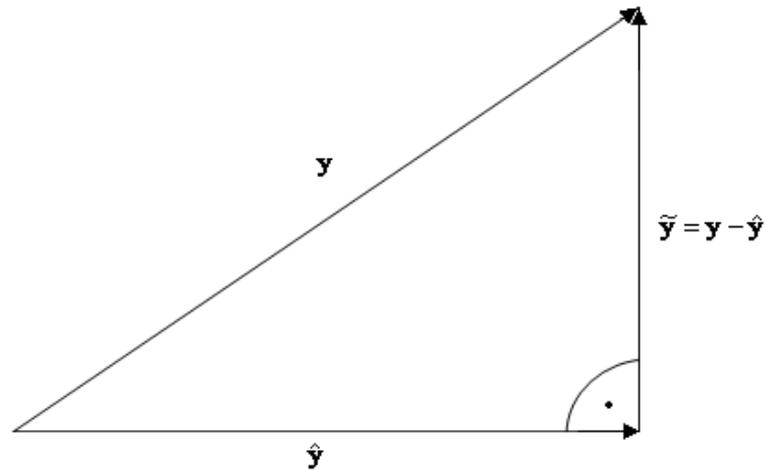


Abbildung 14.12: Geometrische Interpretation von Prognose und Prognosefehler (orthogonale Zufallsvariablen). Daher enthält die Streuungszersetzung von \mathbf{y} keine Korrelationsterme $\text{Cov}(\hat{\mathbf{y}}, \tilde{\mathbf{y}})$.

$$\text{Var}(\mathbf{y}) = E[\tilde{\mathbf{y}}\tilde{\mathbf{y}}'] + E[(\hat{\mathbf{y}} - \boldsymbol{\mu})(\hat{\mathbf{y}} - \boldsymbol{\mu})'] \quad (14.108)$$

$$= E[\text{Var}(\mathbf{y}|\mathbf{z})] + \text{Var}(E[\mathbf{y}|\mathbf{z}]) \quad (14.109)$$

$$MQT = MQR + MQE \quad (14.110)$$

Streuungszerlegung

Die gesamte Streuung ist also aus der gemittelten Residualstreuung (partielle Kovarianz) und der Streuung der Prognose (erklärte Streuung) zusammengesetzt.

Die **partielle Korrelations-Matrix** ergibt sich aus der gemittelten Residualstreuung MQR durch Standardisierung.

14.8.3 Stetige Kontrollvariablen: Linearer Fall

In der Praxis werden häufig nur lineare Prognosen betrachtet, da die genauen nichtlinearen Zusammenhänge zwischen \mathbf{y} und \mathbf{z} nicht bekannt sind. Ein linearer Prädiktor $\boldsymbol{\alpha} + \mathbf{Bz}$ ist dann die einfachste Näherung für $g(\mathbf{z})$.

Weiterhin gilt für gemeinsam Gauß(Normal)-verteilte Variablen $[\mathbf{y}', \mathbf{z}']$, daß die **optimale Prognose linear** ist, d.h.

$$E[\mathbf{y}|\mathbf{z}] = E[\mathbf{y}] + \text{Cov}(\mathbf{y}, \mathbf{z})\text{Cov}(\mathbf{z}, \mathbf{z})^{-1}(\mathbf{z} - E[\mathbf{z}]) \quad (14.111)$$

$$\text{Var}[\mathbf{y}|\mathbf{z}] = \text{Var}(\mathbf{y}) - \text{Cov}(\mathbf{y}, \mathbf{z})\text{Cov}(\mathbf{z}, \mathbf{z})^{-1}\text{Cov}(\mathbf{z}, \mathbf{y}). \quad (14.112)$$

**Satz von der
Normalkorrelation**

Diese **exakte Linearität für Gauß-Variablen** ist die Grundlage der linearen Prognose-Modelle, auch wenn die Variablen nicht normalverteilt sind.

Wenn nur geringe Abweichungen von der Normalverteilung vorliegen, ist eine lineare Prognose ausreichend.

Die **optimale lineare Prognose** im quadratischen Mittel ist durch das lineare Regressionsmodell

**optimale lineare
Prognose**

$$\mathbf{y} = \boldsymbol{\alpha} + \mathbf{B}\mathbf{z} + \boldsymbol{\epsilon} : p \times 1 \quad (14.113)$$

mit Parametern

$$\boldsymbol{\alpha} = E[\mathbf{y}] - \mathbf{B}E[\mathbf{z}] : p \times 1 \quad (14.114)$$

$$\mathbf{B} = \text{Cov}(\mathbf{y}, \mathbf{z})\text{Cov}(\mathbf{z}, \mathbf{z})^{-1} : p \times q \quad (14.115)$$

gegeben.⁶

Dabei wird unterstellt, daß die Regressoren \mathbf{z} und Gleichungsfehler $\boldsymbol{\epsilon}$ unkorreliert sind ($\text{Cov}(\mathbf{z}, \boldsymbol{\epsilon}) = \mathbf{O}$; vgl. 14.107).

Als Prognose (linearer Prädiktor) ergibt sich

$$\hat{\mathbf{y}} = \boldsymbol{\alpha} + \mathbf{B}\mathbf{z} \quad (14.116)$$

$$= E[\mathbf{y}] + \mathbf{B}(\mathbf{z} - E[\mathbf{z}]) \quad (14.117)$$

$$= E[\mathbf{y}] + \text{Cov}(\mathbf{y}, \mathbf{z})\text{Cov}(\mathbf{z}, \mathbf{z})^{-1}(\mathbf{z} - E[\mathbf{z}]) \quad (14.118)$$

$$= E[\mathbf{y}|\mathbf{z}] \text{ bei Normalverteilung}$$

linearer Prädiktor

⁶Man differenziert $S(\boldsymbol{\alpha}, \mathbf{B}) = E[\mathbf{y} - \hat{\mathbf{y}}]^2 = E[\mathbf{y} - \boldsymbol{\alpha} - \mathbf{B}\mathbf{z}]^2$ nach $\boldsymbol{\alpha}$ und \mathbf{B} und erhält so die angegebenen optimalen Parameter. Setzt man die optimalen Parameter ein, so ist der minimale Wert $S(\boldsymbol{\alpha}, \mathbf{B}) = (\mathbf{y}, \mathbf{y}) - (\mathbf{y}, \mathbf{z})(\mathbf{z}, \mathbf{z})^{-1}(\mathbf{z}, \mathbf{y}) = \text{Var}(\hat{\mathbf{y}}) = MQR$.

(vgl. Mardia et al., 1979, Kap. 3.3).

\mathbf{z} ist also nur informativ für \mathbf{y} , wenn die Variablen korreliert sind, also $\text{Cov}(\mathbf{y}, \mathbf{z}) \neq \mathbf{O}$. Falls sie unkorreliert sind, ist die beste Prognose für \mathbf{y} die Konstante $E[\mathbf{y}]$.

erklärte Streuung
MQE

Der lineare Prädiktor ist **erwartungstreu** ($E[\hat{\mathbf{y}}] = E[\mathbf{y}] = \boldsymbol{\mu}$) und hat die Varianz

$$\text{Var}(\hat{\mathbf{y}}) = \text{Cov}(\mathbf{y}, \mathbf{z})\text{Cov}(\mathbf{z}, \mathbf{z})^{-1}\text{Cov}(\mathbf{z}, \mathbf{y}). \quad (14.119)$$

Prognose-Fehler

Die Prognose-Fehler sind wieder die Abweichungen der Prognose vom wahren Wert

$$\tilde{\mathbf{y}} := \mathbf{y} - \hat{\mathbf{y}}. \quad (14.120)$$

Die Korrelation der Komponenten des Prognosefehlers $\tilde{\mathbf{y}} = [\tilde{y}_1, \dots, \tilde{y}_p]'$ wird als **partielle Korrelation** $\rho(\tilde{y}_i, \tilde{y}_j)$ bezeichnet.

Explizit ergibt sich aus der Definition des Prognose-Fehlers die Streuungszersetzung ⁷

Streuungszersetzung

$$(\tilde{\mathbf{y}}, \tilde{\mathbf{y}}) = (\mathbf{y} - \hat{\mathbf{y}}, \mathbf{y} - \hat{\mathbf{y}}) \quad (14.121)$$

$$= (\mathbf{y}, \mathbf{y}) - (\mathbf{y}, \hat{\mathbf{y}}) - (\hat{\mathbf{y}}, \mathbf{y}) + (\hat{\mathbf{y}}, \hat{\mathbf{y}}) \quad (14.122)$$

$$= (\mathbf{y}, \mathbf{y}) - (\hat{\mathbf{y}}, \hat{\mathbf{y}}) \quad (14.123)$$

$$= (\mathbf{y}, \mathbf{y}) - (\mathbf{y}, \mathbf{z})(\mathbf{z}, \mathbf{z})^{-1}(\mathbf{z}, \mathbf{y}) \quad (14.124)$$

$$MQR = MQT - MQE. \quad (14.125)$$

Hierbei wurde die Identität

$$(\mathbf{y}, \hat{\mathbf{y}}) = (\mathbf{y}, \mathbf{Bz}) = (\mathbf{y}, \mathbf{z})(\mathbf{z}, \mathbf{z})^{-1}(\mathbf{z}, \mathbf{y}) = (\hat{\mathbf{y}}, \hat{\mathbf{y}}) \quad (14.126)$$

eingesetzt.

Daraus folgt, daß Prognose-Fehler und Prognose orthogonal (unkorreliert) sind, d.h.

$$(\tilde{\mathbf{y}}, \hat{\mathbf{y}}) = (\mathbf{y} - \hat{\mathbf{y}}, \hat{\mathbf{y}}) \quad (14.127)$$

$$= (\mathbf{y}, \hat{\mathbf{y}}) - (\hat{\mathbf{y}}, \hat{\mathbf{y}}) \quad (14.128)$$

$$= (\hat{\mathbf{y}}, \hat{\mathbf{y}}) - (\hat{\mathbf{y}}, \hat{\mathbf{y}}) = \mathbf{O} \quad (14.129)$$

⁷Um die Notation zu vereinfachen, wird im folgenden die Abkürzung $\text{Cov}(\mathbf{y}, \mathbf{z}) = (\mathbf{y}, \mathbf{z})$, $\text{Var}(\mathbf{y}) = (\mathbf{y}, \mathbf{y})$ benutzt.

Rechenregeln: $(\mathbf{a} + \mathbf{Bx}, \mathbf{c} + \mathbf{Dy}) = \mathbf{B}(\mathbf{x}, \mathbf{y})\mathbf{D}'$ (Bilinear-Form),
da $(\mathbf{x}, \mathbf{y}) = E[\mathbf{xy}'] - E[\mathbf{x}]E[\mathbf{y}']$.

(Abb. 14.12).

Die Korrelation zwischen Prognose $\hat{\mathbf{y}}$ und abhängiger Variable \mathbf{y} wird auch als **multiple Korrelation** bezeichnet. Es gilt explizit für die Komponenten $i = 1, \dots, p$ (vgl. (14.126))

$$\text{Corr}^2(\mathbf{y}_i, \hat{\mathbf{y}}_i) = \frac{(\hat{\mathbf{y}}_i, \hat{\mathbf{y}}_i)^2}{(\mathbf{y}_i, \mathbf{y}_i)(\hat{\mathbf{y}}_i, \hat{\mathbf{y}}_i)} \quad (14.130)$$

$$= \frac{(\hat{\mathbf{y}}_i, \hat{\mathbf{y}}_i)}{(\mathbf{y}_i, \mathbf{y}_i)}. \quad (14.131)$$

Dies ist aber der Quotient der Diagonalelemente aus erklärter und totaler Streuung SQE_{ii}/SQT_{ii} (Determinationskoeffizient). Somit findet man:

Determinationskoeffizient = quadrierte multiple Korrelation.

Zusammenfassend gilt:

**multiple
Korrelation**

$$\begin{aligned} (\tilde{\mathbf{y}}, \tilde{\mathbf{y}}) &= (\mathbf{y}, \mathbf{y}) - (\mathbf{y}, \mathbf{z})(\mathbf{z}, \mathbf{z})^{-1}(\mathbf{z}, \mathbf{y}) & (14.132) \\ MQR &= MQT - MQE \\ &= \text{Var}(\mathbf{y}|\mathbf{z}) \text{ bei Normalverteilung.} \end{aligned}$$

Residualstreuung

Im Beispiel mit $p = 2$ Variablen $\mathbf{y} = [X, Y]'$ findet man für die Residual-Kovarianzen von X und Y :

$$(\tilde{X}, \tilde{Y}) = (X, Y) - (X, \mathbf{z})(\mathbf{z}, \mathbf{z})^{-1}(\mathbf{z}, Y) \quad (14.133)$$

$$(\tilde{X}, \tilde{X}) = (X, X) - (X, \mathbf{z})(\mathbf{z}, \mathbf{z})^{-1}(\mathbf{z}, X) \quad (14.134)$$

$$(\tilde{Y}, \tilde{Y}) = (Y, Y) - (Y, \mathbf{z})(\mathbf{z}, \mathbf{z})^{-1}(\mathbf{z}, Y). \quad (14.135)$$

Standardisierung ergibt die partielle Korrelation

**partielle
Korrelation**

$$\begin{aligned}\rho(\tilde{X}, \tilde{Y}) &:= \rho_{XY.Z} = \frac{(\tilde{X}, \tilde{Y})}{\sqrt{(\tilde{X}, \tilde{X})}\sqrt{(\tilde{Y}, \tilde{Y})}} & (14.136) \\ &= \frac{\rho(X, Y) - \rho(X, \mathbf{z})\rho(\mathbf{z}, \mathbf{z})^{-1}\rho(\mathbf{z}, Y)}{\sqrt{1 - \rho(X, \mathbf{z})\rho(\mathbf{z}, \mathbf{z})^{-1}\rho(\mathbf{z}, X)}\sqrt{1 - \rho(Y, \mathbf{z})\rho(\mathbf{z}, \mathbf{z})^{-1}\rho(\mathbf{z}, Y)}}.\end{aligned}$$

Hat man nur eine Kontrollvariable $\mathbf{z} = Z$, so gilt $\rho(\mathbf{z}, \mathbf{z}) = 1$ und es ergibt sich die bekannte Formel

$$\begin{aligned}\rho(\tilde{X}, \tilde{Y}) &:= \rho_{XY.Z} = \frac{(\tilde{X}, \tilde{Y})}{\sqrt{(\tilde{X}, \tilde{X})}\sqrt{(\tilde{Y}, \tilde{Y})}} & (14.137) \\ &= \frac{\rho(X, Y) - \rho(X, Z)\rho(Z, Y)}{\sqrt{1 - \rho(X, Z)^2}\sqrt{1 - \rho(Y, Z)^2}}.\end{aligned}$$

Ersetzt man die theoretischen Korrelationen durch Stichproben-Korrelationen, etwa r_{xy} , so ergibt sich der **empirische partielle Korrelationskoeffizient**.

Allgemeiner kann man aus der Streuungszerlegung eine partielle Korrelationsmatrix $\rho(\tilde{\mathbf{y}}, \tilde{\mathbf{y}}) : p \times p$ ableiten.

Beispiel 14.5 (Kosmetik (X) und Körpergewicht (Y))

Zwar ist die Variable $Z = \text{Geschlecht}$ nominal, jedoch wurde sie als Indikator-Variable codiert. Daher ist die eben beschriebene Theorie anwendbar. Aus dem SPSS-Output (Abb. 14.8, oben) kann man ablesen:

$$r_{\tilde{x}\tilde{y}} = \frac{r_{xy} - r_{xz}r_{zy}}{\sqrt{1 - r_{xz}^2}\sqrt{1 - r_{yz}^2}} \quad (14.138)$$

$$= \frac{-.588 - (.753)(-.810)}{\sqrt{1 - (.753)^2}\sqrt{1 - (-.810)^2}} = 0.0568 \quad (14.139)$$

($r_{zz} = 1$). Dies ist aber genau der Wert aus Abb. 14.9.



Beispiel 14.6 (Quadratische Einflüsse der Kontrollvariablen Z)

In diesem Beispiel wird gezeigt, daß eine lineare Kontrolle der Variablen X und Y nicht ausreicht, um die exakte partielle Korrelation zu berechnen.

Zunächst wird der (wahre) quadratische Zusammenhang

$$X = Z^2 + \epsilon_1 \quad (14.140)$$

$$Y = Z^2 + \epsilon_2 \quad (14.141)$$

unterstellt (Abb. 14.14, oben).

Weitere Annahmen:

Z und ϵ_i seien unabhängig und es gelte $Z \sim N(0, 1)$, $\epsilon_i \sim N(0, \sigma^2)$.

Dann ist die **exakte Prognose**

$$\hat{X} = E[X|Z] = Z^2 \quad (14.142)$$

$$\hat{Y} = E[Y|Z] = Z^2 \quad (14.143)$$

und somit der Prognosefehler

$$\tilde{X} = X - \hat{X} = \epsilon_1 \quad (14.144)$$

$$\tilde{Y} = Y - \hat{Y} = \epsilon_2. \quad (14.145)$$

Daher verschwindet die **optimale nichtlineare partielle Korrelation**

$$\text{Corr}(\tilde{X}, \tilde{Y}) = 0. \quad (14.146)$$

Die gesamte Kovarianz zerlegt sich in $\text{Cov}(X, Y) = E[\tilde{X}\tilde{Y}] + \text{Cov}(\hat{X}, \hat{Y}) = E[\epsilon_1\epsilon_2] + \text{Cov}(Z^2, Z^2) = 0 + (E[Z^4] - E^2[Z^2]) = 0 + 3 - 1 = 2$.

Nimmt man nur eine **lineare Prädiktion**, so ergibt sich:

$$\hat{X} = E[X] + (X, Z)(Z, Z)^{-1}(Z - E[Z]) \quad (14.147)$$

$$\hat{Y} = E[Y] + (Y, Z)(Z, Z)^{-1}(Z - E[Z]). \quad (14.148)$$

Nun ist $E[X] = E[Z^2] = 1$ und $(X, Z) = (Z^2 + \epsilon_1, Z) = 0$ (analog für Y). Deshalb ist der optimale lineare Prädiktor eine Konstante

$$\hat{X} = 1 \quad (14.149)$$

$$\hat{Y} = 1 \quad (14.150)$$

(eine Parabel kann nur schlecht durch eine Gerade approximiert werden).

Die Prognose-Fehler sind daher

$$\tilde{X} = X - 1 \quad (14.151)$$

$$\tilde{Y} = Y - 1 \quad (14.152)$$

und die Kovarianz ist

$$\begin{aligned} \text{Cov}(\tilde{X}, \tilde{Y}) &= (X - 1, Y - 1) = (X, Y) = (Z^2 + \epsilon_1, Z^2 + \epsilon_2) \\ &= (Z^2, Z^2) = E[Z^4] - E^2[Z^2] = 3 - 1 = 2. \end{aligned} \quad (14.153)$$

Für die Varianzen gilt (analog für Y)

$$\begin{aligned} \text{Cov}(\tilde{X}, \tilde{X}) &= (X, X) = (Z^2 + \epsilon_1, Z^2 + \epsilon_1) = (Z^2, Z^2) + (\epsilon_1, \epsilon_1) \\ &= E[Z^4] - E^2[Z^2] + \sigma^2 = 3 - 1 = 2 + \sigma^2. \end{aligned} \quad (14.154)$$

Daher ist die **lineare partielle Korrelation**

$$\text{Corr}(\tilde{X}, \tilde{Y}) = \frac{2}{2 + \sigma^2}. \quad (14.155)$$

Für kleines σ^2 ist dies nahezu 1, obwohl die exakte partielle Korrelation (14.146) verschwindet. Die lineare Prognose ist nicht in der Lage, den Einfluß der Drittvariable adäquat zu eliminieren. Für große σ^2 hat Z einen vernachlässigbaren Einfluß auf X und Y , sodaß auch die lineare partielle Korrelation gegen 0 geht.

Praktisches Vorgehen:

Im Fall eines nichtlinearen Einflusses der Kontrollvariable sollte man eine zumindest quadratische Regression von Z auf X und Y rechnen und dann die Residuen \tilde{X}, \tilde{Y} korrelieren.

Zu diesem Zweck wurde ein SPSS-Datensatz mit $N = 10$ Einheiten simuliert und das wahre Modell Glg. 14.140 unterstellt. Abb. 14.13 (oben) zeigt die Berechnung der Variable $Y = Z^2 + e2$ aus Z und $e2$. Im unteren Teil ist die Simulation der Störvariable (Gleichungsfehler) $e2$ abgebildet. Abb. 14.14 zeigt oben eine lineare und quadratische Regression von Z auf X und Y . Im unteren Bild sind die Residuen \tilde{X} für das lineare und quadratische Modell aufgetragen. Man beachte die Skalierung der Ordinate (y-Achse). Das quadratische Modell paßt wesentlich besser, wie man der Modellzusammenfassung entnehmen kann (Abb. 14.15 oben, $R^2 = .822 = \text{Erklärte Streuung}/\text{Totale Streuung im quadratischen Modell}$).

Die geschätzte lineare partielle Korrelation $r_{xy.z} = r_{\tilde{x}\tilde{y}} = .738$ ist in Abb. 14.16 (oben) angegeben (theoretischer Wert = $2/3$). Der gleiche

Wert ergibt sich bei Korrelation der Regressionsresiduen (\tilde{x}, \tilde{y}) (Mitte). Die exakte partielle Korrelation ergibt sich nach quadratischer Kontrolle der Z -Variablen (unten). Um die Residuen zu korrelieren, müssen sie zunächst vom Regressionsmodell im Datensatz (`Quadrat.sav`) abgespeichert werden (Variablen $err_1 - err_4$).



Übung 14.3 (Quadratische Einflüsse der Kontrollvariablen Z)

- (a) Wiederholen Sie die Analyse mit anderen Zufallszahlen.
- (b) Vergrößern Sie die Stichprobe auf $N = 20$, $N = 50$.



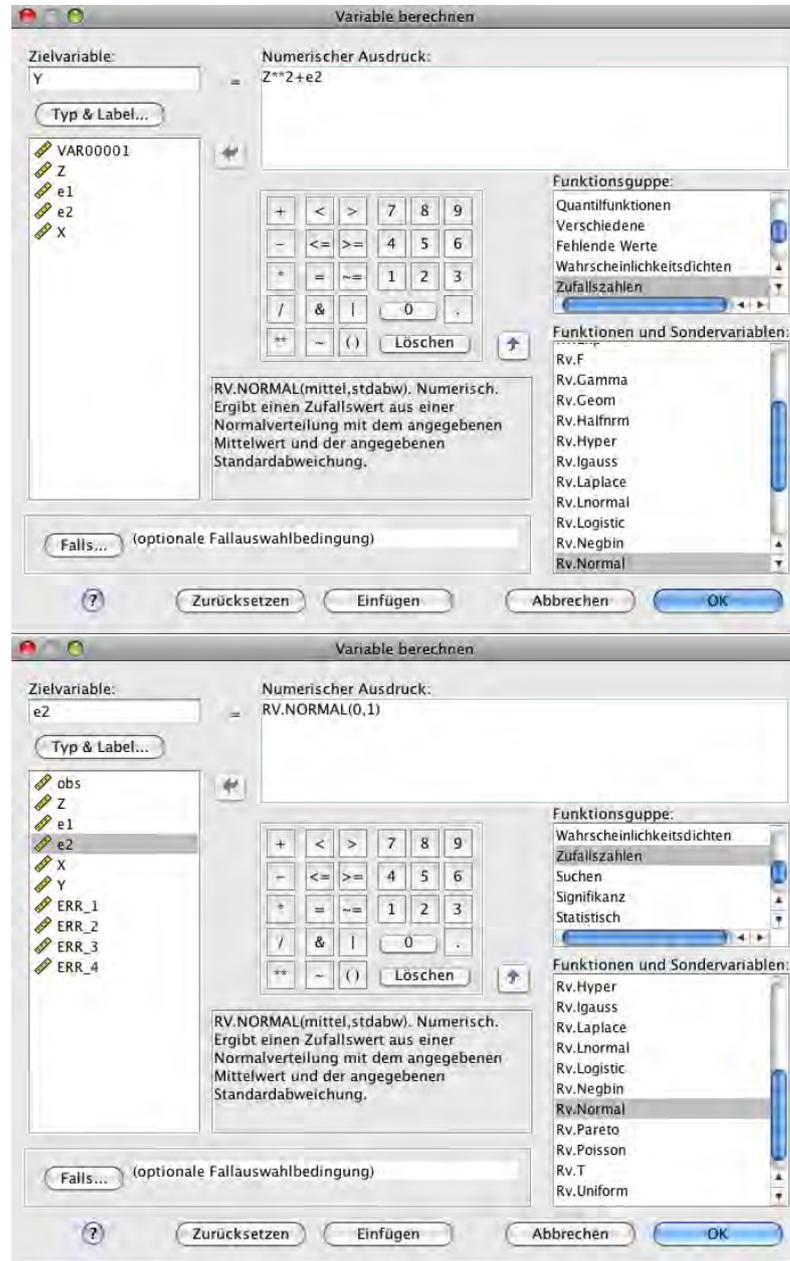


Abbildung 14.13: Berechnung der Variable $Y = Z^2 + e2$. Vorher wurden Z und $e2$ als normalverteilte Variablen $N(0, 1)$ simuliert.

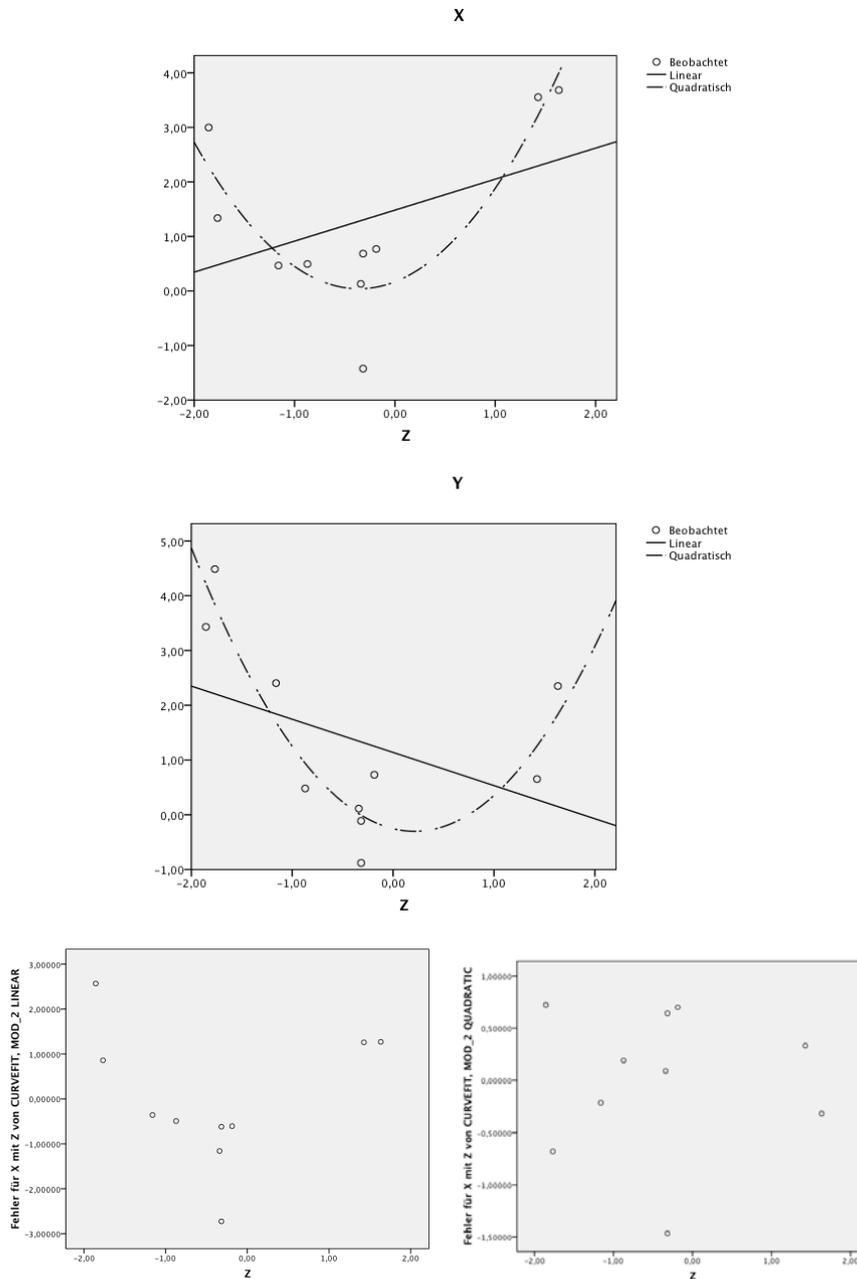


Abbildung 14.14: Lineare und quadratische Regression von Z auf X und Y (oben, mitte). Residuen \hat{X} (lineares Modell: unten, links; quadratisches Modell: unten, rechts). Das quadratische Modell paßt wesentlich besser.

Modellzusammenfassung und Parameterschätzer

Abhängige Variable: X

Gleichung	Modellzusammenfassung					Parameterschätzer		
	R-Quadrat	F	Freiheitsgrad e 1	Freiheitsgrad e 2	Sig.	Konstante	b1	b2
Linear	,163	1,559	1	8	,247	1,483	,569	
Quadratisch	,822	16,211	2	7	,002	,169	,716	,997

Die unabhängige Variable ist Z.

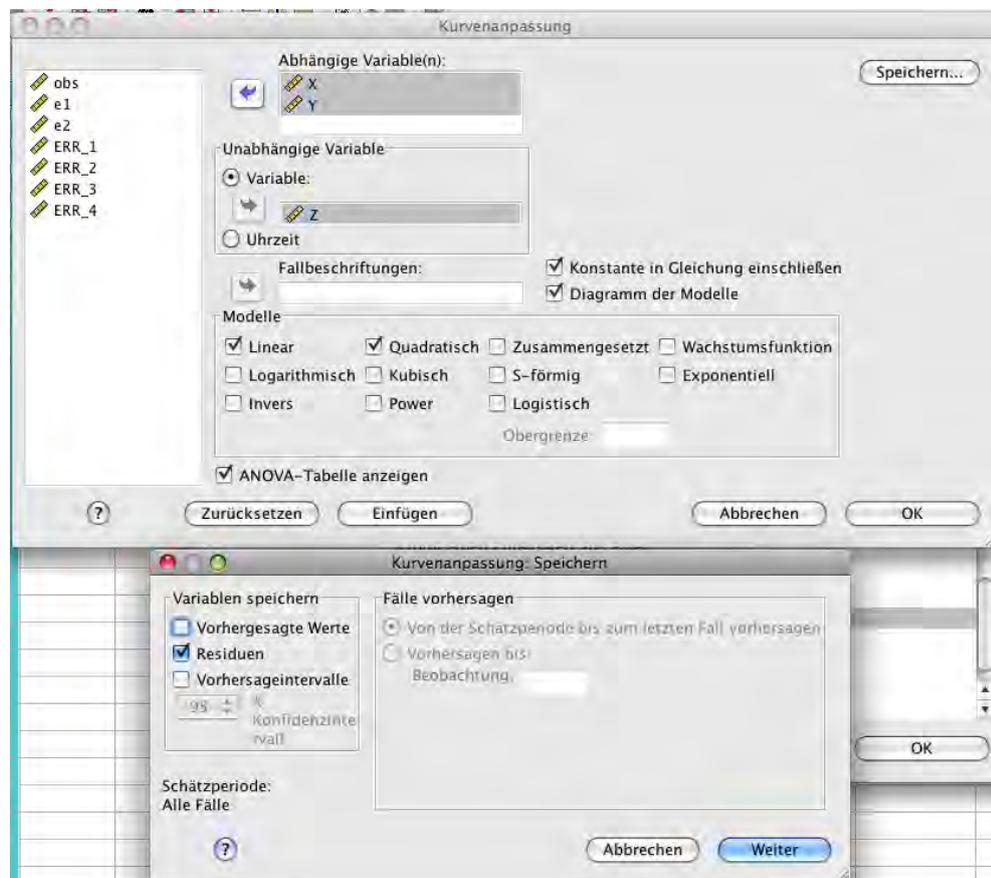


Abbildung 14.15: Lineare und quadratische Regression von Z auf X . Das quadratische Modell paßt wesentlich besser ($R^2 = .822 = \text{Erklärte Streuung/ Totale Streuung}$).

Korrelationen

Kontrollvariablen			X	Y
Z	X	Korrelation	1,000	,738
		Signifikanz (zweiseitig)	.	,023
		Freiheitsgrade	0	7
Y	Y	Korrelation	,738	1,000
		Signifikanz (zweiseitig)	,023	.
		Freiheitsgrade	7	0

Korrelationen

		Fehler für X mit Z von CURVEFIT, MOD_2 LINEAR	Fehler für Y mit Z von CURVEFIT, MOD_2 LINEAR
Fehler für X mit Z von CURVEFIT, MOD_2 LINEAR	Korrelation nach Pearson	1	,738*
	Signifikanz (2-seitig)		,015
	Quadratsummen und Kreuzprodukte	20,437	15,662
	Kovarianz	2,271	1,740
	N	10	10
Fehler für Y mit Z von CURVEFIT, MOD_2 LINEAR	Korrelation nach Pearson	,738*	1
	Signifikanz (2-seitig)	,015	
	Quadratsummen und Kreuzprodukte	15,662	22,029
	Kovarianz	1,740	2,448
	N	10	10

*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Korrelationen

		Fehler für X mit Z von CURVEFIT, MOD_2 QUADRATIC	Fehler für Y mit Z von CURVEFIT, MOD_2 QUADRATIC
Fehler für X mit Z von CURVEFIT, MOD_2 QUADRATIC	Korrelation nach Pearson	1	-,341
	Signifikanz (2-seitig)		,335
	Quadratsummen und Kreuzprodukte	4,336	-1,408
	Kovarianz	,482	-,156
	N	10	10
Fehler für Y mit Z von CURVEFIT, MOD_2 QUADRATIC	Korrelation nach Pearson	-,341	1
	Signifikanz (2-seitig)	,335	
	Quadratsummen und Kreuzprodukte	-1,408	3,931
	Kovarianz	-,156	,437
	N	10	10

Abbildung 14.16: Oben: Partielle Korrelation .738 von X und Y (Z linear kontrolliert). Mitte: Residualkorrelation .738 (lineare Prognose). Unten: Residualkorrelation -.341 (quadratische Prognose). Exakte Populations-Parameter: Nichtlineare partielle Korrelation = 0, lineare partielle Korrelation $\text{Corr}(\tilde{X}, \tilde{Y}) = 2/(2 + \sigma^2) = 2/3$. Große p -Werte aufgrund der geringen Stichprobengröße.

14.8.4 * Stetige Kontrollvariablen: Multivariate Regression

Im letzten Abschnitt wurde die partielle Korrelation als theoretische Größe hergeleitet und die Momente durch Stichprobengrößen ersetzt. Hier soll nun eine Stichprobenversion der Streuungszerlegung abgeleitet werden. Die residuale Kovarianzmatrix enthält dann nach Normierung die partiellen Korrelationen.

Man betrachtet das multivariate Regressionsmodell für die p abhängigen Variablen $\mathbf{y} = [y_1, \dots, y_p]'$

**multivariates
Regressionsmodell**

$$\mathbf{y}_n = \boldsymbol{\alpha} + \mathbf{B}\mathbf{z}_n + \boldsymbol{\epsilon}_n; n = 1, \dots, N. \quad (14.156)$$

Hierbei werden die Daten für alle Personen n als Spalten-Vektoren $\mathbf{y}_n : p \times 1$, $\mathbf{z}_n : q \times 1$ geschrieben und $\boldsymbol{\epsilon}_n : p \times 1$ ist ein Gleichungs-Fehler mit Kovarianz-Matrix $\text{Cov}(\boldsymbol{\epsilon}_n) = E[\boldsymbol{\epsilon}_n \boldsymbol{\epsilon}_n'] = \boldsymbol{\Sigma} : p \times p$.

In Komponenten lautet das Modell

$$y_{ni} = \alpha_i + \beta_{i1}z_{n1} + \dots + \beta_{iq}z_{nq} + \epsilon_{ni}; n = 1, \dots, N; i = 1, \dots, p.$$

Bei Regressionsmodellen wird normalerweise angenommen, daß die unabhängigen Variablen deterministisch sind. Bei empirischen Studien sind die Variablen jedoch im allgemeinen **stochastisch**, sofern kein experimentelles Design vorliegt (vgl. Fahrmeir et al., 1996, Kap. 4, S. 96).

In diesem Fall gelten alle Aussagen nur **bedingt auf die Regressoren**.

In *Matrix-Form* kann das Modell sehr einfach angeschrieben werden:⁸

$$\mathbf{Y} = \mathbf{1}\boldsymbol{\alpha}' + \mathbf{Z}\mathbf{B}' + \boldsymbol{\epsilon} : N \times p \quad (14.157)$$

wobei $\mathbf{1} := [1, \dots, 1] : N \times 1$ ein Vektor von Einsen ist. Man schichtet also alle Daten zeilenweise in die Datenmatrizen $\mathbf{Y}' = [\mathbf{y}_1, \dots, \mathbf{y}_N] : p \times N$ und $\mathbf{Z}' = [\mathbf{z}_1, \dots, \mathbf{z}_N] : q \times N$.

Die Parameter-Matrizen lauten:

- $\boldsymbol{\alpha}' : 1 \times p$ (Achsenabschnitt) und

⁸Die übliche Form des Regressionmodells ist $\mathbf{Y} = \mathbf{Z}_1\mathbf{B}_1 + \boldsymbol{\epsilon}$ mit $\mathbf{Z}_1 = [\mathbf{1}, \mathbf{Z}]$ und $\mathbf{B}_1 = [\boldsymbol{\alpha}, \mathbf{B}]' : (q+1) \times p$. Als KQ-Schätzer ergibt sich $\hat{\mathbf{B}} = (\mathbf{Z}_1'\mathbf{Z}_1)^{-1}\mathbf{Z}_1'\mathbf{Y}$.

- $\mathbf{B}' : q \times p$ (Steigung).

Aus der Quadratsumme

$$S(\boldsymbol{\alpha}, \mathbf{B}) = \sum_n \boldsymbol{\epsilon}'_n \boldsymbol{\epsilon}_n \quad (14.158)$$

$$\boldsymbol{\epsilon}_n = \mathbf{y}_n - \boldsymbol{\alpha} - \mathbf{B}\mathbf{z}_n \quad (14.159)$$

läßt sich durch Minimierung der kleinste-Quadrate-Schätzer (KQ) für die Parameter gewinnen:

$$\hat{\boldsymbol{\alpha}}' = \bar{\mathbf{y}}' - \bar{\mathbf{z}}' \hat{\mathbf{B}}' \quad (14.160)$$

$$\hat{\mathbf{B}}' = (\mathbf{Z}'_* \mathbf{Z}_*)^{-1} \mathbf{Z}'_* \mathbf{Y}_* \quad (14.161)$$

KQ-Schätzer

Hierbei sind

$$\mathbf{Y}_* = \mathbf{Y} - \mathbf{1}\bar{\mathbf{y}}' = \mathbf{H}\mathbf{Y} \quad (14.162)$$

$$\mathbf{Z}_* = \mathbf{Z} - \mathbf{1}\bar{\mathbf{z}}' = \mathbf{H}\mathbf{Z} \quad (14.163)$$

die auf die **Mittelwerte**

$$\bar{\mathbf{y}}' = (1/N)\mathbf{1}'\mathbf{Y} : 1 \times p \quad (14.164)$$

$$\bar{\mathbf{z}}' = (1/N)\mathbf{1}'\mathbf{Z} : 1 \times q \quad (14.165)$$

zentrierten Datenmatrizen.

Bei \mathbf{H} handelt es sich um die **Zentrierungsmatrix**

Zentrierungsmatrix

$$\mathbf{H} = \mathbf{I} - (1/N)\mathbf{1}\mathbf{1}' := \mathbf{I} - \mathbf{M} \quad (14.166)$$

$$\mathbf{H}^2 = \mathbf{H} \quad (14.167)$$

$$\mathbf{M}^2 = \mathbf{M} \quad (14.168)$$

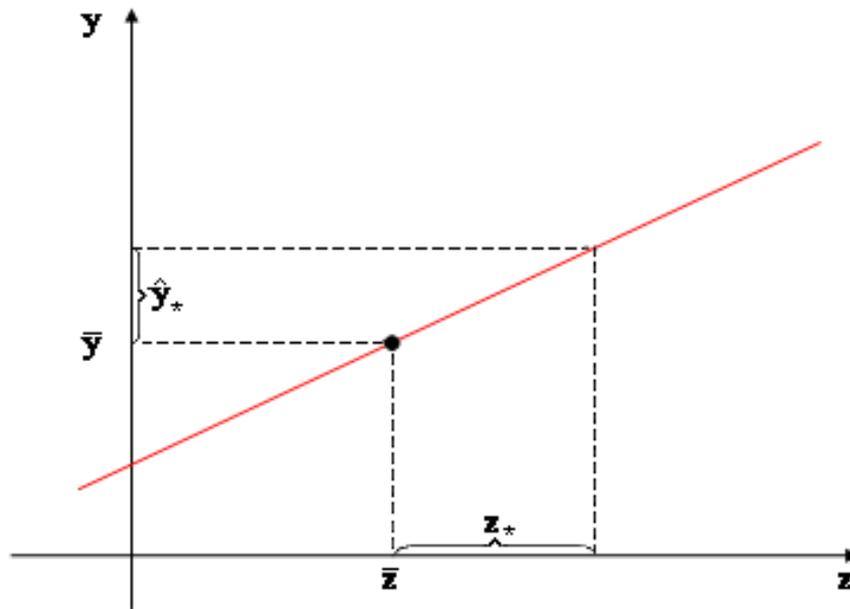
$$\mathbf{H}\mathbf{M} = (\mathbf{I} - \mathbf{M})\mathbf{M} = \mathbf{O} \quad (14.169)$$

(Übung! Außerdem gilt $\mathbf{H}' = \mathbf{H}, \mathbf{M}' = \mathbf{M}$).

Daher sind \mathbf{H} und \mathbf{M} (**Mittelwerts-Matrix**) sogenannte **Projektionsmatrizen** (diese sind *idempotent*, d.h. $\mathbf{H}^2 = \mathbf{H}, \mathbf{M}^2 = \mathbf{M}$).

Projektionsmatrizen

$\mathbf{H}\mathbf{M} = \mathbf{O}$ bedeutet, daß eine zentrierte Matrix und ihr Mittelwert unkorreliert (orthogonal sind), d.h. $(\mathbf{H}\mathbf{Y})'(\mathbf{M}\mathbf{Y}) = \mathbf{Y}'\mathbf{H}\mathbf{M}\mathbf{Y} = \mathbf{O}$. Ausserdem hat eine zentrierte Matrix den Mittelwert $\mathbf{M}(\mathbf{H}\mathbf{Y}) = \mathbf{O}$.

Abbildung 14.17: Zentriertes Regressionsmodell $\hat{Y} = 1\bar{y}' + Z_*\hat{\beta}'$.**prognostizierte
Werte**

Zur Kontrolle der abhängigen Variablen Y durch die Variablen Z benötigt man die prognostizierten Werte

$$\hat{Y} = 1\hat{\alpha}' + Z\hat{B}' \quad (14.170)$$

$$= 1(\bar{y}' - \bar{z}'\hat{B}') + Z\hat{B}' \quad (14.171)$$

$$= 1\bar{y}' + (Z - 1\bar{z}')\hat{B}' \quad (14.172)$$

$$= 1\bar{y}' + Z_*\hat{B}' \quad (14.173)$$

$$= 1\bar{y}' + Z_*(Z_*'Z_*)^{-1}Z_*'Y_* \quad (14.174)$$

(Abb. 14.17).

Mit der Abkürzung

$$\hat{Y} = 1\bar{y}' + PY_* \quad (14.175)$$

$$P = Z_*(Z_*'Z_*)^{-1}Z_*' \quad (14.176)$$

$$P^2 = P \quad (14.177)$$

Hat-Matrix P

kann man die prognostizierten Werte \hat{Y} leicht angeben.

Übung: Zeigen Sie, daß $P^2 = P$ gilt, d.h. P ist idempotent.

Der Mittelwert der Prognose ist $\mathbf{M}\hat{\mathbf{Y}} = \mathbf{M}\mathbf{1}\bar{y}' + \mathbf{M}\mathbf{P}\mathbf{Y}_* = \mathbf{1}\bar{y}'$, da $\mathbf{M}\mathbf{P} = \mathbf{O}$ bzw. $\mathbf{M}\mathbf{Z}_* = \mathbf{O}$ gilt (zentrierte Matrix). Daraus läßt sich ableiten, daß für die zentrierte Prognose

$$\hat{\mathbf{Y}}_* = \mathbf{H}\hat{\mathbf{Y}} = (\mathbf{I} - \mathbf{M})\hat{\mathbf{Y}} = \hat{\mathbf{Y}} - \mathbf{1}\bar{y}' = \mathbf{P}\mathbf{Y}_* \quad (14.178)$$

gilt ($\mathbf{H}\mathbf{P} = \mathbf{P}$, da $\mathbf{H}\mathbf{Z}_* = \mathbf{Z}_*$).

Die **Residuen** $\tilde{\mathbf{Y}} := \mathbf{Y} - \hat{\mathbf{Y}}$ (geschätzte Gleichungsfehler, Prognosefehler) sind dann

**Prognosefehler,
Residuen**

$$\mathbf{Y} - \hat{\mathbf{Y}} = \mathbf{Y} - \mathbf{1}\bar{y}' - \mathbf{P}\mathbf{Y}_* \quad (14.179)$$

$$= \mathbf{Y}_* - \mathbf{P}\mathbf{Y}_* = (\mathbf{I} - \mathbf{P})\mathbf{Y}_* \quad (14.180)$$

$$:= \mathbf{Q}\mathbf{Y}_*. \quad (14.181)$$

Auch für $\mathbf{Q} = \mathbf{I} - \mathbf{P}$ gilt $\mathbf{Q}^2 = \mathbf{Q}$ und $\mathbf{P}\mathbf{Q} = \mathbf{O}$ sowie $\mathbf{P}' = \mathbf{P}$, $\mathbf{Q}' = \mathbf{Q}$. Damit sind auch \mathbf{P} und \mathbf{Q} **Projektionsmatrizen**.

Für weitere Berechnungen hier noch einige Eigenschaften der Projektionsmatrizen in einer Übersicht:

$\mathbf{H} = \mathbf{I} - \mathbf{M}$	$\mathbf{M} = (1/N)\mathbf{1}\mathbf{1}'$	$\mathbf{M}\mathbf{1} = \mathbf{1}$
$\mathbf{H}^2 = \mathbf{H}$	$\mathbf{M}^2 = \mathbf{M}$	$\mathbf{H}\mathbf{M} = \mathbf{M}\mathbf{H} = \mathbf{O}$
$\mathbf{P} = \mathbf{Z}_*(\mathbf{Z}'_*\mathbf{Z}_*)^{-1}\mathbf{Z}'_*$	$\mathbf{Q} = \mathbf{I} - \mathbf{P}$	$\mathbf{P}\mathbf{Q} = \mathbf{O}$
$\mathbf{P}^2 = \mathbf{P}$	$\mathbf{Q}^2 = \mathbf{Q}$	$\mathbf{M}\mathbf{Q} = \mathbf{M}(\mathbf{I} - \mathbf{P}) = \mathbf{M}$
$\mathbf{H}\mathbf{P} = \mathbf{P}$	$\mathbf{M}\mathbf{P} = \mathbf{O}$	

**Projektions-
matrizen**

Die Residuen haben den Mittelwert 0 und stehen senkrecht auf der Prognose $\hat{\mathbf{Y}}_*$:

$$\mathbf{M}\tilde{\mathbf{Y}} = \mathbf{M}(\mathbf{Y} - \hat{\mathbf{Y}}) = \mathbf{1}\mathbf{y}' - \mathbf{1}\mathbf{y}' = \mathbf{O} \quad (14.182)$$

$$\widehat{\text{Cov}}(\hat{\mathbf{Y}}, \tilde{\mathbf{Y}}) := (N-1)^{-1}\hat{\mathbf{Y}}_*'\tilde{\mathbf{Y}} \quad (14.183)$$

$$= (N-1)^{-1}(\mathbf{P}\mathbf{Y}_*)'\mathbf{Q}\mathbf{Y}_* \quad (14.184)$$

$$= (N-1)^{-1}\mathbf{Y}_*'\mathbf{P}\mathbf{Q}\mathbf{Y}_* = \mathbf{O} \quad (14.185)$$

(Abb. 14.12). Hierbei wurde die Definition der **Stichproben-Kovarianz-Matrix** für beliebige Daten-Matrizen $\mathbf{X} : N \times p$ und $\mathbf{Y} : N \times q$

$$\widehat{\text{Cov}}(\mathbf{X}, \mathbf{Y}) = S_{\mathbf{X}\mathbf{Y}} \quad (14.186)$$

$$= (N-1)^{-1}\mathbf{X}'\mathbf{H}\mathbf{Y} \quad (14.187)$$

$$= (N-1)^{-1}\mathbf{X}'_*\mathbf{Y}_* : p \times q \quad (14.188)$$

**Stichproben-
Kovarianz-Matrix**

eingesetzt.

Geometrische Interpretation:

\mathbf{P} projiziert den Vektor \mathbf{Y}_* auf den Unterraum der (zentrierten) prognostierten Werte

$$\hat{\mathbf{Y}} = \mathbf{1}\bar{\mathbf{y}}' + \mathbf{P}\mathbf{Y}_*, \quad (14.189)$$

während $\mathbf{Q} = \mathbf{I} - \mathbf{P}$ die Residualwerte

$$\mathbf{Y} - \hat{\mathbf{Y}} = \tilde{\mathbf{Y}} = \mathbf{Q}\mathbf{Y}_* \quad (14.190)$$

ergibt.

Die beiden Unterräume sind orthogonal (unkorreliert), da

$$(\mathbf{P}\mathbf{Y}_*)'(\mathbf{Q}\mathbf{Y}_*) = \mathbf{Y}_*'\mathbf{P}\mathbf{Q}\mathbf{Y}_* = \mathbf{O}. \quad (14.191)$$

Mit Hilfe der Projektionsmatrizen läßt sich die Streuungszerlegung sehr einfach ausdrücken:

$$SQT := (\mathbf{Y} - \mathbf{1}\bar{y}')'(\mathbf{Y} - \mathbf{1}\bar{y}') \quad (14.192)$$

$$= \mathbf{Y}'_* \mathbf{Y}_* \quad (14.193)$$

$$= \mathbf{Y}'_*(\mathbf{P} + \mathbf{Q})\mathbf{Y}_* \quad (14.194)$$

$$= \mathbf{Y}'_* \mathbf{P} \mathbf{Y}_* + \mathbf{Y}'_* \mathbf{Q} \mathbf{Y}_* \quad (14.195)$$

$$= (\hat{\mathbf{Y}} - \mathbf{1}\bar{y}')'(\hat{\mathbf{Y}} - \mathbf{1}\bar{y}') + (\mathbf{Y} - \hat{\mathbf{Y}})'(\mathbf{Y} - \hat{\mathbf{Y}}) \quad (14.196)$$

$$= SQE + SQR. \quad (14.197)$$

Streuungszerlegung

Hierbei wurde die Eigenschaft der Projektoren $\mathbf{P}^2 = \mathbf{P}$ und $\mathbf{Q}^2 = \mathbf{Q}$ sowie $\mathbf{PQ} = \mathbf{O}$ benutzt. Die gemischten Terme verschwinden (Orthogonalität). Damit ist wieder die Gesamtstreuung aus der Streuung der prognostizierten Werte $\hat{\mathbf{Y}}$ und der Streuung der Residuen $\tilde{\mathbf{Y}} = \mathbf{Y} - \hat{\mathbf{Y}}$ zusammengesetzt. Die Residualstreuung ist die Reststreuung der Variablen \mathbf{Y} nach Korrektur der Kontrollvariablen \mathbf{Z} .

Explizit gilt

$$SQT := \sum_n (\mathbf{y}_n - \mathbf{1}\bar{y}')(\mathbf{y}_n - \mathbf{1}\bar{y}')' \quad (14.198)$$

$$= \sum_n (\hat{\mathbf{y}}_n - \mathbf{1}\bar{y}')(\hat{\mathbf{y}}_n - \mathbf{1}\bar{y}')' + \sum_n (\mathbf{y}_n - \hat{\mathbf{y}}_n)(\mathbf{y}_n - \hat{\mathbf{y}}_n)'$$

$$= SQE + SQR. \quad (14.199)$$

Die residuale Quadratsumme hat die explizite Form

$$SQR = \tilde{\mathbf{Y}}' \tilde{\mathbf{Y}} \quad (14.200)$$

$$= \mathbf{Y}'_* \mathbf{Q} \mathbf{Y}_* \quad (14.201)$$

$$= \mathbf{Y}'_* \mathbf{Y}_* - \mathbf{Y}'_* \mathbf{P} \mathbf{Y}_* \quad (14.202)$$

$$= \mathbf{Y}'_* \mathbf{Y}_* - \mathbf{Y}'_* \mathbf{Z}_* (\mathbf{Z}'_* \mathbf{Z}_*)^{-1} \mathbf{Z}'_* \mathbf{Y}_*. \quad (14.203)$$

Man kann zeigen, daß

residuale
Kovarianzen

$$MQR = \frac{1}{N-q-1} \tilde{\mathbf{Y}}' \tilde{\mathbf{Y}} \quad (14.204)$$

$$= \frac{1}{N-q-1} \mathbf{Y}'_* \mathbf{Q} \mathbf{Y}_* \quad (14.205)$$

$$= \frac{1}{N-q-1} \mathbf{Y}'_* \mathbf{Y}_* - \left(\frac{1}{N-q-1} \mathbf{Y}'_* \mathbf{Z}_* \right) \left(\frac{1}{N-q-1} \mathbf{Z}'_* \mathbf{Z}_* \right)^{-1} \left(\frac{1}{N-q-1} \mathbf{Z}'_* \mathbf{Y}_* \right) \\ := \mathbf{S}_{\mathbf{Y}\mathbf{Y}} - \mathbf{S}_{\mathbf{Y}\mathbf{Z}} \mathbf{S}_{\mathbf{Z}\mathbf{Z}}^{-1} \mathbf{S}_{\mathbf{Z}\mathbf{Y}} \quad (14.206)$$

$$= \hat{\Sigma}. \quad (14.207)$$

ein erwartungstreuer Schätzer der Fehler-Kovarianz Σ ist.⁹

Die Stichproben-Kovarianz der abhängigen Variablen wird also durch die Wirkung der Kontrollvariablen verkleinert. Es handelt sich um die Stichproben-Version der residualen Varianz Glg. 14.132 (lineare Prognose)

$$\text{Var}(\tilde{\mathbf{y}}) = \text{Var}(\mathbf{y}) - \text{Cov}(\mathbf{y}, \mathbf{z}) \text{Cov}(\mathbf{z}, \mathbf{z})^{-1} \text{Cov}(\mathbf{z}, \mathbf{y}) \quad (14.208)$$

Aus der Kovarianzmatrix der Residuen MQR läßt sich durch Standardisierung (vgl. Glg. 14.136) die Matrix der *partiellen Korrelationen* gewinnen.

Multiple Korrelation

Die empirische Korrelation zwischen Prognose $\hat{\mathbf{y}}$ und abhängiger Variable \mathbf{y} wird auch als **multiple Korrelation** bezeichnet. Die Kovarianz hat

⁹Der Nenner $N - q - 1$ ergibt sich aus der Tatsache, daß die Spur der Projektionsmatrix \mathbf{P} durch $\text{tr}[\mathbf{P}] = \text{tr}[\mathbf{Z}_* (\mathbf{Z}'_* \mathbf{Z}_*)^{-1} \mathbf{Z}'_*] = \text{tr}[\mathbf{I}_q] = q$ gegeben ist. Hierbei wurde die Rechenregel $\text{tr}[\mathbf{A}\mathbf{B}] = \text{tr}[\mathbf{B}\mathbf{A}]$ für Spuren benutzt. Außerdem ist $N - 1$ die Summe der Diagonale der Zentrierungsmatrix \mathbf{H} .

Dann findet man für die ij -te Komponente $E[SQR_{ij}] = E[\mathbf{Y}'_{*i} \mathbf{Q} \mathbf{Y}_{*j}] = E[\text{tr}[(\mathbf{Q} \mathbf{Y}_{*j})(\mathbf{Q} \mathbf{Y}_{*i})']] = \text{tr}[\mathbf{Q} \mathbf{H} \mathbf{I}_N \sigma_{ji} \mathbf{H}' \mathbf{Q}'] = \sigma_{ij} \text{tr}[\mathbf{Q} \mathbf{H}] = \sigma_{ij} (N - 1 - q)$, da die einzelnen Komponenten der Spalte $\mathbf{Y}_j = [y_{1j}, \dots, y_{Nj}]'$ voneinander unabhängig sind. Hierbei wurde die Zentrierungsmatrix $\mathbf{Y}_* = \mathbf{H} \mathbf{Y}$ eingesetzt sowie ausgenutzt, daß $E[\tilde{\mathbf{Y}}] = E[\mathbf{Q} \mathbf{Y}_*] = \mathbf{Q} \mathbf{Z}_* \mathbf{B}' = \mathbf{O}$ ist.

Dies folgt aus der Darstellung $\mathbf{Y}_* = \mathbf{Z}_* \mathbf{B}' + \mathbf{H} \boldsymbol{\epsilon}$ des Regressionsmodells sowie $\mathbf{Q} \mathbf{Z}_* = \mathbf{Z}_* - \mathbf{P} \mathbf{Z}_* = \mathbf{Z}_* - \mathbf{Z}_* = \mathbf{O}$. Daher gilt $E[(\mathbf{Q} \mathbf{Y}_{*j})(\mathbf{Q} \mathbf{Y}_{*i})'] = \text{Cov}(\mathbf{Q} \mathbf{Y}_{*j}, \mathbf{Q} \mathbf{Y}_{*i}) = \mathbf{Q} \mathbf{H} \text{Cov}(\mathbf{Y}_j, \mathbf{Y}_i) \mathbf{H} \mathbf{Q} = \mathbf{Q} \mathbf{H} \mathbf{I}_N \sigma_{ij} \mathbf{H} \mathbf{Q}$. Durch Vertauschung unter der Spur und mit $\mathbf{H}^2 = \mathbf{H}$, $\mathbf{Q}^2 = \mathbf{Q}$, $\mathbf{Q} = \mathbf{I}_N - \mathbf{P}$, $\mathbf{P} \mathbf{H} = \mathbf{P}$ ergibt sich der gesuchte Faktor $\text{tr}[\mathbf{Q} \mathbf{H}] = \text{tr}[\mathbf{H} - \mathbf{P} \mathbf{H}] = N - 1 - \text{tr}[\mathbf{P}] = N - 1 - q$.

die Form

$$\hat{\mathbf{Y}}' \mathbf{H} \mathbf{Y} = \hat{\mathbf{Y}}' \mathbf{H} (\hat{\mathbf{Y}} + \tilde{\mathbf{Y}}) \quad (14.209)$$

$$= \hat{\mathbf{Y}}' \mathbf{H} \hat{\mathbf{Y}} + \hat{\mathbf{Y}}' \mathbf{H} \tilde{\mathbf{Y}} \quad (14.210)$$

$$= \hat{\mathbf{Y}}' \mathbf{H} \hat{\mathbf{Y}} = SQE, \quad (14.211)$$

da Prognose und Residualwerte unkorreliert sind.

Damit gilt explizit für die Komponenten $i = 1, \dots, p$

$$R^2(\hat{\mathbf{y}}_i, \mathbf{y}_i) = \frac{(SQE_{ii})^2}{SQT_{ii} SQE_{ii}} \quad (14.212)$$

$$= \frac{SQE_{ii}}{SQT_{ii}} \quad (14.213)$$

**multiple
Korrelation**

Dies ist aber der Quotient der Diagonalelemente aus erklärter und totaler Streuung (Determinationskoeffizient). Somit findet man:

Determinationskoeffizient = quadrierte multiple Korrelation.

14.9 Reliabilität

14.9.1 Meßmodell

In vielen praktisch relevanten Fällen sind die eigentlich interessierenden Variablen von Meßfehlern überlagert, etwa durch Störungen in einer Test-Situation (Ablenkung des Probanden). Dann sind Korrelationen zwischen solchen fehlerbehafteten Größen (Fehler in den Variablen) kleiner als erwartet, da sich die Streuungen durch das Meßrauschen vergrößern (Abschwächung).

Indikatoren

Die gemessenen Werte X werden als Indikatoren (manifeste Variablen) bezeichnet, während die wahren Variablen T true scores, Konstrukte oder latente Variablen heißen.

true score

Führt man noch einen Meßfehler ϵ ein, so ergibt sich das Fehlermodell der klassischen Testtheorie

**klassische
Testtheorie**

$$X = T + \epsilon \quad (14.214)$$

Man nimmt hierbei an, daß $\text{Cov}(T, \epsilon) = 0$, d.h. die Meßfehler hängen nicht von der Ausprägung des wahren Werts ab.

Geht man davon aus, daß Personen mit hoher Intelligenz leichter gestört werden, so ist diese Annahme verletzt.

14.9.2 Reliabilität

Die relative Stärke des wahren Werts im Vergleich zum Gesamtwert wird als Reliabilität

Reliabilität

$$rel = \frac{\text{Var}(T)}{\text{Var}(X)} = \frac{\text{Var}(T)}{\text{Var}(T) + \text{Var}(\epsilon)} \leq 1 \quad (14.215)$$

bezeichnet. Es handelt sich also um das Verhältnis $rel = (\text{wahre Streuung} / \text{Gesamt-Streuung})$.

Hierbei wurde ausgenutzt, daß wahrer Wert und Meßfehler unkorreliert sind. Bei positiver Korrelation ist $\text{Var}(X) > \text{Var}(T) + \text{Var}(\epsilon)$ und die Relibiliatät wird überschätzt.

14.9.3 Parallel-Test, Re-Test

Da der true score T nicht gemessen werden kann, ist zunächst unklar, wie die Reliabilität bestimmt werden soll. Hierzu sind zusätzliche Annahmen erforderlich. Nimmt man an, daß ein Parallel-Test oder eine unabhängige Testwiederholung ohne Änderung des true score möglich sind, so ergibt sich das Meßmodell

$$X_1 = T + \epsilon_1 \quad (14.216)$$

$$X_2 = T + \epsilon_2 \quad (14.217)$$

$$\text{Cov}(X_1, X_2) = \text{Cov}(T, T) \quad (14.218)$$

da die Fehler unkorreliert sind (Annahme).

Man erhält so die Parallel-Test (Re-Test)-Reliabilität

$$rel = \frac{\text{Var}(T)}{\text{Var}(X)} = \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}} \quad (14.219)$$

$$= \text{Corr}(X_1, X_2) = \rho_{12} \quad (14.220)$$

**Parallel-Test
(Re-Test)-
Reliabilität**

Es ist allerdings auch möglich, daß sich der true score im Zeitablauf ändert, d.h. $T_1 \neq T_2$. Dann ist $\text{Cov}(X_1, X_2) = \text{Cov}(T_1, T_2) \neq \text{Cov}(T, T)$. In diesem Fall berechnet man mit der Retest-Methode die Autokovarianz des true scores (Stabilität) und nicht die Reliabilität (vgl. hierzu ausführlich Möbus und Nagl, 1983).

14.9.4 Test-Halbierung

Falls kein Parallel-Test existiert, behilft man sich mit einer Test-Halbierung und dem Meßmodell

$$X = X_1 + X_2 \quad (14.221)$$

$$X_1 = T/2 + \epsilon_1 \quad (14.222)$$

$$X_2 = T/2 + \epsilon_2 \quad (14.223)$$

$$\text{Cov}(X_1, X_2) = \sigma_{12} = \text{Cov}(T, T)/4 \geq 0. \quad (14.224)$$

Die Reliabilität ergibt sich dann wieder aus der Definition $rel = (\text{wahre Streuung} / \text{Gesamt-Streuung})$

$$rel = \frac{\text{Var}(T)}{\text{Var}(X)} = \frac{4\text{Cov}(X_1, X_2)}{\text{Var}(X_1 + X_2)} \quad (14.225)$$

$$= \frac{4\sigma_{12}}{\sigma_1^2 + \sigma_2^2 + 2\sigma_{12}} \quad (14.226)$$

$$= \frac{4\rho_{12}}{\sigma_1/\sigma_2 + \sigma_2/\sigma_1 + 2\rho_{12}} \quad (14.227)$$

$$= \frac{2\rho_{12}}{1 + \rho_{12}} \quad (14.228)$$

**Split-Half-
Reliabilität,
Spearman-Brown-
Formel**

$$rel = \frac{2\rho_{12}}{1 + \rho_{12}} \quad (14.229)$$

bei gleichen Standardabweichungen σ_i der Testhälften. Dies ist die sog. Spearman-Brown-Formel, welche die naive Näherung $rel \approx \rho_{12}$ korrigiert.

14.9.5 Allgemeine Spearman-Brown-Formel (k Test-Teile)

Teilt man den Test in k Teile, so ergibt sich

$$X = \sum X_i \quad (14.230)$$

$$X_i = T/k + \epsilon_i \quad (14.231)$$

$$\text{Cov}(X_i, X_j) := \sigma_{ij} = \text{Cov}(T, T)/(k^2) \geq 0. \quad (14.232)$$

Daraus findet man die Reliabilität

$$rel = \frac{\text{Var}(T)}{\text{Var}(X)} \quad (14.233)$$

$$= \frac{k^2 \sigma_{ij}}{\sum_i \sigma_i^2 + \sum_{i \neq j} \sigma_{ij}} \quad (14.234)$$

$$= \frac{k\rho_{ij}}{1 + (k-1)\bar{\rho}} \quad (14.235)$$

wenn man die Varianz-Zerlegung und die durchschnittliche Subtest-Korrelation

$$\text{Var}(X) = \sum_i \sigma_i^2 + \sum_{i \neq j} \sigma_{ij} \quad (14.236)$$

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j} \quad (14.237)$$

$$\bar{\rho} = \frac{1}{k(k-1)} \sum_{i \neq j} \rho_{ij} \quad (14.238)$$

einsetzt sowie gleiche Standardabweichungen $\sigma_i = \sigma$ oder Standardisierung der Test-Teile annimmt.

14.9.6 Cronbach's Alpha

Aufgrund der Annahmen sind (theoretisch) alle $\rho_{ij} \geq 0$ gleich. Ersetzt man ρ_{ij} durch die durchschnittliche Subtest-Korrelation, so ergibt sich

$$rel = \frac{k\bar{\rho}}{1 + (k-1)\bar{\rho}} \quad (14.239)$$

Cronbach's Alpha

(gleiche Standardabweichungen $\sigma_i = \sigma$ der Test-Teile oder standardisierte Items).

Dies ist jedoch eine Variante von Cronbach's Alpha (interne Konsistenz), d.h. die Spearman-Brown-Formel für die durchschnittliche Subtest-Korrelation

$$\bar{\rho} = \frac{1}{k(k-1)} \sum_{i \neq j} \rho_{ij} \quad (14.240)$$

Alternativ kann man schreiben

$$rel = \frac{\text{Var}(T)}{\text{Var}(X)} = \frac{k}{k-1} \frac{\text{Var}(X) - \sum \sigma_i^2}{\text{Var}(X)} \quad (14.241)$$

da

$$\text{Var}(X) = \sum_i \sigma_i^2 + \sum_{i \neq j} \sigma_{ij} \quad (14.242)$$

$$= \sum_i \sigma_i^2 + \sum_{i \neq j} \text{Cov}(T, T)/(k^2) \quad (14.243)$$

$$= \sum_i \sigma_i^2 + \frac{k-1}{k} \text{Cov}(T, T). \quad (14.244)$$

Bei dieser Formel wird jedoch nicht unterstellt, daß alle Test-Teile gleiche Standardabweichungen σ_i aufweisen.

Beispiel 14.7 (Filialgestaltung von Bankfilialen)

Wir betrachten das Konstrukt Kundenzufriedenheit, das sich aus 4 Items mit Wertebereich 1–5 zusammensetzt (Datensatz `main2.xls.sav`; vgl. Kap. 15):

13. *Wie zufrieden sind Sie insgesamt mit Ihren Besuchen in dieser Geschäftsstelle?*

Sehr zufrieden 1 2 3 4 5 Gar nicht zufrieden

14. *Wie zufrieden sind Sie mit den Wartezeiten?*

Sehr zufrieden 1 2 3 4 5 Gar nicht zufrieden

16. *Wie zufrieden sind Sie mit der Kundenbetreuung?*

Sehr zufrieden 1 2 3 4 5 Gar nicht zufrieden

17. *Wie zufrieden sind Sie mit den Öffnungszeiten?*

Sehr zufrieden 1 2 3 4 5 Gar nicht zufrieden

Zunächst ist festzustellen, daß alle Items positiv korrelieren, also allen Fragen eine gemeinsame Dimension unterliegt (Abb. 14.18, 14.19–14.20). Die Werte für Cronbachs α zeigen, daß ca. 70% der wahren Streuung in den Messungen enthalten ist. Aus der durchschnittlichen Korrelation

$$\bar{r} = \frac{1}{k(k-1)} \sum_{i \neq j} r_{ij} = .405 \quad (14.245)$$

ergibt sich das geschätzte α von

$$\hat{\alpha} = \frac{k\bar{r}}{1 + (k-1)\bar{r}} = 0.7315. \quad (14.246)$$

($k = 4$ items; Glg. 14.239). Bei gleicher durchschnittlicher Korrelation kann man durch Erhöhung der Itemanzahl die Reliabilität erhöhen (Abb. 14.21).

Aus Glg. 14.241 ergibt sich durch Einsetzen der Gesamt-Varianz = 6.661 (Summe der Inter-Item-Kovarianzmatrix) und der Diagonalsumme der Inter-Item-Kovarianzmatrix = 3.06 ein Wert von

$$\hat{\alpha} = \frac{k}{k-1} \frac{\widehat{\text{Var}}(X) - \sum s_i^2}{\widehat{\text{Var}}(X)} = 0.720905. \quad (14.247)$$

Da die item-Standardabweichungen ähnlich sind, geben beide Formeln ein fast identisches Ergebnis.



Übung 14.4 (Filialgestaltung von Bankfilialen)

Überprüfen Sie die Werte von Cronbachs Alpha durch Einsetzen der Werte in der Inter-Item-Korrelations(Kovarianz)matrix.



14.9.7 Abschwächungs-Korrektur

Aufgrund der Meßfehler ist die Korrelation von Indikatoren geringer als die der zugrundeliegenden true scores. Daher wird in vielen Anwendungen ein Meßmodell postuliert (Fehler in den Variablen), das diese Problematik explizit modelliert (etwa in Strukturgleichungsmodellen wie LISREL). Hier soll nur eine einfache Version der Abschwächungs-Korrektur diskutiert werden. Wir betrachten zwei Konstrukte mit zugehörigem Meßmodell

$$X_1 = T_1 + \epsilon_1 \quad (14.248)$$

$$X_2 = T_2 + \epsilon_2. \quad (14.249)$$

Dann ist die Korrelation der wahren Werte

$$\text{Corr}(T_1, T_2) = \text{Corr}(X_1, X_2) / \sqrt{rel_1 rel_2} \quad (14.250)$$

$$> \text{Corr}(X_1, X_2) \quad (14.251)$$

Zusammenfassung der Fallverarbeitung

		N	%
Fälle	Gültig	354	93,4
	Ausgeschlossen ^a	25	6,6
	Gesamt	379	100,0

a. Listenweise Löschung auf der Grundlage aller Variablen in der Prozedur.

Reliabilitätsstatistiken

Cronbachs Alpha	Cronbachs Alpha für standardisierte Items	Anzahl der Items
,721	,732	4

Itemstatistiken

	Mittelwert	Standardabweichung	N
ZGESAMT	1,8220	,79622	354
ZWARTE	2,1836	,94174	354
ZBETREU	1,8051	,82033	354
ZOEFFNUN	1,9209	,93072	354

Inter-Item-Korrelationsmatrix

	ZGESAMT	ZWARTE	ZBETREU	ZOEFFNUN
ZGESAMT	1,000	,493	,628	,348
ZWARTE	,493	1,000	,402	,220
ZBETREU	,628	,402	1,000	,340
ZOEFFNUN	,348	,220	,340	1,000

Inter-Item-Kovarianzmatrix

	ZGESAMT	ZWARTE	ZBETREU	ZOEFFNUN
ZGESAMT	,634	,370	,410	,258
ZWARTE	,370	,887	,311	,193
ZBETREU	,410	,311	,673	,259
ZOEFFNUN	,258	,193	,259	,866

Abbildung 14.19: Cronbachs Alpha: Statistiken für das Konstrukt *Zufriedenheit mit dem Filialbesuch*.

Auswertung der Itemstatistiken

	Mittelwert	Minimum	Maximum	Bereich	Maximum / Minimum	Varianz	Anzahl der Items
Item-Mittelwerte	1,933	1,805	2,184	,379	1,210	,031	4
Item-Varianzen	,765	,634	,887	,253	1,399	,017	4
Inter-Item-Kovarianzen	,300	,193	,410	,217	2,124	,006	4
Inter-Item-Korrelationen	,405	,220	,628	,407	2,850	,018	4

Item-Skala-Statistiken

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Quadrierte multiple Korrelation	Cronbachs Alpha, wenn Item weggelassen
ZGESAMT	5,9096	3,952	,656	,477	,579
ZWARTE	5,5480	4,027	,462	,259	,691
ZBETREU	5,9266	4,029	,595	,420	,611
ZOEFFNUN	5,8107	4,375	,365	,147	,748

Skala-Statistiken

Mittelwert	Varianz	Standardabweichung	Anzahl der Items
7,7316	6,661	2,58099	4

Abbildung 14.20: Cronbachs Alpha: Statistiken für das Konstrukt *Zufriedenheit mit dem Filialbesuch*.

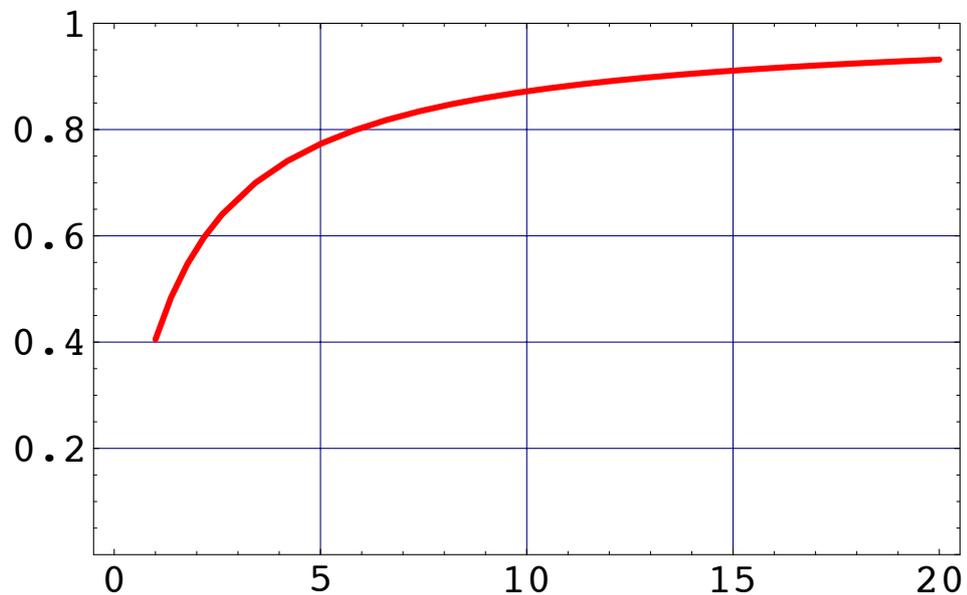


Abbildung 14.21: Cronbachs Alpha als Funktion der Skalenlänge.

wobei die Formel

$$\text{Var}(T_i) = rel_i \text{Var}(X_i) < \text{Var}(X_i) \quad (14.252)$$

eingesetzt wurde. Daher korrelieren die Konstrukte stärker als die Messungen, was als Abschwächungs-Korrektur bezeichnet wird.

**Abschwächungs-
Korrektur**

Übung 14.5 (Filialgestaltung von Bankfilialen)

Berechnen Sie unter Verwendung des Werts von α aus dem letzten Beispiel die Korrelation der true scores $\text{Corr}(T_1, T_2)$, wenn $\text{Corr}(X_1, X_2) = 0.5$ ist.



14.10 Skalen

14.10.1 Likert-Skala

Eine sehr häufig benutzte Methode zur Konstruktion von Skalen ist die von Likert (1932) entwickelte **Methode der summierten Ratings** (Einschätzungen). Hierbei werden Behauptungen aufgestellt und die Versuchsperson gibt an, inwieweit diese zutreffen (vgl. auch Bsp. 14.7):

Likert-Skala

Behauptung (item X_i):

Hagen ist weltoffen

- A: trifft eindeutig zu
- B: trifft zu
- C: trifft weder zu noch nicht zu
- D: trifft nicht zu
- E: trifft eindeutig nicht zu

Unter der Annahme, daß alle Kategorien der Ratingskala äquidistant sind, wird der gewählten Kategorie ein numerischer Wert zwischen $A \rightarrow 1, \dots, E \rightarrow 5$ zugeordnet. Mit diesen numerischen Werten (Codes), die eigentlich nur ordinales Skalenniveau aufweisen, wird üblicherweise wie mit metrischen Variablen gerechnet (etwa in der Itemanalyse).

Eine ausführliche Darstellung von Skalierungsmethoden (Thurstone, Rasch etc.) ist in Bortz/Döring (2006, Kap. 4) enthalten.

14.10.2 Item-Analyse

Item-Analyse

Um die Items X_i einer Skala (etwa Kundenzufriedenheit) in ihrer Qualität zu beurteilen, wird eine Aufgaben(Item)-Analyse durchgeführt. Darunter wird die Analyse der Rohwert-Verteilung, Itemschwierigkeit, Trennschärfe, Homogenität und Dimensionalität verstanden:

- Items einer Skala: X_1, \dots, X_k
- Gesamtwert: $X = \sum_{i=1}^k X_i$
- Item-Schwierigkeit: $0 \leq p_i = \frac{\bar{X}_i - X_{i,min}}{X_{i,max} - X_{i,min}} \leq 1$
- Trennschärfe: $\text{Corr}(X_i, X)$
- Homogenität: $\bar{\rho} = \frac{1}{k(k-1)} \sum_{i \neq j} \rho_{ij}$
- Dimensionalität: Analyse auf latente Faktoren

14.10.2.1 Item-Schwierigkeit

Item-Schwierigkeit

Die **Item-Schwierigkeit** ist eine lineare Transformation $p(X)$ eines Items auf das Intervall $[0, 1] = [p(X_{min}), p(X_{max})]$. Es gilt also $p_i = p(\bar{X}_i)$. Die Terminologie läßt sich durch dichotome Leistungstest-Items motivieren, wo ein Item gelöst ($X_i = 1$) oder nicht gelöst ($X_i = 0$) werden kann. Dies ergibt $\bar{X}_i = N^{-1} \sum_n X_{ni}$, also den Prozentsatz der Personen, die Item i gelöst haben.

Beispielsweise ergibt sich bei einer

- 5-er Skala $(1, \dots, 5)$: $\bar{X}_i = 3 \Rightarrow p_i = \frac{3-1}{5-1} = 2/4 = 50\%$.

14.10.2.2 Trennschärfe

Trennschärfe

Die **Trennschärfe** = $\text{Corr}(X_i, X)$ gibt an, wie gut ein Item mit dem Gesamtwert der Skala korreliert. Meistens wird das Item X_i im Gesamtwert weggelassen (warum?). Trennscharfe Items haben hohe Werte bei hohem Gesamtwert. Items mit geringer Trennschärfe sind schlechte Indikatoren des Konstrukts (Gesamtwert) und werden aus einem eindimensionalen Test entfernt.

Items mit hoher (geringer) Schwierigkeit haben eine geringe Trennschärfe.

14.10.2.3 Homogenität

Unter der **Homogenität** $\bar{\rho} = \frac{1}{k(k-1)} \sum_{i \neq j} \rho_{ij}$ versteht man die durchschnittliche Korrelation der Items. Wie schon diskutiert, wird dieser Durchschnittswert in die Spearman-Brown-Formel eingesetzt (Cronbachs Alpha).

Homogenität

Man kann auch zeilenweise oder spaltenweise mitteln und erhält so die **Itemspezifische Homogenität**

$$\bar{\rho}_{i+} = \frac{1}{k-1} \sum_{j, i \neq j} \rho_{ij}. \quad (14.253)$$

Items mit geringer Homogenität sollten aus dem Test entfernt werden.

14.10.2.4 Dimensionalität

Schließlich kann mit einer Faktorenanalyse oder Clusteranalyse (der Variablen) untersucht werden, ob die Item-Batterie ein- oder mehrdimensional ist. Etwa kann die Korrelationsmatrix der Items solange umgeordnet werden, bis ein oder mehrere Blöcke mit homogener Korrelation entstehen (diagonales Ordnen).

Dimensionalität

Beispiel 14.8 (Filialgestaltung von Bankfilialen)

Wir betrachten wieder das Konstrukt Kundenzufriedenheit, das sich aus 4 Items mit Wertebereich 1–5 zusammensetzt (Datensatz `main2.xls.sav`). Aus den Itemstatistiken (Abb. 14.19) sieht man, daß geringe Itemschwierigkeiten vorliegen ($p_i = (\{1.82, 2.18, 1.81, 1.92\} - 1)/(5 - 1) = \{0.205, 0.295, 0.2025, 0.23\}$). Die Items sind 'leicht', d.h. die Personen sind eher zufrieden.

Die Tabelle Item-Skala-Statistiken (Abb. 14.20) zeigt die korrigierte Item-Skala-Korrelation (Trennschärfe). Item 4 fällt etwas aus dem Rahmen ($r_{X_4X} = .365$). Auch ist die quadrierte multiple Korrelation (= 0.147) von (X_1, \dots, X_3) auf X_4 ziemlich klein (vgl. Abschnitt 14.8.4).

Die Homogenität \bar{r} beträgt 0.405, während man für die itemspezifische Homogenität die Werte $\bar{r}_{i+} = \{0.489752, 0.372043, 0.456485, 0.302553\}$ erhält.

Eine Clusteranalyse der Variablen mit der Item-Korrelationsmatrix als Ähnlichkeitsmaß bestätigt die Außenseiter-Rolle von Item 4 (Abb. 14.22). Eine Faktorenanalyse (Hauptkomponentenanalyse) der Item-Korrelationsmatrix extrahiert einen Faktor (Kriterium: Eigenwerte

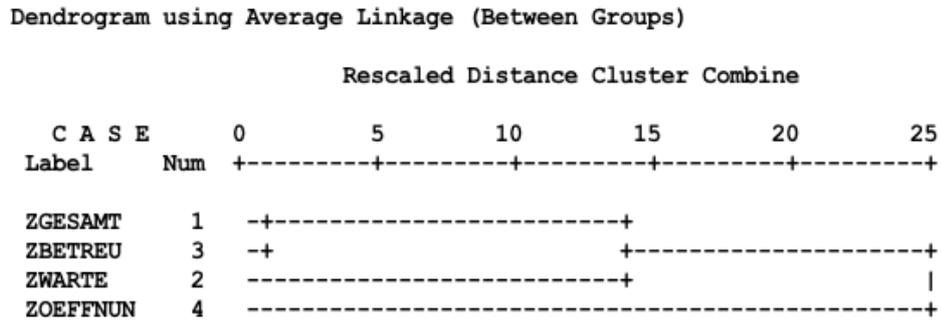


Abbildung 14.22: Clusteranalyse der Items der Skala Zufriedenheit (Item-Korrelationsmatrix als Ähnlichkeitsmaß).

> 1). Schreibt man die geschätzten Faktorwerte $\hat{\xi}$ in den Datensatz (Variable ZUFRIED), so ergibt sich eine sehr hohe Korrelation zwischen Summenscore $X = \text{ZufriedSum}$ und Faktor ZUFRIED (Abb. 14.24).



14.11 Faktoren-Analyse

14.11.1 Modell

Im letzten Abschnitt wurde erwähnt, daß die Faktorenanalyse eine wichtige Methode darstellt, um die Dimensionalität einer Item-Batterie zu erforschen. In der Tat ist das Meßmodell der klassischen Testtheorie eine einfache Version der Faktorenanalyse. Schreibt man für die Items $i = 1, \dots, p$ (Gesamtwert $X = \sum X_i = T + \epsilon$)

$$X_i = T/p + \epsilon_i, \tag{14.254}$$

so ist dies ein Spezialfall des Ein-Faktor-Modells

$$X_i = \lambda_i \xi + \epsilon_i \tag{14.255}$$

mit der Restriktion $\lambda_i = 1/p$ und $T = \xi$. In Matrix-Form gilt für $p = 4$ Items:

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} \xi + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \end{bmatrix} \tag{14.256}$$

Erklärte Gesamtvarianz

Komponente	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion		
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	2,246	56,161	56,161	2,246	56,161	56,161
2	,798	19,949	76,110			
3	,596	14,907	91,017			
4	,359	8,983	100,000			

Extraktionsmethode: Hauptkomponentenanalyse.

Screeplot

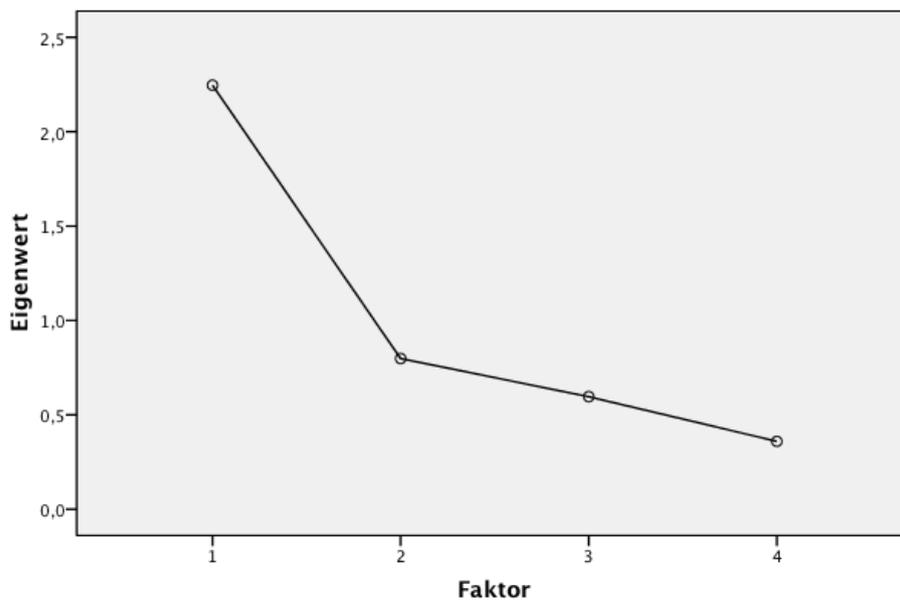
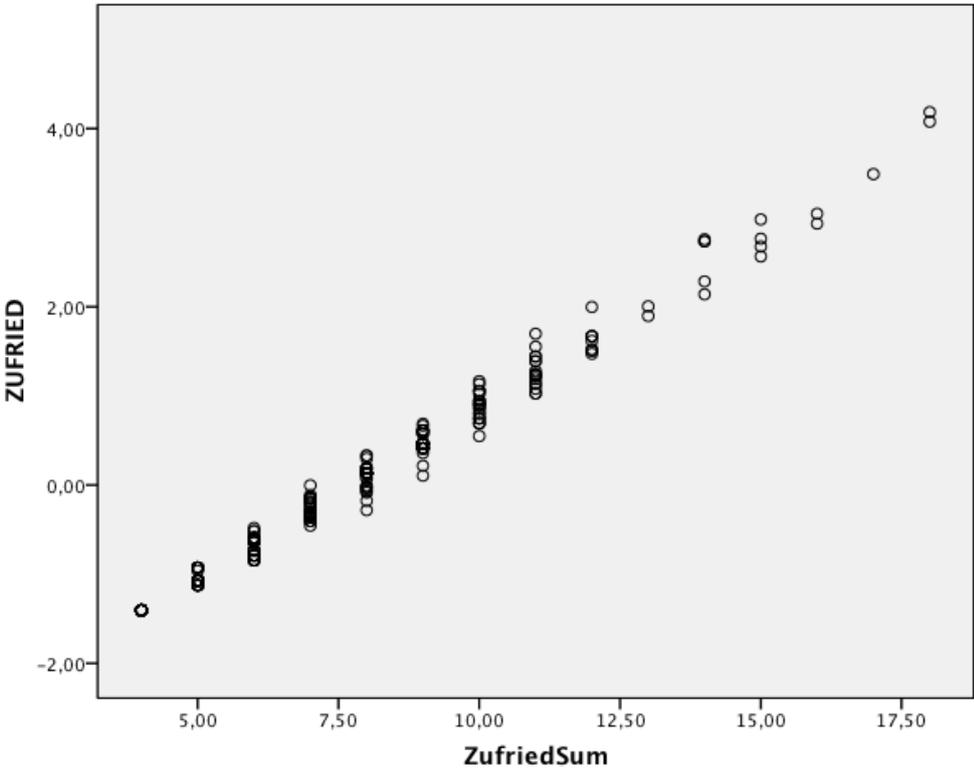


Abbildung 14.23: Faktorenanalyse der Items (Analyse der Item-Korrelationsmatrix). Der erste Faktor erklärt 56% der Gesamt-Varianz (=4).



Korrelationen

		ZufriedSum	ZUFRIED
ZufriedSum	Korrelation nach Pearson	1	,994**
	Signifikanz (2-seitig)		,000
	N	354	354
ZUFRIED	Korrelation nach Pearson	,994**	1
	Signifikanz (2-seitig)	,000	
	N	354	354

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Abbildung 14.24: Streudiagramm und Korrelation zwischen Faktor und Summenscore.

oder in allgemeiner Form

$$\mathbf{x} = \mathbf{\Lambda}\boldsymbol{\xi} + \boldsymbol{\epsilon}. \quad (14.257)$$

**Faktorenanalyse-
Modell**

Die Gewichts-Matrix $\mathbf{\Lambda} : p \times q$ wird als **Faktor-Ladungs-Matrix** bezeichnet und $\boldsymbol{\xi} : q \times 1$ (hier $q = 1$) ist ein Vektor von **Faktoren** (latenten Variablen, true scores).

**Faktor-Ladungs-
Matrix
Faktoren**

Diese können **orthogonal** ($\text{Cov}(\boldsymbol{\xi}, \boldsymbol{\xi}) = \mathbf{I}$) oder schiefwinklig sein ($\text{Cov}(\boldsymbol{\xi}, \boldsymbol{\xi}) = \boldsymbol{\Phi} \neq \mathbf{I}$).

Die üblichen Annahmen und Bezeichnungen sind:

- $\text{Cov}(\boldsymbol{\xi}, \boldsymbol{\epsilon}) = \mathbf{O} : q \times p$ (Faktoren und Fehler sind orthogonal)
- $E[\boldsymbol{\xi}] = \mathbf{0}; E[\boldsymbol{\epsilon}] = E[\mathbf{x}] = \mathbf{0}$ (zentrierte Variablen)
- $\text{Cov}(\mathbf{x}) = \boldsymbol{\Sigma} : p \times p$
- $\text{Cov}(\boldsymbol{\xi}) = \boldsymbol{\Phi} : q \times q$
($\boldsymbol{\Phi} = \mathbf{I}$: Faktoren sind rechtwinklig und standardisiert)
- $\text{Cov}(\boldsymbol{\epsilon}) = \mathbf{V} : p \times p$ (Fehler-Kovarianzmatrix).

Es handelt sich beim Faktorenmodell (14.257) trotz der identischen Form **nicht um ein Regressionsmodell**, da $\boldsymbol{\xi}$ nicht beobachtbar ist. Die beobachteten Variablen werden durch drei nichtbeobachtete Größen $\mathbf{\Lambda}, \boldsymbol{\xi}, \boldsymbol{\epsilon}$ erklärt, was durchaus ambitioniert ist und auch diverse Mehrdeutigkeiten zur Folge hat.

Die Faktoren (Zufallsvariablen) lassen sich schätzen, wenn man die optimale Prognose (Glg. 14.111)

$$\begin{aligned} \hat{\boldsymbol{\xi}} = E[\boldsymbol{\xi}|\mathbf{x}] &= E[\boldsymbol{\xi}] + \text{Cov}(\boldsymbol{\xi}, \mathbf{x})\text{Cov}(\mathbf{x}, \mathbf{x})^{-1}(\mathbf{x} - E[\mathbf{x}]) \quad (14.258) \\ &= \text{Cov}(\boldsymbol{\xi}, \mathbf{x})\text{Cov}(\mathbf{x}, \mathbf{x})^{-1}\mathbf{x} \quad (14.259) \end{aligned}$$

**Faktoren-
Schätzung
(Thompson)**

zugrundelegt (**Regressions-Schätzer, Thompson 1951**).

Alternativ kann man den **GLS-Schätzer (Bartlett)**¹⁰

¹⁰GLS = generalized least squares

**Faktoren-
Schätzung
(Bartlett)**

$$\hat{\xi} = (\Lambda'V^{-1}\Lambda)^{-1}\Lambda'V^{-1}\mathbf{x}. \quad (14.260)$$

verwenden. Dieser ergibt sich aus dem Grundmodell

$$\mathbf{x} = \Lambda\xi + \epsilon. \quad (14.261)$$

durch multiplizieren mit $V^{-1/2}$ und berechnen des KQ-Schätzers (vgl. Mardia et al., Kap. 9.7).

Setzt man noch

$$\text{Cov}(\xi, \mathbf{x}) = \text{Cov}(\xi, \Lambda\xi) = \text{Cov}(\xi, \xi)\Lambda' = \Phi\Lambda' \quad (14.262)$$

$$\text{Cov}(\mathbf{x}, \mathbf{x}) = \Lambda\text{Cov}(\xi, \xi)\Lambda' + \text{Cov}(\epsilon, \epsilon) \quad (14.263)$$

$$= \Lambda\Phi\Lambda' + V = \Sigma \quad (14.264)$$

in (14.258) ein, so können die Faktorwerte als lineare Regression der beobachtbaren Werte \mathbf{x} ausgedrückt werden.¹¹

Die Varianzzerlegung (14.264) wird etwas hochtrabend als **Fundamentaltheorem der Faktorenanalyse** bezeichnet.

Die Modellparameter (Φ, Λ, V) müssen zuvor geschätzt werden.

Im obigen Ein-Faktor-Fall ergibt sich der Regressions-Schätzer $(\Phi = 1)$ als

$$E[\xi|\mathbf{x}] = [\lambda_1, \lambda_2, \lambda_3, \lambda_4]\Sigma^{-1}\mathbf{x}. \quad (14.266)$$

Man erhält also ein **gewichtetes Mittel**

$$\hat{\xi} = E[\xi|\mathbf{x}] = \sum g_i X_i \quad (14.267)$$

¹¹Mit Hilfe der Formel

$$\Phi\Lambda'(\Lambda\Phi\Lambda' + V)^{-1} = (\Lambda'V^{-1}\Lambda + \Phi^{-1})^{-1}\Lambda'V^{-1} \quad (14.265)$$

ergibt sich eine analoge Form zum Bartlett-Schätzer. Für eine nichtinformativ Prior-Verteilung der Faktoren, d.h. $\xi \sim N(\mathbf{0}, \Phi \rightarrow \infty)$ geht der Thompson-Schätzer in den Bartlett-Schätzer über.

**Fundamental-
theorem**

Komponentenmatrix^a

	Komponente...
	1
ZGESAMT	,856
ZWARTE	,707
ZBETREU	,819
ZOEFFNUN	,587

Extraktionsmethode:
Hauptkomponenten
analyse.

a. 1 Komponenten
extrahiert

**Koeffizientenmatrix der
Komponentenwerte**

	Komponente...
	1
ZGESAMT	,381
ZWARTE	,315
ZBETREU	,364
ZOEFFNUN	,261

Extraktionsmethode:
Hauptkomponenten
analyse.

Abbildung 14.25: Hauptkomponentenanalyse: Faktorladungen und Gewichte zur Berechnung der Faktorwerte (ein Faktor).

der items und keineswegs einen Summenwert $\sum_i X_i$. Die Faktorladungen und Gewichte werden vom Programm ausgegeben und zur Berechnung der geschätzten Faktoren benutzt (Abb. 14.25). Eine 2-Faktorlösung (orthogonale Faktoren) ist in Abb. 14.26 gezeigt. Es bestätigt sich das Bild aus der Item-Analyse, daß Item 4 (Wie zufrieden sind Sie mit den Öffnungszeiten?) nicht so gut zu den anderen paßt und eine zweite Dimension eröffnet. Hierbei wurden die Faktorladungen zeilenweise als Koordinaten im Raum der Faktoren (Komponenten) eingezeichnet. Die Gewichtsmatrix zur Berechnung der Faktorenwerte ist im Bild unten angegeben. Man erkennt wieder, daß keineswegs Summenwerte entstehen. Bei der 2. Komponente spielt vor allem Item 4 eine große Rolle. Abb. 14.27 zeigt die Kommunalitäten $h_i^2 = (\mathbf{\Lambda}\mathbf{\Lambda}')_{ii} = \sum_j \lambda_{ij}^2$, also die erklärte Streuung der Faktoren (vgl. Abs. 14.11.2). Schließlich ist ein Dialog zum Speichern der Faktorenwerte gezeigt (Abb. 14.28; vgl. auch (14.258)).

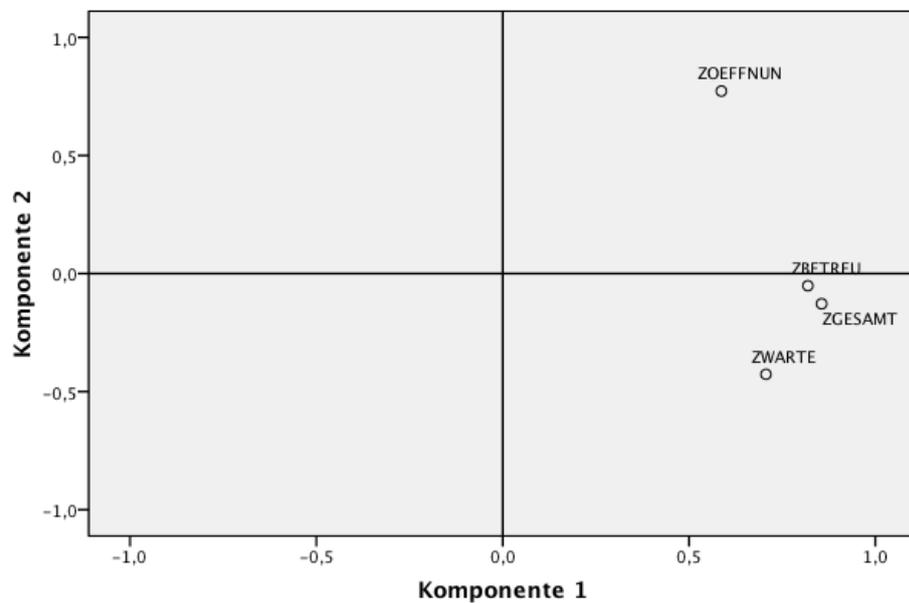
Komponentenmatrix^a

	Komponente	
	1	2
ZGESAMT	,856	-,128
ZWARTE	,707	-,427
ZBETREU	,819	-,052
ZOEFFNUN	,587	,773

Extraktionsmethode:
Hauptkomponentenanalyse.

a. 2 Komponenten
extrahiert

Komponentendiagramm

Koeffizientenmatrix der
Komponentenwerte

	Komponente	
	1	2
ZGESAMT	,381	-,160
ZWARTE	,315	-,535
ZBETREU	,364	-,065
ZOEFFNUN	,261	,968

Extraktionsmethode:
Hauptkomponentenanalyse.

Abbildung 14.26: Hauptkomponentenanalyse: Faktorladungen und Gewichte zur Berechnung der Faktorwerte (zwei Faktoren).

Kommunalitäten

	Anfänglich	Extraktion
ZGESAMT	1,000	,732
ZWARTE	1,000	,499
ZBETREU	1,000	,670
ZOEFFNUN	1,000	,344

Extraktionsmethode:
Hauptkomponentenanalyse.

Erklärte Gesamtvarianz

Komponente	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion		
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	2,246	56,161	56,161	2,246	56,161	56,161
2	,798	19,949	76,110			
3	,596	14,907	91,017			
4	,359	8,983	100,000			

Extraktionsmethode: Hauptkomponentenanalyse.

Abbildung 14.27: Kommunalitäten und erklärte Varianz (zwei Faktoren).

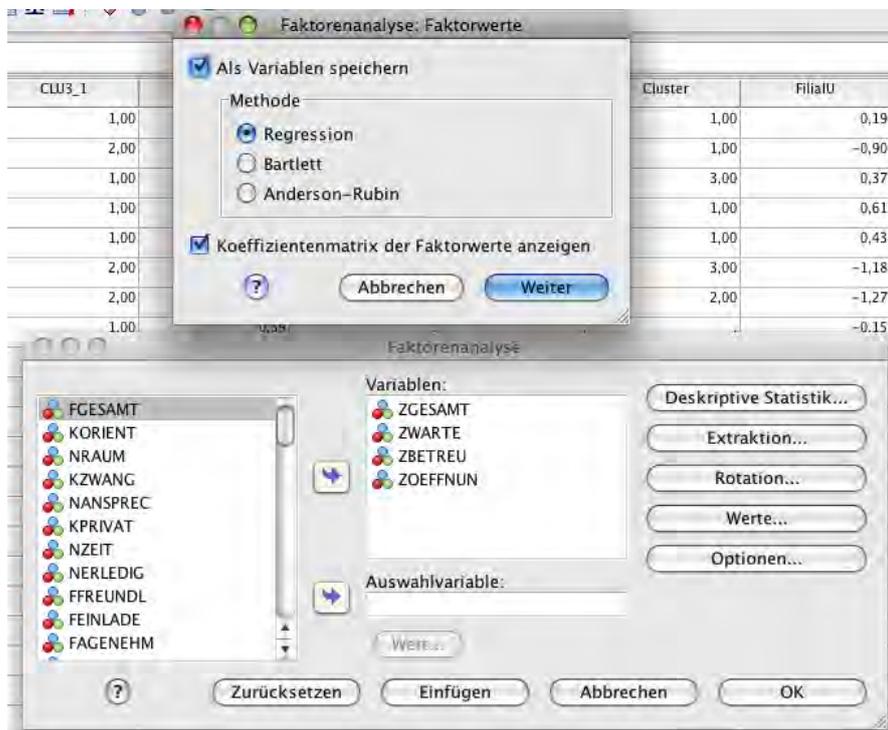


Abbildung 14.28: Dialog zur Berechnung der Faktorwerte (siehe Text).

14.11.2 Hauptkomponentenanalyse

Die Hauptkomponentenanalyse ist eine deskriptive Methode, bei der die beobachtete Kovarianz- oder Korrelationsmatrix so zerlegt wird, daß eine geringere Zahl $q < p$ von Faktoren ausreicht, um die Beobachtungen zu erklären. Sind etwa 2 Variablen stark korreliert, so genügt im Grunde eine Dimension, um deren Variabilität zu beschreiben (Abb. 14.29). Aus dem Faktorenmodell

$$\mathbf{x} = \mathbf{\Lambda}\boldsymbol{\xi} + \boldsymbol{\epsilon} \quad (14.268)$$

ergibt sich die Kovarianzmatrix (Streuungszerlegung)

$$\boldsymbol{\Sigma} = \mathbf{\Lambda}\mathbf{\Lambda}' + \mathbf{V} \quad (14.269)$$

der beobachtbaren Variablen. Hier wurde die Orthogonalität von $\boldsymbol{\xi}$ und $\boldsymbol{\epsilon}$ sowie die Orthonormalität von $\boldsymbol{\xi}$ ($\boldsymbol{\Phi} = \mathbf{I}$) ausgenutzt.

Daher setzt sich $\boldsymbol{\Sigma}$ aus dem Quadrat der sogenannten Ladungsmatrix $\mathbf{\Lambda} : p \times q$ (Gewichtsmatrix der Faktoren) und einer Fehlermatrix \mathbf{V} zusammen.

Die Diagonale der von den Faktoren erklärten Varianz wird als **Kommunalität**

Kommunalität

$$h_i^2 = (\mathbf{\Lambda}\mathbf{\Lambda}')_{ii} = \sum_{j=1}^q \lambda_{ij}^2 \quad (14.270)$$

bezeichnet, während die Diagonale $[v_{11}, \dots, v_{pp}]$ von \mathbf{V} **spezifische oder Einzelvarianz** heißt. Zusammenfassend gilt also die Zerlegung der Diagonale

Einzelvarianz

$$\sigma_{ii} = (\mathbf{\Lambda}\mathbf{\Lambda}')_{ii} + v_{ii} \quad (14.271)$$

$$= \sum_{j=1}^q \lambda_{ij}^2 + v_{ii}. \quad (14.272)$$

Der erste Term enthält gemeinsame Ladungen, während der zweite Term nur für Variable i spezifische Anteile enthält.

Man kennt jedoch zunächst nicht die Zahl q der Faktoren $\boldsymbol{\xi} = [\xi_1, \dots, \xi_q]'$.

Nun kann man sich die sogenannte **Hauptachsentransformation** zunutze machen. Es gilt für jede symmetrische Matrix die **Eigenwert (Spektral)-Zerlegung**

$$\Sigma = \sum_{i=1}^p \mu_i \psi_i \psi_i' = \mathbf{P} \mathbf{M} \mathbf{P}' \quad (14.273)$$

$$\mathbf{P} = [\psi_1, \dots, \psi_p] : p \times p \quad (14.274)$$

$$\mathbf{P}' = \mathbf{P}^{-1} \text{ (orthogonale Matrix)} \quad (14.275)$$

$$\mathbf{M} = \text{Diag}(\mu_1, \dots, \mu_p) : p \times p \text{ (Diagonalmatrix)} \quad (14.276)$$

**Eigenwert
(Spektral)-
Zerlegung**

wobei die reellen Zahlen $\mu_1 \geq \mu_2 \geq \dots \geq \mu_p$ als **Eigenwerte** und die Vektoren ψ_i als **Eigenvektoren** bezeichnet werden. Es handelt sich dabei um solche Vektoren

$$\Sigma \psi_i = \mu_i \psi_i, \quad (14.277)$$

deren Richtung durch Σ nicht verändert wird (Matrizen drehen Vektoren).

Die Eigenwert-Zerlegung hat **zunächst keine statistische Interpretation**, sondern gilt ganz allgemein für quadratische Matrizen (siehe Abs. 14.11.3). Allerdings sind für **positiv definite Matrizen die Eigenwerte positiv**, was z.B. bei Kovarianz-Matrizen der Fall ist.

Vergleicht man nun (14.269) mit (14.273), so kann man die Teilsumme

$$\Sigma_1 = \sum_{i=1}^q \mu_i \psi_i \psi_i' \quad (14.278)$$

**Hauptkomponenten-
Analyse**

mit $\Lambda \Lambda'$ identifizieren und den Rest

$$\Sigma_2 = \sum_{i=q+1}^p \mu_i \psi_i \psi_i' \quad (14.279)$$

mit \mathbf{V} .¹² Die Idee ist, daß sich die Matrix $\Sigma = \Sigma_1 + \Sigma_2$ gut durch Σ_1 approximieren läßt.

¹²Allerdings ist die Residualmatrix Σ_2 i.a. nicht diagonal und auch singular, da sie nur aus $p - q$ Dyaden $\psi_i \psi_i'$ besteht, d.h. $\text{rg}(\Sigma_2) = p - q$. Aufgrund der Nichtdiagonalität lassen sich keine Einzelvarianzen modellieren, die nur die Variable i beeinflussen.

In Matrixform ergibt sich

$$\Sigma_1 = \sum_{i=1}^q \mu_i \psi_i \psi_i' \quad (14.280)$$

$$= \mathbf{P}_1 \mathbf{M}_1 \mathbf{P}_1' = \mathbf{\Lambda} \mathbf{\Lambda}' \quad (14.281)$$

mit den Teilmatrizen

$$\mathbf{P}_1 = [\psi_1, \dots, \psi_q] = \mathbf{P}[\mathbf{I}_q, \mathbf{O}]' : p \times q \quad (14.282)$$

$$\mathbf{M}_1 = \text{Diag}(\mu_1, \dots, \mu_q) = \mathbf{M}[\mathbf{I}_q, \mathbf{O}]' : q \times q. \quad (14.283)$$

Setzt man also

**Faktorladungen
(Hauptkomponenten-
Analyse)**

$$\mathbf{\Lambda} = \mathbf{P}_1 \mathbf{M}_1^{1/2} \quad (14.284)$$

so ist das Produkt

$$\mathbf{\Lambda} \mathbf{\Lambda}' = \mathbf{P}_1 \mathbf{M}_1^{1/2} (\mathbf{P}_1 \mathbf{M}_1^{1/2})' = \mathbf{P}_1 \mathbf{M}_1 \mathbf{P}_1' = \Sigma_1. \quad (14.285)$$

Somit findet man eine Darstellung der Faktorladungen $\mathbf{\Lambda}$ durch die Eigenwerte und Eigenvektoren der Kovarianzmatrix.

scree plot

Die optimale Zahl q kann durch Auftragen der Eigenwerte graphisch bestimmt werden (Abb. 14.23 unten). Diese Graphik wird als **scree plot** bezeichnet. Meistens findet man einige große und viele kleine Eigenwerte, sodaß man am **Knick** des scree plot die optimale Anzahl von Faktoren (hier $q = 1$) ablesen kann.

Die geschätzten **Faktorwerte** ergeben sich durch die optimale lineare Prognose (14.258) nach Thompson

$$E[\boldsymbol{\xi}|\mathbf{x}] = \mathbf{\Lambda}' \Sigma^{-1} \mathbf{x}. \quad (14.286)$$

Durch eine einfache Umrechnung ergibt sich für die optimale lineare Prognose die explizite Form

$$E[\boldsymbol{\xi}|\mathbf{x}] = \boldsymbol{\Lambda}'\boldsymbol{\Sigma}^{-1}\mathbf{x} \quad (14.287)$$

$$= (\mathbf{M}_1^{-1/2}\mathbf{P}'_1)(\mathbf{P}\mathbf{M}^{-1}\mathbf{P}')\mathbf{x} \quad (14.288)$$

$$= \mathbf{M}_1^{-1/2}(\mathbf{P}'_1\mathbf{x}) \quad (14.289)$$

$$= \mathbf{M}_1^{-1/2}\mathbf{y}_1, \quad (14.290)$$

da $\mathbf{P}'_1\mathbf{P} = [\mathbf{I}_q, \mathbf{O}] : q \times p$.

Der Term $\mathbf{y}_1 = \mathbf{P}'_1\mathbf{x}$ ist aber die **Projektion** der Beobachtungen \mathbf{x} auf die ersten q Eigenvektoren $\boldsymbol{\psi}_1, \dots, \boldsymbol{\psi}_q$, die sogenannten **Hauptkomponenten**

$$\mathbf{y}_1 = \mathbf{P}'_1\mathbf{x} = \begin{bmatrix} \boldsymbol{\psi}'_1\mathbf{x} \\ \vdots \\ \boldsymbol{\psi}'_q\mathbf{x} \end{bmatrix}. \quad (14.291)$$

Hauptkomponenten

Die geschätzten Faktoren sind also die mit $1/\sqrt{\mu_i}$ skalierten Hauptkomponenten.

Die Skalierung ist nötig, da \mathbf{y}_1 nur orthogonal ist ($\text{Cov}(\mathbf{y}) = \text{Cov}(\mathbf{P}'_1\mathbf{x}) = \mathbf{P}'_1\boldsymbol{\Sigma}\mathbf{P}_1 = \mathbf{M}_1$), während die Faktoren sogar orthonormiert sind, d.h.

$$\begin{aligned} \text{Cov}(\mathbf{M}_1^{-1/2}\mathbf{P}'_1\mathbf{x}) &= \mathbf{M}_1^{-1/2}\mathbf{P}'_1\boldsymbol{\Sigma}\mathbf{P}_1\mathbf{M}_1^{-1/2} \\ &= \mathbf{M}_1^{-1/2}\mathbf{M}_1\mathbf{M}_1^{-1/2} = \mathbf{I}_q. \end{aligned} \quad (14.292)$$

Die bisherige Argumentation erfolgte auf der Ebene der Kovarianz-Matrizen. Man kann aber auch direkt den Beobachtungsvektor

$$\mathbf{x} = \mathbf{P}\mathbf{y} = [\boldsymbol{\psi}_1, \dots, \boldsymbol{\psi}_p] \begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix} \quad (14.293)$$

als Funktion der orthogonalen Hauptkomponenten $\mathbf{y} = \mathbf{P}'\mathbf{x}$ ausdrücken (Cov(\mathbf{x}) = $\mathbf{P}\text{Cov}(\mathbf{y})\mathbf{P}' = \mathbf{P}\mathbf{M}\mathbf{P}' = \mathbf{\Sigma}$) und somit eine Zerlegung

$$\mathbf{x} = \mathbf{P}_1\mathbf{y}_1 + \mathbf{P}_2\mathbf{y}_2 \quad (14.294)$$

$$= \mathbf{\Lambda}\boldsymbol{\xi} + \boldsymbol{\epsilon} \quad (14.295)$$

induzieren ($\mathbf{y}_1 = [y_1, \dots, y_q]'$, $\mathbf{y}_2 = [y_{q+1}, \dots, y_p]'$; entsprechend für \mathbf{P}). Dann kann man durch Multiplizieren mit \mathbf{P}'_1 direkt nach der orthonormalen Komponente ($\text{Var}(\boldsymbol{\xi}) = \mathbf{I}_q$)

$$\boldsymbol{\xi} = \mathbf{M}_1^{-1/2}\mathbf{P}'_1\mathbf{x} = \mathbf{M}_1^{-1/2}\mathbf{y}_1 \quad (14.296)$$

auflösen, was mit (14.290) übereinstimmt. Somit sind bei der Hauptkomponenten-Analyse die Faktoren und deren Prognose identisch. In obigen Formeln wurden die Identitäten

$$\mathbf{P}'_1\mathbf{P}_1 = \begin{bmatrix} \boldsymbol{\psi}'_1 \\ \vdots \\ \boldsymbol{\psi}'_q \end{bmatrix} [\boldsymbol{\psi}_1, \dots, \boldsymbol{\psi}_q] = \mathbf{I}_q \quad (14.297)$$

$$\mathbf{P}'_1\mathbf{P}_2 = \begin{bmatrix} \boldsymbol{\psi}'_1 \\ \vdots \\ \boldsymbol{\psi}'_q \end{bmatrix} [\boldsymbol{\psi}_{q+1}, \dots, \boldsymbol{\psi}_p] = \mathbf{O}_{q \times p} \quad (14.298)$$

benutzt.

Bemerkung:

Bisher wurden die Eigenvektoren und Eigenwerte aus der theoretischen Kovarianzmatrix $\mathbf{\Sigma}$ berechnet. In der Praxis muß diese aus Daten $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N]'$: $N \times p$ geschätzt werden, d.h.

$$\hat{\boldsymbol{\Sigma}} = \mathbf{S} = \frac{1}{N-1}\mathbf{X}'\mathbf{H}\mathbf{X}. \quad (14.299)$$

Die Berechnungen erfolgen dann analog mit den Stichprobengrößen $\mathbf{S} = \hat{\mathbf{P}}\hat{\mathbf{M}}\hat{\mathbf{P}}'$.

Die Eigenvektoren und Eigenwerte sind somit Zufallsgrößen (vgl. Mardia et al., Kap. 8).

14.11.3 Mathematischer Einschub: Hauptachsentransformation

Im vorigen Abschnitt wurde die **Eigenwertgleichung**

$$\Sigma \boldsymbol{\psi}_i = \mu_i \boldsymbol{\psi}_i \quad (14.300)$$

$$\Sigma \mathbf{P} = \mathbf{P} \mathbf{M} \quad (14.301)$$

zur Zerlegung der symmetrischen Matrix Σ benutzt. Daher muß die Gleichung

$$(\Sigma - \mu_i \mathbf{I}) \boldsymbol{\psi}_i = \mathbf{0} \quad (14.302)$$

gelöst werden, was nur für

$$\det(\Sigma - \mu_i \mathbf{I}) = 0 \quad (14.303)$$

zu nichttrivialen Lösungen $\boldsymbol{\psi}_i \neq \mathbf{0}$ führt. Aus der **Determinante** ergibt sich eine Gleichung p -ten Grades für die Eigenwerte (Säkulargleichung).

Da \mathbf{P} die orthonormalen Eigenvektoren $\boldsymbol{\psi}_i$ enthält, gilt

Säkulargleichung

$$\mathbf{P}' \mathbf{P} = \mathbf{P} \mathbf{P}' = \mathbf{I} \quad (14.304)$$

$$\mathbf{P}' = \mathbf{P}^{-1} \quad (14.305)$$

$$\Sigma = \mathbf{P} \mathbf{M} \mathbf{P}' = \sum_{i=1}^p \mu_i \boldsymbol{\psi}_i \boldsymbol{\psi}_i' \quad (14.306)$$

$\mathbf{P} = [\boldsymbol{\psi}_1, \dots, \boldsymbol{\psi}_p] : p \times p$, $\mathbf{M} = \text{Diag}(\mu_1, \dots, \mu_p) : p \times p$ (Diagonalmatrix).

Die Summendarstellung von Σ wird als **Eigenwertzerlegung** oder **Spektral-Darstellung** bezeichnet. Man spricht auch von **Diagonalisierung** ($\mathbf{P}' \Sigma \mathbf{P} = \mathbf{M}$) oder **Hauptachsentransformation**.

Eigenwertzerlegung

Die Wichtigkeit dieser Formeln kann gar nicht überschätzt werden. Sie erlauben, eine Matrix als Überlagerung von Projektionen $\boldsymbol{\psi}_i \boldsymbol{\psi}_i'$ auf eindimensionale Unterräume darzustellen, mit den Eigenwerten (Spektrum) als Gewicht.

Ganz allgemein gilt für die Spur (= trace) der Matrix

Spur

$$\sum_i \sigma_{ii} := \text{tr}(\Sigma) = \sum_i \mu_i = \text{tr}(\mathbf{M}), \quad (14.307)$$

da $\text{tr}(\mathbf{\Sigma}) = \text{tr}(\mathbf{PMP}') = \text{tr}(\mathbf{MP}'\mathbf{P}) = \text{tr}(\mathbf{M})$.

Übung: Beweisen Sie die zyklische Eigenschaft $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$ der Spur.

Beispiel 14.9 (Eigenwerte einer Korrelationsmatrix)

Für die (theoretische) Korrelationsmatrix¹³

$$\mathbf{R} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \quad (14.308)$$

ergeben sich die Eigenwerte aus der Bedingung

$$\det\left(\begin{bmatrix} 1-\mu & \rho \\ \rho & 1-\mu \end{bmatrix}\right) = 0 = (1-\mu)^2 - \rho^2 \quad (14.309)$$

$$\mu_{1,2} = 1 \pm \rho \quad (14.310)$$

Die Summe der Eigenwerte ist also $p = 2 = \text{tr}(\mathbf{R}) =$ Summe der Diagonale := Spur = trace. Ganz allgemein gilt

$$\sum_i r_{ii} := \text{tr}(\mathbf{R}) = \sum_i \mu_i = p. \quad (14.311)$$

Die Eigenvektoren ergeben sich aus den Bedingungen

$$(\mathbf{R} - \mu_1 \mathbf{I}_2) \boldsymbol{\psi}_1 = \begin{bmatrix} -\rho & \rho \\ \rho & -\rho \end{bmatrix} \begin{bmatrix} \psi_{11} \\ \psi_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (14.312)$$

$$(\mathbf{R} - \mu_2 \mathbf{I}_2) \boldsymbol{\psi}_2 = \begin{bmatrix} \rho & \rho \\ \rho & \rho \end{bmatrix} \begin{bmatrix} \psi_{21} \\ \psi_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (14.313)$$

Etwa löst

$$\begin{bmatrix} \psi_{11} \\ \psi_{12} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (14.314)$$

$$\begin{bmatrix} \psi_{21} \\ \psi_{22} \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (14.315)$$

$$(14.316)$$

obige Gleichungen. Das Betrags-Quadrat der Vektoren ist $[1, 1][1, 1]' = 2$, $[1, -1][1, -1]' = 2$, sodaß man

$$\boldsymbol{\psi}_1 = \begin{bmatrix} \psi_{11} \\ \psi_{12} \end{bmatrix} = 1/\sqrt{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (14.317)$$

$$\boldsymbol{\psi}_2 = \begin{bmatrix} \psi_{21} \\ \psi_{22} \end{bmatrix} = 1/\sqrt{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (14.318)$$

¹³Notation \mathbf{R} , um eine Verwechslung mit \mathbf{P} zu vermeiden.

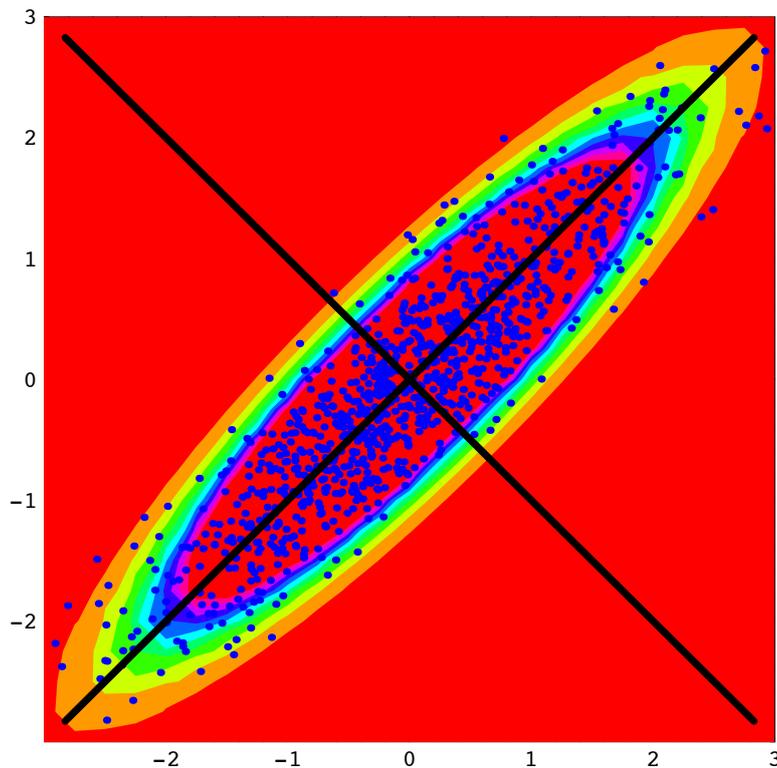


Abbildung 14.29: Simulierte normalverteilte Daten $\mathbf{x}_n, n = 1, \dots, N = 1000$ mit Kovarianz-Matrix $\mathbf{R} = \begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}$. Die Hauptachsen zeigen in Richtung der Winkelhalbierenden.

als orthonormierte Eigenvektoren findet.

Übung: Zeigen Sie, daß ψ_1, ψ_2 orthonormiert sind.

Es ist wichtig, daß die Eigenvektoren gar nicht von der Korrelation ρ abhängen. Sie zeigen in Richtung der Winkelhalbierenden der Quadranten. Abb. 14.29 zeigt simulierte Daten aus einer bivariaten Normalverteilung

$$N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}\right). \quad (14.319)$$

Die Eigenwerte von \mathbf{R} sind $1 \pm 0.9 = 1.9, 0.1$ und die orthogonale Matrix

der Eigenvektoren lautet

$$\mathbf{P} = 2^{-1/2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (14.320)$$

$$\mathbf{P}\mathbf{P}' = \mathbf{P}'\mathbf{P} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}_2. \quad (14.321)$$

Im gedrehten Koordinatensystem gilt daher

$$\mathbf{y} = \mathbf{P}'\mathbf{x} = \begin{bmatrix} \psi'_1\mathbf{x} \\ \psi'_2\mathbf{x} \end{bmatrix} = 2^{-1/2} \begin{bmatrix} x_1 + x_2 \\ x_1 - x_2 \end{bmatrix} \quad (14.322)$$

und $\text{Cov}(\mathbf{y}) = \mathbf{P}'\mathbf{R}\mathbf{P} = \mathbf{M} = \text{Diag}(1.9, 0.1)$.

Daher sind die Koordinaten (**Hauptkomponenten**) y_1, y_2 unkorreliert. Die quadratische Form (Ellipse) der Matrix \mathbf{R}

$$\mathbf{x}'\mathbf{R}\mathbf{x} = \sum_{ij} x_i \rho_{ij} x_j = x_1^2 + 2\rho x_1 x_2 + x_2^2 \quad (14.323)$$

ist diagonal im gedrehten System:

$$\mathbf{x}'\mathbf{R}\mathbf{x} = \mathbf{x}'\mathbf{P}\mathbf{P}'\mathbf{R}\mathbf{P}\mathbf{P}'\mathbf{x} \quad (14.324)$$

$$= \mathbf{y}'\mathbf{M}\mathbf{y} = \mu_1 y_1^2 + \mu_2 y_2^2 = (1 + \rho)y_1^2 + (1 - \rho)y_2^2. \quad (14.325)$$

Die im Bild gezeigte Ellipse ist allerdings

$$\mathbf{x}'\mathbf{R}^{-1}\mathbf{x} = \mathbf{y}'\mathbf{M}^{-1}\mathbf{y} \quad (14.326)$$

$$= \frac{y_1^2}{\mu_1} + \frac{y_2^2}{\mu_2} \quad (14.327)$$

$$= \frac{y_1^2}{1 + \rho} + \frac{y_2^2}{1 - \rho} \quad (14.328)$$

$$= \frac{y_1^2}{1.9} + \frac{y_2^2}{0.1}, \quad (14.329)$$

da die (multivariat) standardisierte Variable

$$\mathbf{z} = \mathbf{R}^{-1/2}\mathbf{x} = \mathbf{M}^{-1/2}\mathbf{P}'\mathbf{x} \quad (14.330)$$

auf die quadratische Form

$$\mathbf{z}'\mathbf{z} = (\mathbf{M}^{-1/2}\mathbf{P}'\mathbf{x})'\mathbf{M}^{-1/2}\mathbf{P}'\mathbf{x} = \mathbf{x}'\mathbf{P}\mathbf{M}^{-1}\mathbf{P}'\mathbf{x} = \mathbf{x}'\mathbf{R}^{-1}\mathbf{x} \quad (14.331)$$

führt.

Im Exponent der bivariaten Normalverteilung ($\mu_i = 0, \sigma_i^2 = 1$)

$$\begin{aligned}
 f(x_1, x_2) &= |2\pi\mathbf{R}|^{-1/2} \exp\left(-\frac{1}{2}\mathbf{x}'\mathbf{R}^{-1}\mathbf{x}\right) & (14.332) \\
 &= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}(x_1^2 - 2\rho x_1 x_2 + x_2^2)\right)
 \end{aligned}$$

**bivariate
Normalverteilung**

steht aber genau der Ausdruck $\mathbf{x}'\mathbf{R}^{-1}\mathbf{x} = \mathbf{z}'\mathbf{z}$.

Übung:

Zeigen Sie, daß $\det(\mathbf{R}) = |\mathbf{R}| = 1 - \rho^2$ und $\mathbf{R}^{-1} = \frac{1}{1-\rho^2} \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}$ gilt.

Berechnen Sie die quadratische Form $\mathbf{x}'\mathbf{R}^{-1}\mathbf{x}$ explizit.

Hauptkomponentenanalyse:

Aus dem bisher gezeigten ergeben sich leicht die Formeln der Hauptkomponentenanalyse. Nimmt man $q = 1$ Faktoren, so gilt

$$\mathbf{\Lambda} = \boldsymbol{\psi}_1 \mu_1^{1/2} = \sqrt{\frac{1+\rho}{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \quad (14.333)$$

Das Produkt

$$\mathbf{\Lambda}\mathbf{\Lambda}' = \frac{1+\rho}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} \quad (14.334)$$

$$= \frac{1+\rho}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad (14.335)$$

ist der erste Teil $\boldsymbol{\Sigma}_1$ der Varianz $\boldsymbol{\Sigma}(= \mathbf{R})$. Die Kommunalitäten sind somit

$$h_i^2 = (\mathbf{\Lambda}\mathbf{\Lambda}')_{ii} = \frac{1+\rho}{2} \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} \right)_{ii}; \quad i = 1, 2 \quad (14.336)$$

$$= \frac{1+\rho}{2} [1, 1] \quad (14.337)$$

$$= [0.95, 0.95] \quad (14.338)$$

im Zahlenbeispiel. Die Gewichte zur Berechnung der Faktoren lauten

$$\boldsymbol{\xi} = \mathbf{M}_1^{-1/2} \mathbf{P}'_1 \mathbf{x} \quad (14.339)$$

$$= (1 + \rho)^{-1/2} \boldsymbol{\psi}'_1 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (14.340)$$

$$= (1 + \rho)^{-1/2} \frac{1}{\sqrt{2}} [1, 1] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (14.341)$$

$$= \frac{1}{\sqrt{2(1 + \rho)}} (x_1 + x_2) \quad (14.342)$$

$$= 0.512989(x_1 + x_2) \quad (14.343)$$

(standardisierte 1. Hauptkomponente).

Übung:

Berechnen Sie die Restvarianz $\Sigma_2 = \mathbf{V} = \mu_2 \boldsymbol{\psi}_2 \boldsymbol{\psi}'_2$.



Beispiel 14.10 (Filialgestaltung von Bankfilialen)

Wir betrachten nochmals das Konstrukt Kundenzufriedenheit (vgl. Bsp. 14.8). Abb. 14.30 zeigt die Resultate einer Hauptkomponentenanalyse (Eigenwerte, Eigenvektoren und 3-D-Streudiagramm mit Hauptachsen). Das Ellipsoid ist durch die quadratische Form $\mathbf{x}' \boldsymbol{\Sigma}^{-1} \mathbf{x} = r^2$ gegeben (siehe Bsp. 14.9). Die Graphik läßt sich mit der Maus interaktiv rotieren.



14.11.4 Maximum-Likelihood-Methode

14.11.4.1 Modell

Faktorenanalyse (Maximum- Likelihood)

Die Maximum-Likelihood(ML)-Methode geht von dem Grundmodell

$$\mathbf{x} = \boldsymbol{\Lambda} \boldsymbol{\xi} + \boldsymbol{\epsilon}. \quad (14.344)$$

aus, wobei nun probabilistische Annahmen gemacht werden. Unter Normalverteilungsannahme für $\boldsymbol{\xi}$ und $\boldsymbol{\epsilon}$ ist auch \mathbf{x} normalverteilt mit Erwartungswert und Kovarianzmatrix

$$E[\mathbf{x}] = \boldsymbol{\mu} = \mathbf{0} \quad (14.345)$$

$$\text{Var}(\mathbf{x}) = \boldsymbol{\Sigma} = \boldsymbol{\Lambda} \boldsymbol{\Phi} \boldsymbol{\Lambda}' + \mathbf{V}. \quad (14.346)$$

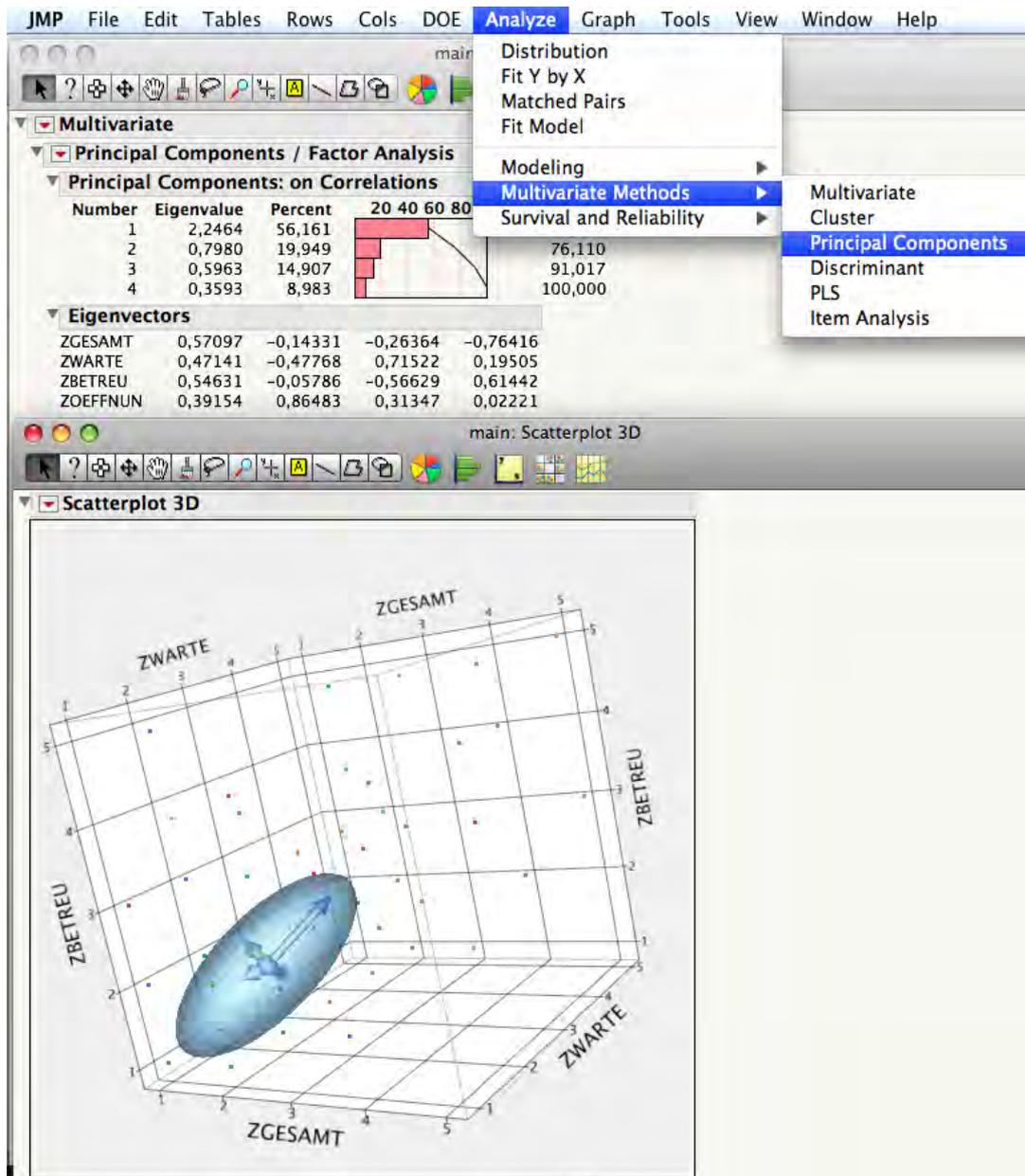


Abbildung 14.30: JMP: Hauptkomponentenanalyse und 3-D-Streudiagramm mit Hauptachsen.

Hierbei wurde wieder die Zentrierung sowie Unkorreliertheit der Faktoren und Fehler angenommen ($\text{Cov}(\boldsymbol{\xi}, \boldsymbol{\epsilon}) = \mathbf{O}$). Daher ist die gemeinsame Verteilung der N Messungen $\mathbf{x}_n, n = 1, \dots, N$ durch

$$f(\mathbf{x}_1, \dots, \mathbf{x}_N) = \prod_{n=1}^N f(\mathbf{x}_n) \quad (14.347)$$

$$= \prod_{n=1}^N |2\pi\boldsymbol{\Sigma}|^{-1/2} \exp\left(-\frac{1}{2}\mathbf{x}_n'\boldsymbol{\Sigma}^{-1}\mathbf{x}_n\right) \quad (14.348)$$

gegeben. Interpretiert man die Wahrscheinlichkeits-Dichte der Daten als Funktion der freien Parameter $\boldsymbol{\theta}$ in $\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \boldsymbol{\Sigma}(\boldsymbol{\Lambda}(\boldsymbol{\theta}), \boldsymbol{\Phi}(\boldsymbol{\theta}), \mathbf{V}(\boldsymbol{\theta}))$, so ergibt sich die **Likelihood-Funktion**¹⁴

$$L(\boldsymbol{\theta}) = f(\mathbf{x}_1, \dots, \mathbf{x}_N; \boldsymbol{\theta}). \quad (14.349)$$

Maximiert man diese bezüglich $\boldsymbol{\theta}$, so ergibt sich der **Maximum-Likelihood(ML)-Schätzer**

Maximum-Likelihood(ML)-Schätzer

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} L(\boldsymbol{\theta}) = \arg \max_{\boldsymbol{\theta}} f(\mathbf{x}_1, \dots, \mathbf{x}_N; \boldsymbol{\theta}). \quad (14.350)$$

Meistens wird anstelle von L die **Log-Likelihood** $l = \log(L)$

$$l(\boldsymbol{\theta}) = \sum_{n=1}^N \log(f(\mathbf{x}_n)) \quad (14.351)$$

$$= -\frac{N}{2} \log |2\pi\boldsymbol{\Sigma}| - \frac{1}{2} \sum_{n=1}^N \text{tr}[\boldsymbol{\Sigma}^{-1}\mathbf{x}_n\mathbf{x}_n'] \quad (14.352)$$

$$= -\frac{N}{2} \log |2\pi\boldsymbol{\Sigma}(\boldsymbol{\theta})| - \frac{N}{2} \text{tr}[\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\mathbf{S}] \quad (14.353)$$

benutzt [$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta}); \mathbf{x}'\mathbf{A}\mathbf{x} = \text{tr}(\mathbf{x}'\mathbf{A}\mathbf{x}) = \text{tr}(\mathbf{A}\mathbf{x}\mathbf{x}')$ (Übung)]. Die Daten wurden hierbei in der Statistik

$$\mathbf{S} = \frac{1}{N} \sum_{n=1}^N \mathbf{x}_n\mathbf{x}_n' \quad (14.354)$$

¹⁴Die immer wiederkehrende Behauptung, daß die ML-Methode nur bei Normalverteilung anwendbar sei, ist falsch. Es ergibt sich nur eine andere Form als (14.348).

komprimiert (Stichproben-Kovarianz bei zentrierten Variablen). Bei nichtzentrierten Größen $E[\mathbf{x}] = \boldsymbol{\mu}$ nimmt man

$$\mathbf{S} = \frac{1}{N} \sum_{n=1}^N (\mathbf{x}_n - \bar{\mathbf{x}})(\mathbf{x}_n - \bar{\mathbf{x}})' = \frac{1}{N} \mathbf{X}'\mathbf{H}\mathbf{X}. \quad (14.355)$$

da $E(\mathbf{x}_n - \bar{\mathbf{x}}) = \mathbf{0}$, wie gefordert.¹⁵

Im klassischen Modell der orthonormierten Faktoren kann zudem $\boldsymbol{\Phi} = \text{Cov}(\boldsymbol{\xi}) = \mathbf{I}_q$ gesetzt werden. Dann enthält das Modell nur die freien Parameter in $\boldsymbol{\Lambda}$ und \mathbf{V} . Die Berechnung der ML-Schätzer muß numerisch erfolgen, da es sich um ein nichtlineares Schätzproblem handelt.

14.11.4.2 Identifikation

Zuvor muß aber geklärt werden, ob die Parameter identifizierbar sind (**Identifikationsproblem**). Im Fall von $p = 2$ manifesten Variablen und einem latenten Faktor ($q = 1$) gilt bei diagonaler Fehler-Varianz \mathbf{V}

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}' + \mathbf{V} \quad (14.356)$$

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} = \begin{bmatrix} \lambda_{11} \\ \lambda_{21} \end{bmatrix} \begin{bmatrix} \lambda_{11} & \lambda_{21} \end{bmatrix} + \begin{bmatrix} v_{11} & 0 \\ 0 & v_{22} \end{bmatrix} \quad (14.357)$$

$$= \begin{bmatrix} \lambda_{11}^2 & \lambda_{11}\lambda_{21} \\ \lambda_{21}\lambda_{11} & \lambda_{21}^2 \end{bmatrix} + \begin{bmatrix} v_{11} & 0 \\ 0 & v_{22} \end{bmatrix}. \quad (14.358)$$

Aus den Parametern in $\boldsymbol{\Sigma}$ (Parameter der beobachtbaren Größen) müssen sich die freien Parameter der latenten Variablen **eindeutig** berechnen lassen, damit das Modell identifiziert ist. Aufgrund der Symmetrie hat man 3 Parameter in $\boldsymbol{\Sigma}$, jedoch 4 Parameter auf der rechten Seite.

Identifikation

Daher kann keine eindeutige Lösung gefunden werden – das Modell ist **nicht identifizierbar** und kann daher auch nicht geschätzt werden. Man muß **Restriktionen** setzen (etwa $\lambda_{11} = 1$) oder mehr Indikatoren messen. Dann ergibt sich

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} = \begin{bmatrix} 1 \\ \lambda_{21} \end{bmatrix} \begin{bmatrix} 1 & \lambda_{21} \end{bmatrix} + \begin{bmatrix} v_{11} & 0 \\ 0 & v_{22} \end{bmatrix} \quad (14.359)$$

$$= \begin{bmatrix} 1 & \lambda_{21} \\ \lambda_{21} & \lambda_{21}^2 \end{bmatrix} + \begin{bmatrix} v_{11} & 0 \\ 0 & v_{22} \end{bmatrix}. \quad (14.360)$$

¹⁵Glg. (14.351) ist dann die sog. konzentrierte Likelihood nach Einsetzen des ML-Schätzers $\hat{\boldsymbol{\mu}} = \bar{\mathbf{x}}$.

$s_1(p, q)$	1	2	3	4
1	-1	-2	-3	-4
2	-1	-3	-5	-7
3	0	-3	-6	-9
4	2	-2	-6	-10
5	5	0	-5	-10
6	9	3	-3	-9

$s_2(p, q)$	1	2	3	4
1	-1	-1	0	2
2	-1	-2	-2	-1
3	0	-2	-3	-3
4	2	-1	-3	-4
5	5	1	-2	-4
6	9	4	0	-3

Tabelle 14.10: Zahl der Gleichungen minus Zahl der freien Parameter. Oben: ohne Restriktionen. Unten: mit Restriktionen (vgl. Haupttext)

woraus alle Parameter identifiziert sind.

Übung: Lösen Sie die Gleichungen nach λ und v auf.

Im allgemeinen hat man $p(p+1)/2$ freie Parameter in Σ sowie $pq+p$ freie Parameter in Λ und \mathbf{V} (falls diagonal). Die Zahl der Gleichungen $p(p+1)/2$ muß also auf jeden Fall größer oder gleich $pq+p$ sein. Die **Differenz** $s_1(p, q)$ beider Größen ist als Tabelle aufgetragen ($p = 1, \dots, 6; q = 1, \dots, 4$), Tab. 14.10, oben. Führt man $q(q-1)$ Restriktionen durch

$$\Lambda \operatorname{diag}(\sigma_{11}, \dots, \sigma_{pp})^{-1} \Lambda' = \text{diagonal} \quad (14.361)$$

ein (vgl. Mardia et al., Kap. 9.2), so ergibt sich die untere Tabelle für $s_2(p, q) = p(p+1)/2 - (pq+p) + q(q-1)$. s darf auf jeden Fall nicht negativ sein.

Hat man $s > 0$, so ergibt das Faktor-Modell eine sparsamere Erklärung der Daten, als dies $\operatorname{Var}(\mathbf{x}) = \Sigma$ leisten würde.

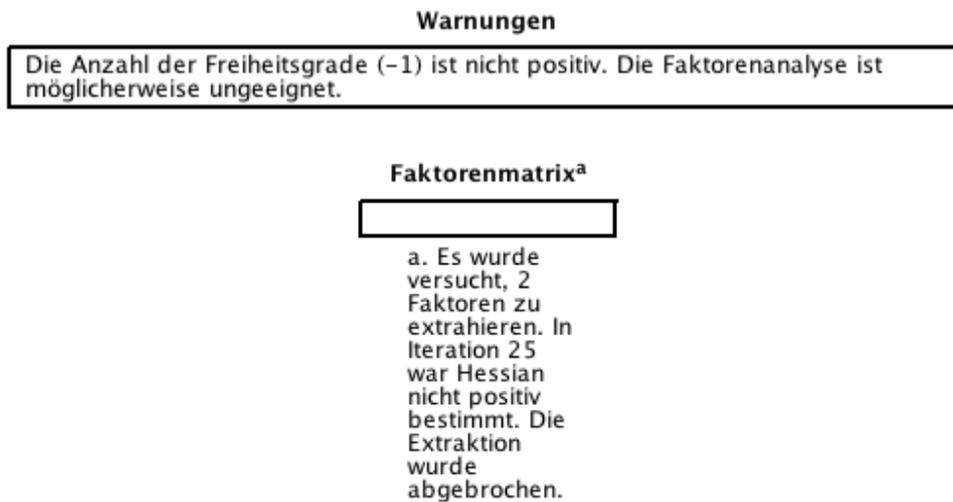


Abbildung 14.31: ML-Faktorenanalyse (zwei Faktoren). Das Modell ist nicht identifizierbar.

Beispiel 14.11 (Filialgestaltung von Bankfilialen)

In Bsp. 14.8 wurden 1-Faktor und 2-Faktor-Lösungen für 4 Items betrachtet (Hauptkomponentenanalyse). Schätzt man die Faktoren mit ML, so ergibt sich für $q = 2$ ein Problem (Abb. 14.31). Blickt man in Tabelle 14.10, unten, so ist $s_2(4, 2) = -1$. Daher ist das Modell **nicht identifizierbar**.

Die Bemerkung von SPSS,

... Die Faktorenanalyse ist möglicherweise ungeeignet.

ist für den Anwender irreführend.

Genauso ist der Text

Es wurde versucht, 2 Faktoren zu extrahieren. In Iteration 25 war Hessian nicht positiv bestimmt. Die Extraktion wurde abgebrochen.

weder gutes Deutsch noch für den Benutzer sehr hilfreich (*not positive definite* wurde offenbar falsch in *nicht positiv bestimmt* übersetzt).

In der Tat spiegelt sich die Nichtidentifizierbarkeit des Modells darin wider, daß die geschätzte Fisher-Informationsmatrix (negative Hesse-Matrix) singulär ist

	Faktor
	1
ZGESAMT	,855
ZWARTE	,565
ZBETREU	,735
ZOEFFNUN	,420

Extraktionsmethode:
Maximum-Likelihood.

a. 1 Faktoren extrahiert. Es werden 4 Iterationen benötigt.

	Faktor
	1
ZGESAMT	,570
ZWARTE	,149
ZBETREU	,287
ZOEFFNUN	,091

Extraktionsmethode:
Maximum-Likelihood.

Abbildung 14.32: ML-Faktorenanalyse: Faktorladungen und Gewichte zur Berechnung der Faktorwerte ($q = 1$ Faktor; vgl. Abb. 14.25).

$[\mathbf{H} = \partial^2 l / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}' = \text{Hesse-Matrix} = 2. \text{ Ableitung der log-Likelihood } l(\boldsymbol{\theta}); \text{ vgl. Rothenberg (1971)}]$.

Leider lassen sich in SPSS keine zusätzlichen Restriktionen setzen, um eine Identifikation zu ermöglichen (dies ist in LISREL, Amos oder anderen SEM-Modellen möglich).¹⁶

Die Fisher-Informationsmatrix $\mathbf{F} = -E[\mathbf{H}] \approx -\mathbf{H}$ erlaubt es, die Standardfehler der ML-Schätzungen für die Parameter $\boldsymbol{\theta}$ zu berechnen. Die asymptotische Kovarianzmatrix der Schätzer ist durch $\text{Cov}(\hat{\boldsymbol{\theta}}) = \mathbf{F}^{-1}$ gegeben, deren Diagonalelemente die Standardfehler ergeben (vgl. Mardia et al., loc. cit., Kap. 4.2). In Programmen wie LISREL etc. werden diese Schätzfehler ausgegeben.

Nimmt man nur einen Faktor ($q = 1$) bei der ML-Analyse, so ergibt sich die Lösung Abb. 14.32. Auch hier hat Item 4 ein relativ kleines Gewicht (vgl. Abb. 14.25).

■

¹⁶SEM = Structural Equations Models = Strukturgleichungsmodelle; vgl. http://www.fernuni-hagen.de/imperia/md/content/ls_statistik/publikationen/semarchive.exe

14.12 Methode der kleinsten Quadrate

Die Maximum-Likelihood-Methode setzt die Kenntnis der Verteilung von \mathbf{x} voraus. Wenn diese nicht bekannt ist, kann man anstatt der möglicherweise inkorrekten gaußschen Likelihood-Funktion

$$l(\boldsymbol{\theta}) = -\frac{N}{2} \log |2\pi \boldsymbol{\Sigma}(\boldsymbol{\theta})| - \frac{N}{2} \text{tr}[\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} \mathbf{S}] \quad (14.362)$$

eine Diskrepanzfunktion

$$F(\boldsymbol{\theta}) = \text{tr}[(\boldsymbol{\Sigma}(\boldsymbol{\theta}) - \mathbf{S})^2] \quad (14.363)$$

$$= \|\boldsymbol{\Sigma}(\boldsymbol{\theta}) - \mathbf{S}\|^2 = \sum_{ij} (\sigma_{ij} - s_{ij})^2 \geq 0 \quad (14.364)$$

**Diskrepanzfunktion
(ungewichtete
kleinste Quadrate)**

betrachten (quadrierter euklidischer Abstand von $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ und der Stichprobenkovarianzmatrix \mathbf{S}) und diese minimieren.¹⁷

Man spricht von einer ULS-Diskrepanzfunktion (ULS = unweighted least squares = ungewichtete kleinste Quadrate).

Allgemeiner definiert man eine Diskrepanzfunktion mit gewichteten kleinsten Quadraten (Fahrmeir et al., 1996, Kap. 11, S. 746–748 f.)

$$F(\boldsymbol{\theta}) = \text{tr}\{[(\boldsymbol{\Sigma}(\boldsymbol{\theta}) - \mathbf{S})\mathbf{G}]^2\} \quad (14.365)$$

**Diskrepanzfunktion
(gewichtete
kleinste Quadrate)**

und Gewichtungsmatrix $\mathbf{G} > 0$ (positiv definit und symmetrisch). F ist eine positiv definite quadratische Form.¹⁸ Die Wahl $\mathbf{G} = \mathbf{S}^{-1}$ wird als GLS-Diskrepanzfunktion bezeichnet, vgl. Abb. 14.33 (GLS = generalized least squares = verallgemeinerte KQ-Methode).

¹⁷Der Ausdruck $\|\mathbf{A}\|^2 = \text{tr}(\mathbf{A}\mathbf{A}') = \sum_{ij} a_{ij}^2$ ist die quadrierte euklidische Matrixnorm von \mathbf{A} . Hier gilt $\boldsymbol{\Sigma} - \mathbf{S} := \mathbf{A} = \mathbf{A}'$.

¹⁸Schreibt man für die symmetrischen Matrizen $a_{ij} = a_{ji}, g_{ij} = g_{ji}$ in Komponenten $\text{tr}[\mathbf{A}\mathbf{G}\mathbf{A}\mathbf{G}] = \sum_{ijkl} a_{ij}g_{jk}a_{kl}g_{li} = \sum_{ijk} a_{ij}g_{il}g_{jk}a_{ik} = \mathbf{a}'(\mathbf{G} \otimes \mathbf{G})\mathbf{a}$, $\mathbf{a} = [a_{11}, a_{12}, \dots, a_{pp}]' := \text{row}\mathbf{A}$ (zeilenweiser Vektor aus \mathbf{A}), so ist dies eine quadratische Form der positiv definiten Matrix $\boldsymbol{\Omega} = \mathbf{G} \otimes \mathbf{G}$. Man kann auch eine allgemeine positiv definite Matrix $\boldsymbol{\Omega}$ als Gewichtsmatrix verwenden (Diskrepanzfunktionen von quadratischer Form).

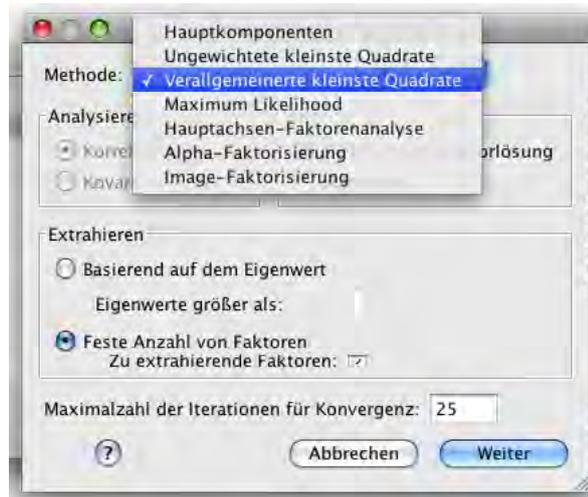


Abbildung 14.33: GLS-Faktorenanalyse (Auswahldialog).

Auch in diesem Fall stellt sich das Identifikationsproblem, nämlich ob sich die Parameter in $\Sigma(\theta)$ eindeutig aus Σ berechnen lassen. Die Schätzung mit SPSS führt auf die analogen Probleme wie bei ML.

Kapitel 15

Fallstudie: Filialgestaltung und Kundenzufriedenheit

Im Rahmen eines Präsenzseminars wurde die **Wirkung der Filialgestaltung von Banken auf das Kundenverhalten** empirisch untersucht.¹

15.1 Fragestellung

15.1.1 Ausgangslage

Banken verlieren an Kundenbindung aufgrund von

- Online-Banking
- Variety Seeking / Risikostreuung
- Preisbewusstsein

Zielsetzung der Banken:

- Kunden in die Filiale an den sog. Point-of-Sale ziehen
→ Einflussnahme durch persönliche Kontakte auf die Entscheidungen des Kunden

¹Das Seminar wurde in Zusammenarbeit mit dem Lehrstuhl für Dienstleistungsmanagement (Univ.-Prof. Dr. S. Fließ) durchgeführt. Die folgenden Ausführungen wurden aus einem Diskussionspapier des Lehrstuhls zusammengestellt.

- Erhöhung der Kontaktwahrscheinlichkeit
→ Information über Angebote, Einflussnahme auf das Entscheidungsverhalten des Kunden

Ein Mittel hierzu stellt die Umgestaltung der Filialen dar.

15.1.1.1 Merkmale des neuen Filialkonzeptes

Raumkonzept: 3 Zonen-Konzept:

- SB-Bereich
- Schnellberatungs-Bereich
- Beratungsbereich

Personalkonzept: Qualifikationsabstufungen

15.1.2 Untersuchungsziel

Welchen Einfluss hat die Gestaltung der Filiale auf die Wahrnehmung der Kunden, die Nutzung und Zielerreichung in der Filiale, die Zufriedenheit mit dem Filialbesuch und die Kundenbindung?

- Wie nehmen die Kunden die Filialgestaltung wahr?
- In welchem Maße unterstützt oder behindert die räumliche Gestaltung die Kunden bei der Erledigung ihrer Bankgeschäfte?
- Wie wirken sich Raumgestaltung und Beratungskonzept auf die Kundenzufriedenheit aus?
- Sind die Kunden mit dem Besuch insgesamt und mit einzelnen Aspekten zufrieden?
- Wie beeinflusst die Filialgestaltung die Kundenloyalität?
- Unterscheiden sich die Kunden hinsichtlich ihrer Reaktionen auf die Filialgestaltung und in ihrem Verhalten in der Filiale?

15.1.3 Stand der Forschung

Empirische Untersuchungen basieren auf der Umweltpsychologie. Das umweltpsychologische Grundmodell folgt einem S-O-R-Ansatz (Stimulus-Organismus-Reaktion), also einem modifizierten behavioristischen Ansatz durch die Integration von Organismus-Variablen.

15.1.4 Wirkungsmodell

In Abb. 15.1 ist ein hypothetisches Wirkungsmodell (simultanes Gleichungssystem) dargestellt, das im folgenden überprüft werden soll.

Die einzelnen Konstrukte sind

1. Wahrgenommene Filialgestaltung (FilialU)
2. Nutzung und Zielerreichung (NutzZielU)
3. Wahrgenommene Kontrolle (KontrollU)
4. Zufriedenheit mit dem Filialbesuch (ZufriedU)
5. Kundenloyalität (LoyalU)

Die in den Datensätzen benutzten Variablenamen sind zu Referenzzwecken schon in Klammer angegeben ('U' für umgepolt – hohe Werte = besser). Die Variable *Filiale* steht für das Filialkonzept (alt/neu) sowie die Filialgröße (klein/groß). Es handelt sich hierbei um Moderatorvariablen, die zur Gruppierung des Wirkungszusammenhangs dienen. Etwa lassen sich die Pfeile (Regressionskoeffizienten) in Abb. 15.1 in den Filialen mit altem/neuem Filialkonzept getrennt schätzen (vgl. Abb. 15.27).

Moderatorvariablen

Alternativ kann ein gemeinsames Modell mit Konstrukten und Moderatorvariablen (auch Geschlecht, Alter, Berufsgruppe etc.) formuliert werden (vgl. Abs. 15.7.5).

Die einzelnen Variablen lassen sich genauer charakterisieren:

15.1.4.1 Wahrgenommene Filialgestaltung

- Mensch und Umwelt stehen miteinander in Interaktion

Die Wahrnehmung der Umwelt beeinflusst das Verhalten

Das Verhalten in der Filiale wird durch die Filialgestaltung beeinflusst. Hierbei bestimmt die Wahrnehmung der Filiale das Verhalten in der Filiale.

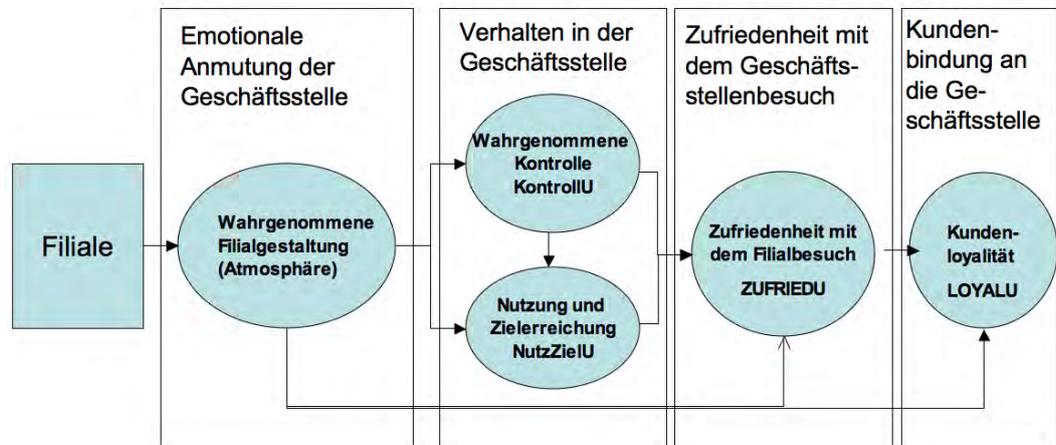


Abbildung 15.1: hypothetisches Wirkungsmodell.

- **Ambience:**
Wahrnehmung von Gestaltungsfaktoren wie Licht, Geräuschpegel, Temperatur u.ä.
- **Atmosphäre:**
Wahrgenommener Gesamteindruck der Umgebung.
- **Informationsrate:**
Wahrgenommene Informationsmenge der Umwelt.

15.1.4.2 Nutzung und Zielerreichung

- Besucher einer Filiale halten sich nicht zufällig in einer Filiale auf, sondern kommen mit einer bestimmten Absicht, z.B. Geld abheben, beraten lassen etc.
- Die Zufriedenheit hängt davon ab, ob es dem Besucher gelingt, seine Ziele zu erreichen.
- Hierbei kann ihn die Filialgestaltung unterstützen, z.B. indem sie eine einfache Orientierung ermöglicht.
- **Informationsrate:** Wahrgenommene Informationsmenge der Umwelt.

15.1.4.3 Wahrgenommene Kontrolle

- Theoretisches Konstrukt aus der Sozialpsychologie.
- Überzeugung, dass der Mensch seine Umwelt beeinflussen kann, auf eine Art und Weise, die darüber entscheidet, ob sich in der Folge Positives oder Negatives ergibt.
- 4 Formen der Kontrolle:
Verhaltenskontrolle, kognitive Kontrolle, Informationskontrolle, Retrospektive Kontrolle.
- Wahrgenommene Kontrolle entscheidet darüber, ob ein Besucher bei seinem Filialbesuch seine Ziele erreicht. Ein Besucher, der eine geringe wahrgenommene Kontrolle spürt, fühlt sich ausgeliefert oder dominiert und daher nicht wohl.
- Die Filialgestaltung beeinflusst die wahrgenommene Kontrolle maßgeblich.
- Wahrgenommene Kontrolle liegt vor, wenn der Kunde das Gefühl hat, dass er sich frei bewegen kann, dass er in seinen geplanten Verhaltensweisen nicht behindert oder eingeschränkt wird, dass die Umgebung ihn in seinen Verhaltensweisen unterstützt, wenn er sich gut orientieren kann.

15.1.4.4 Zufriedenheit mit dem Filialbesuch

- Kunden zufriedenzustellen, ist eine wichtige Aufgabe des Marketing, denn zufriedene Kunden kommen wieder und tätigen weitere Geschäfte mit der Bank.
- Verschiedene theoretische Ansätze, meist als Erfüllung von Erwartungen verstanden.
- Die Zufriedenheit mit dem Filialbesuch hängt in hohem Maße davon ab, ob der Kunde seine Ziele erreicht hat.
- Die Zufriedenheit mit dem Filialbesuch hängt in hohem Maße davon ab, ob der Kunde das Gefühl von Kontrolle hat.
- Die Zufriedenheit hat auch eine emotionale Komponente, d.h. die Zufriedenheit mit dem Filialbesuch wird auch von der Atmosphäre in der Filiale beeinflusst.

15.1.4.5 Kundenloyalität

- Loyale Kunden sind eine Zielgröße des Marketing, denn loyale Kunden senken die Kosten der Kundenneugewinnung und führen zu höheren Erlösen.
- Kundenloyalität beruht – im Gegensatz zu Kundenbindung – auf der freiwilligen Bindung von Kunden an ein Unternehmen.
- Gelingt es, durch die Filialgestaltung die Kunden an die Filiale und/oder die Bank zu binden, so besuchen sie die Filiale häufiger. Dies führt zu einer höheren Zahl von Kunden- und damit auch Werbekontakte, so dass die Bank leichter neue Produkte vermarkten kann.
- Ob ein Kunde loyal ist oder nicht, wird entscheidend von seiner Zufriedenheit beeinflusst.

15.2 Operationalisierung der Konstrukte

Die im vorigen Abschnitt beschriebenen Charakterisierungen der Begriffe wurden in Fragen übersetzt. Damit sind die Operationen festgelegt, mit denen Daten erhoben werden können.² Natürlich sind die Begriffe in Wahrheit reichhaltiger, als sich in den Likert-Skalierungen der Items ausdrückt.

15.2.1 Wahrgenommene Filialgestaltung

Im Folgenden möchten wir zunächst von Ihnen wissen, wie die Raumgestaltung auf Sie wirkt. Geben Sie mir bitte auf einer Notenskala von 1 bis 5 an, wie gut Ihnen die Gestaltung gefällt.

1 bedeutet 'sehr gut', 5 bedeutet 'sehr schlecht', 2,3 und 4 sind Abstufungen dazwischen.

1. Wie gefällt Ihnen die Gestaltung in dieser Geschäftsstelle insgesamt?

Sehr gut 1 2 3 4 5 Sehr schlecht

Ich lese Ihnen jetzt Paare von Eigenschaftswörtern vor, z.B. vertraut – unvertraut. Bitte sagen Sie mir auf einer Skala von 1 bis 5, wie die Raumgestaltung auf Sie wirkt.

1 steht für vertraut, 5 steht für unvertraut. 2, 3 und 4, geben die Abstufungen dazwischen an.

Die Raumgestaltung wirkt auf mich

- | | | |
|--------------------|-----------|-------------------|
| 9. Sehr freundlich | 1 2 3 4 5 | Sehr unfreundlich |
| 10. Sehr einladend | 1 2 3 4 5 | Sehr abstossend |
| 11. Sehr angenehm | 1 2 3 4 5 | Sehr unangenehm |
| 12. Sehr ruhig | 1 2 3 4 5 | Sehr hektisch |

²In einer Extremform des Operationalismus (Bridgeman) kann man sagen: Intelligenz ist, was der Intelligenztest mißt.

15.2.2 Wahrgenommene Kontrolle

Nun lese ich Ihnen mehrere Aussagen vor und Sie sagen mir auf einer Skala von 1 bis 5, ob Sie dieser Aussage zustimmen oder nicht.

1 bedeutet 'stimme voll und ganz zu', 5 bedeutet 'stimme ganz und gar nicht zu', 2, 3 und 4 sind Abstufungen dazwischen.

2. Die Umgebung lässt eine sehr gute Orientierung zu.

Stimme ich voll und ganz zu 1 2 3 4 5 Ich stimme ganz und gar nicht zu

4. Jeder Kunde kann sich hier ungezwungen bewegen.

Stimme ich voll und ganz zu 1 2 3 4 5 Ich stimme ganz und gar nicht zu

6. Die Räumlichkeiten bieten eine angemessene Privatsphäre.

Stimme ich voll und ganz zu 1 2 3 4 5 Ich stimme ganz und gar nicht zu

15.2.3 Nutzung und Zielerreichung

3. Die Raumgestaltung unterstützt mich bei der Erledigung meiner Bankgeschäfte.

Stimme ich voll und ganz zu 1 2 3 4 5 Ich stimme ganz und gar nicht zu

5. Wechselnde Ansprechpartner stören mich nicht.

Stimme ich voll und ganz zu 1 2 3 4 5 Ich stimme ganz und gar nicht zu

7. Wenn ich Beratung in Anspruch nehme, hat der Berater genügend Zeit für mich.

Stimme ich voll und ganz zu 1 2 3 4 5 Ich stimme ganz und gar nicht zu

8. In dieser Geschäftsstelle schaffe ich es immer, meine Bankgeschäfte so zu erledigen, wie ich es mir vorstelle.

Stimme ich voll und ganz zu 1 2 3 4 5 Ich stimme ganz und gar nicht zu

15.2.4 Zufriedenheit

Bei den nächsten Fragen geht es um Ihre Zufriedenheit mit dieser Geschäftsstelle. Bitte beantworten Sie wieder auf einer Notenskala von 1 bis 5, wie zufrieden Sie sind.

1 bedeutet 'sehr zufrieden', 5 bedeutet 'gar nicht zufrieden', 2, 3 und 4 sind Abstufungen dazwischen.

13. Wie zufrieden sind Sie insgesamt mit Ihren Besuchen in dieser Geschäftsstelle?

Sehr zufrieden 1 2 3 4 5 Gar nicht zufrieden

14. Wie zufrieden sind Sie mit den Wartezeiten?

Sehr zufrieden 1 2 3 4 5 Gar nicht zufrieden

16. Wie zufrieden sind Sie mit der Kundenbetreuung?

Sehr zufrieden 1 2 3 4 5 Gar nicht zufrieden

17. Wie zufrieden sind Sie mit den Öffnungszeiten?

Sehr zufrieden 1 2 3 4 5 Gar nicht zufrieden

15.2.5 Loyalität

18. Wie häufig kommen Sie in diese Geschäftsstelle?

Sehr häufig 1 2 3 4 5 Sehr selten

19. Wie gerne kommen Sie in diese Geschäftsstelle?

Sehr gern 1 2 3 4 5 Sehr ungern

20. Würden Sie diese Geschäftsstelle an Freunde und Bekannte weiterempfehlen?

Ja, ganz sicher 1 2 3 4 5 Nein, ganz sicher nicht

21. Wie verbunden fühlen Sie sich dieser Geschäftsstelle?

Sehr verbunden 1 2 3 4 5 Sehr wenig verbunden

15.2.6 Soziodemographische Fragen

22. Sind Sie heute zum ersten Mal hier?

ja/nein

23. In welchem Jahr sind Sie geboren?

24. Welcher Berufsgruppe gehören Sie an?

Schüler/Student/Azubi/Zivi

Arbeiter/Angestellter/Beamter

Hausfrau/-mann

Selbstständig

Rentner/Pensionär

ohne Beschäftigung

25. Geschlecht

m w

15.2.7 Objektive Filialgestaltung

26. Alt/Neu

Alt 0 Neu 1

27. Klein/Groß

Klein 0 Groß 1

15.3 Durchführung der Untersuchung

Die Befragung wurde an insgesamt zwei Tagen (30. und 31.05.08) an 4 Geschäftsstellen einer Bank durchgeführt. Die Filialen wurden im Sinne eines faktoriellen Designs (vgl. Varianzanalyse) so ausgewählt, daß es sich um die 4 Kombinationen der Ausprägungen (alt/neu) und (klein/groß) handelt (vgl. Fragebogen, S. 2, unten).

Am 30.05.08 wurde ein Pre-Test vorgenommen, um die Komplexität des Fragebogens, sowie die zeitliche Inanspruchnahme zu überprüfen. Nach diesem Pre-Test wurde der Fragebogen geringfügig angepaßt.

Die Hauptuntersuchung fand am 31.05.2008 in der Zeit zwischen 9 und

Interviewer/Fragebogen-Nummer

--

Im Folgenden möchten wir zunächst von Ihnen wissen, wie die Raumgestaltung auf Sie wirkt. Geben Sie mir bitte auf einer Notenskala von 1 bis 5 an, wie gut Ihnen die Gestaltung gefällt. 1 bedeutet „sehr gut“, 5 bedeutet „sehr schlecht“, 2,3 und 4 sind Abstufung dazwischen.

1. Wie gefällt Ihnen die Gestaltung in dieser Geschäftsstelle insgesamt?

Sehr gut

1	2	3	4	5
---	---	---	---	---

 Sehr schlecht

Nun lese ich Ihnen mehrere Aussagen vor und Sie sagen mir auf einer Skala von 1 bis 5, ob Sie dieser Aussage zustimmen oder nicht. 1 bedeutet „stimme voll und ganz zu“, 5 bedeutet „stimme ganz und gar nicht zu“, 2,3 und 4 sind Abstufungen dazwischen.

2. Die Umgebung lässt eine sehr gute Orientierung zu.

Stimme ich voll und ganz zu

1	2	3	4	5
---	---	---	---	---

 Ich stimme ganz und gar nicht zu

3. Die Raumgestaltung unterstützt mich bei der Erledigung meiner Bankgeschäfte.

Stimme ich voll und ganz zu

1	2	3	4	5
---	---	---	---	---

 Ich stimme ganz und gar nicht zu

4. Jeder Kunde kann sich hier ungezwungen bewegen.

Stimme ich voll und ganz zu

1	2	3	4	5
---	---	---	---	---

 Ich stimme ganz und gar nicht zu

5. Wechselnde Ansprechpartner stören mich nicht.

Stimme ich voll und ganz zu

1	2	3	4	5
---	---	---	---	---

 Ich stimme ganz und gar nicht zu

6. Die Räumlichkeiten bieten eine angemessene Privatsphäre.

Stimme ich voll und ganz zu

1	2	3	4	5
---	---	---	---	---

 Ich stimme ganz und gar nicht zu

7. Wenn ich Beratung in Anspruch nehme, hat der Berater genügend Zeit für mich.

Stimme ich voll und ganz zu

1	2	3	4	5
---	---	---	---	---

 Ich stimme ganz und gar nicht zu

8. In dieser Geschäftsstelle schaffe ich es immer, meine Bankgeschäfte so zu erledigen, wie ich es mir vorstelle.

Stimme ich voll und ganz zu

1	2	3	4	5
---	---	---	---	---

 Ich stimme ganz und gar nicht zu

Ich lese Ihnen jetzt Paare von Eigenschaftswörtern vor, z.B. vertraut – unvertraut. Bitte sagen Sie mir auf einer Skala von 1 bis 5, wie die Raumgestaltung auf Sie wirkt. 1 steht für vertraut, 5 steht für unvertraut. 2, 3 und 4, geben die Abstufungen dazwischen an.

Die Raumgestaltung wirkt auf mich

9.	Sehr freundlich	<table border="1" style="display: inline-table;"><tr><td style="width: 20px; height: 20px;">1</td><td style="width: 20px; height: 20px;">2</td><td style="width: 20px; height: 20px;">3</td><td style="width: 20px; height: 20px;">4</td><td style="width: 20px; height: 20px;">5</td></tr></table>	1	2	3	4	5	Sehr unfreundlich
1	2	3	4	5				
10.	Sehr einladend	<table border="1" style="display: inline-table;"><tr><td style="width: 20px; height: 20px;">1</td><td style="width: 20px; height: 20px;">2</td><td style="width: 20px; height: 20px;">3</td><td style="width: 20px; height: 20px;">4</td><td style="width: 20px; height: 20px;">5</td></tr></table>	1	2	3	4	5	Sehr abstoßend
1	2	3	4	5				
11.	Sehr angenehm	<table border="1" style="display: inline-table;"><tr><td style="width: 20px; height: 20px;">1</td><td style="width: 20px; height: 20px;">2</td><td style="width: 20px; height: 20px;">3</td><td style="width: 20px; height: 20px;">4</td><td style="width: 20px; height: 20px;">5</td></tr></table>	1	2	3	4	5	Sehr unangenehm
1	2	3	4	5				
12.	Sehr ruhig	<table border="1" style="display: inline-table;"><tr><td style="width: 20px; height: 20px;">1</td><td style="width: 20px; height: 20px;">2</td><td style="width: 20px; height: 20px;">3</td><td style="width: 20px; height: 20px;">4</td><td style="width: 20px; height: 20px;">5</td></tr></table>	1	2	3	4	5	Sehr hektisch
1	2	3	4	5				

Bei den nächsten Fragen geht es um Ihre Zufriedenheit mit dieser Geschäftsstelle. Bitte beantworten Sie wieder auf einer Notenskala von 1 bis 5, wie zufrieden Sie sind. 1 bedeutet „sehr zufrieden“, 5 bedeutet „gar nicht zufrieden“, 2, 3 und 4 sind Abstufungen dazwischen.

13. Wie zufrieden sind Sie insgesamt mit Ihren Besuchen in dieser Geschäftsstelle?

Sehr zufrieden

1	2	3	4	5
---	---	---	---	---

 Gar nicht zufrieden

Abbildung 15.2: Fragebogen Seite 1.

14. Wie zufrieden sind Sie mit den Wartezeiten? 15. _____

Sehr zufrieden

1	2	3	4	5
---	---	---	---	---

 Gar nicht zufrieden

16. Wie zufrieden sind Sie mit der Kundenbetreuung?

Sehr zufrieden

1	2	3	4	5
---	---	---	---	---

 Gar nicht zufrieden

17. Wie zufrieden sind Sie mit den Öffnungszeiten?

Sehr zufrieden

1	2	3	4	5
---	---	---	---	---

 Gar nicht zufrieden

18. Wie häufig kommen Sie in diese Geschäftsstelle?

Sehr häufig

1	2	3	4	5
---	---	---	---	---

 Sehr selten

19. Wie gerne kommen Sie in diese Geschäftsstelle?

Sehr gern

1	2	3	4	5
---	---	---	---	---

 Sehr ungern

20. Würden Sie diese Geschäftsstelle an Freunde und Bekannte weiterempfehlen?

Ja, ganz sicher

1	2	3	4	5
---	---	---	---	---

 Nein, ganz sicher nicht

21. Wie verbunden fühlen Sie sich dieser Geschäftsstelle?

Sehr verbunden

1	2	3	4	5
---	---	---	---	---

 Sehr wenig verbunden

22. Sind Sie heute zum ersten Mal hier?

ja nein

23. In welchem Jahr sind Sie geboren?

.....

24. Welcher Berufsgruppe gehören Sie an?

Schüler/Student/Azubi/Zivi Arbeiter/Angestellter/Beamter Hausfrau/-mann
 Selbstständig Rentner/Pensionär ohne Beschäftigung

25. Geschlecht

m w

Filiale	Groß 1	klein 0
Alt 0	Filiale 1	Filiale 2
Neu 1	Filiale 3	Filiale 4

Beobachtungsbogen Nr. _____

Abbildung 15.3: Fragebogen Seite 2.

16 Uhr statt. Es wurden insgesamt $N = 378$ Personen befragt (vgl. Abb. 15.8).

Die Zuordnung der Probanden zu den Geschäftstellen ergab sich rein faktisch, d.h. im Rahmen einer Survey-Studie konnte keine zufällige (randomisierte) Zuweisung auf die 4 Filialtypen (alt/neu) \times (klein/groß) erfolgen. Dies wäre nur in einem geplanten Experiment möglich. Daher sind die Stichproben in den Teilpopulationen sicherlich unterschiedlich strukturiert (vgl. Abb. 15.9).

15.4 Variablen in SPSS

Die items des Fragebogens wurden in der Reihenfolge pro Person zeilenweise in einen SPSS-Datensatz eingegeben. Dazu öffnet man ein neues Daten-Fenster mit

SPSS/Datei/Neu/Daten

Die Variablennamen sind kurz und prägnant – etwa steht der erste Buchstabe für das Konstrukt (Abb. 15.4). Dabei wurden dichotome Variablen mit 0-1 codiert, damit Mittelwerte etc. interpretierbar sind (Kap. 14). Die ursprünglichen Variablen reichen bis Spalte 29, anschließend kommen neu berechnete Größen, die sich im Laufe der Analyse ergeben haben. Es ist empfehlenswert, die ursprüngliche Datei unverändert zu lassen, also mit einer Kopie zu arbeiten.

Überhaupt gelten die 3 Regeln des Computerwesens:

1. Sichern
2. Sichern
3. Sichern.

	Name	Typ	
1	FGESAMT	Numerisch	1
2	KORIENT	Numerisch	1
3	NRAUM	Numerisch	1
4	KZWANG	Numerisch	1
5	NANSPREC	Numerisch	1
6	KPRIVAT	Numerisch	1
7	NZEIT	Numerisch	1
8	NERLEDIG	Numerisch	1
9	FFREUNDL	Numerisch	1
10	FEINLADE	Numerisch	1
11	FAGENEHM	Numerisch	1
12	FRUHIG	Numerisch	1
13	ZGESAMT	Numerisch	1
14	ZWARTE	Numerisch	1
15	ZK01	Numerisch	1
16	ZBETREU	Numerisch	1
17	ZOEFFNUN	Numerisch	1
18	LHAEUFIG	Numerisch	1
19	LGERNE	Numerisch	1
20	LEMPFEHL	Numerisch	1
21	LVERBUND	Numerisch	1
22	NEU	Numerisch	1
23	JAHR	Numerisch	1
24	BERUF	Numerisch	1
25	GESCHLECHT	Numerisch	1
26	ALTNEU	Numerisch	1
27	KLEINGR	Numerisch	1
28	INTERVIE	Numerisch	1
29	FRAGEBOG	Numerisch	1
30	FILIAL	Numerisch	1
31	KONTROLL	Numerisch	1
32	NUTZZIEL	Numerisch	1
33	ZUFRIED	Numerisch	1
34	LOYAL	Numerisch	1

Abbildung 15.4: Variablen (Spalten) des Datensatzes *main2.xls.sav*.

15.5 Überblick

Zunächst verschafft man sich einen Überblick der Daten, zuerst univariat, dann bivariat:

15.5.1 Univariate Analyse

Häufigkeitsdiagramme, Histogramme, einfache Statistiken

SPSS/Analysieren/Deskriptive Statistiken/Häufigkeiten
(Abb. 15.5).

Übersichtlicher sind Häufigkeitsdiagramme (Stabdiagramme) oder Histogramme (bei stetigen Variablen): Abb.: 15.6–15.7.

Sehr instruktiv zum Gruppenvergleich ist die Bevölkerungspyramide (Abb. 15.7)

**Bevölkerungs-
pyramide**

SPSS/Diagramme/veraltete Dialogfelder/Bevölkerungspyramide

Hier erkennt man eine Mittelwertsverschiebung (mehr jüngere Frauen in der Stichprobe). Dies kann mit einem t -Test überprüft werden (Übung). Sehr informativ sind die Diagramme von SAS/JMP (Abb. 15.8). Beim sog. Stem and Leaf-Diagramm bilden die Originaldaten ein Histogramm (John Tukey).

**Stem and
Leaf-Diagramm**

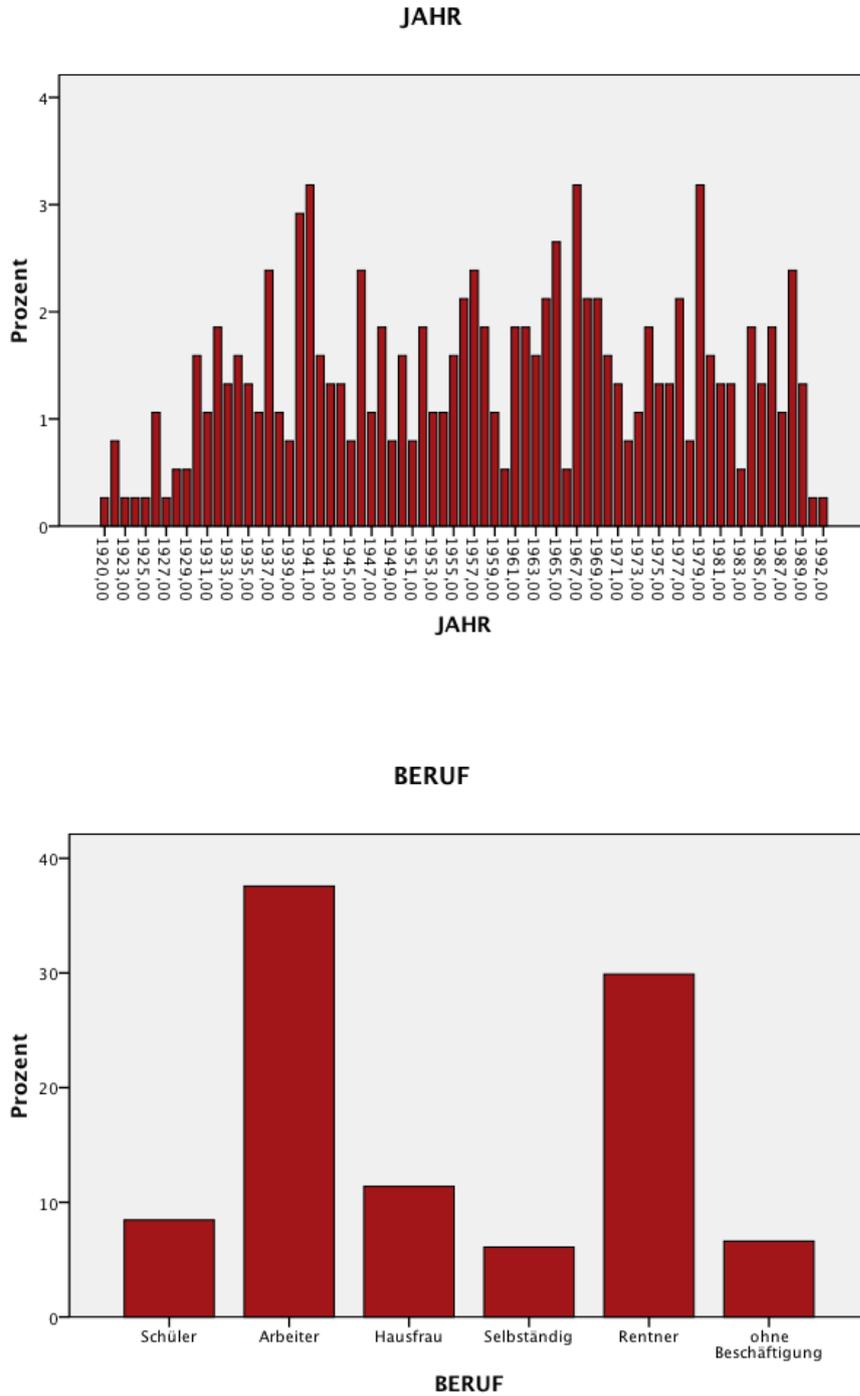


Abbildung 15.6: Häufigkeitsdiagramme (Stabdiagramme): Geburtsjahr und Beruf.

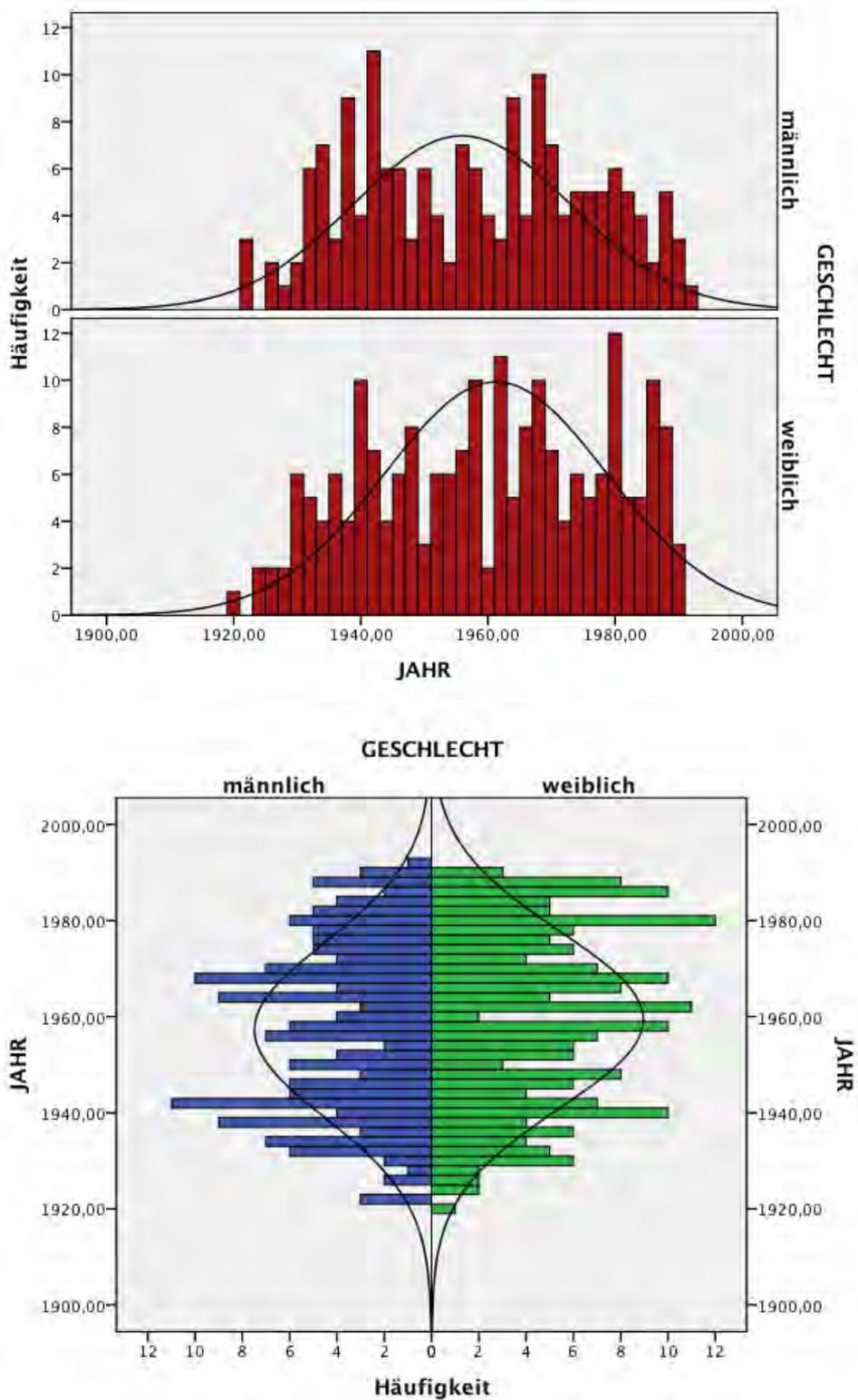


Abbildung 15.7: Histogramme Geburtsjahr (gruppiert nach Geschlecht und als Pyramide).

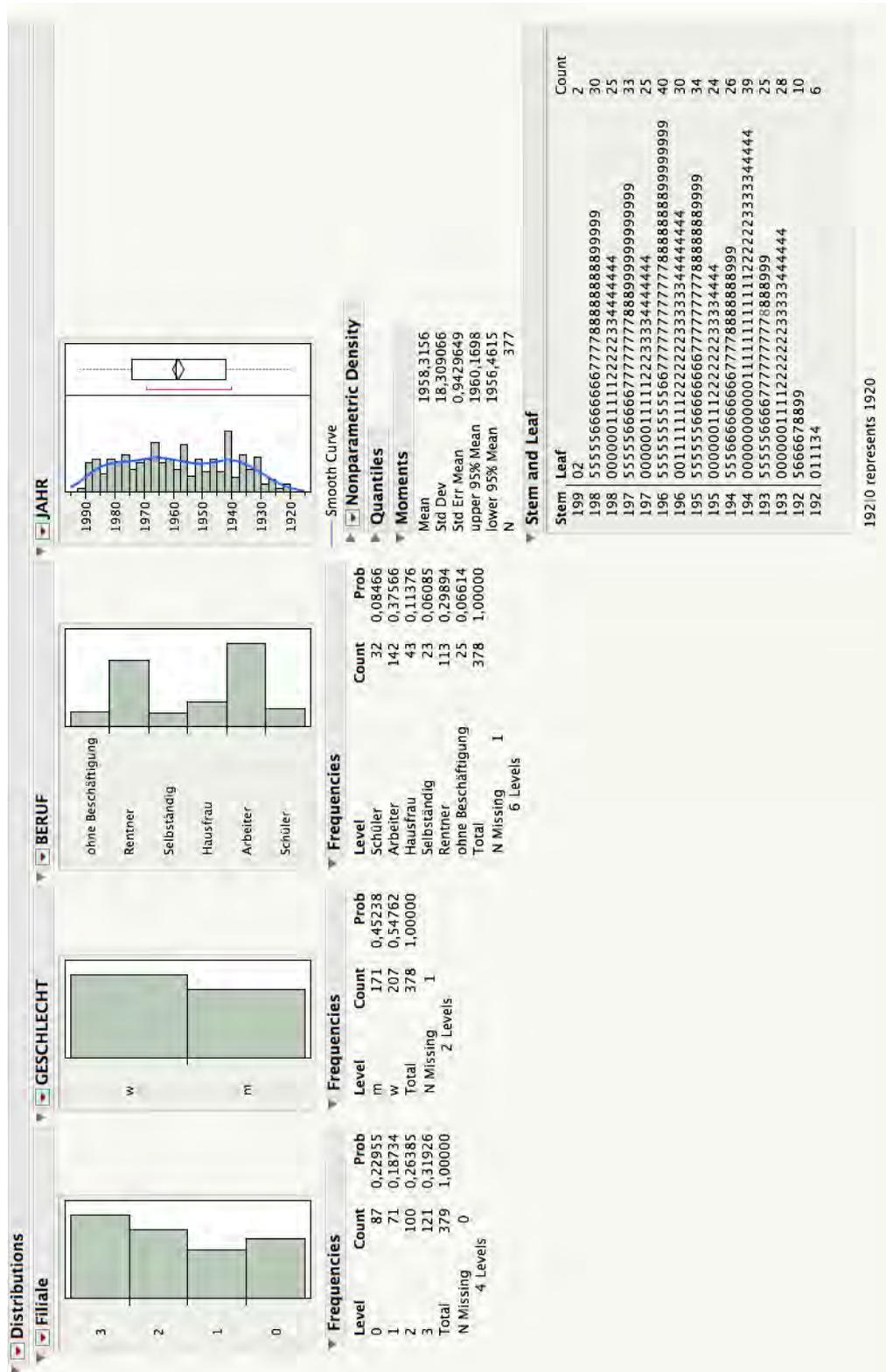


Abbildung 15.8: SAS/JMP: Verteilungen und Statistiken der Variablen Filiale, Geschlecht, Beruf und Geburtsjahr (auch als Stem and Leaf-Diagramm).

15.5.2 Bivariate Analyse: Kreuztabellen, Korrelationen, Streudiagramme

Zusammenhänge lassen sich je nach Skalenniveau in Kreuztabellen oder Korrelationstabellen darstellen. Etwa gibt der Mosaik-Plot in SAS/JMP eine graphische Darstellung der bedingten Häufigkeiten $f(\text{Beruf}|\text{Filiale})$ oder $f(\text{Filiale}|\text{Beruf})$ (Abb. 15.9).

Mosaik-Plot

Diese Darstellungen zeigen, daß sich die Filialen (0-3) in der Verteilung der Berufsgruppen unterscheiden, etwa sind in Filiale 2 (klein, neu) sehr viele Personen ohne Beschäftigung vertreten. Die Variable **Filiale** wurde aus den Indikator-Variablen **kleingross** (0/1) und **altneu** (0/1) durch die Formel

$$\text{Filiale} = 2 \times \text{altneu} + \text{kleingross}$$

(Wertebereich 0, . . . , 3) erzeugt.

Übung 15.1 (Neue Variable)

Berechnen Sie in SPSS oder JMP eine neue Variable aus den bestehenden Variablen **kleingross** (0/1) und **altneu** (0/1) nach obiger Formel oder durch Konditionale.



Die Graphiken und Statistiken lassen sich interaktiv verändern (rote Dreiecke). Am rechten Rand der Graphiken sind die Randverteilungen $f(\text{Beruf})$ und $f(\text{Filiale})$ aufgetragen. Zu beachten ist, daß die Bedingungen (Gruppen) spaltenweise aufgetragen sind (d.h. jede Spalte addiert sich zu 100%). Beispielsweise ist in Filiale 3 ein relativ hoher Schüleranteil, in Filiale 0 sind sehr viele Rentner (oberes Bild). Geht man in die Spalte Rentner (unteres Bild), so sieht man die große Häufigkeit von Filiale 3 in dieser Gruppe.

Dagegen sind Beschäftigungslose sind in Filiale 2 überproportional vertreten.

Interessant, aber zunächst etwas unübersichtlich ist eine Korrelationsmatrix aller Variablen (15.11).

Korrelationsmatrix

Es lohnt sich, davon einen möglichst großen Ausdruck zu erstellen und über alle möglichen Kombinationen zu meditieren.

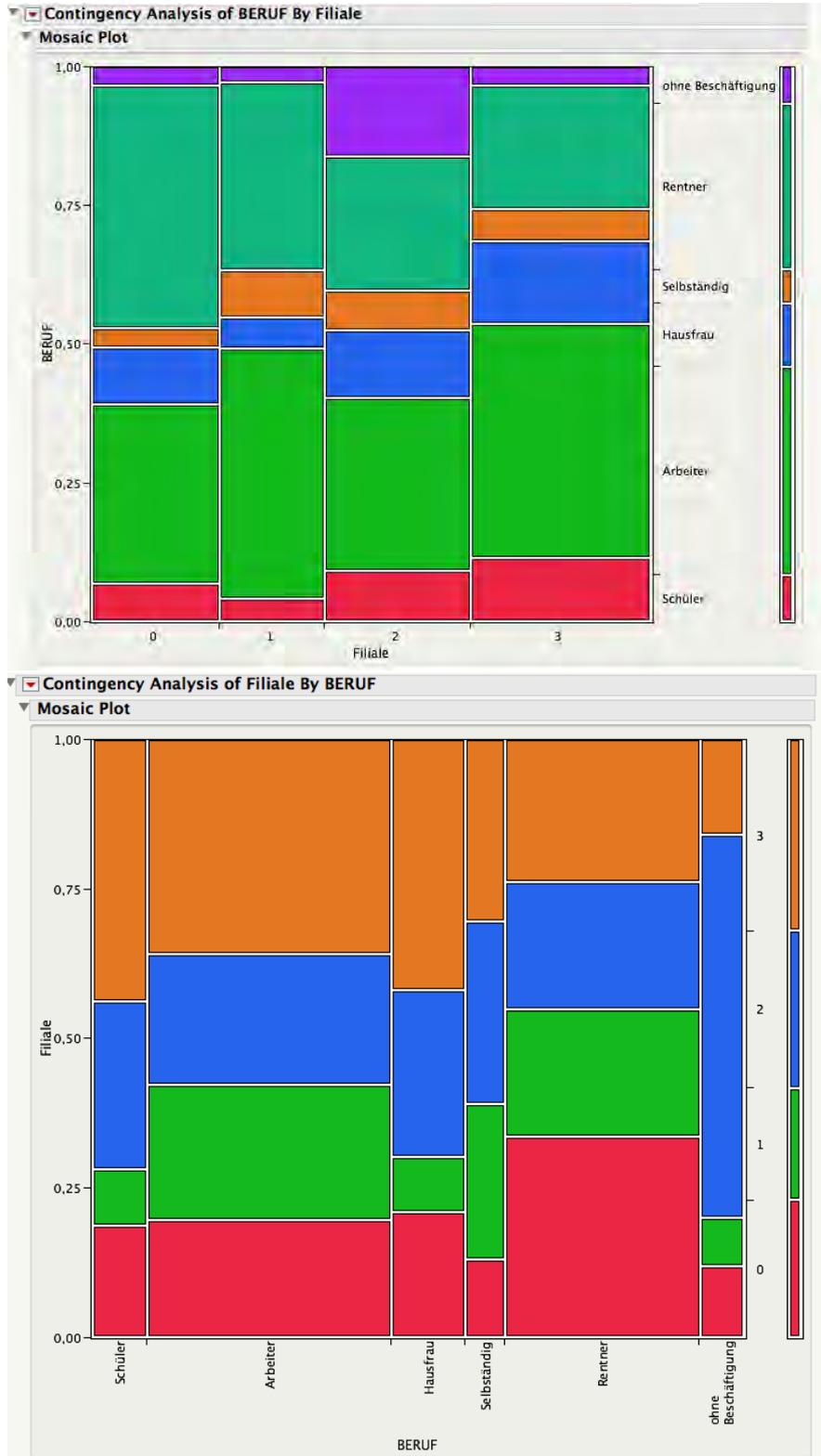


Abbildung 15.9: SAS/JMP: Mosaik-Plot (bedingte Häufigkeiten) der Variablen *Beruf|Filiale* (oben) und *Filiale|Beruf* (unten).

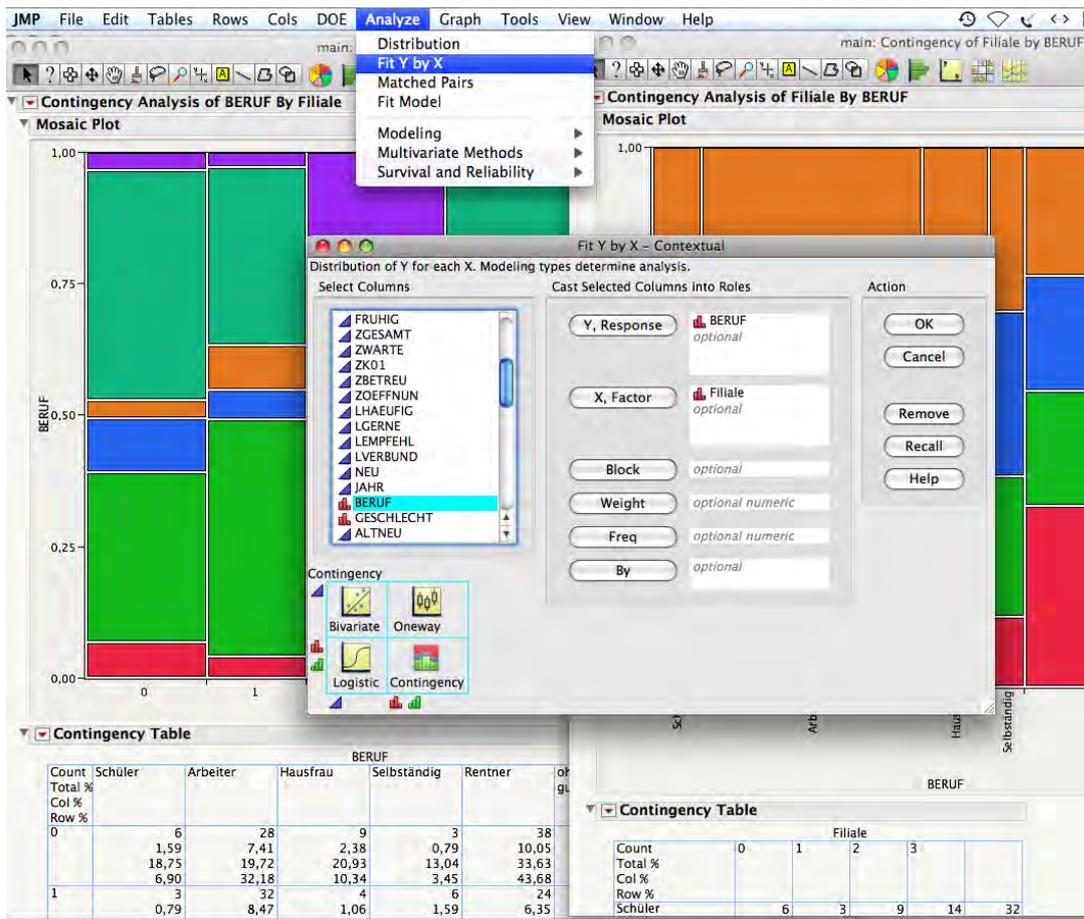


Abbildung 15.10: SAS/JMP: Mosaik-Plot: Dialog in JMP.

**Cluster der
Korrelationen**

Einerseits sieht man die Interkorrelationen der theoretisch erwogenen Konstrukte, andererseits die Zusammenhänge zwischen Konstrukten und anderen (etwa sozioökonomischen) Variablen. Hier gewinnt man erste Hypothesen über empirische Zusammenhänge.

Da die **dichotomen Variablen als Indikatoren codiert** wurden, lassen sich auch deren Korrelationen sinnvoll interpretieren. Etwas übersichtlicher ist eine farblich codierte Matrix, die nach Blöcken geordnet wurde (Cluster der Korrelationen; Abb. 15.12).

Da die Korrelation nur lineare Assoziationen mißt, sollte man auch Streudiagramme erstellen (Übung), die auch nichtlineare Zusammenhänge abbilden.

Ein Ausschnitt der Variablen ist in Abb. 15.13 gezeigt. Filialgestaltung (FGESAMT ; 1 = sehr gut, 5 = sehr schlecht) korreliert negativ mit ALTNEU (1 = neu), also werden neue Filialen als eher gut gestaltet bewertet. Dagegen ist die Korrelation mit Zufriedenheit (ZGESAMT) positiv, sodaß die Kunden in neuen Filialen tendenziell unzufriedener sind. Zufriedenheit und Filialgestaltung korrelieren dagegen positiv.

Ist dies ein Widerspruch?

Teilt man die Stichprobe in die Gruppen (alt/neu) ein, so ist die Korrelation zwischen Filialgestaltung und Zufriedenheit in beiden Gruppen positiv, jedoch sind die Mittelwerte für Zufriedenheit in Richtung schlechterer (höherer) Werte (= 1.93) in den neuen Filialen verschoben (Abb. 15.14). Umgekehrt sind die Mittelwerte für Filialgestaltung in Richtung besserer (kleinerer) Werte (= 2.04) in den neuen Filialen verschoben.

Dieser Effekt wird besonders deutlich in Abb. 15.15. Obwohl beide Linien einen positiven Zusammenhang angeben, ist in den alten Filialen die Zufriedenheit höher (kleinere Werte; rote Linie).

Eine etwas andere Darstellung (Abb. 15.16) vermittelt den Sachverhalt nochmals:

In den alten Filialen klaffen die Mittelwerte für FGESAMT, ZGESAMT weit auseinander, während sie in den neuen Filialen näher beisammen (dazwischen) liegen. FGESAMT ist besser, ZGESAMT ist deutlich schlechter.

Correlations

	Natrium	Kalium	Magnesium	Kalzium	Chlorid	Sulfat	Hydrogen-Karbonat	Nitrat	Nitrit	Mangan	Anionen Fluorid	Gesamt-Mineralien in mg/Liter
Natrium	1	,837**	,572	-,063	,334	-,097	,873**	-,434	.	,864	-,814	,909**
		,003	,084	,862	,346	,790	,001	,210	.	,337	,394	,000
	10	10	10	10	10	10	10	10	0	3	3	10
Kalium	,837**	1	,522	-,071	,243	,204	,737*	-,416	.	,160	-,070	,785**
	,003	,10	,121	,846	,499	,572	,015	,232	.	,898	,956	,007
	10	10	10	10	10	10	10	10	0	3	3	10
Magnesium	,572	,522	1	,529	,438	,366	,302	,110	.	,897	-,853	,525
	,084	,121	,10	,116	,206	,299	,397	,763	.	,291	,349	,119
	10	10	10	10	10	10	10	10	0	3	3	10
Kalzium	-,063	-,071	,529	1	,238	,490	,011	,145	.	-,971	,946	,204
	,862	,846	,116	,507	,507	,151	,977	,690	.	,153	,211	,571
	10	10	10	10	10	10	10	10	0	3	3	10
Chlorid	,334	,243	,438	,238	1	,674*	,083	-,382	.	,976	-,952	,376
	,346	,499	,206	,507	,820	,033	,820	,276	.	,141	,199	,284
	10	10	10	10	10	10	10	10	0	3	3	10
Sulfat	,097	,204	,366	,490	,674*	1	,031	-,524	.	-,621	,547	,332
	,790	,572	,299	,151	,033	,10	,932	,120	.	,574	,632	,348
	10	10	10	10	10	10	10	10	0	3	3	10
Hydrogen-Karbonat	,873**	,737**	,302	,011	,083	,031	1	-,558	.	,481	-,400	,936**
	,001	,015	,397	,977	,820	,932	,094	,094	.	,680	,738	,000
	10	10	10	10	10	10	10	10	0	3	3	10
Nitrat	-,434	-,416	,110	,145	-,382	-,524	-,558	1	.	-,734	,792	-,605
	,210	,232	,763	,690	,276	,120	,094	,094	.	,476	,418	,064
	10	10	10	10	10	10	10	10	0	3	3	10
Nitrit

	0	0	0	0	0	0	0	0	0	0	0	0
Mangan	,864	,160	,897	-,971	,976	-,621	,481	-,734	.	1	-,996	,525
	,337	,898	,291	,153	,141	,574	,680	,476	.	.	,058	,648
	3	3	3	3	3	3	3	3	0	3	3	3
Anionen Fluorid	-,814	-,070	-,853	,946	-,952	,547	-,400	,792	.	-,996	1	-,445
	,394	,956	,349	,211	,199	,632	,738	,418	.	,058	.	,706
	3	3	3	3	3	3	3	3	0	3	3	3
Gesamt-Mineralien in mg/Liter	,909**	,785**	,525	,204	,376	,332	,936**	-,605	.	,525	-,445	1
	,000	,007	,119	,571	,284	,348	,000	,064	.	,648	,706	,000
	10	10	10	10	10	10	10	10	0	3	3	10

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

a. Cannot be computed because at least one of the variables is constant.

Abbildung 15.11: SPSS: Korrelationsmatrix aller Variablen (Ausschnitt).

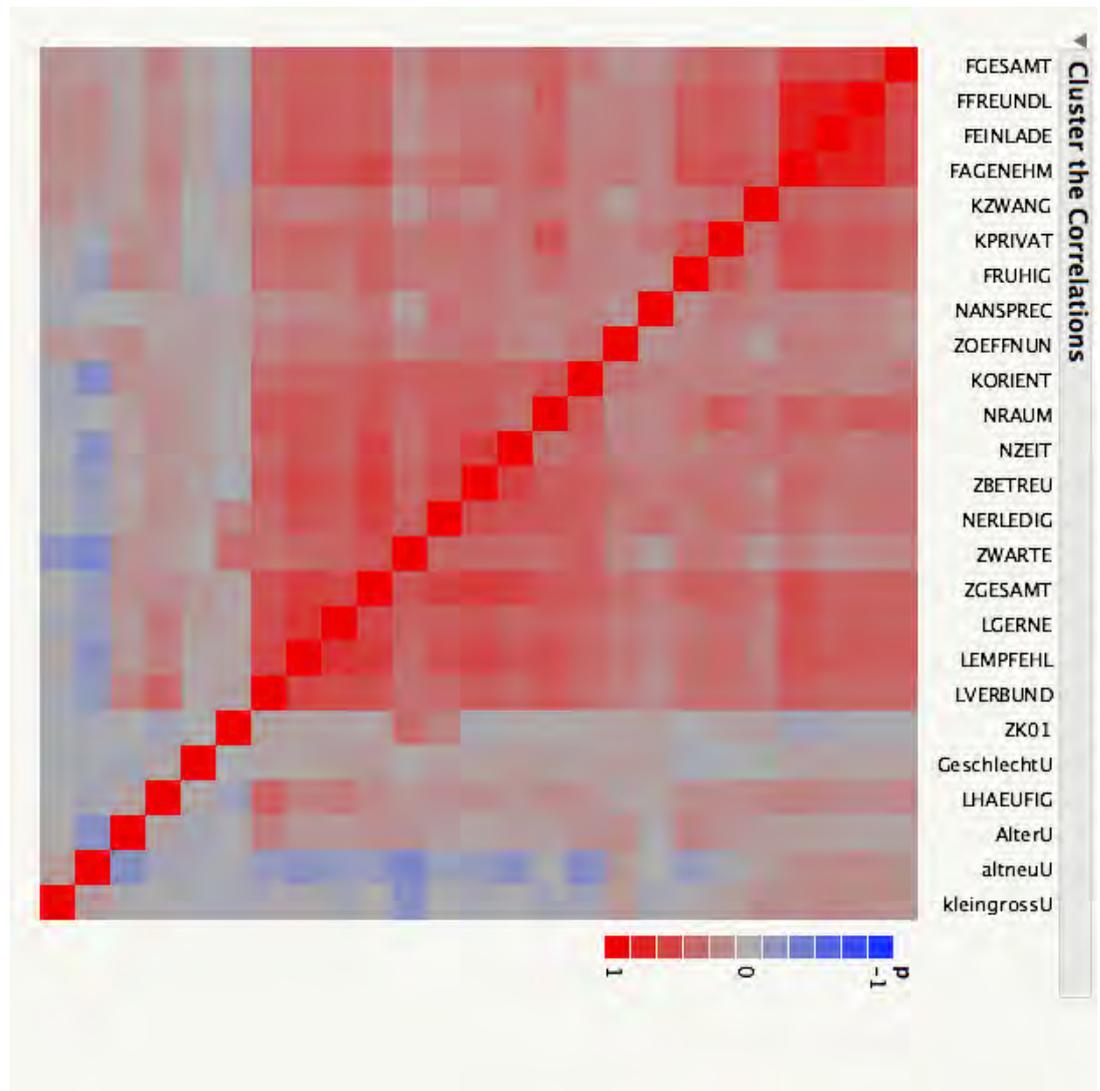


Abbildung 15.12: JMP: Korrelationsmatrix aller Variablen (Cluster der Korrelationen). Die Stärke der Korrelation ist durch die Farbe markiert (rot: $r > 0$, blau: $r < 0$). Die items der Konstrukte bilden einen positiv korrelierten Block.

Deskriptive Statistiken

	Mittelwert	Standardabweichung	N
FGESAMT	2,1319	,89295	379
ZGESAMT	1,8391	,80380	373
ALTNEU	,5831	,49370	379
KLEINGR	,5066	,50062	379
GESCHLECHT	,5476	,49839	378

Korrelationen

		FGESAMT	ZGESAMT	ALTNEU	KLEINGR	GESCHLECHT
FGESAMT	Korrelation nach Pearson	1	,450**	-,121*	-,144**	-,068
	Signifikanz (2-seitig)		,000	,018	,005	,189
	N	379	373	379	379	378
ZGESAMT	Korrelation nach Pearson	,450**	1	,142**	,000	-,115*
	Signifikanz (2-seitig)	,000		,006	,992	,027
	N	373	373	373	373	372
ALTNEU	Korrelation nach Pearson	-,121*	,142**	1	,097	-,033
	Signifikanz (2-seitig)	,018	,006		,060	,527
	N	379	373	379	379	378
KLEINGR	Korrelation nach Pearson	-,144**	,000	,097	1	-,033
	Signifikanz (2-seitig)	,005	,992	,060		,517
	N	379	373	379	379	378
GESCHLECHT	Korrelation nach Pearson	-,068	-,115*	-,033	-,033	1
	Signifikanz (2-seitig)	,189	,027	,527	,517	
	N	378	372	378	378	378

** . Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

* . Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Abbildung 15.13: SPSS: Korrelationsmatrix der Variablen FGESAMT, ZGESAMT, ALTNEU, KLEINGROSS und GESCHLECHT.

Deskriptive Statistiken

ALTNEU		Mittelwert	Standardabweichung	N
,00	FGESAMT	2,2595	,75865	158
	ZGESAMT	1,7051	,63467	156
1,00	FGESAMT	2,0407	,96915	221
	ZGESAMT	1,9355	,89519	217

Korrelationen

ALTNEU			FGESAMT	ZGESAMT
,00	FGESAMT	Korrelation nach Pearson	1	,368**
		Signifikanz (2-seitig)		,000
		N	158	156
	ZGESAMT	Korrelation nach Pearson	,368**	1
		Signifikanz (2-seitig)	,000	
		N	156	156
1,00	FGESAMT	Korrelation nach Pearson	1	,517**
		Signifikanz (2-seitig)		,000
		N	221	217
	ZGESAMT	Korrelation nach Pearson	,517**	1
		Signifikanz (2-seitig)	,000	
		N	217	217

** . Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Abbildung 15.14: SPSS: Korrelationsmatrix der Variablen FGESAMT, ZGESAMT gruppiert nach ALTNEU.

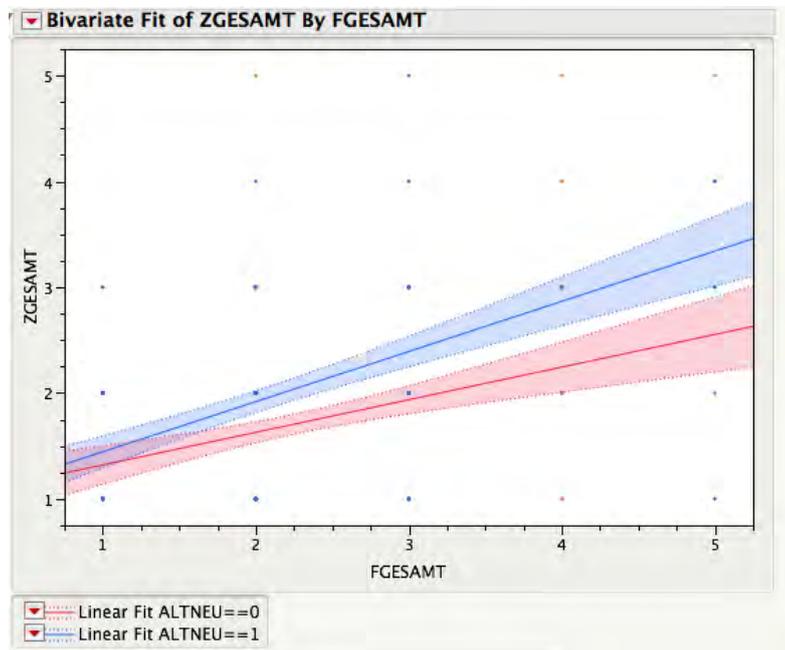


Abbildung 15.15: JMP: Regressionslinie und Konfidenzintervalle der Variablen FGESAMT, ZGESAMT gruppiert nach ALTNEU.

Die bisherige Analyse legt nahe, die items so umzucodieren, daß hohe numerische Werte auch hohe Konstruktwerte symbolisieren. Außerdem sollten zuerst die Konstrukte analysiert werden (item-Analyse), damit Summenwerte oder Faktorwerte berechnet werden können.

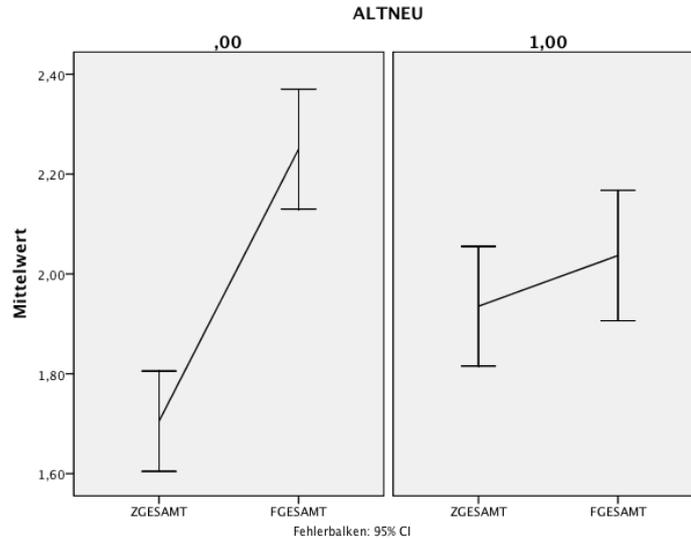


Abbildung 15.16: SPSS: Mittelwerte und Konfidenzintervalle der Variablen FGESAMT, ZGESAMT gruppiert nach ALTNEU.

15.6 Item-Analyse

Die theoretisch abgeleiteten Konstrukte werden nun analysiert (vgl. Kap. 14.10.2).

15.6.1 Wahrgenommene Kontrolle

Die 3 Items für Wahrgenommene Kontrolle interkorrelieren nur gering, so daß auch nach der Spearman-Brown-Korrektur nur ein α von ca. 0.55 erreichbar ist. Die Hauptkomponentenanalyse ergibt *einen* Faktor (Kaiser-Kriterium: $\lambda_1 > 1$). Siehe Abb. 15.17.

Übung 15.2 (Wahrgenommene Kontrolle)

Berechnen Sie bitte die

- Item-Schwierigkeit: $0 \leq p_i = \frac{\bar{X}_i - X_{i,\min}}{X_{i,\max} - X_{i,\min}} \leq 1$
- Trennschärfe: $\text{Corr}(X_i, X)$
- Homogenität: $\bar{\rho} = \frac{1}{k(k-1)} \sum_{i \neq j} \rho_{ij}$
- itemspezifische Homogenität $\bar{\rho}_{i+} = \frac{1}{k-1} \sum_{j, i \neq j} \rho_{ij}$

Reliabilitätsstatistiken

Cronbachs Alpha	Cronbachs Alpha für standardisierte Items	Anzahl der Items
,545	,552	3

Itemstatistiken

	Mittelwert	Standardabweichung	N
KORIENT	2,0296	,91214	372
KZWANG	1,7769	,92652	372
KPRIVAT	2,6290	1,22728	372

Inter-Item-Korrelationsmatrix

	KORIENT	KZWANG	KPRIVAT
KORIENT	1,000	,260	,258
KZWANG	,260	1,000	,356
KPRIVAT	,258	,356	1,000

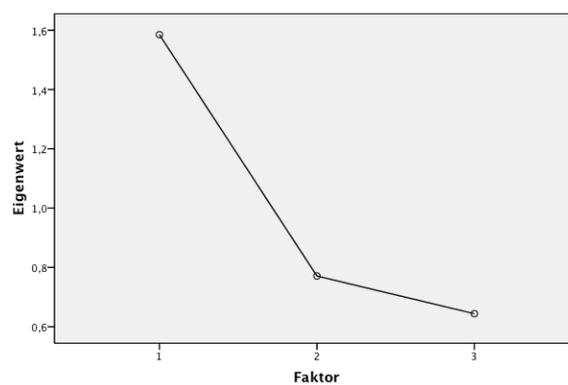
Screepplot

Abbildung 15.17: *Wahrgenommene Kontrolle*: Cronbachs alpha (oben), scree plot (unten).

der Skala und interpretieren Sie die Werte.

Führen Sie bitte eine Hauptkomponentenanalyse durch und schreiben Sie den 1. Faktor in die SPSS-Datei. Vergleichen Sie mit der bestehenden Variable KONTROLL.



15.6.2 Nutzung und Zielerreichung

Die 4 Items für das Konstrukt Nutzung und Zielerreichung interkorrelieren etwas höher, sodaß ein α von ca. 0.65 erreicht wird. Die Hauptkomponentenanalyse ergibt *einen Faktor* (Kaiser-Kriterium: $\lambda_1 > 1$). Siehe Abb. 15.18.

Übung 15.3 (Nutzung und Zielerreichung)

Berechnen Sie bitte die

- Item-Schwierigkeit
- Trennschärfe
- Homogenität
- itemspezifische Homogenität

der Skala.

Führen Sie eine Hauptkomponentenanalyse durch und schreiben Sie den 1. Faktor in die SPSS-Datei. Vergleichen Sie mit der bestehenden Variable NUTZZIEL.

Berechnen Sie einen Summenscore und vergleichen Sie diesen mit NUTZZIEL.



15.6.3 Wahrgenommene Kontrolle + Nutzung und Zielerreichung

Die Items für Wahrgenommene Kontrolle und für das Konstrukt Nutzung und Zielerreichung wurden nun zusammen analysiert, da zwischen den items inhaltliche Überschneidungen bestehen.

Die Reliabilität ist nun wesentlich höher (7 items) und es findet sich nur *ein Faktor* ($\lambda > 1$) (Abb. 15.19).

Reliabilitätsstatistiken

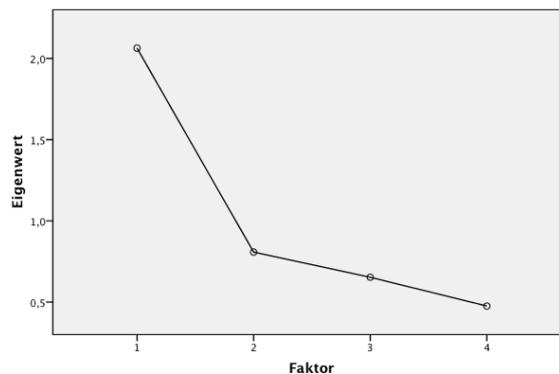
Cronbachs Alpha	Cronbachs Alpha für standardisierte Items	Anzahl der Items
,669	,681	4

Itemstatistiken

	Mittelwert	Standardabweichung	N
NRAUM	2,1994	1,01576	336
NANSPREC	2,0446	1,20735	336
NZEIT	1,7262	,95388	336
NERLEDIG	1,8036	,98804	336

Inter-Item-Korrelationsmatrix

	NRAUM	NANSPREC	NZEIT	NERLEDIG
NRAUM	1,000	,236	,352	,441
NANSPREC	,236	1,000	,231	,333
NZEIT	,352	,231	1,000	,494
NERLEDIG	,441	,333	,494	1,000

Screepplot**Koeffizientenmatrix der Komponentenwerte**

	Komponente...
	1
NRAUM	,346
NANSPREC	,279
NERLEDIG	,397
NZEIT	,360

Extraktionsmethode:
Hauptkomponentenanalyse.
Komponentenwerte.

Abbildung 15.18: *Nutzung und Zielerreichung*: Cronbachs alpha (oben), scree plot (mitte) und Gewichte der 1. Komponente.

Extrahiert man 2 Faktoren, so ergibt sich das interessante Resultat, daß eine Mischung der beiden Konstrukte stattfindet: es ergibt sich die Gruppe (NANSPREC KZWANG KPRIVAT) und (NRAUM NERLEDIG NZEIT KORIENT) (Abb. 15.20).

Eine schiefwinklige Rotation der Faktoren ($\Phi \neq I$) ist im Bild 15.20 unten gezeigt. Die Item-Gruppen liegen nun auf den Achsen der nichtorthogonalen Faktoren ZIELOBLI und KONTROLOBLI.

Deren Korrelation ist $r = 0.484$ (Abb. 15.21) , während die Konstrukte NUTZZIEL und KONTROLL zu .665 korreliert waren (Abb. 15.22).

15.6.4 Andere Items

Übung 15.4 (Item-Analyse)

- Berechnen Sie bitte
 - Cronbachs alpha
 - Item-Schwierigkeit
 - Trennschärfe
 - Homogenität
 - itemspezifische Homogenität

der anderen Skalen.

- Führen Sie eine Hauptkomponentenanalyse durch und schreiben Sie jeweils den 1. Faktor in die SPSS-Datei.
- Vergleichen Sie die Resultate mit den bereits bestehenden Konstrukt-Variablen.
- Berechnen Sie Summenscores und vergleichen Sie diese mit den entsprechenden Hauptkomponenten.
- Erstellen Sie bitte eine Korrelationsmatrix aller Konstrukte (Abb. 15.22).



Reliabilitätsstatistiken

Cronbachs Alpha	Cronbachs Alpha für standardisierte Items	Anzahl der Items
,776	,783	7

Itemstatistiken

	Mittelwert	Standardabweichung	N
NRAUM	2,1952	1,01244	333
NANSPREC	2,0330	1,20320	333
NZEIT	1,7267	,95711	333
NERLEDIG	1,8048	,99140	333
KORIENT	1,9820	,89492	333
KZWANG	1,7477	,92946	333
KPRIVAT	2,6517	1,23892	333

Inter-Item-Korrelationsmatrix

	NRAUM	NANSPREC	NZEIT	NERLEDIG	KORIENT	KZWANG	KPRIVAT
NRAUM	1,000	,227	,351	,440	,453	,318	,475
NANSPREC	,227	1,000	,230	,334	,252	,261	,303
NZEIT	,351	,230	1,000	,493	,416	,325	,298
NERLEDIG	,440	,334	,493	1,000	,437	,293	,359
KORIENT	,453	,252	,416	,437	1,000	,255	,250
KZWANG	,318	,261	,325	,293	,255	1,000	,379
KPRIVAT	,475	,303	,298	,359	,250	,379	1,000

Screeplot

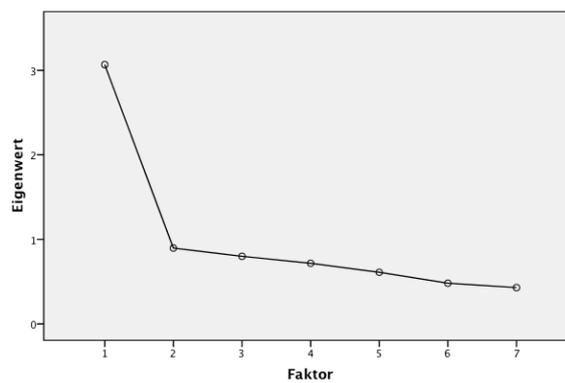
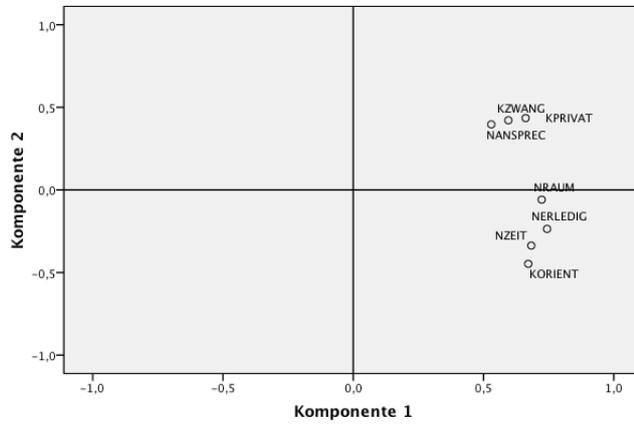


Abbildung 15.19: *Kontrolle + Nutzung und Zielerreichung*: Cronbachs alpha (oben), scree plot (unten).

Komponentendiagramm

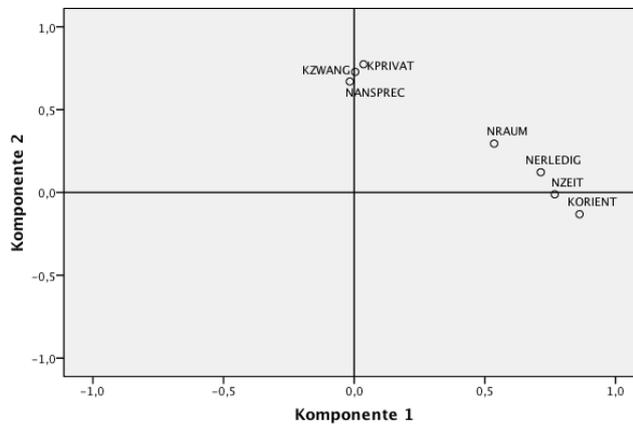


Koeffizientenmatrix der Komponentenwerte

	Komponente	
	1	2
NRAUM	,236	-,066
NANSPREC	,173	,442
NERLEDIG	,243	-,263
NZEIT	,223	-,376
KORIENT	,219	-,499
KZWANG	,194	,470
KPRIVAT	,216	,485

Extraktionsmethode:
Hauptkomponentenanalyse.

Komponentendiagramm im rotierten Raum



Koeffizientenmatrix der Komponentenwerte

	Komponente	
	1	2
NRAUM	,241	,154
NANSPREC	-,034	,398
NERLEDIG	,332	,044
NZEIT	,363	-,037
KORIENT	,412	-,112
KZWANG	-,027	,431
KPRIVAT	-,014	,457

Extraktionsmethode:
Hauptkomponentenanalyse.
Rotationsmethode: Oblimin
mit Kaiser-Normalisierung.

Abbildung 15.20: *Kontrolle + Nutzung und Zielerreichung*: PCA (oben), PCA mit Rotation (Oblimin, unten).

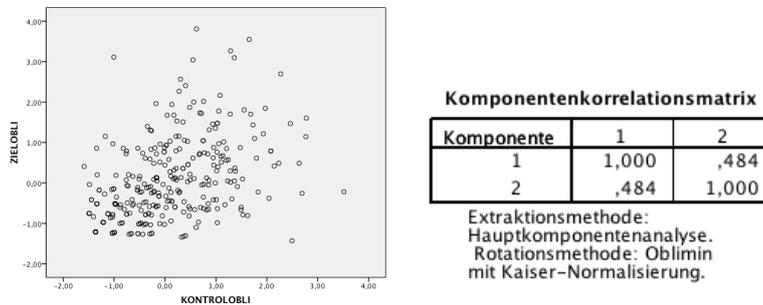


Abbildung 15.21: Kontrolle + Nutzung und Zielerreichung: PCA mit Rotation (Oblimin): Korrelation der Faktoren ZIELOBLI und KONTROLOBLI.

Korrelationen

		FILIAL	KONTROLOBLI	ZIELOBLI	KONTROLL	NUTZZIEL	ZUFRIED	LOYAL
FILIAL	Korrelation nach Pearson	1	,593**	,486**	,612**	,524**	,502**	,587**
	Signifikanz (2-seitig)		,000	,000	,000	,000	,000	,000
	N	375	329	329	368	332	350	364
KONTROLOBLI	Korrelation nach Pearson	,593**	1	,484**	,845**	,692**	,434**	,445**
	Signifikanz (2-seitig)	,000		,000	,000	,000	,000	,000
	N	329	333	333	333	333	327	330
ZIELOBLI	Korrelation nach Pearson	,486**	,484**	1	,696**	,903**	,746**	,602**
	Signifikanz (2-seitig)	,000	,000		,000	,000	,000	,000
	N	329	333	333	333	333	327	330
KONTROLL	Korrelation nach Pearson	,612**	,845**	,696**	1	,665**	,568**	,544**
	Signifikanz (2-seitig)	,000	,000	,000		,000	,000	,000
	N	368	333	333	372	333	350	362
NUTZZIEL	Korrelation nach Pearson	,524**	,692**	,903**	,665**	1	,712**	,585**
	Signifikanz (2-seitig)	,000	,000	,000	,000		,000	,000
	N	332	333	333	333	336	330	333
ZUFRIED	Korrelation nach Pearson	,502**	,434**	,746**	,568**	,712**	1	,670**
	Signifikanz (2-seitig)	,000	,000	,000	,000	,000		,000
	N	350	327	327	350	330	354	349
LOYAL	Korrelation nach Pearson	,587**	,445**	,602**	,544**	,585**	,670**	1
	Signifikanz (2-seitig)	,000	,000	,000	,000	,000	,000	
	N	364	330	330	362	333	349	367

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Abbildung 15.22: Korrelation aller Konstrukte (inklusive ZIELOBLI und KONTROLOBLI). Paarweiser Fallausschluss.

15.7 Zusammenhänge der Konstrukte

15.7.1 Umcodierung

Wie schon erwähnt, wurden die items als **Schulnoten** skaliert, d.h. niedrige Werte entsprechen guten Beurteilungen. Es ist jedoch intuitiver, daß hohe Werte guten Beurteilungen entsprechen. Daher wurden die Skalenergebnisse umcodiert (linear transformiert; $X \rightarrow X' = -X$; vgl. Abb. 15.23). Die umcodierten Konstrukte wurden mit der Bezeichnung U (etwa FilialU) von den bestehenden Konstrukten (die aus einer Hauptkomponentenanalyse entstanden) unterschieden.

Im folgenden wird mit den neuen Variablen gearbeitet.

15.7.2 Korrelation und gruppierte Streudiagramme

Die Korrelation der Konstrukte mit anderen Variablen ist in Abb. 15.24–15.25 zu sehen. Alle Konstrukte interkorrelieren positiv (Abb. 15.26; roter Block rechts oben).

Wie schon vermutet besteht eine negative Korrelation zwischen Zufriedenheit (umcodiert) und dem Filialtyp (ALTNEU), obwohl Filialgestaltung positiv sowohl mit Zufriedenheit als auch Filialtyp korreliert.

Betrachtet man die Streudiagramme, so ist im Diagramm Zufriedenheit/Filialtyp(ALTNEU) bereits eine Mittelwertsverschiebung der roten Punkte (Filialtyp=alt) nach oben zu sehen.

Die genauen Verhältnisse zeigt ein nach Filialtyp (ALTNEU = alt/neu) gruppiertes Streudiagramm der Variablen Zufriedenheit mit den anderen Konstrukten (Abb. 15.27–15.28). Die Zufriedenheitswerte in den alten Filialen (rot) sind nach oben verschoben

$$r(\text{Zufried}U, \text{ALTNEU}) = -.2048,$$

die Werte für Filialgestaltung etc. sind in den neuen Filialen nach rechts verschoben

$$r(\text{Filial}U, \text{ALTNEU}) = +.1317,$$

trotzdem korrelieren alle Konstrukte positiv miteinander.

Schereneffekt

Außerdem ist ein Schereneffekt zu sehen: Die Steigungen der Geraden sind unterschiedlich – bei hohen Konstruktwerten (etwa wahrgenommene Kontrolle) holen die neuen Filialen in der Zufriedenheit auf.

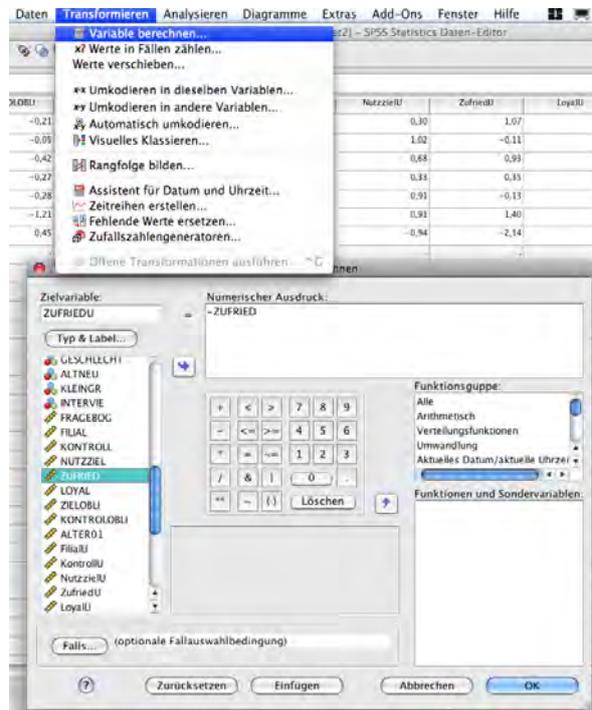


Abbildung 15.23: Transformation der Konstrukte $X \rightarrow X' = -X$.

Correlations								
	FiliaIU	KontrollIU	NutzzielIU	ZufriedU	LoyalU	ALTNEU	KLEINGRGESCHLECHT	ALTER01
FiliaIU	1,0000	0,6169	0,5299	0,5222	0,6049	0,1317	0,1556	0,0945
KontrollIU	0,6169	1,0000	0,6647	0,5626	0,5266	-0,0833	0,1309	0,0366
NutzzielIU	0,5299	0,6647	1,0000	0,7146	0,5803	-0,1256	-0,0273	0,1119
ZufriedU	0,5222	0,5626	0,7146	1,0000	0,6449	-0,2048	-0,0214	0,1231
LoyalU	0,6049	0,5266	0,5803	0,6449	1,0000	-0,1947	0,0130	0,0631
ALTNEU	0,1317	-0,0833	-0,1256	-0,2048	-0,1947	1,0000	0,1057	-0,0148
KLEINGR	0,1556	0,1309	-0,0273	-0,0214	0,0130	0,1057	1,0000	0,0007
GESCHLECHT	0,0945	0,0366	0,1119	0,1231	0,0631	-0,0148	0,0007	1,0000
ALTER01	0,1061	0,1539	0,0767	0,1817	0,1838	-0,2296	0,0214	-0,0521

61 rows not used due to missing or excluded values or frequency or weight variables missing, negative or less than one.

Partial Corr								
	FiliaIU	KontrollIU	NutzzielIU	ZufriedU	LoyalU	ALTNEU	KLEINGRGESCHLECHT	ALTER01
FiliaIU	.	0,3413	0,0342	0,0998	0,3860	0,3473	0,1201	0,0688
KontrollIU	0,3413	.	0,4004	0,0453	0,0153	-0,0797	0,1389	-0,0739
NutzzielIU	0,0342	0,4004	.	0,4471	0,1101	-0,0006	-0,1121	0,0511
ZufriedU	0,0998	0,0453	0,4471	.	0,2872	-0,1224	-0,0386	0,0732
LoyalU	0,3860	0,0153	0,1101	0,2872	.	-0,2070	-0,0354	-0,0389
ALTNEU	0,3473	-0,0797	-0,0006	-0,1224	-0,2070	.	0,0548	-0,0287
KLEINGR	0,1201	0,1389	-0,1121	-0,0386	-0,0354	0,0548	.	0,0092
GESCHLECHT	0,0688	-0,0739	0,0511	0,0732	-0,0389	-0,0287	0,0092	.
ALTER01	0,0316	0,0777	-0,1210	0,0970	0,0534	-0,1952	0,0192	-0,0678

partialled with respect to all other variables

Abbildung 15.24: (Partielle) Korrelationsmatrix der Konstrukte mit anderen Variablen (Listenweiser Fallausschluss).

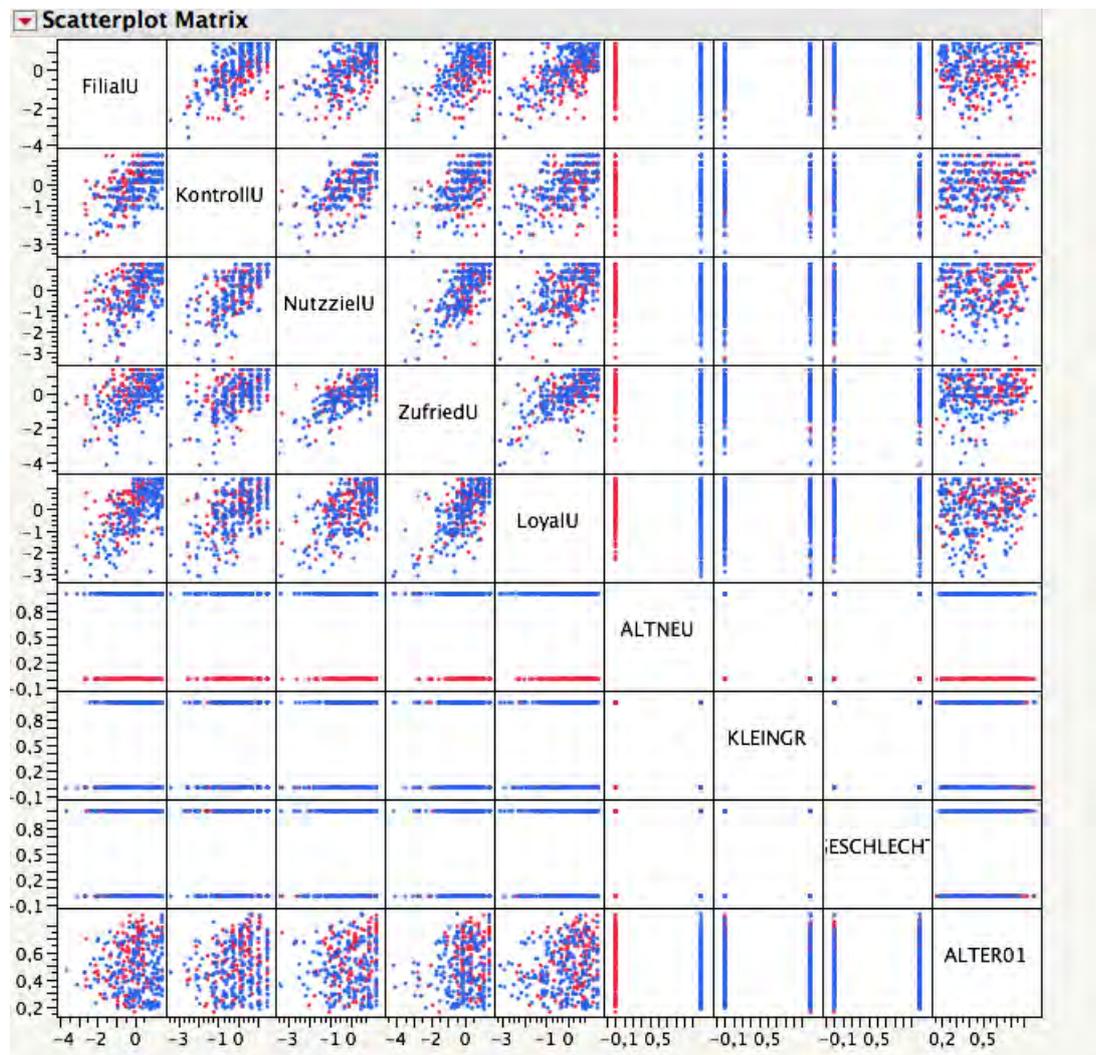


Abbildung 15.25: Streudiagramm der Konstrukte mit anderen Variablen.
Farbe der Punkte: Rot = alt, blau = neu. Listenweiser Fallausschluss.

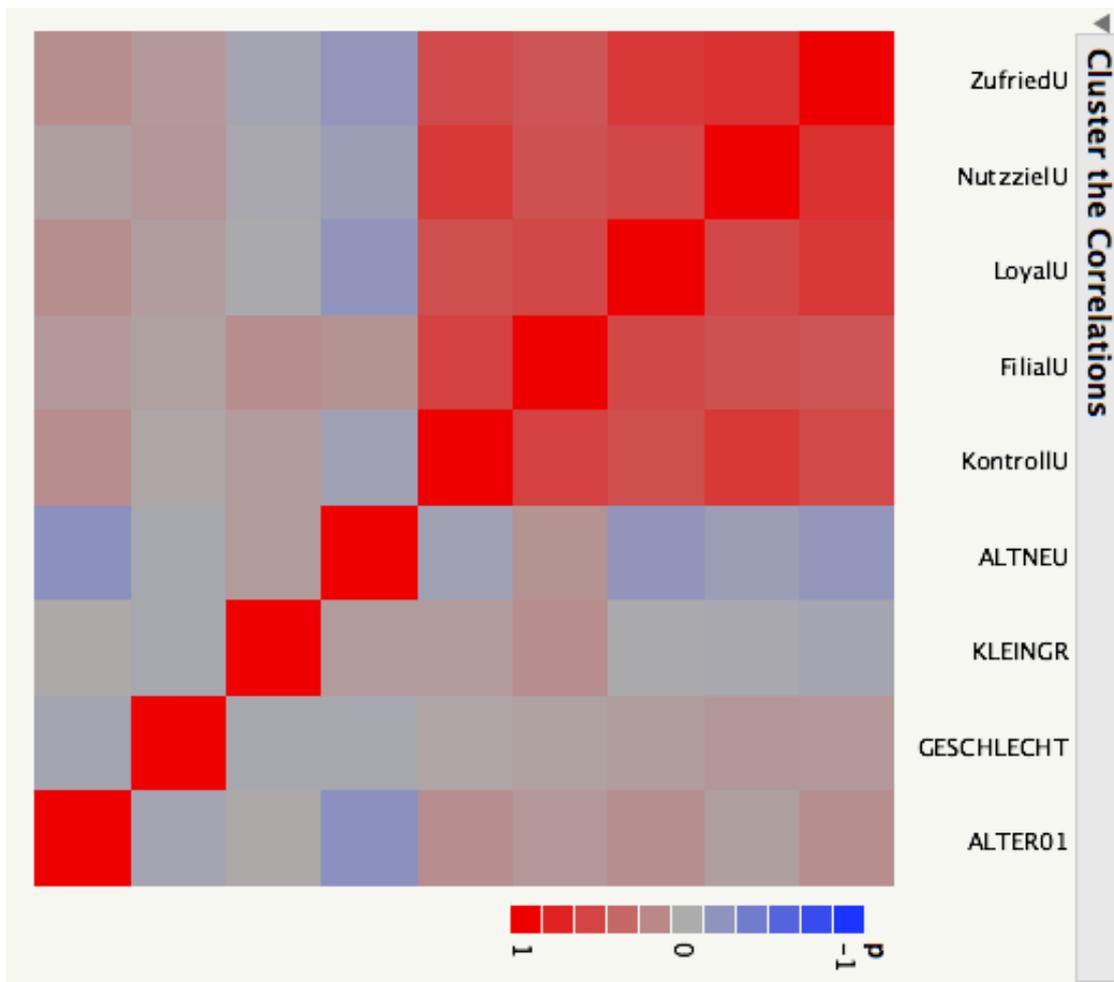


Abbildung 15.26: Diagonales Ordnen der Korrelationsmatrix (Cluster der Korrelationen). Deutlich ist der rote Block der Konstrukte sowie die negative Korrelation mit ALTNEU zu erkennen (Ausnahme: FilialU).

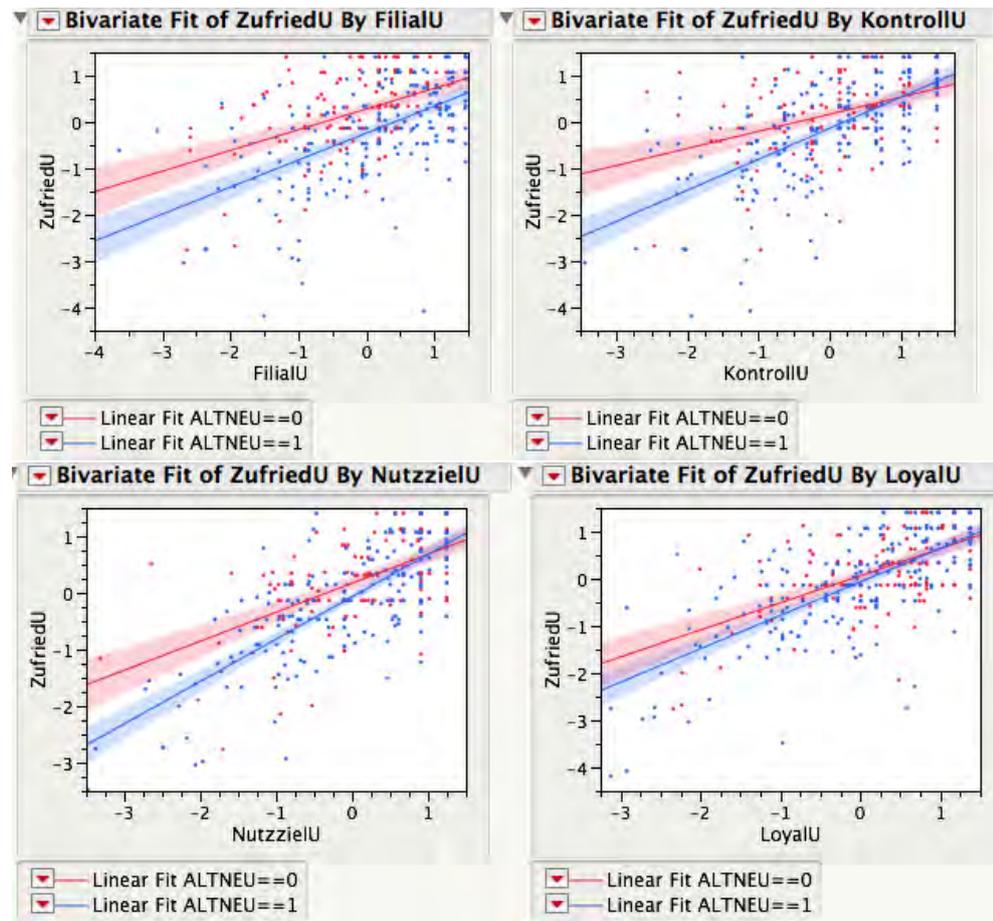


Abbildung 15.27: Gruppiertes Streudiagramm und Regression der Variable *Zufriedenheit* mit den anderen Konstrukten. Die Zufriedenheitswerte in den alten Filialen (rot) sind nach oben verschoben [$r(\text{Zufriedenheit}, \text{ALTNEU}) = -.2048$], die Werte für Filialgestaltung etc. sind in den neuen Filialen nach rechts verschoben [$r(\text{FilialU}, \text{ALTNEU}) = +.1317$].

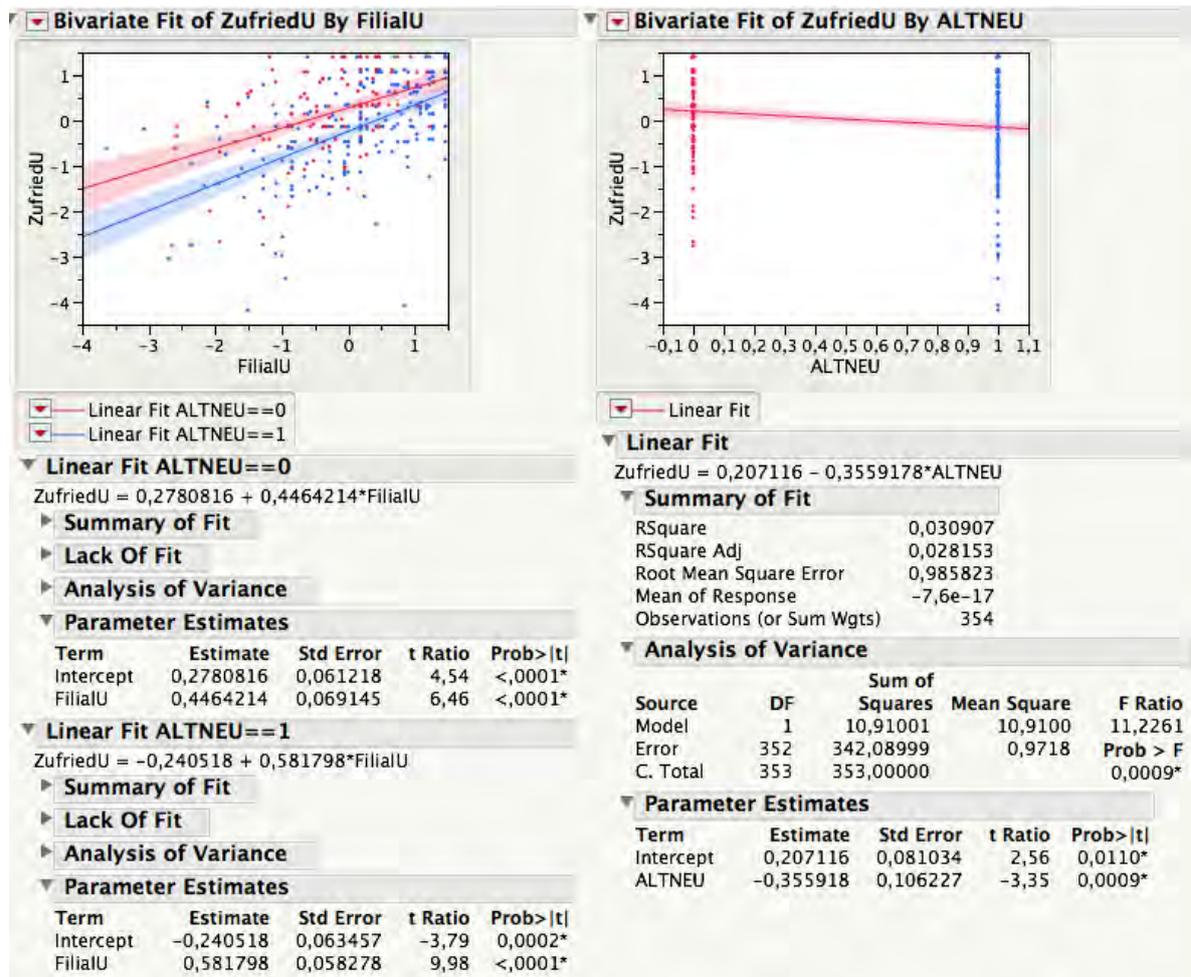


Abbildung 15.28: Gruppiertes Streudiagramm und Regression der Variable *Zufriedenheit* mit *Filialgestaltung* (links) und mit *ALTNEU* (rechts). Man erkennt deutlich die negative Korrelation (höhere Zufriedenheitswerte in den alten Filialen).

15.7.3 Mittelwerte

Bekanntlich enthalten Korrelationen die Mittelwerte und Streuungen der Variablen nicht mehr, da diese in der Definition

$$\begin{aligned}
 R(X, Y) &= \frac{S(X, Y)}{S(X)S(Y)} \\
 S(X, Y) &= \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})(Y_n - \bar{Y}) \\
 S(X)^2 &= \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})^2 = S(X, X) \\
 S(Y)^2 &= \frac{1}{N-1} \sum_{n=1}^N (Y_n - \bar{Y})^2 = S(Y, Y)
 \end{aligned}$$

schon herauskorrigiert werden.

Eine lineare Transformation $X \rightarrow X' = a + bX$, $Y \rightarrow Y' = c + dY$ läßt daher den Korrelationskoeffizienten unverändert, d.h.

$$R(X, Y) = R(X', Y') \quad (15.1)$$

Übung 15.5 (Invarianz des Korrelationskoeffizienten)

Zeigen Sie, daß bei einer linearen Transformation

$$R(X, Y) = R(X', Y')$$

gilt.



Daher ist es sinnvoll, auch die Mittelwerte und Standardabweichungen der Konstrukte in gruppierter Form zu betrachten.

Profilplot

Ein Profilplot der Konstruktmittelwerte gibt die gewünschte Information. So ist zu sehen, daß in den alten Filialen höhere Zufriedenheits-, Zielerreichungs- und Loyalitätswerte zu finden sind – in den neuen Filialen ist das Profil gegenläufig.

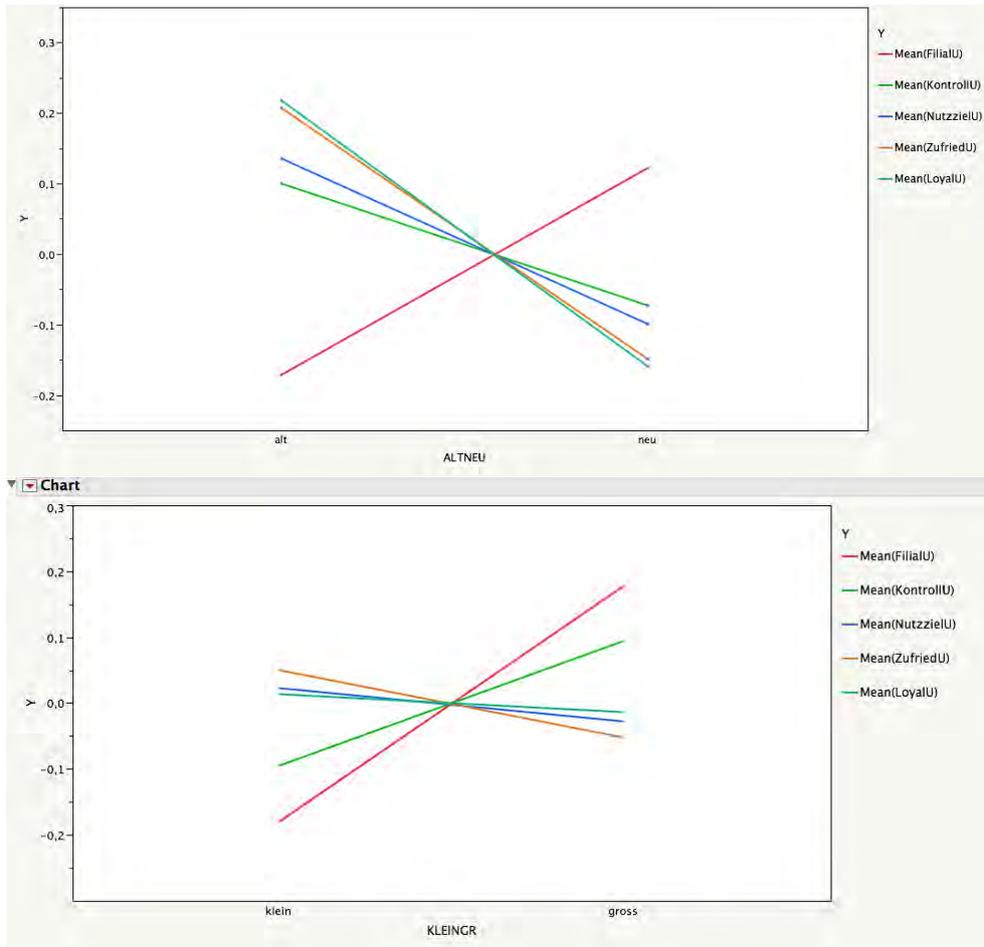


Abbildung 15.29: Mittelwerte und 95%-Konfidenzintervalle der Konstrukte gruppiert nach Filialkonzept (oben) und Filialgröße (unten).

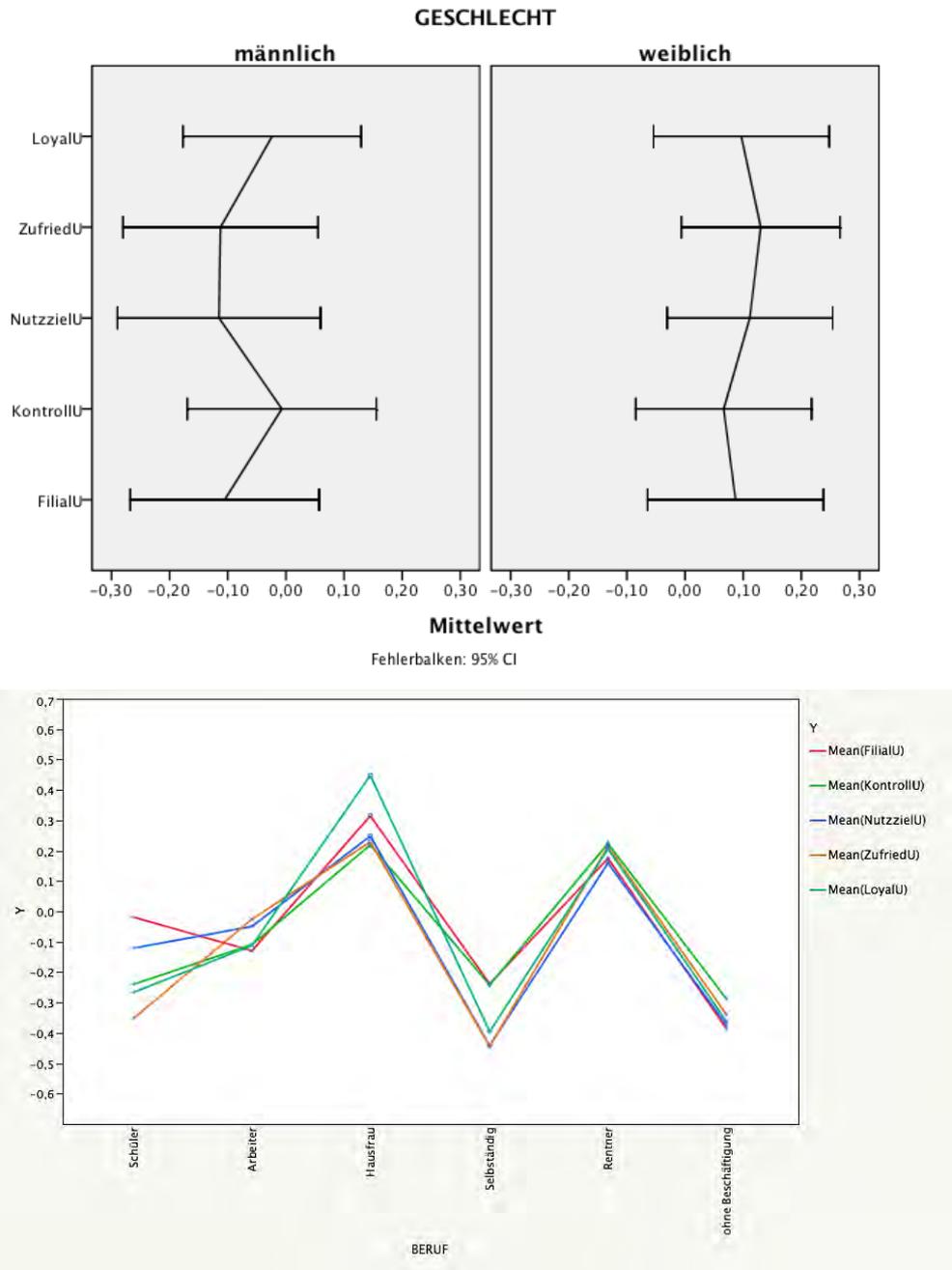


Abbildung 15.30: Mittelwerte und 95%-Konfidenzintervalle der Konstrukte gruppiert nach Geschlecht (oben) und Berufsgruppe (unten).

15.7.4 Statistische Überprüfung

15.7.5 Regression

Die in den Streudiagrammen sowie in den Profil-Plots gefundenen Zusammenhänge lassen sich auch statistisch überprüfen.

In Abb. 15.28 ist der Zusammenhang zwischen den Variablen ZufriedU und FilialU (gruppiert nach ALTNEU (Moderatorvariable)) sowie ZufriedU und ALTNEU dargestellt.

Die gruppierte Regression ist signifikant (die p -Werte der t -Statistik sind sehr klein). Die Steigungen der Regressionsgeraden

$$\hat{\beta}_0 = .446(.069), \hat{\beta}_1 = .582(.058)$$

(Standardabweichung in Klammern) sind unterschiedlich (Schereneffekt). Auch die Achsenabschnitte (intercepts)

$$\hat{\alpha}_0 = .278(.061), \hat{\alpha}_1 = -.241(.063)$$

unterscheiden sich deutlich (höhere Zufriedenheit in alten Filialen).

Um die gruppenspezifischen Achsenabschnitte und Steigungen gegeneinander zu testen, können t -Tests gerechnet werden.

Übung 15.6 (Unterschiede der Regressionskoeffizienten)

Prüfen Sie, ob sich die Achsenabschnitte und Steigungen in den durch ALTNEU = 0/1 definierten Gruppen signifikant unterscheiden.

Die Stichprobengrößen der Gruppen sind $N_0 = 146$, $N_1 = 204$, $N_0 + N_1 = 350$.

Führen Sie t -Tests für unabhängige Stichproben durch.



Eine Alternative ist, die Variablen FilialU und ALTNEU (Indikator) als unabhängige Variablen in eine multiple Regressionsgleichung mit Interaktionsterm FilialU*ALTNEU einzugeben (vgl. Kap. 13.4)

$$\begin{aligned} \text{ZufriedU} &= (\alpha + \beta \text{FilialU}) \\ &+ (\alpha_1 \text{ALTNEU} + \beta_1 \text{FilialU} * \text{ALTNEU}) + \epsilon. \end{aligned}$$

Der erste Teil der Gleichung ist die Gerade für ALTNEU = 0 (alte Filialen), da in diesem Fall der zweite Teil der Gleichung wegfällt.

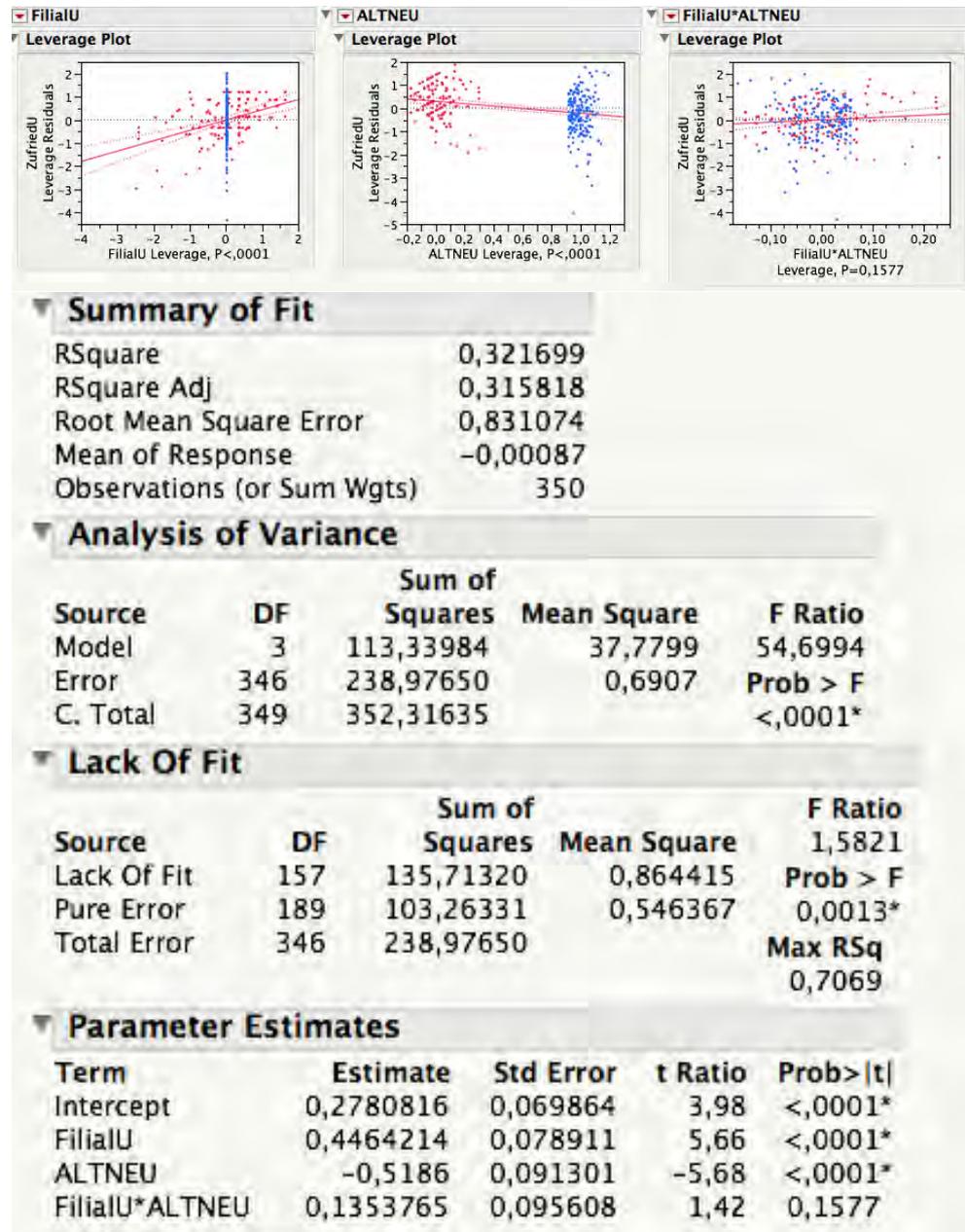


Abbildung 15.31: Multiple Regression mit Indikator-Variablen (0/1) ALT-NEU.

In den neuen Filialen ($ALTNEU = 1$) ist die Gleichung

$$(\alpha + \alpha_1) + (\beta + \beta_1)FilialU.$$

Man erhält also einen anderen Achsenabschnitt und eine andere Steigung.

Man könnte auch schreiben

$$\begin{aligned} ZufriedU &= (\alpha + \alpha_1 ALTNEU) \\ &+ (\beta + \beta_1 ALTNEU) FilialU + \epsilon \\ &:= \alpha(ALTNEU) + \beta(ALTNEU) FilialU + \epsilon, \end{aligned}$$

also in Form gruppenspezifischer Achsenabschnitte und Steigungen.

Abb. 15.31 zeigt die Abhängigkeit der Zufriedenheit von den Variablen Filialgestaltung und Filialkonzept.

Die Tabellen der Effektschätzungen (unten) zeigen, daß der Interaktionseffekt β_1 (FilialU*ALTNEU im output) nicht signifikant ist ($p = .1577 > .05$).

Die Schätzungen $\hat{\alpha} = .2781$ (Intercept) und $\hat{\beta} = .4464$ (Steigung) stimmen mit den Schätzungen aus Abb. 15.28 überein ($ALTNEU = 0$).

Der Achsenabschnitt für $ALTNEU = 1$ ergibt sich aus

$$(\hat{\alpha} + \hat{\alpha}_1 * ALTNEU) = .2781 - .5186 = -0.2405$$

(Abb. 15.28, unten).

Der Steigung für $ALTNEU = 1$ ergibt sich aus

$$(\hat{\beta} + \hat{\beta}_1 * ALTNEU) = .4464 + .1353 = 0.5817$$

(Abb. 15.28, unten).

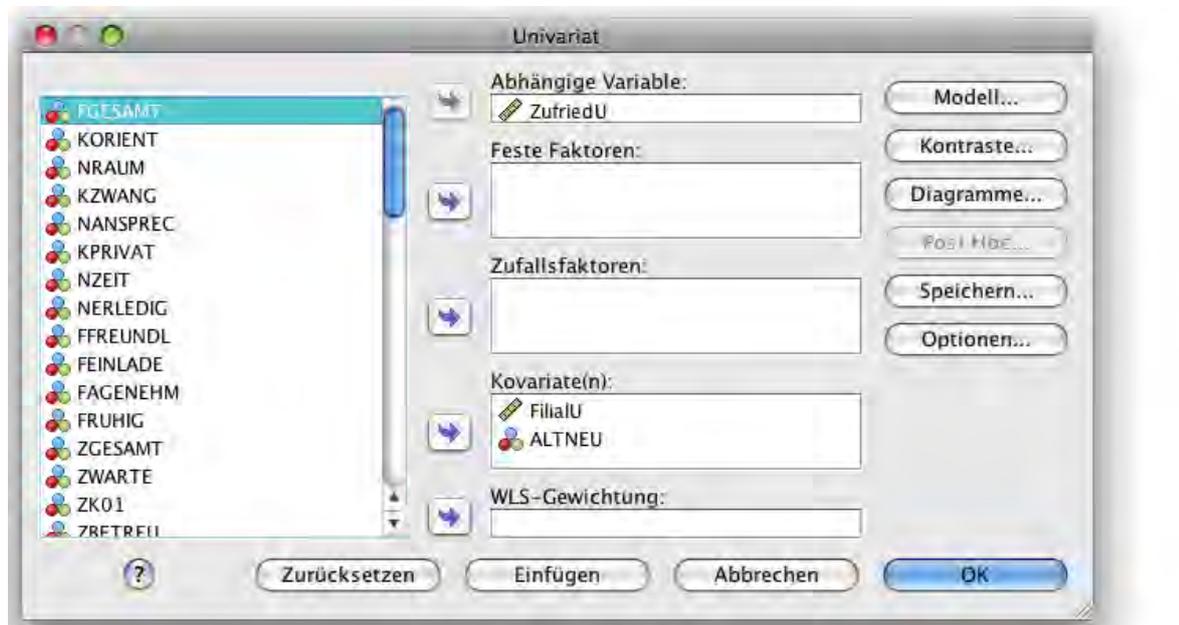
Also sind die Mittelwertsverschiebungen signifikant ($\alpha_1 = 0$ ablehnen), während der Schereneffekt nicht bestätigt werden kann ($\beta_1 = 0$ beibehalten).

Der entsprechende SPSS-Output ist in Abb. 15.32 zu sehen.

Menü-Befehl:

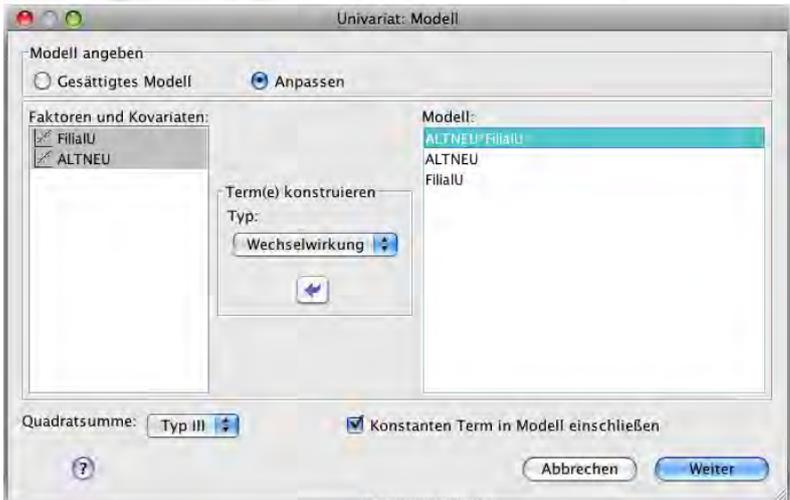
SPSS/Analysieren/Allgemeines lineares Modell/univariat

Das in Abb. 15.1 gezeigte Wirkungsmodell ist wesentlich komplexer als die eben behandelte Regression, da ZufriedU im Diagramm auch als unabhängige Variable auf Loyalität wirkt (entsprechend für KontrollU und



The screenshot shows the SPSS 'Univariate' dialog box. The dependent variable is 'ZufriedU'. The fixed factors are empty. The random factors are empty. The covariate is 'FiliaU' and 'ALTNEU'. The WLS weighting is empty. The 'Parameters Estimator' table is shown below the dialog box.

Parameter	Regressionskoeffizient B	Standardfehler	T	Sig.	95%-Konfidenzintervall		Partielles Eta-Quadrat
					Untergrenze	Obergrenze	
Konstanter Term	,278	,070	3,980	,000	,141	,415	,044
FiliaU + ALTNEU	,135	,096	1,416	,158	-,053	,323	,006
ALTNEU	-,519	,091	-5,680	,000	-,698	-,339	,085
FiliaU	,446	,079	5,657	,000	,291	,602	,085



The screenshot shows the 'Univariate: Modell' dialog box. The model is set to 'Anpassen'. The factors and covariates are 'FiliaU' and 'ALTNEU'. The model is 'ALTNEU*FiliaU', 'ALTNEU', and 'FiliaU'. The quadratic sum is 'Typ III'. The constant term is included in the model.

Abbildung 15.32: Varianzanalyse (Allgemeines lineares Modell/univariat) mit Kovariaten FiliaU und ALTNEU.

NutzZielU). Die Konstrukte fungieren also sowohl als unabhängige als auch als abhängige Variable.

Man spricht auch von simultanen Gleichungssystemen oder Strukturgleichungsmodellen.

Das Modell lautet explizit für die Konstrukte

Strukturgleichungsmodell

1. Wahrgenommene Filialgestaltung (FilialU)
2. Nutzung und Zielerreichung (NutzZielU)
3. Wahrgenommene Kontrolle (KontrollU)
4. Zufriedenheit mit dem Filialbesuch (ZufriedU)
5. Kundenloyalität (LoyalU)

(Moderator ALTNEU weggelassen):

$$\begin{aligned}
 \text{FilialU} &= \alpha_1 + \epsilon_1 \\
 \text{NutzZielU} &= \alpha_2 + \beta_{21}\text{FilialU} + \beta_{23}\text{KontrollU} + \epsilon_2 \\
 \text{KontrollU} &= \alpha_3 + \beta_{31}\text{FilialU} + \epsilon_3 \\
 \text{ZufriedU} &= \alpha_4 + \beta_{41}\text{FilialU} + \beta_{42}\text{NutzZielU} + \beta_{43}\text{KontrollU} + \epsilon_4 \\
 \text{LoyalU} &= \alpha_5 + \beta_{51}\text{FilialU} + \beta_{54}\text{ZufriedU} + \epsilon_5.
 \end{aligned}$$

Man kann auf die Idee kommen, 5 einzelne Regressionsanalysen zu rechnen. Dies ist aber falsch, da die Regressoren auf der linken **und** auf der rechten Seite stehen und somit mit den Fehlern korreliert sind. Dann erhält man inkorrekte Schätzungen.

Zur Schätzung simultaner Gleichungssysteme müssen spezielle Programme wie LISREL, AMOS, Mx, MECOSA etc. benutzt werden, die auf diese Problemstellung zugeschnitten sind.

15.7.6 Mittelwerts-Vergleiche

In den Profilen Abb. 15.29-15.30 wurden die Mittelwerte der Konstrukte FilialU, KontrollU, NutzzielU, ZufriedU, LoyalU als Funktion der Variablen ALTNEU, KLEINGROSS, Geschlecht oder Beruf aufgetragen.

Im Rahmen eines (multiplen) t -Tests oder einer Varianzanalyse kann überprüft werden, ob die Unterschiede signifikant sind. Abb. 15.33 zeigt den Auswahldialog für den t -Test (unabhängige Stichproben), während in Abb. 15.34–15.35 die einzelnen (univariaten) t -Tests tabelliert sind.

Bis auf KontrollU sind alle Unterschiede bzgl. der Variablen ALTNEU signifikant ($p < .05$). Dagegen finden sich Geschlechtsunterschiede bei den Variablen NutzzielU und ZufriedU ($p < .05$).

Eine Varianzanalyse der Unterschiede als Funktion der Berufsgruppe (Abb. 15.30, 15.36) erbringt signifikante Ergebnisse auf allen Konstrukten.

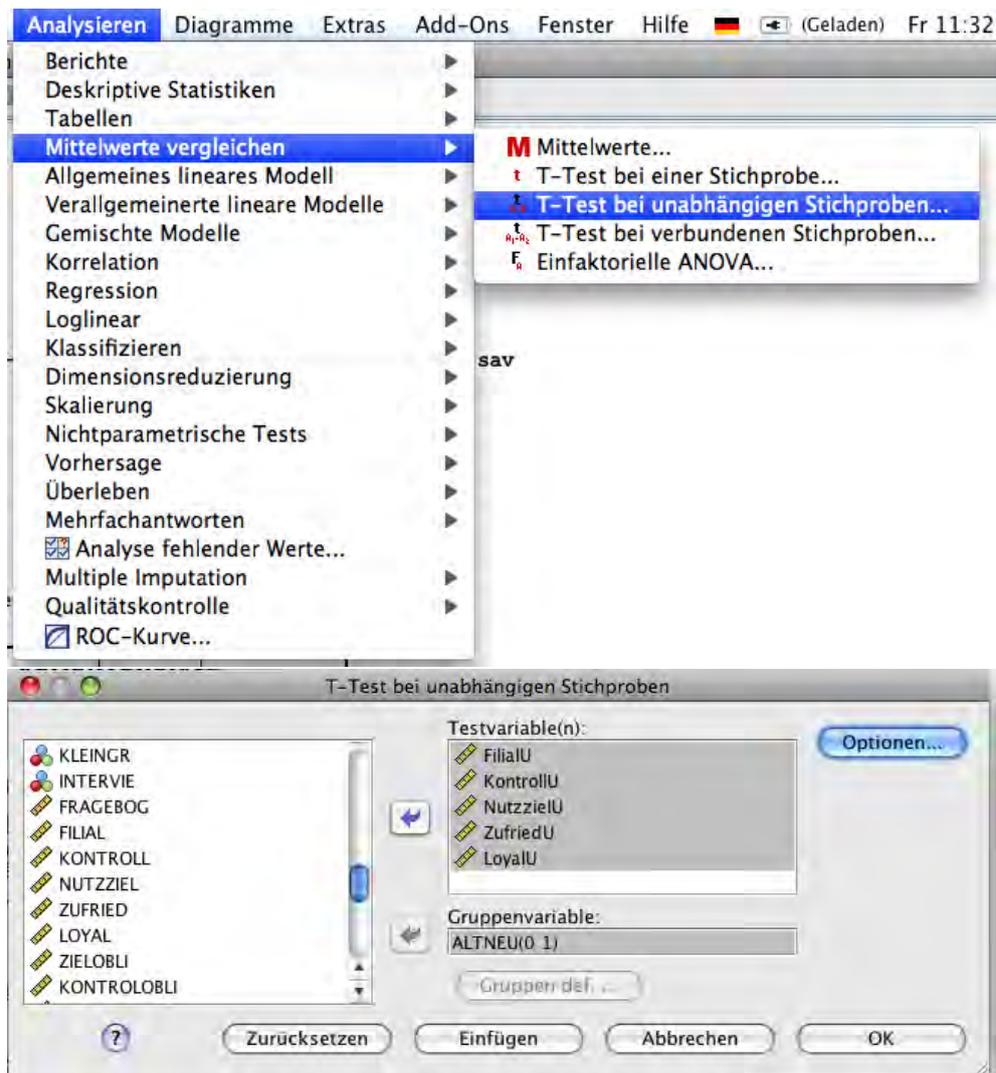
Welche Ausprägungen der unabhängigen Variablen dies genau verursachen, kann mit Hilfe von Post-Hoc-Tests (etwa mit der Bonferroni-Methode) geklärt werden. Obwohl der globale F -Test signifikant ist, erreicht kein einziger Bonferroni- t -Vergleich einen Wert von $p < .05$ (Abb. 15.37). Der Einzelvergleich Selbständig–Rentner ergibt gerade noch $p = 0.066$.

Es ist allerdings so, daß die globale Signifikanz ($H_0 : \mu_i = \mu$ ablehnen) nicht die Signifikanz von mindestens einem Differenz-Wert $\mu_i - \mu_j$, sondern nur die Signifikanz irgendeines linearen Kontrasts

Kontrast

$$\sum_i c_i \mu_i$$

impliziert (simultane Scheffé-Konfidenzintervalle; vgl. Fahrmeir et al., 1996, Kap. 5.1, S. 178 ff.).

Abbildung 15.33: *t*-Test der Konstrukte. Auswahldialog.

Gruppenstatistiken

ALTNEU	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
FilialU	156	-,1713	,86972	,06963
KontrollU	219	,1220	1,06856	,07221
NutzzielU	157	,1000	,92114	,07351
ZufriedU	215	-,0730	1,04997	,07161
LoyalU	142	,1355	,90508	,07595
	194	-,0992	1,05541	,07577
	148	,2071	,81898	,06732
	206	-,1488	1,08985	,07593
	155	,2177	,83889	,06738
	212	-,1592	1,07736	,07399

Test bei unabhängigen Stichproben

	Levene-Test der Varianzgleichheit		T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
	F	Signifikanz						Untere	Obere
	T-Test für die Mittelwertgleichheit								
FilialU	9,514	,002	-2,826	373	,005	-,29333	,10380	-,49745	-,08922
KontrollU	4,719	,030	-2,924	366,366	,004	-,29333	,10031	-,49059	-,09607
NutzzielU	4,541	,034	1,651	370	,100	,17295	,10474	-,03300	,37890
ZufriedU	8,691	,003	1,685	357,713	,093	,17295	,10263	-,02888	,37477
LoyalU	10,844	,001	2,137	334	,033	,23473	,10986	,01863	,45082
			2,188	325,663	,029	,23473	,10729	,02366	,44579
			3,351	352	,001	,35592	,10623	,14700	,56484
			3,507	351,271	,001	,35592	,10148	,15634	,55550
			3,625	365	,000	,37689	,10397	,17243	,58135
			3,766	363,529	,000	,37689	,10008	,18009	,57369

Abbildung 15.34: t-Test der Konstrukte: Gruppenvariable ALTNEU.

Gruppenstatistiken

GESCHLECHT	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
FilialU männlich	168	-,0707	,94352	,07279
FilialU weiblich	206	,0572	1,04487	,07280
KontrollU männlich	166	-,0367	,94442	,07330
KontrollU weiblich	205	,0242	1,04374	,07290
NutzzielU männlich	147	-,1337	1,03174	,08510
NutzzielU weiblich	188	,1037	,96714	,07054
ZufriedU männlich	156	-,1243	1,00848	,08074
ZufriedU weiblich	197	,0969	,98716	,07033
LoyalU männlich	162	-,0986	,94284	,07408
LoyalU weiblich	204	,0759	1,04077	,07287

Test bei unabhängigen Stichproben

	Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						95% Konfidenzintervall der Differenz	
	F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	Untere	Obere	
	FilialU	1,639	,201	-1,230 -1,243	372 368,124	,219 ,215	-,12795 -,12795	,10402 ,10295	-,33249 -,33039	,07660 ,07450
KontrollU	2,971	,086	-,583 -,589	369 364,442	,560 ,556	-,06093 -,06093	,10447 ,10338	-,26636 -,26423	,14450 ,14236	
NutzzielU	,468	,495	-2,164 -2,147	333 303,635	,031 ,033	-,23734 -,23734	,10966 ,11053	-,45305 -,45484	-,02164 -,01984	
ZufriedU	,041	,840	-2,071 -2,066	351 329,454	,039 ,040	-,22123 -,22123	,10681 ,10708	-,43131 -,43188	-,01115 -,01058	
LoyalU	2,318	,129	-1,661 -1,680	364 357,701	,098 ,094	-,17456 -,17456	,10509 ,10391	-,38122 -,37891	,03211 ,02979	

Abbildung 15.35: t-Test der Konstrukte: Gruppenvariable Geschlecht.

```

ONEWAY FilialU KontrollU NutzzielU ZufriedU LoyalU BY BERUF
/PLOT MEANS
/MISSING ANALYSIS
/POSTHOC=BONFERRONI ALPHA(0.05).

```

→ **Univariat**

```
[DatenSet1] /HermannTeX/Multi/MasterA-Modul/Master A-Modul/Daten/SPSS-Datensätze/main2.xls.sav
```

ONEWAY ANOVA

		Quadratsumme	df	Mittel der Quadrate	F	Signifikanz
FilialU	Zwischen den Gruppen	15,021	5	3,004	3,098	,009
	Innerhalb der Gruppen	356,842	368	,970		
	Gesamt	371,863	373			
KontrollU	Zwischen den Gruppen	14,460	5	2,892	2,971	,012
	Innerhalb der Gruppen	355,285	365	,973		
	Gesamt	369,745	370			
NutzzielU	Zwischen den Gruppen	12,783	5	2,557	2,612	,025
	Innerhalb der Gruppen	321,998	329	,979		
	Gesamt	334,781	334			
ZufriedU	Zwischen den Gruppen	17,657	5	3,531	3,662	,003
	Innerhalb der Gruppen	334,605	347	,964		
	Gesamt	352,262	352			
LoyalU	Zwischen den Gruppen	23,782	5	4,756	5,018	,000
	Innerhalb der Gruppen	341,270	360	,948		
	Gesamt	365,052	365			

Abbildung 15.36: Univariate Varianzanalyse der Konstrukte mit Gruppvariable (UV) Beruf ($J = 6$ Ausprägungen).

Post-Hoc-Tests

Mehrfachvergleiche

ZufriedU
Bonferroni

(I) BERUF	(J) BERUF	Mittlere Differenz (I-J)	Standardfehler	Signifikanz	95%-Konfidenzintervall	
					Untergrenze	Obergrenze
Schüler	Arbeiter	-,32681	,20139	1,000	-,9220	,2684
	Hausfrau	-,58120	,24078	,245	-1,2929	,1305
	Selbständig	,09227	,27764	1,000	-,7283	,9129
	Rentner	-,56606	,20538	,092	-1,1731	,0410
	ohne Beschäftigung	-,01298	,27418	1,000	-,8234	,7974
Arbeiter	Schüler	,32681	,20139	1,000	-,2684	,9220
	Hausfrau	-,25439	,17897	1,000	-,7834	,2746
	Selbständig	,41908	,22613	,970	-,2493	1,0874
	Rentner	-,23925	,12741	,919	-,6158	,1373
	ohne Beschäftigung	,31382	,22188	1,000	-,3420	,9696
Hausfrau	Schüler	,58120	,24078	,245	-,1305	1,2929
	Arbeiter	,25439	,17897	1,000	-,2746	,7834
	Selbständig	,67347	,26183	,158	-,1004	1,4474
	Rentner	,01514	,18345	1,000	-,5271	,5574
	ohne Beschäftigung	,56821	,25817	,426	-,1948	1,3313
Selbständig	Schüler	-,09227	,27764	1,000	-,9129	,7283
	Arbeiter	-,41908	,22613	,970	-1,0874	,2493
	Hausfrau	-,67347	,26183	,158	-1,4474	,1004
	Rentner	-,65833	,22969	,066	-1,3372	,0206
	ohne Beschäftigung	-,10525	,29284	1,000	-,9708	,7603
Rentner	Schüler	,56606	,20538	,092	-,0410	1,1731
	Arbeiter	,23925	,12741	,919	-,1373	,6158
	Hausfrau	-,01514	,18345	1,000	-,5574	,5271
	Selbständig	,65833	,22969	,066	-,0206	1,3372
	ohne Beschäftigung	,55307	,22551	,220	-,1135	1,2196
ohne Beschäftigung	Schüler	,01298	,27418	1,000	-,7974	,8234
	Arbeiter	-,31382	,22188	1,000	-,9696	,3420
	Hausfrau	-,56821	,25817	,426	-1,3313	,1948
	Selbständig	,10525	,29284	1,000	-,7603	,9708
	Rentner	-,55307	,22551	,220	-1,2196	,1135

Mittelwert-Diagramme

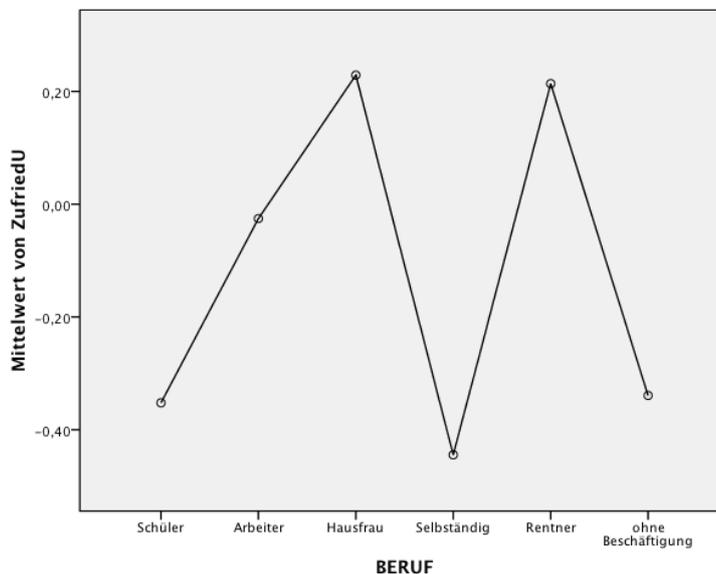


Abbildung 15.37: Bonferroni-t-Tests der Variablen ZufriedU.

Teil IV
Anhang

Kapitel 16

Aufgaben

16.1 Aufgaben KE 1

16.1.1 Punkt- und Intervallschätzung

Aufgabe 1

Bei der Politbarometer-Umfrage erhielt man bei der „Sonntagsfrage“ im Mai 2009 folgende Ergebnisse:

(Quelle: <http://www.forschungsgruppe.de/Aktuelles/Politbarometer/>)

Union bleibt stärkste Partei

Nach leichten Zugewinnen im März und April verliert die CDU/CSU jetzt etwas in der politischen Stimmung und kommt nur noch auf 39 Prozent (minus 2). Die SPD hingegen legt zum dritten Mal in Folge zu und kommt jetzt auf 30 Prozent (plus 3). Abermals Verluste verzeichnet die FDP mit 10 Prozent (minus 1), die Linke mit 8 Prozent (minus 1), während sich die Grünen mit 9 Prozent (plus 1) leicht verbessern können.

Die Umfragen zum Politbarometer wurden wie immer von der Mannheimer Forschungsgruppe Wahlen durchgeführt. Die Interviews wurden in der Zeit vom 05. bis 07. Mai 2009 bei 1.343 zufällig ausgewählten Wahlberechtigten telefonisch erhoben. Die Befragung ist repräsentativ für die wahlberechtigte Bevölkerung in ganz Deutschland. Der Fehlerbereich beträgt bei einem Parteianteil von 40 Prozent rund +/- drei Prozentpunkte und bei einem Parteianteil von 10 Prozent rund +/- zwei Prozentpunkte.

- a) Welche Grundgesamtheit (GG) hat man mit dieser Frage im Auge?
- b) Welche der Grundgesamtheitsparameter entsprechen den obigen Prozentsätzen? Schätzen Sie diese GG-Parameter.
- c) Betrachten Sie nur die Dichotomie

A : „Jetzige Regierungskoalition CDU/CSU/SPD“

\bar{A} : „Jetzige Opposition“

und geben Sie einen Punktschätzer und ein näherungsweise 95%-Vertrauensintervall für den Anteil der jetzigen Regierungskoalition an.

- d) Welche Näherungen muss man bei der Berechnung des Vertrauensintervalls in c) machen?
- e) Ist die Aussage:

Der Fehlerbereich beträgt bei einem Parteianteil von 40 Prozent rund +/- drei Prozentpunkte und bei einem Parteianteil von 10 Prozent rund +/- zwei Prozentpunkte.

tatsächlich korrekt?

Aufgabe 2

Gegeben sei eine GG, in der jedes Element die Eigenschaft A oder \bar{A} besitzt, mit den Wahrscheinlichkeiten $P(A) = P(\bar{A}) = 1/2$. Die relative Häufigkeit vom Umfang N ist ein Schätzer für $P(A)$. In Abhängigkeit vom Stichprobenumfang bezeichnen wir sie mit $f_N(A)$.

- a) Berechnen Sie im Modell mit Zurücklegen die Wahrscheinlichkeit dafür, daß die relative Häufigkeit $f_N(A)$ die zu schätzende Wahrscheinlichkeit $P(A)$ genau trifft, und zwar für die Stichprobenumfänge $N = 2$, $N = 10$, $N = 15$.
- b) In einer Stichprobe vom Umfang $N = 30$ hatten 12 Elemente die Eigenschaft A . Geben Sie ein 95%-Konfidenzintervall für $P(A)$ an.

Aufgabe 3

Eine neue Werbestrategie wird an 300 Probanden erprobt. Sie bringt 210 Erfolge, 90 Mißerfolge. Geben Sie einen Punkt- und einen 95%-Intervallschätzer für die Erfolgswahrscheinlichkeit der Werbestrategie an.

Aufgabe 4

Ein stetiges, metrisch skaliertes Merkmal X (Kundenzufriedenheit) wurde an N Kunden erhoben. Es ergab sich $\bar{x} = 105$.

- a) Bestimmen Sie ein Konfidenzintervall für den Erwartungswert, wenn $N = 10$ und
- X normalverteilt ist und $\sigma^2 = 400$
 - X normalverteilt ist mit unbekanntem σ^2 ($s^2 = 400$).

Vergleichen Sie die beiden Konfidenzintervalle.

- b) Bestimmen Sie Konfidenzintervalle für μ , wenn $N = 50$ und $\sigma^2 = 300$ bekannt ist, bzw. nur $s^2 = 300$ gegeben ist.

Aufgabe 5

Der AStA einer Universität will schätzen, wie viel Geld die 10000 Studenten der Universität durchschnittlich im Monat zur Verfügung haben. Dazu werden 250 zufällig ausgewählte Studenten am Haupteingang befragt. Diese Stichprobe ergibt einen Schätzwert für das mittlere Einkommen von $\bar{x} = 590$ Euro und eine Stichprobenvarianz von $\sigma^2 = 2500$ Euro².

- a) Man berechne ein Konfidenzintervall für das durchschnittliche Einkommen der Studenten der Universität ($\alpha = 0.05$).
- b) Welcher Stichprobenumfang N muß gewählt werden, um für das in a) berechnete Konfidenzintervall einen Sicherheitsgrad von 0.99 zu erreichen?
- c) Kann man davon ausgehen, daß die für das Konfidenzintervall notwendigen Annahmen erfüllt sind?

Aufgabe 6

Im Rahmen einer Marktforschungs-Studie wird jedem Teilnehmer ein Fragebogen vorgelegt. Bei $N = 6$ Personen ergeben sich folgende Testwerte auf den Skalen Kundenzufriedenheit und Loyalität:

Kundenzufriedenheit	46	48	31	32	43	68
Loyalität	102	106	88	90	100	122

- Geben Sie Punkt- und Intervallschätzer für die unbekanntes Populationsparameter μ und σ^2 an (jeweils für beide Variablen).
- Schätzen Sie den unbekanntes Korrelationskoeffizienten ρ und berechnen Sie ein Konfidenzintervall (diskutieren Sie, ob die Annahmen erfüllt sind).
- Was ist der Unterschied zwischen den Populationsparametern (μ, σ^2, ρ) und den Punktschätzern?

Skizzieren Sie die empirischen Verteilungen der Punktschätzer, wenn die Stichprobe ($N = 6$) $M = 10$ mal gezogen wird. (Insgesamt 60 Personen; Unabhängigkeitsannahme).

Aufgabe 7

An einer Stichprobe von $N = 5$ Personen wurde ein Merkmal X erhoben, von dem bekannt ist, dass es Erwartungswert $\mu = 100$ und Varianz $\sigma^2 = 100$ aufweist.

- Welchen Erwartungswert hat die Zufallsvariable $\bar{X} = \frac{1}{N} \sum X_n$ (Stichprobenmittelwert) und welche Varianz weist die auf? (Annahme: X_i sind voneinander unabhängig)
- Skizzieren Sie die Dichtefunktion der Zufallsvariablen X_n und \bar{X} , wenn X_i als normalverteilt $X_n \sim N(\mu = 100, \sigma^2 = 100)$ angenommen wird.
- Wie sieht die Dichtefunktion von \bar{X} aus, wenn sehr große Stichproben $N \rightarrow \infty$ gezogen werden?

Aufgabe 8

- a) Definieren Sie die Begriffe Grundgesamtheit, Stichprobe, Parameter, Schätzer, Schätzung, Population, Populationsparameter und grenzen Sie diese voneinander ab.
- b) Warum ist ein Schätzer eine Zufallsvariable? Welche Resultate erhält man für einen Schätzer (z.B. S^2), wenn die Stichprobe wiederholt gezogen wird?

Aufgabe 9

Wie lauten die kritischen Werte, oberhalb derer 5% der möglichen Werte liegen (95%-Quantil):

- a) bei der Normalverteilung $N(0, 1)$:
- b) bei der t -Verteilung mit $df = 27$:
- c) bei der χ^2 -Verteilung mit $df = 3$:
- d) bei der F -Verteilung mit df (Zähler) = 2, df (Nenner) = 16:

Skizzieren Sie die Verteilungen und schraffieren Sie die entsprechenden Flächen (Wahrscheinlichkeiten).

16.1.2 Tests für Anteilswerte**Aufgabe 10**

Eine Brauerei produziert ein neues alkoholfreies Bier. In einem Geschmackstest erhalten 150 Personen je ein Glas alkoholfreies bzw. gewöhnliches Bier, und sie sollen versuchen, das alkoholfreie Bier zu identifizieren.

- a) Das gelingt 98 Personen. Testen Sie anhand dieser Daten die Hypothese, alkoholfreies und gewöhnliches Bier seien geschmacklich nicht zu unterscheiden ($\alpha = 0.1$).
- b) Unter den befragten Personen waren 15 Beschäftigte der Brauerei. Von diesen gelingt 9 die richtige Identifizierung. Überprüfen Sie die Hypothese aus a) für diese Subpopulation.

Aufgabe 11

- a) In einem Land regnet es auf lange Sicht an 100 Tagen im Jahr. 2005 regnete es nur an 80 Tagen. Hat sich das Klima signifikant verändert ($\alpha = 0.05$), Binominaltest)?
- b) Diskutieren Sie die Annahmen des Binominaltests für dieses Beispiel.

Aufgabe 12

Der Hersteller eines Medikaments behauptet, daß sie in 90% der Fälle eine Allergie wirksam erleichtere. In einer Stichprobe von 200 Personen, die unter Allergien litten, brachte die Medizin 160 Personen Erleichterung. Ist die Behauptung des Herstellers berechtigt? - Man wähle $\alpha = 0.01$.

Aufgabe 13

Eine Münze wird 12 Mal geworfen. Dabei erscheint 8 Mal Kopf. Testen Sie, ob die Münze symmetrisch ist ($\alpha = 0.05$).

16.1.3 Anpassungstests/Unabhängigkeitstests**Aufgabe 14**

Mendel erhielt bei einem seiner Kreuzungsversuche an Erbsenpflanzen folgende Werte:

315	runde gelbe Erbsen,
108	runde grüne Erbsen,
101	kantige gelbe Erbsen,
32	kantige grüne Erbsen.

Spricht dies für oder gegen die Theorie, dass das Verhältnis der 4 Zahlen 9:3:3:1 sein müsste ($\alpha = 0.05$)?

Aufgabe 15

In einer Statistikklausur erhielt man folgende Kontingenztabelle

	Note				
	1	2	3	4	5
weiblich	14	20	20	4	2
männlich	7	20	11	2	5

- Geben Sie Schätzer an für die bedingte Wahrscheinlichkeit der Noten in den Subpopulationen und vergleichen Sie diese deskriptiv.
- Überprüfen Sie mit einem geeigneten Testverfahren, ob zwischen den Merkmalen Geschlecht und Klausurergebnis ein Zusammenhang besteht.

Aufgabe 16

In einer Umfrage werden 50 Studenten und 50 Nichtstudenten befragt, wie viele Faschingsbälle sie besuchten. Man erhielt

Ballbesuch	viele	wenig	gar keinen	
Studenten	25	15	10	50
Nicht-Studenten	15	20	15	50
				100

Testen Sie die Hypothese, daß sich die Häufigkeit des Ballbesuchs für die beiden Gruppen nicht unterscheiden ($\alpha = 0.1$).

Aufgabe 17

Eine Variable X ist normalverteilt mit unbekanntem Erwartungswert μ und der bekannten Varianz $\sigma^2 = 25$. Zu prüfen ist die Nullhypothese $H_0 : \mu \geq \mu_0 = 1000$ beim Signifikanzniveau $\alpha = 0.01$ und bei einem Stichprobenumfang von $N = 64$.

- Man ermittle die kritische Region des hier zu verwendenden Standardtests.

- b) Man ermittle die Wahrscheinlichkeit eines Fehlers zweiter Art, wenn $\mu = \mu_1 = 998$ ist.
- c) Man ermittle den Stichprobenumfang, der im vorliegenden Fall nötig ist, damit die in b) genannte Wahrscheinlichkeit höchstens 0.05 beträgt.

Aufgabe 18

Bei einer Umfrage zur Kompetenzeinschätzung der Politiker A und B werden folgende Zufallsvariablen betrachtet

$$X = \begin{cases} 1 & A \text{ ist kompetent} \\ 0 & A \text{ ist nicht kompetent} \end{cases}$$
$$Y = \begin{cases} 1 & B \text{ ist kompetent} \\ 0 & B \text{ ist nicht kompetent} \end{cases}$$

Es wird eine Stichprobe von $N = 100$ befragt. 60 Personen halten A für kompetent, 40 Personen halten B für kompetent, 35 Personen halten beide für kompetent.

- a) Geben Sie in einer Kontingenztafel (Kreuztabelle) die gemeinsame (absolute) Häufigkeitsverteilung der Zufallsvariablen X und Y an.
- b) Testen Sie die Hypothese der Unabhängigkeit von X und Y ($\alpha = 0.05$)
- c) Bestimmen Sie die bedingten Häufigkeiten, daß B für kompetent bzw. inkompetent gehalten wird, wenn bekannt ist, dass A für kompetent gehalten wurde.

16.1.4 Tests auf zentrale Tendenz

Aufgabe 19

Die Wirksamkeit einer neuen Unterrichtsmethode soll getestet werden. Kriterium sind die Punktwerte X in einem Leistungstest. Für 10 Personen erhielt man nach Unterricht mit der neuen Methode die Werte

50, 82, 73, 65, 64, 59, 72, 84, 69, 75

- a) Schätzen Sie $E(X)$, $\text{Var}(X)$
- b) Bei Unterrichtung mit der alten Methode war $\mu_0 = 50$ der durchschnittliche Punktwert in der GG. Testen Sie unter Annahme, daß die Werte normalverteilt sind mit bekannter Varianz $\sigma = 9$ die Hypothese, dass die neue Methode eine Verbesserung bringt ($\alpha = 0.05$).
- c) Bearbeiten Sie dasselbe Problem wie unter b), ohne die Annahme, dass die Varianz bekannt ist.
- d) Geben Sie für $E(X)$ ein 90%-Konfidenzintervall an
- d_1) unter den Annahmen von b)
- d_2) unter den Annahmen von c)
- und vergleichen Sie die beiden Konfidenzintervalle.

Aufgabe 20

Bei empirischen Untersuchungen wurde festgestellt, daß das Merkmal X „Körpergröße (gemessen in cm)“ approximativ normalverteilt ist. In einer Stichprobe von $N = 26$ ergab sich ein Mittelwert von $\bar{X} = 181$ und für die Stichprobenvarianz $S^2 = 5.1^2$.

- a) Man prüfe die Hypothese, die durchschnittliche Körpergröße betrage höchstens 180 cm unter der Voraussetzung, daß das Merkmal zwar normalverteilt, die Varianz jedoch nicht bekannt ist ($\alpha = 0.05$).
- Lösungshinweis:* Der Wert der Prüfgröße beträgt 0.9998.
- b) Wie wirken sich eine Verringerung des Signifikanzniveaus α bzw. eine Vergrößerung des Stichprobenumfangs n auf die Wahrscheinlichkeit eines Fehlers zweiter Art bei einem statistischen Test aus (kurze Begründung erforderlich)?

Aufgabe 21

Angenommen, ein Untersucher zieht 3 Hypothesen in Erwägung. Nur eine der 3 kann wahr sein:

$$H_0 : \mu = 200$$

$$H_1 : \mu = 210$$

$$H_2 : \mu = 220$$

Er weiß, daß die Population normalverteilt ist mit $\sigma_X = 20$. Er zieht eine Zufallsstichprobe vom Umfang $N = 25$ und berechnet \bar{X} . Dann formuliert er die Entscheidungsregel: Wenn $\bar{X} \leq 205$ nehme H_0 an; wenn $205 < \bar{X} < 215$ nehme H_1 an; wenn $215 \geq \bar{X}$ nehme H_2 an. Finden Sie die Wahrscheinlichkeiten einer richtigen bzw. falschen Entscheidung unter dieser Regel.

Aufgabe 22

In einer Firma soll festgestellt werden, wieviel Zeit die Mitarbeiter pro Monat aufzuwenden bereit sind, um sich über betriebliche Belange zu informieren. Die Untersuchung soll unter anderem Aufschluss bringen über eventuelle Unterschiede zwischen männlichen und weiblichen Mitarbeitern. Bei $N_1 = 45$ zufällig ausgewählten männlichen Beschäftigten ergaben sich $\bar{x} = 4.3$ (Std.) und $s_x^2 = 1.1^2$, bei $N_2 = 50$ zufällig ausgewählten weiblichen Beschäftigten ergaben sich $\bar{y} = 3.5$ (Std.) und $s_y^2 = 0.9^2$

- Man prüfe die Hypothese, die Erwartungswerte der Untersuchungsmerkmale unterscheiden sich nicht ($\alpha = 0.05$).
Hinweis: Der Wert der Prüfgröße beträgt 3.85.
- Man prüfe die Hypothese, männliche Mitarbeiter informieren sich im Durchschnitt um mindestens 1 Stunde pro Monat länger über betriebliche Belange als weibliche Belegschaftsmitglieder ($\alpha = 0.05$).
Hinweis: Der Wert der Prüfgröße beträgt -0.96.
- Genügen die obigen Angaben, um Teilaufgabe a) mit einem verteilungsfreien Verfahren zu lösen?

Aufgabe 23

Zur Überprüfung der Gedächtnisleistung von Mitarbeitern wurden 30 Begriffe den Personen jeweils zweimal vorgelesen und dann das Untersuchungsmerkmal X : „Anzahl der reproduzierten Begriffe“ registriert. Bei $N = 10$ gleichaltrigen Personen ergaben sich folgende Werte:

17, 12, 13, 16, 9, 19, 21, 12, 18, 17.

An weiteren 12 Mitarbeitern wurde dasselbe Experiment durchgeführt, wobei die Personen jedoch während des gesamten Experiments einer ständigen Lärmbelästigung ausgesetzt waren. Hier ergaben sich die Werte:

10, 6, 15, 9, 8, 11, 8, 16, 13, 7, 5, 14.

- a) Man überprüfe mit Hilfe des Wilcoxon-Rangsummen-Tests die Hypothese, die Lärmbelastigung habe keinen Einfluss auf die Anzahl der reproduzierten Begriffe ($\alpha = 0.05$; zweiseitig).
- b) Prüfen Sie dieselbe Hypothese unter Normalverteilungsannahme für X .

Aufgabe 24

Für den Vergleich zweier Verteilungen sind verschiedene Tests besprochen worden. Diskutieren Sie Gemeinsamkeiten und Unterschiede dieser Tests. Geben Sie für jeden Test eine typische Datensituation an, für die der Test angemessen erscheint.

Aufgabe 25

$N = 10$ zufällig ausgewählte Mitarbeiter sollen zwei Geschicklichkeitsaufgaben A und B vom gleichen Schwierigkeitsgrad lösen. Zuerst wird jedem Mitarbeiter die Aufgabe A vorgelegt und jeweils die Lösungszeit (in Minuten) gemessen. Danach erhält die gesamte Gruppe ein Training, in dessen Verlauf sie lernt, mit ähnlichen, jedoch nicht identischen Geschicklichkeitsaufgaben fertig zu werden. Es soll untersucht werden, ob das Vortraining einen auf die nachfolgende Arbeitsleistung (Lösung der Geschicklichkeitsaufgabe B) begünstigenden Einfluss hat, so daß eine Übertragung der durch Training erworbenen Fähigkeiten auf die nachfolgende Aufgabe B möglich ist (positiver Transfer). Die gemessenen Werte sind in der folgenden Tabelle festgehalten:

Person	Lösungszeit für Aufgabe A (vor dem Training)	Lösungszeit für Aufgabe B (nach dem Training)
1	6.24	6.10
2	5.80	5.75
3	5.57	5.60
4	6.04	5.91
5	5.56	5.30
6	5.92	5.62
7	6.17	6.29
8	5.48	5.28
9	6.11	5.70
10	6.20	6.01

- a) Man prüfe mit Hilfe des Tests von Wilcoxon die Hypothese, das Training habe keinen Einfluss auf das Merkmal Lösungszeit ($\alpha = 0.05$).
- b) Prüfen Sie die gleiche Hypothese mit dem t -Test.

Aufgabe 26

Zu einer Diskussion über die Einführung einer neuen Verkaufsmethode wurden $N = 70$ zufällig ausgewählte Kundenberater eingeladen und vor bzw. nach der Diskussion zu ihrer Einstellung (dafür oder dagegen) über die neue Methode befragt. Es ergab sich, dass vorher 40, nachher 56 Personen für die neue Methode waren, während 36 Berater vor und nach der Diskussion für die neue Methode waren.

- a) Man ermittle mit diesen Angaben eine Kontingenztabelle der Struktur

		<i>Einstellung nach der Diskussion</i>	
		dafür	dagegen
<i>Einstellung vor der Diskussion</i>	dafür		
	dagegen		

- b) Man prüfe die Hypothese, die Einstellung zur neuen Verkaufsmethode habe sich durch die Diskussion nicht geändert ($\alpha = 0.05$).
(McNemar-Test)

16.1.5 Zusammenhangsanalyse

Aufgabe 27

Bei $N = 10$ Kunden wurde die Zufriedenheit mit den Produkten (Variable X) und die Loyalität mit der Firma (Variable Y) ermittelt. Man erhielt die Wertepaare:

X	124	79	118	102	86	89	109	128	114	95
Y	100	94	101	112	76	98	91	73	90	84

- a) Testen Sie die Hypothese der Unabhängigkeit von X und Y unter der Verwendung des Bravais-Pearsonschen Korrelationskoeffizienten ($\alpha = 0.05$).

Hinweis: $\sum x_i^2 = 111548$, $\sum y_i^2 = 85727$, $\sum x_i y_i = 95929$.

- b) Lösen Sie Aufgabe a) unter Verwendung des Rangkorrelationskoeffizienten von Spearman.

Aufgabe 28

Vergleichen Sie in den folgenden Teilaufgaben den

- χ^2 -Unabhängigkeitstest
 - Test auf $\rho_{XY} = 0$ unter Verwendung des Stichprobenkorrelationskoeffizienten.
- a) Geben Sie für jeden der beiden Tests Null- und Alternativhypothesen an. Unter welcher Bedingung sind die beiden Nullhypothesen äquivalent?
- b) Charakterisieren Sie hinsichtlich des Messniveaus der Daten die unterschiedlichen Anwendungsmöglichkeiten der beiden Tests. Lassen sich bei intervallskalierten Daten beide Tests verwenden? Geben Sie einen fiktiven Datensatz an, der sich mit dem einen Test behandeln ließe, mit dem anderen nicht.
- c) Charakterisieren Sie informell den Unterschied zwischen exakter und asymptotischer Verteilung einer Prüfgröße. Welchen Sinn haben hier Faustregeln wie „ $n > 30$ “, „ $np_j \geq 1$ für alle Klassen j “ etc.?

Aufgabe 29

Illustrieren Sie anhand einiger selbstgewählter Beispiele, daß man Kausalzusammenhänge nicht allein auf signifikante Korrelationskoeffizienten stützen kann.

Aufgabe 30

Mit einer Befragungsaktion in Holstein-Unterwasser soll die politische Durchsetzbarkeit eines Kraftwerkbaus überprüft werden. In dieser Region wurden $N = 2000$ zufällig ausgewählte Personen befragt. Für den Kernkraftwerksbau sprachen sich 400 Befragte aus, von denen 350 parteilos waren. Unter den insgesamt 200 parteigebundenen Befragten waren 130 Kernkraftsgegner und 20 hatten „keine Meinung“. Von allen Befragten sprachen sich 1400 gegen das Projekt aus. Es soll statistisch bewiesen werden, daß eine Abhängigkeit zwischen den statistischen Größen

- X : Parteizugehörigkeit mit den Ausprägungen x_1 : „partei- gebunden“ und x_2 : „parteilos“ und
 Y : Einstellung zum Kernkraftwerksbau mit den Aus- prägungen y_1 : „dafür“, x_2 : „dagegen“ und y_3 : „keine Meinung“

besteht.

- Stellen Sie das Befragungsergebnis in einer Kontingenztafel dar.
- Formulieren Sie das Hypothesenpaar und begründen Sie die Wahl der Nullhypothese.
- Führen Sie den Test mit einer Irrtumswahrscheinlichkeit von $\alpha = 0.01$ durch und interpretieren Sie das Ergebnis.

16.1.6 Regressionsanalyse**Aufgabe 31**

Zur Überprüfung der Güte eines Schuleignungstests X zur Vorhersage der Schulreife ermittelt ein Schulpsychologe an $N = 500$ Vorschulkindern die Werte $\bar{x} = 40$ und $s_x = 5$. Nach Ablauf des ersten Schuljahres werden

mit einem Schulleistungstest Y die tatsächlichen Leistungen der Kinder gemessen und man erhielt die folgenden Kennwerte: $\bar{y} = 30$, $s_y = 4$. Die Kovarianz in der Stichprobe zwischen dem Schuleignungstest und dem Schulleistungstest betrug:

$$s_{xy} := \widehat{\text{Cov}}(x, y) = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) = 10$$

- Berechnen Sie einen Korrelationskoeffizienten zwischen Schuleignungs- und Schulleistungstest.
- Prüfen Sie die Alternativhypothese, der Korrelationskoeffizienten sei größer als 0.6 ($\alpha = 0.05$).
- Man ermittle die Stichproben-Regressionsgerade zur Vorhersage der tatsächlichen Leistung nach dem ersten Schuljahr aufgrund des Schuleignungstests.
- Man prüfe, ob der Regressionskoeffizient signifikant von 0 verschieden ist ($\alpha = 0.05$; $\hat{\sigma} = 1.75$).
- Ein Kind erreicht im Eignungstest einen Wert von $x_0 = 48$. Mit welcher schulischen Leistung ist bei ihm zu rechnen?
- Man ermittle ein 95%-Prognoseintervall zu Vorhersage der schulischen Leistung des Kindes.

Aufgabe 32

$N = 10$ Vpn sollen verschiedene Entfernungen schätzen. Die wahren Entfernungen seien X_i , die geschätzten Entfernungen Y_i , $i = 1, \dots, 10$. Man erhielt folgende Wertpaare:

x_i	20	20	30	40	40	60	70	90	90	90
y_i	16	26	25	35	55	62	95	70	110	118

- Zeichnen Sie die Wertpaare (x_i, y_i) in ein Streudiagramm ein.
- Schätzen Sie die Parameter α und β des Regressionsmodells

$$Y_i = \alpha + \beta X_i + \epsilon_i, \quad i = 1, \dots, 10$$

und tragen Sie die Stichproben-Regressionsgerade in das Streudiagramm ein.

Hinweis: $\sum x_i^2 = 37700$, $\sum y_i^2 = 49600$, $\sum x_i y_i = 42380$

- c) Erläutern Sie an diesem Beispiel den Unterschied zwischen Stichprobengrößen und Grundgesamtheitsgrößen.
- d) Ist in diesem Beispiel die übliche Annahme plausibel, daß die Varianz der Abweichung ε_i für alle Entfernungen X_i gleich ist?
- e) Was bedeutet die Hypothese $\alpha = 0, \beta = 1$, also

$$Y_i = X_i + \varepsilon_i, E(\varepsilon_i) = 0, i = 1, \dots, 10?$$

Aufgabe 33

Es soll überprüft werden, ob die sensomotorische Koordinationsfähigkeit von Arbeitnehmern durch Training verbessert werden kann. Das unabhängige Merkmal Trainingszeit an einem Reaktionsgerät wird variiert von 0 Std. bis 5 Std. Bei $N = 30$ Vpn wurden in einem abschließenden Test folgende Fehlerzahlen registriert:

X_i (Std.)	Y_i	X_i (Std.)	Y_i
0	8	3	5
0	10	3	6
0	10	3	6
0	11	3	6
0	9	3	4
1	11	4	6
1	9	4	3
1	8	4	3
1	9	4	4
1	7	4	2
2	8	5	4
2	6	5	2
2	4	5	3
2	6	5	3
2	7	5	2

Rechenhilfen: $\sum x_i = 75, \sum x_i^2 = 275, \sum y_i = 182, \sum x_i y_i = 331, \hat{\sigma} = 1.23.$

- a) Man ermittle die Stichprobenregressionsgerade.
- b) Man zeichne die Stichprobenregressionsgerade.

- c) Man ermittle ein 95%-Konfidenzintervall für β . Ist β signifikant von Null verschieden? (Begründung!)
- d) Man ermittle ein 95%-Konfidenzintervall (Prognoseintervall) für die Fehlerzahl eines Probanden, der 2.5 Std. trainiert.
- e) Geben Sie das Bestimmtheitsmaß an.

Aufgabe 34

Sechs Austauschstudenten, die einen Studienaufenthalt in Deutschland verbrachten, wurden vor und nach dem Aufenthalt über die Einschätzung der Arbeitsintensität der Deutschen im akademischen Bereich gefragt. Angegeben sind im folgenden Testergebnisse, die eine Zusammenfassung diverser Einzelaussagen enthalten. Man erhielt folgende Testergebnisse, wobei X für die Befragungsergebnisse vorher, Y für die Ergebnisse nachher steht:

X	10	15	20	19	30	26
Y	8	10	15	12	25	20

- a) Schätzen Sie die Koeffizienten des linearen Regressionsmodells $Y_i = \alpha + \beta X_i + \epsilon_i$.
Zeichnen Sie die Gerade und die Daten in ein Diagramm.
- b) Überprüfen Sie die Nullhypothese $H_0 : \beta = 1$ in einer zweiseitigen Fragestellung. (Benutzen Sie $\hat{\sigma}^2 = 2.4$).
- c) Berechnen Sie aus den Resultaten in a) die Korrelation zwischen X und Y .
- d) Wieviel Prozent der Varianz von Y läßt sich durch Kenntnis der Werte in der 1. Befragung erklären?
- e) Prognostizieren Sie das Resultat Y eines weiteren Austauschstudenten, wenn er vor dem Studienaufenthalt einen Wert $X = 25$ angegeben hat.
- f) Wie groß ist das 95%-Prognoseintervall für diesen vorhergesagten Testwert?

16.1.7 Varianzanalyse

Aufgabe 35

Es soll untersucht werden, ob die Arbeitsleistung in einem Betrieb durch den am Arbeitsplatz herrschenden Lärm beeinflusst wird. Der Faktor „Lärm“ als unabhängige Variable wird durch drei Bedingungskategorien (Faktorstufen) repräsentiert. Eine Stichprobe von $n = 24$ Personen wurden zufällig auf die 3 Faktorstufen aufgeteilt, so daß jede Gruppe aus 8 Personen bestand. Für die abhängige Variable „Arbeitsleistung“ wurde ein Leistungsindex entwickelt, welcher in guter Näherung normalverteilte Werte liefert. Die Ergebnisse des Versuchs sind in der folgenden Tabelle enthalten:

Faktor „Lärm“		
I	II	III
54	40	18
46	30	27
42	35	18
50	38	22
46	36	20
45	32	22
40	43	21
48	40	15

Übt der Faktor „Lärm“ einen Einfluss auf die abhängige Variable „Arbeitsleistung“ aus ($\alpha = 0.05$)?

Aufgabe 36

Bei einem Medikament hat man den Verdacht, daß es als Nebenwirkung zu verlangsamten Reaktionen führen könnte. Um dies zu prüfen, werden $N = 30$ Vpn in drei Gruppen A, B, C eingeteilt. Gruppe A erhält das Medikament, Gruppe B ein Placebo, Gruppe C gar nichts. In einem Experiment erhielt man als Reaktionszeit (in sec.):

A	1.2	0.7	0.8	0.9	0.6	0.7	0.8	0.9	1.0	0.7
B	1.0	0.5	0.6	0.4	0.7	0.5	0.8	0.4	0.4	0.8
C	0.7	0.4	0.5	0.5	1.0	0.6	0.5	0.5	0.8	0.3

- a) Zunächst beschränken wir uns auf den Vergleich der Gruppen A und C.
Vergleichen Sie mit einem einseitigen Test die Mittelwerte dieser beiden Gruppen ($\alpha = 0.05$).
Hinweis: $s_1^2 = 0.031$, $s_2^2 = 0.042$
Unterstellen Sie Varianzhomogenität und Normalverteilung. Geben Sie Null- und Alternativ-Hypothese an und legen Sie diese so, dass Sie einen sinnvollen Test zur Klärung des obigen Verdachts erhalten.
- b) Prüfen Sie die Hypothese, die Populations-Mittelwerte der Reaktionszeiten seien für alle drei Gruppen gleich ($\alpha = 0.05$).

Aufgabe 37

Bei einer Waldschadensuntersuchung wird in 15 Waldstücken gleicher Größe die Anzahl stark geschädigter Bäume erhoben. Die Waldstücke sind so gewählt, daß sie sich hinsichtlich der klimatischen Bedingungen unterscheiden. Man erhielt folgende Daten:

Klima I	7	8	12	13	10
Klima II	10	7	8	13	12
Klima III	14	18	13	19	16

- a) Überprüfen Sie mit einem geeigneten Testverfahren, ob die klimatischen Bedingungen einen Einfluss auf die Anzahl stark geschädigter Bäume haben ($\alpha = 0.05$). Geben Sie an, welche Voraussetzungen in Ihre Analyse eingehen.
- b) Prüfen Sie bei Einhaltung eines gemeinsamen Signifikanzniveaus von ($\alpha = 0.05$), welche Klimabedingungen sich voneinander unterscheiden.
- c) Angenommen, Sie wissen nur, ob die Anzahl geschädigter Bäume an den drei Standorten jeweils unter 12 oder mindestens 12 beträgt. Wie lässt sich ein möglicher Unterscheid der Standorte untersuchen? Formulieren Sie eine Nullhypothese und geben Sie eine geeignete Teststatistik an. Welcher Nachteil entsteht im Verhältnis zur Auswertungsmethode in a)?

16.2 Aufgaben KE 2

Aufgabe 1

- a) Ein Fragebogen besteht aus 3 items A, B, C. Wie viele verschiedene Fragebögen lassen sich daraus bilden?
- b) Ergeben diese das gleiche Ergebnis, wenn sie jeweils einer vergleichbaren Stichprobe von Probanden vorgelegt werden?
- c) Diskutieren Sie das Problem der Kontextabhängigkeit von Antworten (responses) sowie das Problem der Reaktivität von Messungen. Sind analoge Phänomene auch in der Naturwissenschaft bekannt?

Aufgabe 2

- a) Welche Phasen werden beim Forschungsprozeß der (quantifizierenden) empirischen Forschung unterschieden?
- a) Erklären Sie die Begriffe Indikatoren, Operationalisierung, Falsifikatoren, Tautologie und Kontradiktion.
- b) Was ist der Unterschied zwischen einer Beobachtungs(Survey)-Studie und einem geplanten Experiment. Was ist bei der Analyse von Wirkungszusammenhängen zu beachten, insbesondere bei der Modellierung der unabhängigen Variablen (UV).
- c) Was sind Störvariablen und wie wirken sie sich aus?
- d) Kann die sog. Validität einer Variable auch empirisch überprüft bzw. quantifiziert werden?

Aufgabe 3

Welche der folgenden Aussagen sind richtig?

- A Der empirische Gehalt einer Hypothese hängt von dem Grad ihrer Falsifizierbarkeit ab.
- B Die Falsifizierbarkeit einer Hypothese ist unabhängig von ihrem Wahrheitsgehalt.
- C Die Operationalisierung ist eine Methode, die durch Ausprobieren aller oder vieler möglicher Fälle Lösungen von Problemen sucht.
- D Ein Interviewer erhält eine nach seiner zugrundegelegten Theorie unerwartete Antwort. Mittels des Exhaustionsprinzips kann der Interviewer eine nach seiner Theorie erwartete Antwort erhalten, indem er Zusatzbedingungen einfügt.
- E Keine der Aussagen A - D ist richtig.

Aufgabe 4

- a) Erläutern Sie bitte das Kontrapositionsgesetz

$$(A \Rightarrow B) \iff (\bar{B} \Rightarrow \bar{A})$$

an einem Beispiel.

- b) Berechnen Sie $\overline{A \wedge B}$ (\bar{A} = nicht A).

Interpretation am Beispiel aus a).

- c) Was ergibt sich aus der Falsifikation der Aussage eines Satzsystems

$$T \wedge A \wedge R \Rightarrow O,$$

wobei T die Theorie, A Zusatzhypothesen und R Randbedingungen sind.

- d) Kann ein Theorie wirklich falsifiziert werden?

Aufgabe 5

Gegeben sei folgendes Gesetz:

Wenn ein Verhalten X_1 unterdrückt wird, dann erhöht sich die Auftretenswahrscheinlichkeit eines bestimmten Ereignisses Y_1 .

Welche der folgenden Aussagen widerlegt das Gesetz? (x aus 5)

- A Ein Verhalten X_2 wird nicht unterdrückt und die Auftretenswahrscheinlichkeit des Ereignisses Y_1 erhöht sich.
- B Ein Verhalten X_2 wird nicht unterdrückt und die Auftretenswahrscheinlichkeit des Ereignisses Y_1 erhöht sich nicht.
- C Das Verhalten X_1 wird unterdrückt und die Auftretenswahrscheinlichkeit des Ereignisses Y_1 erhöht sich nicht.
- D Das Verhalten X_1 wird nicht unterdrückt und die Auftretenswahrscheinlichkeit des Ereignisses Y_1 erhöht sich nicht.
- E Keine der Aussagen A - D ist richtig.

Aufgabe 6

Gegeben sind drei Hypothesen:

- I Je höher der Schulabschluß ist, desto höher ist das Einkommen und die Firmenposition.
- II Je höher der Schulabschluß ist, desto höher ist das Einkommen.
- III Je höher der Schulabschluß und das Engagement ist, desto höher ist das Einkommen und die Firmenposition.

Welche der folgenden Aussagen sind richtig?

- A Allgemein gilt: Der Informationsgehalt einer Hypothese steigt, wenn weitere Elemente in die Dann-Komponente aufgenommen werden (Wenn A, dann B und C).
- B Allgemein gilt: Der Informationsgehalt einer Hypothese steigt, wenn weitere Elemente in die Wenn-Komponente aufgenommen werden (Wenn A und B, dann C).
- C Hypothese III besitzt den größten Informationsgehalt.
- D Hypothese II besitzt den größten Informationsgehalt.

E Hypothese I besitzt den größten Informationsgehalt.

Aufgabe 7

Welche der folgenden Aussagen sind richtig?

- A Die Reliabilität ist der Anteil der geschätzten Varianz an der Gesamtvarianz.
- B Die Reliabilität setzt die Varianz des wahren Wertes mit der Varianz des beobachteten Wertes ins Verhältnis.
- C Die Reliabilität entspricht dem Quadrat der Korrelation des beobachteten Wertes mit dem wahren Wert.
- D Da der wahre Wert nicht bekannt ist, kann die Reliabilität nicht berechnet, sondern lediglich geschätzt werden.
- E Keine der Aussagen A - D ist richtig.

Aufgabe 8

Welche der folgenden Aussagen sind richtig?

- A Die Reliabilität stellt neben der Validität und der Repräsentativität eines der drei wichtigsten Gütekriterien für empirische Untersuchungen dar.
- B Die Höhe der Validität ist unabhängig von der Höhe der Reliabilität.
- C Unter der Voraussetzung von gleichen Bedingungen ist die Reliabilität ein Maß für die Replizierbarkeit von Meßwerten.
- D Die Reliabilität nimmt Werte zwischen -1 und 1 an.
- E Keine der Aussagen A - D ist richtig.

Aufgabe 9

Welche der folgenden Aussagen sind richtig?

- A Die Validität ist ein Maß für die inhaltliche Gültigkeit einer Untersuchung.
- B Eine hohe Validität impliziert eine hohe Reliabilität und umgekehrt.
- C Mit Hilfe der Reliabilität können systematische Fehler identifiziert werden.
- D Nimmt die Reliabilität den Wert 1 an, so liegt kein Meßfehler vor.
- E Keine der Aussagen A - D ist richtig.

Aufgabe 10

Welche der folgenden Aussagen sind richtig?

- A Ein statistisches Maß für die Interrater-Reliabilität ist der Kappa(κ)-Koeffizient von Cohen, der die Vergleichbarkeit verschiedener Probanden misst.
- B Kappa ist die Differenz der Nichtübereinstimmungen der Beurteiler (zufällige minus tatsächliche) bezogen auf die zufällige Nichtübereinstimmung.

C

$$\kappa = \frac{G(+)-G(-)}{1-G(-)} = \frac{\sum_{i=1}^I f_{ii} - \sum_{i=1}^I f_{i.} \cdot f_{.i}}{1 - \sum_{i=1}^I f_{i.} \cdot f_{.i}}$$

- D Kappa nimmt Werte in dem Bereich von -1 bis 1 an.
- E Keine der Aussagen A - D ist richtig.

16.3 Aufgaben KE 3

Aufgabe 1

a) Was ist der Unterschied zwischen der Datenansicht und der Variablenansicht eines SPSS-Datensatzes **Daten.sav**?

b) Erklären Sie den Zweck der Spalteneinträge:

Name, Typ, Variablenlabel, Spaltenlabel, fehlende Werte, Messniveau
in der Variablenansicht

Aufgabe 2

a) Was ist eine Datenmatrix, welche Dimensionen hat sie und wie wird sie in Statistik-Programmen repräsentiert?

b) Welche statistische Annahme wird im allgemeinen für verschiedene Zeilen der Datenmatrix gemacht?

c) Wie kann ein SPSS-Datensatz durch manuelle Eingabe erzeugt werden?

d) Die Variable **Alter** wurde als stetige (kontinuierliche Variable) erhoben. Sie interessieren sich aber nur für die Altersklassen (jung/mittel/alt). Wie gehen Sie vor, um eine neue Variable **Alterklassiert** zu erzeugen und im Datensatz abzuspeichern?

Aufgabe 3

a) Welche graphischen Darstellungen und Tabellen werden bei einer explorativen Analyse erstellt?

b) Was ist der Unterschied zwischen einem Stabdiagramm und einem Histogramm, auch im Hinblick auf das Skalenniveau der Variable.

- c) Warum ist es sinnvoll, dichotome Variablen mit den Werten 0/1 zu codieren (Indikatoren, Dummy-Variablen, 0-1-Variablen).
- d) Wie läßt sich der Mittelwert einer Indikatorvariable interpretieren?

Aufgabe 4

- a) Nennen Sie Zusammenhangsmaße für nominale Variablen X und Y .
- b) Wie wird der λ -Koeffizient (Goodman-Kruskal) in einer Kreuztabelle berechnet. Warum ist dies ein a posteriori-Maß?
- c) Was sind a priori-Zusammenhangsmaße?
- d) Berechnen Sie Cohen- κ aus der folgenden Kreuztabelle f_{ij} der Urteile A, B, C von 2 Beobachtern, die ein Ereignis klassifizieren sollten (Hypothese: exakte Übereinstimmung der Beobachter):

f_{ij}	A	B	C
A	.3	.1	0
B	.2	.2	0
C	0	0	.1

- e) Berechnen Sie aus obiger Kreuztabelle die Koeffizienten $\lambda(x \rightarrow y)$, $\lambda(y \rightarrow x)$ sowie $\lambda(x \leftrightarrow y)$. Vergleichen Sie mit κ .

Aufgabe 5

- a) Wie lassen sich die Parameter der Regressionsgleichung

$$Y_n = \alpha + \beta X_n + \epsilon_n$$

interpretieren, wenn die Variable X_n als Indikator (0/1) codiert wird. Setzen Sie die Werte ein und schreiben Sie die beiden Gleichungen getrennt auf.

- b) Interpretieren Sie die Parameter der Interaktionsterme in

$$Y_n = \alpha + \beta_1 X_{1n} + \beta_2 X_{2n} + \beta_{12} X_{1n} X_{2n} + \epsilon_n$$

wenn X_{1n} eine Indikatorvariable und X_{2n} eine stetige Regressorvariable ist. Stellen Sie sich vor, daß die Variable X_{1n} zwei Gruppen unterscheidet.

c) kann das Problem auch mit einem t -Test bearbeitet werden (warum)?

Aufgabe 6

a) Geben Sie Beispiele für die Skalenniveaus *Nominal*, *Ordinal*, *Intervall*, *Verhältnis*.

b) Welche Transformationen lassen die Skala invariant?

c) Welche mathematischen Operationen sind für derartige Skalen sinnvoll?

Aufgabe 7

a) Wie unterscheiden sich die Gesichtspunkte *qualitativ/quantitativ*, *diskret/metrisch* und *kategorial/kontinuierlich* voneinander?

b) Welche Statistiken lassen sich bei nominalem Skalenniveau *sinnvollerweise* berechnen?

c) Schulnoten werden üblicherweise zu Durchschnittswerten gemittelt. Unter welchen Voraussetzungen ist dies sinnvoll?

Aufgabe 8

a) Erläutern Sie das Kontrapositionsgesetz

$$(A \Rightarrow B) \iff (\bar{B} \Rightarrow \bar{A})$$

anhand der Aussage:

Korrelierte Variablen sind abhängig.

- b) Was gilt bei unkorrelierten Variablen? Sind diese unabhängig?
- c) Wie sind die Verhältnisse bei bivariat normalverteilten $(X, Y) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$?

Aufgabe 9

- a) Klassifizieren Sie das asymmetrische Modell $X \rightarrow Y$ nach dem Skalenniveau der Variablen.
- b) Wie läßt sich der Quotient aus erklärter und totaler Streuung SQE/SQT beim linearen Regressionsmodell alternativ berechnen und wie kann er interpretiert werden?
- c) Was wird beim globalen F -Test der Regressionsanalyse überprüft, insbesondere bei mehreren Regressoren X_1, \dots, X_q ?

Aufgabe 10

- a) Was versteht man unter einer Scheinkorrelation zwischen den Variablen X und Y ?
- b) Wie kann man den ‘tatsächlichen’ Zusammenhang zwischen den Variablen (näherungsweise) ermitteln?
- c) Was ist eine Kontrollvariable und eine partielle Korrelation?
- d) Berechnen Sie die partielle Korrelation zwischen X und Y bei Kenntnis der Korrelationen $r_{xy} = .7, r_{xz} = .6, r_{yz} = .3$ mit der Kontrollvariablen Z .
- e) Wie kann die Variable Y bei Kenntnis von Z optimal (im Quadratmittel) prognostiziert werden?
- f) Wann ist die lineare Prognose optimal?
Prognostizieren Sie Y linear aus der Kenntnis von $E[Y] = 1, \text{Corr}(X, Y) = 0.7$ und $S_X = S_Y = 1$.

Aufgabe 11

- a) Wie hängen die allgemeine Spearman-Brown-Formel und Cronbachs α zusammen?
- b) Berechnen Sie aus der durchschnittlichen Korrelation $\bar{r} = 0.5$ von 5 items die interne Konsistenz der Skala. Welchen Wert hätte man bei 10 items erhalten?
- c) Was ist bei Berechnung der Retest-Reliabilität zu berücksichtigen?
- d) Was verstehen Sie unter der Abschwächungskorrektur?
Berechnen Sie die Korrelation der Konstrukte T_1 und T_2 , wenn die Korrelation der Messungen 0.3 beträgt und eine Reliabilität von 0.7 vorliegt.

Aufgabe 12

- a) Berechnen Sie die item-Schwierigkeit bei einer 5er-Skala, die zwischen 1 und 5 codiert ist und deren item-Mittelwert 4.7 beträgt.
- b) Welche Trennschärfe hat ein item, das mit dem Gesamtwert eine Korrelation von 0.71 aufweist?
- c) Was kann man tun, wenn ein item eine negative Trennschärfe aufweist?
- d) Wie kann die Dimensionalität einer Item-Batterie untersucht werden?
Wann kann ein Gesamtwert sinnvoll berechnet werden?

Aufgabe 13

- a) Was ist der Unterschied zwischen dem Meßmodell der klassischen Testtheorie und dem Modell der Faktorenanalyse?
- b) Schreiben Sie das Modell der Faktorenanalyse auf und interpretieren Sie die Bestandteile (Bezeichnungen, stochastische Annahmen).

- c) Wie können die (latenten) Faktorwerte geschätzt werden?
- d) Berechnen Sie die Kommunalitäten aus der Ladungsmatrix

$$\mathbf{\Lambda} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ .5 & 1 \end{bmatrix}$$

- e) Was wird bei einer Hauptkomponentenanalyse durchgeführt? Erklären Sie die Begriffe Eigenwert, Eigenvektor und Spektral(Eigenwert)zerlegung.
- f) Was sind die Hauptkomponenten eines Zufallsvektors $\mathbf{x} = (x_1, \dots, x_p)$ und wie können Sie interpretiert werden?
- g) Was ist der Unterschied zwischen Hauptkomponentenanalyse und Maximum-Likelihood(ML)-Faktorenanalyse?
- h) Was ist bei einer ML-Faktorenanalyse zu beachten im Hinblick auf die Zahl der manifesten Variablen und der Faktoren?

Aufgabe 14

- a) Was verstehen Sie unter einem Mosaik-Diagramm und welche Größen werden tabelliert?
- b) Was ist ein stem-leaf-Diagramm und wie unterscheidet es sich von einem Histogramm?
- c) Wie lassen sich gruppierte Daten (etwa nach Geschlecht) übersichtlich graphisch darstellen?

Aufgabe 15

- a) Warum können 0/1-codierte und stetige Variablen in einer gemeinsamen Korrelationsmatrix dargestellt werden?

- b) Wie ist die Korrelation $\text{Corr}(X, Y)$ zwischen einer Indikatorvariablen $X = (0/1)$ und einer stetigen Variablen Y zu interpretieren?
- c) Lassen sich Indikatorvariablen und stetige Variablen gemeinsam in einer Regressionsanalyse verwenden?
- d) Wie gehen Sie vor, wenn die nominale Variable X mehr als 2 Ausprägungen hat (etwa Berufsgruppe mit den Ausprägungen Schüler, Arbeiter/Angestellte, Hausfrau/mann, ...)

Kapitel 17

Tabellen

17.1 Binomialverteilung

Man betrachtet die n -malige unabhängige Wiederholung eines Zufallsexperiments, bei dem man sich für das Auftreten (bzw. Nichtauftreten) des Ereignisses A interessiert ($P(A) = \pi$; $P(\bar{A}) = 1 - \pi$).

Die Zufallsvariable X gibt die Anzahl des Auftretens von A bei n -maliger unabhängiger Durchführung des Zufallsexperiments an ($P(X = x) = f_X(x)$).

Wahrscheinlichkeitsfunktion

$$f_X(x) = \binom{n}{x} \pi^x (1 - \pi)^{n-x} \quad x = 0, \dots, n$$

Verteilungsfunktion

$$F_X(x) = \sum_{k=0}^x \binom{n}{k} \pi^k (1 - \pi)^{n-k} \quad x = 0, 1, \dots, n$$

Eine binomialverteilte Zufallsvariable

$$X \sim B(n, \pi)$$

besitzt den Erwartungswert und die Varianz

Erwartungswert

$$E(X) = n\pi$$

Varianz

$$\text{Var}(X) = n\pi(1 - \pi)$$

Additionseigenschaft

Ist $X \sim B(n, \pi)$ und $Y \sim B(m, \pi)$ so ist $X + Y \sim B(n + m, \pi)$ (man beachte, daß dabei für X und Y die Wahrscheinlichkeiten π gleich sein müssen).

Für $n\pi \geq 5$ und $n(1 - \pi) \geq 5$ ist eine $B(n, \pi)$ -verteilte Zufallsvariable näherungsweise normalverteilt mit den Parametern $\mu = n\pi$ und $\sigma^2 = n\pi(1 - \pi)$.

Die Tabelle der Binomialverteilung enthält für $n = 1$ bis $n = 20$ die Werte der Wahrscheinlichkeitsfunktion für $\pi = 0,05$; $\pi = 0,1$; $\pi = 0,15$; ...; $\pi = 0,5$.

Tabelle der Binomialverteilung

Für Werte $\pi > 0,5$ erhält man die Werte der Wahrscheinlichkeitsfunktion aus der Tabelle für $\pi^* = 1 - \pi$, indem man anstelle von x für die Werte $n - x$ verwendet.

n	x	$\pi = 0,05$		$\pi = 0,10$		$\pi = 0,15$		$\pi = 0,20$		$\pi = 0,25$		$\pi = 0,30$		$\pi = 0,35$		$\pi = 0,40$		$\pi = 0,45$		$\pi = 0,50$	
		f(x)	F(x)																		
1	0	0,9500	0,9500	0,9000	0,9000	0,8500	0,8500	0,8000	0,8000	0,7500	0,7500	0,7000	0,7000	0,6500	0,6500	0,6000	0,6000	0,5500	0,5500	0,5000	0,5000
1	1	0,0500	1,0000	0,1000	1,0000	0,1500	1,0000	0,2000	1,0000	0,2500	1,0000	0,3000	1,0000	0,3500	1,0000	0,4000	1,0000	0,4500	1,0000	0,5000	1,0000
2	0	0,9025	0,9025	0,8100	0,8100	0,7225	0,7225	0,6400	0,6400	0,5625	0,5625	0,4900	0,4900	0,4225	0,4225	0,3600	0,3600	0,3025	0,3025	0,2500	0,2500
2	1	0,0950	0,9975	0,1800	0,9900	0,2550	0,9775	0,3200	0,9600	0,3750	0,9375	0,4200	0,9100	0,4550	0,8775	0,4800	0,8400	0,4950	0,7975	0,5000	0,7500
2	2	0,0025	1,0000	0,0100	1,0000	0,0225	1,0000	0,0400	1,0000	0,0625	1,0000	0,0900	1,0000	0,1225	1,0000	0,1600	1,0000	0,2025	1,0000	0,2500	1,0000
3	0	0,8574	0,8574	0,7290	0,7290	0,6141	0,6141	0,5120	0,5120	0,4219	0,4219	0,3430	0,3430	0,2746	0,2746	0,2160	0,2160	0,1664	0,1664	0,1250	0,1250
3	1	0,1354	0,9928	0,2430	0,9720	0,3251	0,9393	0,3840	0,8960	0,4219	0,8438	0,4410	0,7840	0,4436	0,7183	0,4320	0,6480	0,4084	0,5748	0,3750	0,5000
3	2	0,0071	0,9999	0,0270	0,9990	0,0574	0,9966	0,0960	0,9920	0,1406	0,9844	0,1890	0,9730	0,2389	0,9571	0,2880	0,9360	0,3341	0,9089	0,3750	0,8750
3	3	0,0001	1,0000	0,0010	1,0000	0,0034	1,0000	0,0080	1,0000	0,0156	1,0000	0,0270	1,0000	0,0429	1,0000	0,0640	1,0000	0,0911	1,0000	0,1250	1,0000
4	0	0,8145	0,8145	0,6561	0,6561	0,5220	0,5220	0,4096	0,4096	0,3164	0,3164	0,2401	0,2401	0,1785	0,1785	0,1296	0,1296	0,0915	0,0915	0,0625	0,0625
4	1	0,1715	0,9860	0,2916	0,9477	0,3685	0,8905	0,4096	0,8192	0,4219	0,7383	0,4116	0,6517	0,3845	0,5630	0,3456	0,4752	0,2995	0,3910	0,2500	0,3125
4	2	0,0135	0,9995	0,0486	0,9963	0,0975	0,9880	0,1536	0,9728	0,2109	0,9492	0,2646	0,9163	0,3105	0,8735	0,3456	0,8208	0,3675	0,7585	0,3750	0,6875
4	3	0,0005	1,0000	0,0036	0,9999	0,0115	0,9995	0,0256	0,9984	0,0469	0,9961	0,0756	0,9919	0,1115	0,9850	0,1536	0,9744	0,2005	0,9590	0,2500	0,9375
4	4	0,0000	1,0000	0,0001	1,0000	0,0005	1,0000	0,0016	1,0000	0,0039	1,0000	0,0081	1,0000	0,0150	1,0000	0,0256	1,0000	0,0410	1,0000	0,0625	1,0000
5	0	0,7738	0,7738	0,5905	0,5905	0,4437	0,4437	0,3277	0,3277	0,2373	0,2373	0,1681	0,1681	0,1160	0,1160	0,0778	0,0778	0,0503	0,0503	0,0313	0,0313
5	1	0,2036	0,9774	0,3281	0,9185	0,3915	0,8352	0,4096	0,7373	0,3955	0,6328	0,3602	0,5282	0,3124	0,4284	0,2592	0,3370	0,2059	0,2562	0,1563	0,1875
5	2	0,0214	0,9988	0,0729	0,9914	0,1382	0,9734	0,2048	0,9421	0,2637	0,8965	0,3087	0,8369	0,3364	0,7648	0,3456	0,6826	0,3369	0,5931	0,3125	0,5000
5	3	0,0011	1,0000	0,0081	0,9995	0,0244	0,9978	0,0512	0,9933	0,0879	0,9844	0,1323	0,9692	0,1811	0,9460	0,2304	0,9130	0,2757	0,8688	0,3125	0,8125
5	4	0,0000	1,0000	0,0005	1,0000	0,0022	0,9999	0,0064	0,9997	0,0146	0,9990	0,0284	0,9976	0,0488	0,9947	0,0768	0,9898	0,1128	0,9815	0,1563	0,9688
5	5	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0003	1,0000	0,0010	1,0000	0,0024	1,0000	0,0053	1,0000	0,0102	1,0000	0,0185	1,0000	0,0313	1,0000
6	0	0,7351	0,7351	0,5314	0,5314	0,3771	0,3771	0,2621	0,2621	0,1780	0,1780	0,1176	0,1176	0,0754	0,0754	0,0467	0,0467	0,0277	0,0277	0,0156	0,0156
6	1	0,2321	0,9672	0,3543	0,8857	0,3993	0,7765	0,3932	0,6554	0,3560	0,5339	0,3025	0,4202	0,2437	0,3191	0,1866	0,2333	0,1359	0,1636	0,0938	0,1094
6	2	0,0305	0,9978	0,0984	0,9842	0,1762	0,9527	0,2458	0,9011	0,2966	0,8306	0,3241	0,7443	0,3280	0,6471	0,3110	0,5443	0,2780	0,4415	0,2344	0,3438
6	3	0,0021	0,9999	0,0146	0,9987	0,0415	0,9941	0,0819	0,9830	0,1318	0,9624	0,1852	0,9295	0,2355	0,8826	0,2765	0,8208	0,3032	0,7447	0,3125	0,6563
6	4	0,0001	1,0000	0,0012	0,9999	0,0055	0,9996	0,0154	0,9984	0,0330	0,9954	0,0595	0,9891	0,0951	0,9777	0,1382	0,9590	0,1861	0,9308	0,2344	0,8906
6	5	0,0000	1,0000	0,0001	1,0000	0,0004	1,0000	0,0015	0,9999	0,0044	0,9998	0,0102	0,9993	0,0205	0,9982	0,0369	0,9959	0,0609	0,9917	0,0938	0,9844
6	6	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0002	1,0000	0,0007	1,0000	0,0018	1,0000	0,0041	1,0000	0,0083	1,0000	0,0156	1,0000
7	0	0,6983	0,6983	0,4783	0,4783	0,3206	0,3206	0,2097	0,2097	0,1335	0,1335	0,0824	0,0824	0,0490	0,0490	0,0280	0,0280	0,0152	0,0152	0,0078	0,0078
7	1	0,2573	0,9556	0,3720	0,8503	0,3960	0,7166	0,3670	0,5767	0,3115	0,4449	0,2471	0,3294	0,1848	0,2338	0,1306	0,1586	0,0872	0,1024	0,0547	0,0625
7	2	0,0406	0,9962	0,1240	0,9743	0,2097	0,9262	0,2753	0,8520	0,3115	0,7564	0,3177	0,6471	0,2985	0,5323	0,2613	0,4199	0,2140	0,3164	0,1641	0,2266
7	3	0,0036	0,9998	0,0230	0,9973	0,0617	0,9879	0,1147	0,9667	0,1730	0,9294	0,2269	0,8740	0,2679	0,8002	0,2903	0,7102	0,2918	0,6083	0,2734	0,5000
7	4	0,0002	1,0000	0,0026	0,9998	0,0109	0,9988	0,0287	0,9953	0,0577	0,9871	0,0972	0,9712	0,1442	0,9444	0,1935	0,9037	0,2388	0,8471	0,2734	0,7734
7	5	0,0000	1,0000	0,0002	1,0000	0,0012	0,9999	0,0043	0,9996	0,0115	0,9987	0,0250	0,9962	0,0466	0,9910	0,0774	0,9812	0,1172	0,9643	0,1641	0,9375
7	6	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0004	1,0000	0,0013	0,9999	0,0036	0,9998	0,0084	0,9994	0,0172	0,9984	0,0320	0,9963	0,0547	0,9922
7	7	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0002	1,0000	0,0006	1,0000	0,0016	1,0000	0,0037	1,0000	0,0078	1,0000

n	$\pi = 0,05$		$\pi = 0,10$		$\pi = 0,15$		$\pi = 0,20$		$\pi = 0,25$		$\pi = 0,30$		$\pi = 0,35$		$\pi = 0,40$		$\pi = 0,45$		$\pi = 0,50$		
	$f(x)$	$F(x)$																			
8	0	0,6634	0,4305	0,4305	0,2725	0,2725	0,1678	0,1678	0,1001	0,1001	0,0576	0,0576	0,0319	0,0319	0,0168	0,0168	0,0084	0,0084	0,0039	0,0039	
8	1	0,2793	0,9428	0,8131	0,3847	0,6572	0,3355	0,5033	0,2670	0,3671	0,1977	0,2553	0,1373	0,1691	0,0896	0,1064	0,0548	0,0632	0,0313	0,0352	
8	2	0,0515	0,9942	0,1488	0,9619	0,2376	0,8948	0,2936	0,7969	0,3115	0,6785	0,5518	0,2587	0,4278	0,2090	0,3154	0,1569	0,2201	0,1094	0,1445	
8	3	0,0054	0,9996	0,0331	0,9950	0,0839	0,9786	0,1468	0,9437	0,2076	0,8862	0,2541	0,8059	0,2786	0,7064	0,2568	0,4770	0,2188	0,3633	0,6333	
8	4	0,0004	1,0000	0,0046	0,9996	0,0185	0,9971	0,0459	0,9896	0,0865	0,9727	0,1361	0,9420	0,1875	0,8939	0,2827	0,7396	0,2734	0,6367	0,9667	
8	5	0,0000	1,0000	0,0004	1,0000	0,0026	0,9998	0,0092	0,9988	0,0231	0,9958	0,0467	0,9887	0,0808	0,9747	0,1239	0,9502	0,2188	0,8555	0,9655	
8	6	0,0000	1,0000	0,0000	1,0000	0,0002	1,0000	0,0011	0,9999	0,0038	0,9996	0,0100	0,9987	0,0217	0,9964	0,0413	0,9915	0,1094	0,9648	0,9961	
8	7	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0004	1,0000	0,0012	0,9999	0,0033	0,9998	0,0079	0,9993	0,0313	0,9961	0,9961	
8	8	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0002	1,0000	0,0007	1,0000	0,0039	1,0000	1,0000	
9	0	0,6302	0,6302	0,3874	0,3874	0,2316	0,2316	0,1342	0,1342	0,0751	0,0751	0,0404	0,0404	0,0207	0,0207	0,0101	0,0101	0,0046	0,0046	0,0020	0,0020
9	1	0,2985	0,9288	0,3874	0,7748	0,3679	0,5995	0,3020	0,4362	0,2253	0,3003	0,1556	0,1960	0,1004	0,1211	0,0605	0,0339	0,0385	0,0176	0,0195	
9	2	0,0629	0,9916	0,1722	0,9470	0,2597	0,8591	0,3020	0,7382	0,3003	0,6007	0,2668	0,4628	0,2162	0,3373	0,1612	0,2318	0,1495	0,0703	0,0898	
9	3	0,0077	0,9994	0,0446	0,9917	0,1069	0,9661	0,1762	0,9144	0,2336	0,8343	0,2668	0,7297	0,2716	0,6089	0,2508	0,4826	0,2119	0,1641	0,2539	
9	4	0,0006	1,0000	0,0074	0,9991	0,0283	0,9944	0,0661	0,9804	0,1168	0,9511	0,1715	0,9012	0,2194	0,8283	0,2508	0,7334	0,2600	0,2461	0,5000	
9	5	0,0000	1,0000	0,0008	0,9999	0,0050	0,9994	0,0165	0,9969	0,0389	0,9900	0,0735	0,9747	0,1181	0,9464	0,1672	0,9006	0,2128	0,2461	0,7461	
9	6	0,0000	1,0000	0,0001	1,0000	0,0006	1,0000	0,0028	0,9997	0,0087	0,9987	0,0210	0,9957	0,0424	0,9888	0,0743	0,9750	0,1160	0,1641	0,9102	
9	7	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0003	1,0000	0,0012	0,9999	0,0039	0,9996	0,0098	0,9986	0,0212	0,9962	0,0407	0,0703	0,9805	
9	8	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0004	1,0000	0,0013	0,9999	0,0035	0,9997	0,0083	0,0176	0,9980	
9	9	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0003	1,0000	0,0008	0,0020	1,0000	
10	0	0,5987	0,5987	0,3487	0,3487	0,1969	0,1969	0,1074	0,1074	0,0563	0,0563	0,0282	0,0282	0,0135	0,0135	0,0060	0,0060	0,0025	0,0010	0,0010	
10	1	0,3151	0,9139	0,3874	0,7361	0,3474	0,5443	0,2684	0,3758	0,1877	0,2440	0,1211	0,1493	0,0725	0,0860	0,0403	0,0464	0,0207	0,0233	0,0107	
10	2	0,0746	0,9885	0,1937	0,9298	0,2759	0,8202	0,3020	0,6778	0,2816	0,5256	0,2335	0,3828	0,1757	0,2616	0,1209	0,1673	0,0763	0,0439	0,0547	
10	3	0,0105	0,9990	0,0574	0,9872	0,1298	0,9500	0,2013	0,8791	0,2503	0,7759	0,2668	0,6496	0,2522	0,5138	0,2150	0,3823	0,1665	0,1172	0,1719	
10	4	0,0010	0,9999	0,0112	0,9984	0,0401	0,9901	0,0881	0,9672	0,1460	0,9219	0,2001	0,8497	0,2377	0,7515	0,2508	0,6331	0,2384	0,2051	0,3770	
10	5	0,0001	1,0000	0,0015	0,9999	0,0085	0,9986	0,0264	0,9936	0,0584	0,9803	0,1029	0,9527	0,1536	0,9051	0,2007	0,8338	0,2340	0,2461	0,6230	
10	6	0,0000	1,0000	0,0001	1,0000	0,0012	0,9999	0,0055	0,9991	0,0162	0,9965	0,0368	0,9894	0,0689	0,9740	0,1115	0,9452	0,1596	0,2051	0,8281	
10	7	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0008	0,9999	0,0031	0,9996	0,0090	0,9984	0,0212	0,9952	0,0425	0,9877	0,0746	0,1172	0,9453	
10	8	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0004	1,0000	0,0014	0,9999	0,0043	0,9995	0,0106	0,9983	0,0229	0,0439	0,9893	
10	9	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0005	1,0000	0,0016	0,9999	0,0042	0,0098	0,9990	
10	10	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0003	0,0010	1,0000	
11	0	0,5688	0,5688	0,3138	0,3138	0,1673	0,1673	0,0859	0,0859	0,0422	0,0422	0,0198	0,0198	0,0088	0,0088	0,0036	0,0036	0,0014	0,0005	0,0005	
11	1	0,3293	0,8981	0,3835	0,6974	0,3248	0,4922	0,2362	0,3221	0,1549	0,1971	0,0932	0,1130	0,0518	0,0606	0,0266	0,0302	0,0125	0,0139	0,0059	
11	2	0,0867	0,9848	0,2131	0,9104	0,2866	0,7788	0,2953	0,6174	0,2581	0,4552	0,1998	0,3127	0,1395	0,2001	0,0887	0,1189	0,0513	0,0652	0,0327	
11	3	0,0137	0,9984	0,0710	0,9815	0,1517	0,9306	0,2215	0,8389	0,2581	0,7133	0,2568	0,5696	0,2254	0,4256	0,1774	0,2963	0,1259	0,1911	0,1133	
11	4	0,0014	0,9999	0,0158	0,9972	0,0536	0,9841	0,1107	0,9496	0,1721	0,8854	0,2201	0,7897	0,2428	0,6683	0,2365	0,5328	0,2060	0,3971	0,2744	

n	x	$\pi = 0,05$		$\pi = 0,10$		$\pi = 0,15$		$\pi = 0,20$		$\pi = 0,25$		$\pi = 0,30$		$\pi = 0,35$		$\pi = 0,40$		$\pi = 0,45$		$\pi = 0,50$	
		$f(x)$	$F(x)$																		
11	5	0,0001	1,0000	0,0025	0,9997	0,0132	0,9973	0,0388	0,9883	0,0803	0,9657	0,1321	0,9218	0,1830	0,8513	0,2207	0,7535	0,2360	0,6331	0,2256	0,5000
11	6	0,0000	1,0000	0,0003	1,0000	0,0023	0,9997	0,0097	0,9980	0,0268	0,9924	0,0566	0,9784	0,0985	0,9499	0,1471	0,9006	0,1931	0,8262	0,2256	0,7256
11	7	0,0000	1,0000	0,0000	1,0000	0,0003	1,0000	0,0017	0,9998	0,0064	0,9988	0,0173	0,9957	0,0379	0,9878	0,0701	0,9707	0,1128	0,9390	0,1611	0,8867
11	8	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0002	1,0000	0,0011	0,9999	0,0037	0,9994	0,0102	0,9980	0,0234	0,9941	0,0462	0,9852	0,0806	0,9673
11	9	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0005	1,0000	0,0018	0,9998	0,0052	0,9993	0,0126	0,9978	0,0269	0,9941
11	10	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0002	1,0000	0,0007	1,0000	0,0021	0,9998	0,0054	0,9995
11	11	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0002	1,0000	0,0005	1,0000
12	0	0,5404	0,5404	0,2824	0,2824	0,1422	0,1422	0,0687	0,0687	0,0317	0,0317	0,0138	0,0138	0,0057	0,0057	0,0022	0,0022	0,0008	0,0008	0,0002	0,0002
12	1	0,3413	0,8816	0,3766	0,6590	0,3012	0,4435	0,2062	0,2749	0,1267	0,1584	0,0712	0,0850	0,0368	0,0424	0,0174	0,0196	0,0075	0,0083	0,0029	0,0032
12	2	0,0988	0,9804	0,2301	0,8891	0,2924	0,7358	0,2835	0,5583	0,2323	0,3907	0,1678	0,2528	0,1088	0,1513	0,0639	0,0834	0,0339	0,0421	0,0161	0,0193
12	3	0,0173	0,9978	0,0852	0,9744	0,1720	0,9078	0,2362	0,7946	0,2581	0,6488	0,2397	0,4925	0,1954	0,3467	0,1419	0,2253	0,0923	0,1345	0,0537	0,0730
12	4	0,0021	0,9998	0,0213	0,9957	0,0683	0,9761	0,1329	0,9274	0,1936	0,8424	0,2311	0,7237	0,2367	0,5833	0,2128	0,4382	0,1700	0,3044	0,1208	0,1938
12	5	0,0002	1,0000	0,0038	0,9995	0,0193	0,9954	0,0532	0,9806	0,1032	0,9456	0,1585	0,8822	0,2039	0,7873	0,2270	0,6652	0,2225	0,5269	0,1934	0,3872
12	6	0,0000	1,0000	0,0005	0,9999	0,0040	0,9993	0,0155	0,9961	0,0401	0,9857	0,0792	0,9614	0,1281	0,9154	0,1766	0,8418	0,2124	0,7393	0,2256	0,6128
12	7	0,0000	1,0000	0,0000	1,0000	0,0006	0,9999	0,0033	0,9994	0,0115	0,9972	0,0291	0,9905	0,0591	0,9745	0,1009	0,9427	0,1489	0,8883	0,1934	0,8062
12	8	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0005	0,9999	0,0024	0,9996	0,0078	0,9983	0,0199	0,9944	0,0420	0,9847	0,0762	0,9644	0,1208	0,9270
12	9	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0004	1,0000	0,0015	0,9998	0,0048	0,9992	0,0125	0,9972	0,0277	0,9921	0,0537	0,9807
12	10	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0002	1,0000	0,0008	0,9999	0,0025	0,9997	0,0068	0,9989	0,0161	0,9968
12	11	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0003	1,0000	0,0010	0,9999	0,0029	0,9998
12	12	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0002	1,0000
13	0	0,5133	0,5133	0,2542	0,2542	0,1209	0,1209	0,0550	0,0550	0,0238	0,0238	0,0097	0,0097	0,0037	0,0037	0,0013	0,0013	0,0004	0,0004	0,0001	0,0001
13	1	0,3512	0,8646	0,3672	0,6213	0,2774	0,3983	0,1787	0,2336	0,1029	0,1267	0,0540	0,0637	0,0259	0,0296	0,0113	0,0126	0,0045	0,0049	0,0016	0,0017
13	2	0,1109	0,9755	0,2448	0,8661	0,2937	0,6920	0,2680	0,5017	0,2059	0,3326	0,1388	0,2025	0,0836	0,1132	0,0453	0,0579	0,0220	0,0269	0,0095	0,0112
13	3	0,0214	0,9969	0,0997	0,9658	0,1900	0,8820	0,2457	0,7473	0,2517	0,5843	0,2181	0,4206	0,1651	0,2783	0,1107	0,1686	0,0660	0,0929	0,0349	0,0461
13	4	0,0028	0,9997	0,0277	0,9935	0,0838	0,9658	0,1535	0,9009	0,2097	0,7940	0,2337	0,6543	0,2222	0,5005	0,1845	0,3530	0,1350	0,2279	0,0873	0,1334
13	5	0,0003	1,0000	0,0055	0,9991	0,0266	0,9925	0,0691	0,9700	0,1258	0,9198	0,1803	0,8346	0,2154	0,7159	0,2214	0,5744	0,1989	0,4268	0,1571	0,2905
13	6	0,0000	1,0000	0,0008	0,9999	0,0063	0,9987	0,0230	0,9930	0,0559	0,9757	0,1030	0,9376	0,1546	0,8705	0,1968	0,7712	0,2169	0,6437	0,2095	0,5000
13	7	0,0000	1,0000	0,0001	1,0000	0,0011	0,9998	0,0058	0,9988	0,0186	0,9944	0,0442	0,9818	0,0833	0,9538	0,1312	0,9023	0,1775	0,8212	0,2095	0,7095
13	8	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0011	0,9998	0,0047	0,9990	0,0142	0,9960	0,0336	0,9874	0,0656	0,9679	0,1089	0,9302	0,1571	0,8666
13	9	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0009	0,9999	0,0034	0,9993	0,0101	0,9975	0,0243	0,9922	0,0495	0,9797	0,0873	0,9539
13	10	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0006	0,9999	0,0022	0,9997	0,0065	0,9987	0,0162	0,9959	0,0349	0,9888
13	11	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0003	1,0000	0,0012	0,9999	0,0036	0,9995	0,0095	0,9983
13	12	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0005	1,0000	0,0016	0,9999
13	13	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000
14	0	0,4877	0,4877	0,2288	0,2288	0,1028	0,1028	0,0440	0,0440	0,0178	0,0178	0,0068	0,0068	0,0024	0,0024	0,0008	0,0008	0,0002	0,0002	0,0001	0,0001
14	1	0,3593	0,8470	0,3559	0,5846	0,2539	0,3567	0,1539	0,1979	0,0832	0,1010	0,0407	0,0475	0,0181	0,0205	0,0073	0,0081	0,0027	0,0029	0,0009	0,0009
14	2	0,1229	0,9699	0,2570	0,8416	0,2912	0,6479	0,2501	0,4481	0,1802	0,2811	0,1134	0,1608	0,0634	0,0839	0,0317	0,0398	0,0141	0,0170	0,0056	0,0065

n	$\pi = 0,05$		$\pi = 0,10$		$\pi = 0,15$		$\pi = 0,20$		$\pi = 0,25$		$\pi = 0,30$		$\pi = 0,35$		$\pi = 0,40$		$\pi = 0,45$		$\pi = 0,50$	
	$f(x)$	$F(x)$																		
14	0,0259	0,9958	0,1142	0,9559	0,2056	0,8535	0,2501	0,6982	0,2402	0,5213	0,1943	0,3552	0,1366	0,2205	0,0845	0,1243	0,0462	0,0632	0,0222	0,0287
14	0,0037	0,9996	0,0349	0,9908	0,0998	0,9533	0,1720	0,8702	0,2202	0,7415	0,2290	0,5842	0,2022	0,4227	0,1549	0,2793	0,1040	0,1672	0,0611	0,0898
14	0,0004	1,0000	0,0078	0,9985	0,0352	0,9885	0,0860	0,9561	0,1468	0,8883	0,1963	0,7805	0,2178	0,6405	0,2066	0,4859	0,1701	0,3373	0,1222	0,2120
14	0,0000	1,0000	0,0013	0,9998	0,0093	0,9978	0,0322	0,9884	0,0734	0,9617	0,1262	0,9067	0,1759	0,8164	0,2066	0,6925	0,2088	0,5461	0,1833	0,3953
14	0,0000	1,0000	0,0002	1,0000	0,0019	0,9997	0,0092	0,9976	0,0280	0,9897	0,0618	0,9685	0,1082	0,9247	0,1574	0,8499	0,1952	0,7414	0,2095	0,6047
14	0,0000	1,0000	0,0000	1,0000	0,0003	1,0000	0,0020	0,9996	0,0082	0,9978	0,0232	0,9917	0,0510	0,9757	0,0918	0,9417	0,1398	0,8811	0,1833	0,7880
14	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0003	1,0000	0,0018	0,9997	0,0066	0,9983	0,0183	0,9940	0,0408	0,9825	0,0762	0,9574	0,1222	0,9102
14	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0003	1,0000	0,0000	0,9998	0,0049	0,9989	0,0136	0,9961	0,0312	0,9886	0,0611	0,9713
14	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0002	1,0000	0,0010	0,9999	0,0033	0,9994	0,0093	0,9978	0,0222	0,9935
14	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0005	0,9999	0,0019	0,9997	0,0056	0,9991
14	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0002	1,0000	0,0009	0,9999
14	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000
15	0,4633	0,2059	0,2059	0,0874	0,0874	0,0874	0,0352	0,0352	0,0134	0,0134	0,0047	0,0047	0,0016	0,0016	0,0005	0,0005	0,0001	0,0001	0,0000	0,0000
15	0,3658	0,8290	0,3432	0,5490	0,2312	0,3186	0,1319	0,1671	0,0668	0,0802	0,0305	0,0353	0,0126	0,0142	0,0047	0,0052	0,0016	0,0017	0,0005	0,0005
15	0,1348	0,9638	0,2669	0,8159	0,2856	0,6042	0,2309	0,3980	0,1559	0,2361	0,0916	0,1268	0,0476	0,0617	0,0219	0,0271	0,0090	0,0107	0,0032	0,0037
15	0,0307	0,9945	0,1285	0,9444	0,2184	0,8227	0,2501	0,6482	0,2252	0,4613	0,1700	0,2969	0,1110	0,1727	0,0634	0,0905	0,0318	0,0424	0,0139	0,0176
15	0,0049	0,9994	0,0428	0,9873	0,1156	0,9383	0,1876	0,8358	0,2252	0,6865	0,2186	0,5155	0,1792	0,3519	0,1268	0,2173	0,0780	0,1204	0,0417	0,0592
15	0,0006	0,9999	0,0105	0,9978	0,0449	0,9832	0,1032	0,9389	0,1651	0,8216	0,2061	0,7216	0,2123	0,5643	0,1859	0,4032	0,1404	0,2608	0,0916	0,1509
15	0,0000	1,0000	0,0019	0,9997	0,0132	0,9964	0,0430	0,9819	0,0917	0,9434	0,1472	0,8689	0,1906	0,7548	0,2066	0,6098	0,1914	0,4522	0,1527	0,3036
15	0,0000	1,0000	0,0003	1,0000	0,0030	0,9994	0,0138	0,9958	0,0393	0,9827	0,0811	0,9500	0,1319	0,8868	0,1771	0,7869	0,2013	0,6535	0,1964	0,5000
15	0,0000	1,0000	0,0000	1,0000	0,0005	0,9999	0,0035	0,9992	0,0131	0,9958	0,0348	0,9848	0,0710	0,9578	0,1181	0,9050	0,1647	0,8182	0,1964	0,6964
15	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0007	0,9999	0,0034	0,9992	0,0116	0,9963	0,0298	0,9876	0,0612	0,9662	0,1048	0,9231	0,1527	0,8491
15	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0007	0,9999	0,0030	0,9993	0,0096	0,9972	0,0245	0,9907	0,0515	0,9745	0,0916	0,9408
15	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0006	0,9999	0,0024	0,9995	0,0074	0,9981	0,0191	0,9937	0,0417	0,9824
15	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0004	0,9999	0,0016	0,9997	0,0052	0,9989	0,0139	0,9963
15	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0003	1,0000	0,0010	0,9999	0,0032	0,9995
15	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0005	1,0000
16	0,4401	0,1853	0,1853	0,0743	0,0743	0,0743	0,0281	0,0281	0,0100	0,0100	0,0033	0,0033	0,0010	0,0010	0,0003	0,0003	0,0001	0,0001	0,0000	0,0000
16	0,3706	0,8108	0,3294	0,5147	0,2097	0,2839	0,1126	0,1407	0,0535	0,0635	0,0228	0,0261	0,0087	0,0098	0,0030	0,0033	0,0009	0,0010	0,0002	0,0003
16	0,1463	0,9571	0,2745	0,7892	0,2775	0,5614	0,2111	0,3518	0,1336	0,1971	0,0732	0,0994	0,0353	0,0451	0,0150	0,0183	0,0056	0,0066	0,0018	0,0021
16	0,0359	0,9930	0,1423	0,9316	0,2285	0,7899	0,2463	0,5981	0,2079	0,4050	0,1465	0,2459	0,0888	0,1339	0,0468	0,0651	0,0215	0,0281	0,0085	0,0106
16	0,0061	0,9991	0,0514	0,9830	0,1311	0,9209	0,2001	0,7982	0,2252	0,6302	0,2040	0,4499	0,1553	0,2892	0,1014	0,1666	0,0572	0,0853	0,0278	0,0384
16	0,0008	0,9999	0,0137	0,9967	0,0555	0,9765	0,1201	0,9183	0,1802	0,8103	0,2099	0,6598	0,2008	0,4900	0,1623	0,3288	0,1123	0,1976	0,0667	0,1051
16	0,0001	1,0000	0,0028	0,9995	0,0180	0,9944	0,0650	0,9733	0,1101	0,9204	0,1649	0,8247	0,1982	0,6881	0,1983	0,5272	0,1684	0,3660	0,1222	0,2272
16	0,0000	1,0000	0,0004	0,9999	0,0045	0,9989	0,0197	0,9930	0,0524	0,9729	0,1010	0,9256	0,1524	0,8406	0,1889	0,7161	0,1969	0,5629	0,1746	0,4018
16	0,0000	1,0000	0,0001	1,0000	0,0009	0,9998	0,0055	0,9985	0,0197	0,9925	0,0487	0,9743	0,0923	0,9329	0,1417	0,8577	0,1812	0,7441	0,1964	0,5982
16	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0012	0,9998	0,0058	0,9984	0,0185	0,9929	0,0442	0,9771	0,0840	0,9417	0,1318	0,8759	0,1746	0,7728

n	$\pi = 0,05$		$\pi = 0,10$		$\pi = 0,15$		$\pi = 0,20$		$\pi = 0,25$		$\pi = 0,30$		$\pi = 0,35$		$\pi = 0,40$		$\pi = 0,45$		$\pi = 0,50$	
	f(x)	F(x)																		
16	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0002	1,0000	0,0014	0,9997	0,0056	0,9984	0,0167	0,9938	0,0392	0,9809	0,0755	0,9514	0,1222	0,8949
16	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0002	1,0000	0,0013	0,9997	0,0049	0,9987	0,0142	0,9951	0,0337	0,9851	0,0667	0,9616
16	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0002	1,0000	0,0011	0,9998	0,0040	0,9991	0,0115	0,9965	0,0278	0,9894
16	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0002	1,0000	0,0008	0,9999	0,0029	0,9994	0,0085	0,9979
16	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0005	0,9999	0,0018	0,9997
16	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0002	1,0000
16	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000
17	0,4181	0,1668	0,1668	0,0631	0,0631	0,0225	0,0225	0,0225	0,0075	0,0075	0,0023	0,0023	0,0007	0,0007	0,0002	0,0002	0,0000	0,0000	0,0000	0,0000
17	0,3741	0,7922	0,3150	0,4818	0,1893	0,2525	0,0957	0,1182	0,0426	0,0501	0,0169	0,0193	0,0060	0,0067	0,0019	0,0021	0,0005	0,0006	0,0001	0,0001
17	0,1575	0,9497	0,2800	0,7618	0,2673	0,5198	0,1914	0,3096	0,1136	0,1637	0,0581	0,0774	0,0260	0,0327	0,0102	0,0123	0,0035	0,0041	0,0010	0,0012
17	0,0415	0,9912	0,1556	0,9174	0,2359	0,7556	0,2393	0,5489	0,1893	0,3530	0,1245	0,2019	0,0701	0,1028	0,0341	0,0464	0,0144	0,0184	0,0052	0,0064
17	0,0076	0,9988	0,0605	0,9779	0,1457	0,9013	0,2093	0,7582	0,2209	0,5739	0,1868	0,3887	0,1320	0,2348	0,0796	0,1260	0,0411	0,0596	0,0182	0,0245
17	0,0010	0,9999	0,0175	0,9953	0,0668	0,9681	0,1361	0,8943	0,1914	0,7653	0,2081	0,5968	0,1849	0,4197	0,1379	0,2639	0,0875	0,1471	0,0472	0,0717
17	0,0001	1,0000	0,0039	0,9992	0,0236	0,9917	0,0680	0,9623	0,1276	0,8929	0,1784	0,7752	0,1991	0,6188	0,1839	0,4478	0,1432	0,2902	0,0944	0,1662
17	0,0000	1,0000	0,0007	0,9999	0,0065	0,9983	0,0267	0,9891	0,0668	0,9598	0,1201	0,8954	0,1685	0,7872	0,1927	0,6405	0,1841	0,4743	0,1484	0,3145
17	0,0000	1,0000	0,0001	1,0000	0,0014	0,9997	0,0084	0,9974	0,0279	0,9876	0,0644	0,9597	0,1134	0,9006	0,1606	0,8011	0,1883	0,6626	0,1855	0,5000
17	0,0000	1,0000	0,0000	1,0000	0,0003	1,0000	0,0021	0,9995	0,0093	0,9969	0,0276	0,9873	0,0611	0,9617	0,1070	0,9081	0,1540	0,8166	0,1855	0,6855
17	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0004	0,9999	0,0025	0,9994	0,0095	0,9968	0,0263	0,9880	0,0571	0,9652	0,1008	0,9174	0,1484	0,8338
17	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0005	0,9999	0,0026	0,9993	0,0090	0,9970	0,0242	0,9894	0,0525	0,9699	0,0944	0,8283
17	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0006	0,9999	0,0024	0,9994	0,0081	0,9975	0,0215	0,9914	0,0472	0,9755
17	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0005	0,9999	0,0021	0,9995	0,0068	0,9981	0,0182	0,9936
17	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0004	0,9999	0,0016	0,9997	0,0052	0,9988
17	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0003	1,0000	0,0010	0,9999
17	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000
17	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000
18	0,3972	0,3972	0,1501	0,0536	0,0536	0,0180	0,0180	0,0180	0,0056	0,0056	0,0016	0,0016	0,0004	0,0004	0,0001	0,0001	0,0000	0,0000	0,0000	0,0000
18	0,3763	0,7735	0,3002	0,4503	0,1704	0,2241	0,0811	0,0991	0,0338	0,0395	0,0126	0,0142	0,0042	0,0046	0,0012	0,0013	0,0003	0,0003	0,0001	0,0001
18	0,1683	0,9419	0,2835	0,7338	0,2556	0,4797	0,1723	0,2713	0,0958	0,1353	0,0458	0,0600	0,0190	0,0236	0,0069	0,0082	0,0022	0,0025	0,0006	0,0007
18	0,0473	0,9891	0,1680	0,9018	0,2406	0,7202	0,2297	0,5010	0,1704	0,3057	0,1046	0,1646	0,0547	0,0783	0,0246	0,0328	0,0095	0,0120	0,0031	0,0038
18	0,0093	0,9985	0,0700	0,9718	0,1592	0,8794	0,2153	0,7164	0,2130	0,5187	0,1681	0,3327	0,1104	0,1886	0,0614	0,0942	0,0291	0,0411	0,0117	0,0154
18	0,0014	0,9998	0,0218	0,9936	0,0787	0,9581	0,1507	0,8671	0,1988	0,7175	0,2017	0,5344	0,1664	0,3550	0,1146	0,2088	0,0666	0,1077	0,0327	0,0481
18	0,0002	1,0000	0,0052	0,9988	0,0301	0,9882	0,0816	0,9487	0,1436	0,8610	0,1873	0,7217	0,1941	0,5491	0,1655	0,3743	0,1181	0,2258	0,0708	0,1189
18	0,0000	1,0000	0,0010	0,9998	0,0091	0,9973	0,0350	0,9837	0,0820	0,9431	0,1376	0,8593	0,1792	0,7283	0,1892	0,5634	0,1657	0,3915	0,1214	0,2403
18	0,0000	1,0000	0,0002	1,0000	0,0022	0,9995	0,0120	0,9957	0,0376	0,9807	0,0811	0,9404	0,1327	0,8609	0,1734	0,7368	0,1864	0,5778	0,1669	0,4073
18	0,0000	1,0000	0,0000	1,0000	0,0004	0,9999	0,0033	0,9991	0,0139	0,9946	0,0386	0,9790	0,0794	0,9403	0,1284	0,8653	0,1694	0,7473	0,1855	0,5927
18	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0008	0,9998	0,0042	0,9988	0,0149	0,9939	0,0385	0,9788	0,0771	0,9424	0,1248	0,8720	0,1669	0,7597
18	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0010	0,9998	0,0046	0,9986	0,0151	0,9938	0,0374	0,9797	0,0742	0,9463	0,1214	0,8811

n	x	$\pi = 0,05$		$\pi = 0,10$		$\pi = 0,15$		$\pi = 0,20$		$\pi = 0,25$		$\pi = 0,30$		$\pi = 0,35$		$\pi = 0,40$		$\pi = 0,45$		$\pi = 0,50$	
		$f(x)$	$F(x)$																		
18	12	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0002	1,0000	0,0012	0,9997	0,0047	0,9986	0,0145	0,9942	0,0354	0,9817	0,0708	0,9519
18	13	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0002	1,0000	0,0012	0,9997	0,0045	0,9987	0,0134	0,9951	0,0327	0,9846
18	14	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0002	1,0000	0,0011	0,9998	0,0039	0,9990	0,0117	0,9962
18	15	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0002	1,0000	0,0009	0,9999	0,0031	0,9993
18	16	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0006	0,9999
18	17	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000
18	18	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000
19	0	0,3774	0,3774	0,1351	0,1351	0,0456	0,0456	0,0144	0,0144	0,0042	0,0042	0,0011	0,0011	0,0003	0,0003	0,0001	0,0001	0,0000	0,0000	0,0000	0,0000
19	1	0,3774	0,7547	0,2852	0,4203	0,1529	0,1985	0,0685	0,0829	0,0268	0,0310	0,0093	0,0104	0,0029	0,0031	0,0008	0,0008	0,0002	0,0002	0,0000	0,0000
19	2	0,1787	0,9335	0,2852	0,7054	0,2428	0,4413	0,1540	0,2369	0,0803	0,1113	0,0358	0,0462	0,0138	0,0170	0,0046	0,0055	0,0013	0,0015	0,0003	0,0004
19	3	0,0533	0,9868	0,1796	0,8850	0,2428	0,6841	0,2182	0,4551	0,1517	0,2631	0,0869	0,1332	0,0422	0,0591	0,0175	0,0230	0,0062	0,0077	0,0018	0,0022
19	4	0,0112	0,9980	0,0798	0,9648	0,1714	0,8556	0,2182	0,6733	0,2023	0,4654	0,1491	0,2822	0,0909	0,1500	0,0467	0,0696	0,0203	0,0280	0,0074	0,0096
19	5	0,0018	0,9998	0,0266	0,9914	0,0907	0,9463	0,1636	0,8369	0,2023	0,6678	0,1916	0,4739	0,1468	0,2968	0,0933	0,1629	0,0497	0,0777	0,0222	0,0318
19	6	0,0002	1,0000	0,0069	0,9983	0,0374	0,9837	0,0955	0,9324	0,1574	0,8251	0,1916	0,6655	0,1844	0,4812	0,1451	0,3081	0,0949	0,1727	0,0518	0,0835
19	7	0,0000	1,0000	0,0014	0,9997	0,0122	0,9959	0,0443	0,9767	0,0974	0,9225	0,1525	0,8180	0,1844	0,6656	0,1797	0,4878	0,1443	0,3169	0,0961	0,1796
19	8	0,0000	1,0000	0,0002	1,0000	0,0032	0,9992	0,0166	0,9933	0,0487	0,9713	0,0981	0,9161	0,1489	0,8145	0,1797	0,6675	0,1771	0,4940	0,1442	0,3238
19	9	0,0000	1,0000	0,0000	1,0000	0,0007	0,9999	0,0051	0,9984	0,0198	0,9911	0,0514	0,9674	0,0980	0,9125	0,1464	0,8139	0,1771	0,6710	0,1762	0,5000
19	10	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0013	0,9997	0,0066	0,9977	0,0220	0,9895	0,0528	0,9653	0,0976	0,9115	0,1449	0,8159	0,1762	0,6762
19	11	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0003	1,0000	0,0018	0,9995	0,0077	0,9972	0,0233	0,9886	0,0532	0,9648	0,0970	0,9129	0,1442	0,8204
19	12	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0004	0,9999	0,0022	0,9994	0,0083	0,9969	0,0237	0,9884	0,0529	0,9658	0,0961	0,9165
19	13	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0005	0,9999	0,0024	0,9993	0,0085	0,9969	0,0233	0,9891	0,0518	0,9682
19	14	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0006	0,9999	0,0024	0,9994	0,0082	0,9972	0,0222	0,9904
19	15	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0005	0,9999	0,0222	0,9995	0,0074	0,9978
19	16	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0005	0,9999	0,0018	0,9996
19	17	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0001	1,0000	0,0003	1,0000
19	18	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000
19	19	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000	1,0000
20	0	0,3585	0,3585	0,1216	0,1216	0,0388	0,0388	0,0115	0,0115	0,0032	0,0032	0,0008	0,0008	0,0002	0,0002	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
20	1	0,3774	0,7358	0,2702	0,3917	0,1368	0,1756	0,0576	0,0692	0,0211	0,0243	0,0068	0,0076	0,0020	0,0021	0,0005	0,0005	0,0001	0,0001	0,0000	0,0000
20	2	0,1887	0,9245	0,2852	0,6769	0,2293	0,4049	0,1369	0,2061	0,0669	0,0913	0,0278	0,0355	0,0100	0,0121	0,0031	0,0036	0,0008	0,0009	0,0002	0,0002
20	3	0,0596	0,9841	0,1901	0,8670	0,2428	0,6477	0,2054	0,4114	0,1339	0,2252	0,0716	0,1071	0,0323	0,0444	0,0123	0,0160	0,0040	0,0049	0,0011	0,0013
20	4	0,0133	0,9974	0,0898	0,9568	0,1821	0,8298	0,1742	0,6296	0,1897	0,4148	0,1304	0,2375	0,0738	0,1182	0,0350	0,0510	0,0139	0,0189	0,0046	0,0059
20	5	0,0022	0,9997	0,0319	0,9887	0,1028	0,9327	0,2182	0,8042	0,2023	0,6172	0,1789	0,4164	0,1272	0,2454	0,0746	0,1256	0,0365	0,0553	0,0148	0,0207
20	6	0,0003	1,0000	0,0089	0,9976	0,0454	0,9781	0,1091	0,9133	0,1686	0,7858	0,1916	0,6080	0,1712	0,4166	0,1244	0,2500	0,0746	0,1299	0,0370	0,0577
20	7	0,0000	1,0000	0,0020	0,9996	0,0160	0,9941	0,0545	0,9679	0,1124	0,8982	0,1643	0,7723	0,1844	0,6010	0,1659	0,4159	0,1221	0,2520	0,0739	0,1316
20	8	0,0000	1,0000	0,0004	0,9999	0,0046	0,9987	0,0222	0,9900	0,0609	0,9591	0,1144	0,8867	0,1614	0,7624	0,1797	0,5956	0,1623	0,4143	0,1201	0,2517
20	9	0,0000	1,0000	0,0001	1,0000	0,0011	0,9998	0,0074	0,9974	0,0271	0,9861	0,0654	0,9520	0,1158	0,8782	0,1597	0,7553	0,1771	0,5914	0,1602	0,4119

17.2 Normalverteilung

Dichtefunktion $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$

Verteilungsfunktion $F_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{\xi-\mu^2}{2\sigma^2}\right) d\xi$

Erwartungswert $E(X) = \mu$

Varianz $\text{Var}(X) = \sigma^2$

Parameter der Normalverteilung μ, σ^2

Man schreibt häufig $(X \sim N(\mu, \sigma^2))$ und spricht von einer $N(\mu, \sigma^2)$ -verteilten Zufallsvariablen.

- (1) Ist die Zufallsvariable $X \sim N(\mu, \sigma^2)$, so ist die Zufallsvariable $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$. Es gilt also: Eine lineare Funktion einer normalverteilten Zufallsvariablen ist wieder eine normalverteilte Zufallsvariable. Aus dieser Eigenschaft folgt für $a = \frac{1}{\sigma}$ und $b = -\frac{\mu}{\sigma}$ speziell:

Standard-normalverteilung

Ist die Zufallsvariable $X \sim N(\mu, \sigma^2)$, so ist

$$Z = \frac{X}{\sigma} - \frac{\mu}{\sigma} = \frac{X - \mu}{\sigma}$$

$N(0, 1)$ -verteilt.

Eine Normalverteilung mit den Parametern $\mu = 0$ und $\sigma^2 = 1$ heißt **Standardnormalverteilung**.

Ist die Zufallsvariable $X \sim N(\mu, \sigma^2)$ und ist $Z \sim N(0, 1)$, so folgt

$$P(x_1 \leq X \leq x_2) = P\left(\frac{x_1 - \mu}{\sigma} \leq Z \leq \frac{x_2 - \mu}{\sigma}\right).$$

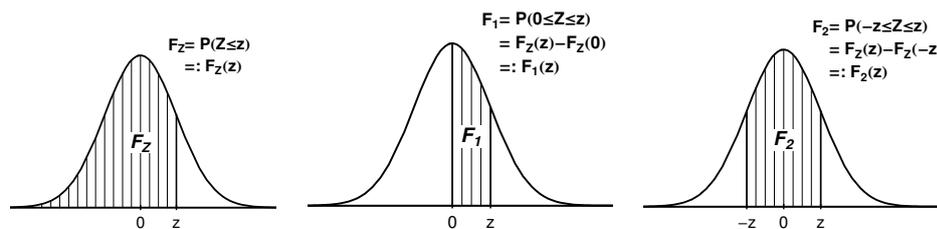
- (2) Sind X_1 und X_2 unabhängige Zufallsvariablen und $N(\mu_1, \sigma_1^2)$ - bzw. $N(\mu_2, \sigma_2^2)$ -verteilt, so ist die Summe der Zufallsvariablen $X = a_1X_1 + a_2X_2 + b \sim N(a_1\mu_1 + a_2\mu_2 + b, a_1^2\sigma_1^2 + a_2^2\sigma_2^2)$ wieder normalverteilt.

Um die Handhabung und Anwendung von Tabellen der Normalverteilung zu erleichtern, gibt es verschiedene Tabellen, in denen zu gegebenen z -Werten folgende Wahrscheinlichkeiten angegeben sind ($z = \frac{x-\mu}{\sigma}$):

Tabelle der Normalverteilung

- (1) $F_Z : P(Z \leq z)$ bzw. $P(X \leq \mu + z\sigma)$
- (2) $F_1 : P(0 \leq Z \leq z)$ bzw. $P(\mu \leq X \leq \mu + z\sigma)$
- (3) $F_2 : P(-z \leq Z \leq z)$ bzw. $P(\mu - z\sigma \leq X \leq \mu + z\sigma)$

Durch die letzte Tabelle werden die Werte der Verteilungsfunktion gegeben. Die folgende Abbildung gibt eine Veranschaulichung der Tabellenwerte:



In der angeführten Tabelle sind F_Z , F_1 und F_2 -Werte dargestellt.

Wegen der Symmetrie der Normalverteilung reicht es aus, die F_1 , F_2 und F_Z -Werte nur für positive Z -Werte zu tabellieren.

http://www.fernuni-hagen.de/ls_statistik/lehre/

z	F_Z	F_1	F_2	z	F_Z	F_1	F_2	z	F_Z	F_1	F_2
0,01	0,5040	0,0040	0,0080	0,51	0,6950	0,1950	0,3899	1,01	0,8438	0,3438	0,6875
0,02	0,5080	0,0080	0,0160	0,52	0,6985	0,1985	0,3969	1,02	0,8461	0,3461	0,6923
0,03	0,5120	0,0120	0,0239	0,53	0,7019	0,2019	0,4039	1,03	0,8485	0,3485	0,6970
0,04	0,5160	0,0160	0,0319	0,54	0,7054	0,2054	0,4108	1,04	0,8508	0,3508	0,7017
0,05	0,5199	0,0199	0,0399	0,55	0,7088	0,2088	0,4177	1,05	0,8531	0,3531	0,7063
0,06	0,5239	0,0239	0,0478	0,56	0,7123	0,2123	0,4245	1,06	0,8554	0,3554	0,7109
0,07	0,5279	0,0279	0,0558	0,57	0,7157	0,2157	0,4313	1,07	0,8577	0,3577	0,7154
0,08	0,5319	0,0319	0,0638	0,58	0,7190	0,2190	0,4381	1,08	0,8599	0,3599	0,7199
0,09	0,5359	0,0359	0,0717	0,59	0,7224	0,2224	0,4448	1,09	0,8621	0,3621	0,7243
0,10	0,5398	0,0398	0,0797	0,60	0,7257	0,2257	0,4515	1,10	0,8643	0,3643	0,7287
0,11	0,5438	0,0438	0,0876	0,61	0,7291	0,2291	0,4581	1,11	0,8665	0,3665	0,7330
0,12	0,5478	0,0478	0,0955	0,62	0,7324	0,2324	0,4647	1,12	0,8686	0,3686	0,7373
0,13	0,5517	0,0517	0,1034	0,63	0,7357	0,2357	0,4713	1,13	0,8708	0,3708	0,7415
0,14	0,5557	0,0557	0,1113	0,64	0,7389	0,2389	0,4778	1,14	0,8729	0,3729	0,7457
0,15	0,5596	0,0596	0,1192	0,65	0,7422	0,2422	0,4843	1,15	0,8749	0,3749	0,7499
0,16	0,5636	0,0636	0,1271	0,66	0,7454	0,2454	0,4907	1,16	0,8770	0,3770	0,7540
0,17	0,5675	0,0675	0,1350	0,67	0,7486	0,2486	0,4971	1,17	0,8790	0,3790	0,7580
0,18	0,5714	0,0714	0,1428	0,68	0,7517	0,2517	0,5035	1,18	0,8810	0,3810	0,7620
0,19	0,5753	0,0753	0,1507	0,69	0,7549	0,2549	0,5098	1,19	0,8830	0,3830	0,7660
0,20	0,5793	0,0793	0,1585	0,70	0,7580	0,2580	0,5161	1,20	0,8849	0,3849	0,7699
0,21	0,5832	0,0832	0,1663	0,71	0,7611	0,2611	0,5223	1,21	0,8869	0,3869	0,7737
0,22	0,5871	0,0871	0,1741	0,72	0,7642	0,2642	0,5285	1,22	0,8888	0,3888	0,7775
0,23	0,5910	0,0910	0,1819	0,73	0,7673	0,2673	0,5346	1,23	0,8907	0,3907	0,7813
0,24	0,5948	0,0948	0,1897	0,74	0,7704	0,2704	0,5407	1,24	0,8925	0,3925	0,7850
0,25	0,5987	0,0987	0,1974	0,75	0,7734	0,2734	0,5467	1,25	0,8944	0,3944	0,7887
0,26	0,6026	0,1026	0,2051	0,76	0,7764	0,2764	0,5527	1,26	0,8962	0,3962	0,7923
0,27	0,6064	0,1064	0,2128	0,77	0,7794	0,2794	0,5587	1,27	0,8980	0,3980	0,7959
0,28	0,6103	0,1103	0,2205	0,78	0,7823	0,2823	0,5646	1,28	0,8997	0,3997	0,7995
0,29	0,6141	0,1141	0,2282	0,79	0,7852	0,2852	0,5705	1,29	0,9015	0,4015	0,8029
0,30	0,6179	0,1179	0,2358	0,80	0,7881	0,2881	0,5763	1,30	0,9032	0,4032	0,8064
0,31	0,6217	0,1217	0,2434	0,81	0,7910	0,2910	0,5821	1,31	0,9049	0,4049	0,8098
0,32	0,6255	0,1255	0,2510	0,82	0,7939	0,2939	0,5878	1,32	0,9066	0,4066	0,8132
0,33	0,6293	0,1293	0,2586	0,83	0,7967	0,2967	0,5935	1,33	0,9082	0,4082	0,8165
0,34	0,6331	0,1331	0,2661	0,84	0,7995	0,2995	0,5991	1,34	0,9099	0,4099	0,8198
0,35	0,6368	0,1368	0,2737	0,85	0,8023	0,3023	0,6047	1,35	0,9115	0,4115	0,8230
0,36	0,6406	0,1406	0,2812	0,86	0,8051	0,3051	0,6102	1,36	0,9131	0,4131	0,8262
0,37	0,6443	0,1443	0,2886	0,87	0,8078	0,3078	0,6157	1,37	0,9147	0,4147	0,8293
0,38	0,6480	0,1480	0,2961	0,88	0,8106	0,3106	0,6211	1,38	0,9162	0,4162	0,8324
0,39	0,6517	0,1517	0,3035	0,89	0,8133	0,3133	0,6265	1,39	0,9177	0,4177	0,8355
0,40	0,6554	0,1554	0,3108	0,90	0,8159	0,3159	0,6319	1,40	0,9192	0,4192	0,8385
0,41	0,6591	0,1591	0,3182	0,91	0,8186	0,3186	0,6372	1,41	0,9207	0,4207	0,8415
0,42	0,6628	0,1628	0,3255	0,92	0,8212	0,3212	0,6424	1,42	0,9222	0,4222	0,8444
0,43	0,6664	0,1664	0,3328	0,93	0,8238	0,3238	0,6476	1,43	0,9236	0,4236	0,8473
0,44	0,6700	0,1700	0,3401	0,94	0,8264	0,3264	0,6528	1,44	0,9251	0,4251	0,8501
0,45	0,6736	0,1736	0,3473	0,95	0,8289	0,3289	0,6579	1,45	0,9265	0,4265	0,8529
0,46	0,6772	0,1772	0,3545	0,96	0,8315	0,3315	0,6629	1,46	0,9279	0,4279	0,8557
0,47	0,6808	0,1808	0,3616	0,97	0,8340	0,3340	0,6680	1,47	0,9292	0,4292	0,8584
0,48	0,6844	0,1844	0,3688	0,98	0,8365	0,3365	0,6729	1,48	0,9306	0,4306	0,8611
0,49	0,6879	0,1879	0,3759	0,99	0,8389	0,3389	0,6778	1,49	0,9319	0,4319	0,8638
0,50	0,6915	0,1915	0,3829	1,00	0,8413	0,3413	0,6827	1,50	0,9332	0,4332	0,8664

z	F_Z	F_1	F_2	z	F_Z	F_1	F_2	z	F_Z	F_1	F_2
1,51	0,9345	0,4345	0,8690	2,01	0,9778	0,4778	0,9556	2,51	0,9940	0,4940	0,9879
1,52	0,9357	0,4357	0,8715	2,02	0,9783	0,4783	0,9566	2,52	0,9941	0,4941	0,9883
1,53	0,9370	0,4370	0,8740	2,03	0,9788	0,4788	0,9576	2,53	0,9943	0,4943	0,9886
1,54	0,9382	0,4382	0,8764	2,04	0,9793	0,4793	0,9586	2,54	0,9945	0,4945	0,9889
1,55	0,9394	0,4394	0,8789	2,05	0,9798	0,4798	0,9596	2,55	0,9946	0,4946	0,9892
1,56	0,9406	0,4406	0,8812	2,06	0,9803	0,4803	0,9606	2,56	0,9948	0,4948	0,9895
1,57	0,9418	0,4418	0,8836	2,07	0,9808	0,4808	0,9615	2,57	0,9949	0,4949	0,9898
1,58	0,9429	0,4429	0,8859	2,08	0,9812	0,4812	0,9625	2,58	0,9951	0,4951	0,9901
1,59	0,9441	0,4441	0,8882	2,09	0,9817	0,4817	0,9634	2,59	0,9952	0,4952	0,9904
1,60	0,9452	0,4452	0,8904	2,10	0,9821	0,4821	0,9643	2,60	0,9953	0,4953	0,9907
1,61	0,9463	0,4463	0,8926	2,11	0,9826	0,4826	0,9651	2,61	0,9955	0,4955	0,9909
1,62	0,9474	0,4474	0,8948	2,12	0,9830	0,4830	0,9660	2,62	0,9956	0,4956	0,9912
1,63	0,9484	0,4484	0,8969	2,13	0,9834	0,4834	0,9668	2,63	0,9957	0,4957	0,9915
1,64	0,9495	0,4495	0,8990	2,14	0,9838	0,4838	0,9676	2,64	0,9959	0,4959	0,9917
1,65	0,9505	0,4505	0,9011	2,15	0,9842	0,4842	0,9684	2,65	0,9960	0,4960	0,9920
1,66	0,9515	0,4515	0,9031	2,16	0,9846	0,4846	0,9692	2,66	0,9961	0,4961	0,9922
1,67	0,9525	0,4525	0,9051	2,17	0,9850	0,4850	0,9700	2,67	0,9962	0,4962	0,9924
1,68	0,9535	0,4535	0,9070	2,18	0,9854	0,4854	0,9707	2,68	0,9963	0,4963	0,9926
1,69	0,9545	0,4545	0,9090	2,19	0,9857	0,4857	0,9715	2,69	0,9964	0,4964	0,9929
1,70	0,9554	0,4554	0,9109	2,20	0,9861	0,4861	0,9722	2,70	0,9965	0,4965	0,9931
1,71	0,9564	0,4564	0,9127	2,21	0,9864	0,4864	0,9729	2,71	0,9966	0,4966	0,9933
1,72	0,9573	0,4573	0,9146	2,22	0,9868	0,4868	0,9736	2,72	0,9967	0,4967	0,9935
1,73	0,9582	0,4582	0,9164	2,23	0,9871	0,4871	0,9743	2,73	0,9968	0,4968	0,9937
1,74	0,9591	0,4591	0,9181	2,24	0,9875	0,4875	0,9749	2,74	0,9969	0,4969	0,9939
1,75	0,9599	0,4599	0,9199	2,25	0,9878	0,4878	0,9756	2,75	0,9970	0,4970	0,9940
1,76	0,9608	0,4608	0,9216	2,26	0,9881	0,4881	0,9762	2,76	0,9971	0,4971	0,9942
1,77	0,9616	0,4616	0,9233	2,27	0,9884	0,4884	0,9768	2,77	0,9972	0,4972	0,9944
1,78	0,9625	0,4625	0,9249	2,28	0,9887	0,4887	0,9774	2,78	0,9973	0,4973	0,9946
1,79	0,9633	0,4633	0,9265	2,29	0,9890	0,4890	0,9780	2,79	0,9974	0,4974	0,9947
1,80	0,9641	0,4641	0,9281	2,30	0,9893	0,4893	0,9786	2,80	0,9974	0,4974	0,9949
1,81	0,9649	0,4649	0,9297	2,31	0,9896	0,4896	0,9791	2,81	0,9975	0,4975	0,9950
1,82	0,9656	0,4656	0,9312	2,32	0,9898	0,4898	0,9797	2,82	0,9976	0,4976	0,9952
1,83	0,9664	0,4664	0,9328	2,33	0,9901	0,4901	0,9802	2,83	0,9977	0,4977	0,9953
1,84	0,9671	0,4671	0,9342	2,34	0,9904	0,4904	0,9807	2,84	0,9977	0,4977	0,9955
1,85	0,9678	0,4678	0,9357	2,35	0,9906	0,4906	0,9812	2,85	0,9978	0,4978	0,9956
1,86	0,9686	0,4686	0,9371	2,36	0,9909	0,4909	0,9817	2,86	0,9979	0,4979	0,9958
1,87	0,9693	0,4693	0,9385	2,37	0,9911	0,4911	0,9822	2,87	0,9979	0,4979	0,9959
1,88	0,9699	0,4699	0,9399	2,38	0,9913	0,4913	0,9827	2,88	0,9980	0,4980	0,9960
1,89	0,9706	0,4706	0,9412	2,39	0,9916	0,4916	0,9832	2,89	0,9981	0,4981	0,9961
1,90	0,9713	0,4713	0,9426	2,40	0,9918	0,4918	0,9836	2,90	0,9981	0,4981	0,9963
1,91	0,9719	0,4719	0,9439	2,41	0,9920	0,4920	0,9840	2,91	0,9982	0,4982	0,9964
1,92	0,9726	0,4726	0,9451	2,42	0,9922	0,4922	0,9845	2,92	0,9982	0,4982	0,9965
1,93	0,9732	0,4732	0,9464	2,43	0,9925	0,4925	0,9849	2,93	0,9983	0,4983	0,9966
1,94	0,9738	0,4738	0,9476	2,44	0,9927	0,4927	0,9853	2,94	0,9984	0,4984	0,9967
1,95	0,9744	0,4744	0,9488	2,45	0,9929	0,4929	0,9857	2,95	0,9984	0,4984	0,9968
1,96	0,9750	0,4750	0,9500	2,46	0,9931	0,4931	0,9861	2,96	0,9985	0,4985	0,9969
1,97	0,9756	0,4756	0,9512	2,47	0,9932	0,4932	0,9865	2,97	0,9985	0,4985	0,9970
1,98	0,9761	0,4761	0,9523	2,48	0,9934	0,4934	0,9869	2,98	0,9986	0,4986	0,9971
1,99	0,9767	0,4767	0,9534	2,49	0,9936	0,4936	0,9872	2,99	0,9986	0,4986	0,9972
2,00	0,9772	0,4772	0,9545	2,50	0,9938	0,4938	0,9876	3,00	0,9987	0,4987	0,9973

17.3 χ^2 -Verteilung

Sind X_1, X_2, \dots, X_ν standardnormalverteilte, stochastisch unabhängige Zufallsvariablen, so ist die Zufallsvariable

Definition

$$Y = \sum_{i=1}^{\nu} X_i^2$$

Dichtefunktion

χ^2 -verteilt mit ν Freiheitsgraden. Die Dichtefunktion lautet

$$f_Y(y) = \frac{1}{2^{\frac{\nu}{2}} \cdot \Gamma(\frac{\nu}{2})} y^{(\frac{\nu}{2}-1)} e^{(-\frac{y}{2})} \quad \text{für } y > 0.$$

Dabei ist $\Gamma(\frac{\nu}{2})$ die Eulersche **Gammafunktion** mit folgenden Eigenschaften:

- (1) $\Gamma\left(\frac{\nu}{2}\right) = \left(\frac{\nu}{2} - 1\right) \Gamma\left(\frac{\nu}{2} - 1\right), \quad \nu > 2$
- (2) $\Gamma\left(\frac{\nu}{2}\right) = \left(\frac{\nu}{2} - 1\right)! \quad \nu = 2n, \quad n \in \mathbb{N}$
- (3) $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

Erwartungswert

$$E(Y) = \nu$$

Varianz

$$\text{Var}(Y) = 2\nu$$

Parameter der χ^2 -Verteilung

ν (Anzahl der Freiheitsgrade)
 Y heißt auch $\chi^2(\nu)$ -verteilt

Approximation

Für $\nu > 100$ ist eine $\chi^2(\nu)$ -verteilte Zufallsvariable näherungsweise normalverteilt mit den Parametern ν und 2ν ($N(\nu, 2\nu)$ -verteilt). Die Zufallsvariable $Z = \sqrt{2Y} - \sqrt{2\nu - 1}$ ist für $\nu > 30$ näherungsweise standardnormalverteilt.

Tabelle der χ^2 -Verteilung

Für verschiedene Freiheitsgrade und für verschiedene Werte von $F_Y(y)$ (kumulierte Verteilung) sind die zugehörigen y -Werte (Quantile) tabelliert.

http://www.fernuni-hagen.de/ls_statistik/lehre/

ν	0,001	0,005	0,01	0,02	0,025	0,05	0,1	0,25	0,5	0,75	0,9	0,95	0,975	0,98	0,99	0,995	0,999
1	-	-	-	-	-	0,004	0,016	0,102	0,455	1,323	2,706	3,841	5,024	5,412	6,635	7,879	10,827
2	0,002	0,010	0,020	0,040	0,051	0,103	0,211	0,575	1,386	2,773	4,605	5,991	7,378	7,824	9,210	10,597	13,815
3	0,024	0,072	0,115	0,185	0,216	0,352	0,584	1,213	2,366	4,108	6,251	7,815	9,348	9,837	11,345	12,838	16,266
4	0,091	0,207	0,297	0,429	0,484	0,711	1,064	1,923	3,357	5,385	7,779	9,488	11,143	11,668	13,277	14,860	18,466
5	0,210	0,412	0,554	0,752	0,831	1,145	1,610	2,675	4,351	6,626	9,236	11,070	12,832	13,388	15,086	16,750	20,515
6	0,381	0,676	0,872	1,134	1,237	1,635	2,204	3,455	5,348	7,841	10,645	12,592	14,449	15,033	16,812	18,548	22,457
7	0,599	0,989	1,239	1,564	1,690	2,167	2,833	4,255	6,346	9,037	12,017	14,067	16,013	16,622	18,475	20,278	24,321
8	0,857	1,344	1,647	2,032	2,180	2,733	3,490	5,071	7,344	10,219	13,362	15,507	17,535	18,168	20,090	21,955	26,124
9	1,152	1,735	2,088	2,532	2,700	3,325	4,168	5,899	8,343	11,389	14,684	16,919	19,023	19,679	21,666	23,589	27,877
10	1,479	2,156	2,558	3,059	3,247	3,940	4,865	6,737	9,342	12,549	15,987	18,307	20,483	21,161	23,209	25,188	29,588
11	1,834	2,603	3,053	3,609	3,816	4,575	5,578	7,584	10,341	13,701	17,275	19,675	21,920	22,618	24,725	26,757	31,264
12	2,214	3,074	3,571	4,178	4,404	5,226	6,304	8,438	11,340	14,845	18,549	21,026	23,337	24,054	26,217	28,300	32,909
13	2,617	3,565	4,107	4,765	5,009	5,892	7,041	9,299	12,340	15,984	19,812	22,362	24,736	25,471	27,688	29,819	34,527
14	3,041	4,075	4,660	5,368	5,629	6,571	7,790	10,165	13,339	17,117	21,064	23,685	26,119	26,873	29,141	31,319	36,124
15	3,483	4,601	5,229	5,985	6,262	7,261	8,547	11,037	14,339	18,245	22,307	24,996	27,488	28,259	30,578	32,801	37,698
16	3,942	5,142	5,812	6,614	6,908	7,962	9,312	11,912	15,338	19,369	23,542	26,296	28,845	29,633	32,000	34,267	39,252
17	4,416	5,697	6,408	7,255	7,564	8,672	10,085	12,792	16,338	20,489	24,769	27,587	30,191	30,995	33,409	35,718	40,791
18	4,905	6,265	7,015	7,906	8,231	9,390	10,865	13,675	17,338	21,605	25,989	28,869	31,526	32,346	34,805	37,156	42,312
19	5,407	6,844	7,633	8,567	8,907	10,117	11,651	14,562	18,338	22,718	27,204	30,144	32,852	33,687	36,191	38,582	43,819
20	5,921	7,434	8,260	9,237	9,591	10,851	12,443	15,452	19,337	23,828	28,412	31,410	34,170	35,020	37,566	39,997	45,314
21	6,447	8,034	8,897	9,915	10,283	11,591	13,240	16,344	20,337	24,935	29,615	32,671	35,479	36,343	38,932	41,401	46,796
22	6,983	8,643	9,542	10,600	10,982	12,338	14,041	17,240	21,337	26,039	30,813	33,924	36,781	37,659	40,289	42,796	48,268
23	7,529	9,260	10,196	11,293	11,689	13,091	14,848	18,137	22,337	27,141	32,007	35,172	38,076	38,968	41,638	44,181	49,728
24	8,085	9,886	10,856	11,992	12,401	13,848	15,659	19,037	23,337	28,241	33,196	36,415	39,364	40,270	42,980	45,558	51,179
25	8,649	10,520	11,524	12,697	13,120	14,611	16,473	19,939	24,337	29,339	34,382	37,652	40,646	41,566	44,314	46,928	52,619
26	9,222	11,160	12,198	13,409	13,844	15,379	17,292	20,843	25,336	30,435	35,563	38,885	41,923	42,856	45,642	48,290	54,051
27	9,803	11,808	12,878	14,125	14,573	16,151	18,114	21,749	26,336	31,528	36,741	40,113	43,195	44,140	46,963	49,645	55,475
28	10,391	12,461	13,565	14,847	15,308	16,928	18,939	22,657	27,336	32,620	37,916	41,337	44,461	45,419	48,278	50,994	56,892
29	10,986	13,121	14,256	15,574	16,047	17,708	19,768	23,567	28,336	33,711	39,087	42,557	45,722	46,693	49,588	52,335	58,301
30	11,588	13,787	14,953	16,306	16,791	18,493	20,599	24,478	29,336	34,800	40,256	43,773	46,979	47,962	50,892	53,672	59,702
31	12,196	14,458	15,655	17,042	17,539	19,281	21,434	25,390	30,336	35,887	41,422	44,985	48,232	49,226	52,191	55,002	61,098
32	12,810	15,134	16,362	17,783	18,291	20,072	22,271	26,304	31,336	36,973	42,585	46,194	49,480	50,487	53,486	56,328	62,487
33	13,431	15,815	17,073	18,527	19,047	20,867	23,110	27,219	32,336	38,058	43,745	47,400	50,725	51,743	54,775	57,648	63,869

ν	$F_Y(y)$																
	0,001	0,005	0,01	0,02	0,025	0,05	0,1	0,25	0,5	0,75	0,9	0,95	0,975	0,98	0,99	0,995	0,999
34	14,057	16,501	17,789	19,275	19,806	21,664	23,952	28,136	33,336	39,141	44,903	48,602	51,966	52,995	56,061	58,964	65,247
35	14,688	17,192	18,509	20,027	20,569	22,465	24,797	29,054	34,336	40,223	46,059	49,802	53,203	54,244	57,342	60,275	66,619
36	15,324	17,887	19,233	20,783	21,336	23,269	25,643	29,973	35,336	41,304	47,212	50,998	54,437	55,489	58,619	61,581	67,985
37	15,965	18,586	19,960	21,542	22,106	24,075	26,492	30,893	36,336	42,383	48,363	52,192	55,668	56,730	59,893	62,883	69,348
38	16,611	19,289	20,691	22,304	22,878	24,884	27,343	31,815	37,335	43,462	49,513	53,384	56,895	57,969	61,162	64,181	70,704
39	17,261	19,996	21,426	23,069	23,654	25,695	28,196	32,737	38,335	44,539	50,660	54,572	58,120	59,204	62,428	65,475	72,055
40	17,917	20,707	22,164	23,838	24,433	26,509	29,051	33,660	39,335	45,616	51,805	55,758	59,342	60,436	63,691	66,766	73,403
41	18,576	21,421	22,906	24,609	25,215	27,326	29,907	34,585	40,335	46,692	52,949	56,942	60,561	61,665	64,950	68,053	74,744
42	19,238	22,138	23,650	25,383	25,999	28,144	30,765	35,510	41,335	47,766	54,090	58,124	61,777	62,892	66,206	69,336	76,084
43	19,905	22,860	24,398	26,159	26,785	28,965	31,625	36,436	42,335	48,840	55,230	59,304	62,990	64,116	67,459	70,616	77,418
44	20,576	23,584	25,148	26,939	27,575	29,787	32,487	37,363	43,335	49,913	56,369	60,481	64,201	65,337	68,710	71,892	78,749
45	21,251	24,311	25,901	27,720	28,366	30,612	33,350	38,291	44,335	50,985	57,505	61,656	65,410	66,555	69,957	73,166	80,078
46	21,929	25,041	26,657	28,504	29,160	31,439	34,215	39,220	45,335	52,056	58,641	62,830	66,616	67,771	71,201	74,437	81,400
47	22,610	25,775	27,416	29,291	29,956	32,268	35,081	40,149	46,335	53,127	59,774	64,001	67,821	68,985	72,443	75,704	82,720
48	23,294	26,511	28,177	30,080	30,754	33,098	35,949	41,079	47,335	54,196	60,907	65,171	69,023	70,197	73,683	76,969	84,037
49	23,983	27,249	28,941	30,871	31,555	33,930	36,818	42,010	48,335	55,265	62,038	66,339	70,222	71,406	74,919	78,231	85,350
50	24,674	27,991	29,707	31,664	32,357	34,764	37,689	42,942	49,335	56,334	63,167	67,505	71,420	72,613	76,154	79,490	86,660
51	25,368	28,735	30,475	32,459	33,162	35,600	38,560	43,874	50,335	57,401	64,295	68,669	72,616	73,818	77,386	80,746	87,967
52	26,065	29,481	31,246	33,256	33,968	36,437	39,433	44,807	51,335	58,468	65,422	69,832	73,810	75,021	78,616	82,001	89,272
53	26,765	30,230	32,019	34,055	34,776	37,276	40,308	45,741	52,335	59,534	66,548	70,993	75,002	76,223	79,843	83,253	90,573
54	27,467	30,981	32,793	34,856	35,586	38,116	41,183	46,676	53,335	60,600	67,673	72,153	76,192	77,422	81,069	84,502	91,871
55	28,173	31,735	33,571	35,659	36,398	38,958	42,060	47,610	54,335	61,665	68,796	73,311	77,380	78,619	82,292	85,749	93,167
56	28,881	32,491	34,350	36,464	37,212	39,801	42,937	48,546	55,335	62,729	69,919	74,468	78,567	79,815	83,514	86,994	94,462
57	29,592	33,248	35,131	37,270	38,027	40,646	43,816	49,482	56,335	63,793	71,040	75,624	79,752	81,009	84,733	88,237	95,750
58	30,305	34,008	35,914	38,078	38,844	41,492	44,696	50,419	57,335	64,857	72,160	76,778	80,936	82,201	85,950	89,477	97,038
59	31,021	34,770	36,698	38,888	39,662	42,339	45,577	51,356	58,335	65,919	73,279	77,930	82,117	83,391	87,166	90,715	98,324
60	31,738	35,534	37,485	39,699	40,482	43,188	46,459	52,294	59,335	66,981	74,397	79,082	83,298	84,580	88,379	91,952	99,608
61	32,458	36,300	38,273	40,512	41,303	44,038	47,342	53,232	60,335	68,043	75,514	80,232	84,476	85,767	89,591	93,186	100,887
62	33,181	37,068	39,063	41,327	42,126	44,889	48,226	54,171	61,335	69,104	76,630	81,381	85,654	86,953	90,802	94,419	102,165
63	33,905	37,838	39,855	42,143	42,950	45,741	49,111	55,110	62,335	70,165	77,745	82,529	86,830	88,137	92,010	95,649	103,442
64	34,632	38,610	40,649	42,960	43,776	46,595	49,996	56,050	63,335	71,225	78,860	83,675	88,004	89,320	93,217	96,878	104,717
65	35,362	39,383	41,444	43,779	44,603	47,450	50,883	56,990	64,335	72,285	79,973	84,821	89,177	90,501	94,422	98,105	105,988
66	36,092	40,158	42,240	44,599	45,431	48,305	51,770	57,931	65,335	73,344	81,085	85,965	90,349	91,681	95,626	99,330	107,257

ν	$F_Y(y)$																
	0,001	0,005	0,01	0,02	0,025	0,05	0,1	0,25	0,5	0,75	0,9	0,95	0,975	0,98	0,99	0,995	0,999
67	36,826	40,935	43,038	45,421	46,261	49,162	52,659	58,872	66,335	74,403	82,197	87,108	91,519	92,860	96,828	100,554	108,525
68	37,561	41,714	43,838	46,244	47,092	50,020	53,548	59,814	67,335	75,461	83,308	88,250	92,688	94,037	98,028	101,776	109,793
69	38,298	42,493	44,639	47,068	47,924	50,879	54,438	60,756	68,334	76,519	84,418	89,391	93,856	95,213	99,227	102,996	111,055
70	39,036	43,275	45,442	47,893	48,758	51,739	55,329	61,698	69,334	77,577	85,527	90,531	95,023	96,387	100,425	104,215	112,317
71	39,776	44,058	46,246	48,720	49,592	52,600	56,221	62,641	70,334	78,634	86,635	91,670	96,189	97,561	101,621	105,432	113,577
72	40,520	44,843	47,051	49,548	50,428	53,462	57,113	63,585	71,334	79,690	87,743	92,808	97,353	98,733	102,816	106,647	114,834
73	41,263	45,629	47,858	50,377	51,265	54,325	58,006	64,528	72,334	80,747	88,850	93,945	98,516	99,904	104,010	107,862	116,092
74	42,009	46,417	48,666	51,208	52,103	55,189	58,900	65,472	73,334	81,803	89,956	95,081	99,678	101,074	105,202	109,074	117,347
75	42,757	47,206	49,475	52,039	52,942	56,054	59,795	66,417	74,334	82,858	91,061	96,217	100,839	102,243	106,393	110,285	118,599
76	43,507	47,996	50,286	52,872	53,782	56,920	60,690	67,362	75,334	83,913	92,166	97,351	101,999	103,410	107,582	111,495	119,850
77	44,257	48,788	51,097	53,705	54,623	57,786	61,586	68,307	76,334	84,968	93,270	98,484	103,158	104,576	108,771	112,704	121,101
78	45,011	49,581	51,910	54,540	55,466	58,654	62,483	69,252	77,334	86,022	94,374	99,617	104,316	105,742	109,958	113,911	122,347
79	45,764	50,376	52,725	55,376	56,309	59,522	63,380	70,198	78,334	87,077	95,476	100,749	105,473	106,906	111,144	115,116	123,595
80	46,520	51,172	53,540	56,213	57,153	60,391	64,278	71,145	79,334	88,130	96,578	101,879	106,629	108,069	112,329	116,321	124,839
81	47,276	51,969	54,357	57,051	57,998	61,262	65,176	72,091	80,334	89,184	97,680	103,010	107,783	109,231	113,512	117,524	126,084
82	48,036	52,767	55,174	57,890	58,845	62,132	66,076	73,038	81,334	90,237	98,780	104,139	108,937	110,393	114,695	118,726	127,324
83	48,795	53,567	55,993	58,730	59,692	63,004	66,976	73,985	82,334	91,289	99,880	105,267	110,090	111,553	115,876	119,927	128,565
84	49,558	54,368	56,813	59,570	60,540	63,876	67,876	74,933	83,334	92,342	100,980	106,395	111,242	112,712	117,057	121,126	129,802
85	50,320	55,170	57,634	60,412	61,389	64,749	68,777	75,881	84,334	93,394	102,079	107,522	112,393	113,871	118,236	122,324	131,043
86	51,084	55,973	58,456	61,255	62,239	65,623	69,679	76,829	85,334	94,446	103,177	108,648	113,544	115,028	119,414	123,522	132,276
87	51,850	56,777	59,279	62,098	63,089	66,498	70,581	77,777	86,334	95,497	104,275	109,773	114,693	116,184	120,591	124,718	133,511
88	52,617	57,582	60,103	62,943	63,941	67,373	71,484	78,726	87,334	96,548	105,372	110,898	115,841	117,340	121,767	125,912	134,746
89	53,385	58,389	60,928	63,788	64,793	68,249	72,387	79,675	88,334	97,599	106,469	112,022	116,989	118,495	122,942	127,106	135,977
90	54,156	59,196	61,754	64,635	65,647	69,126	73,291	80,625	89,334	98,650	107,565	113,145	118,136	119,648	124,116	128,299	137,208
91	54,925	60,005	62,581	65,482	66,501	70,003	74,196	81,574	90,334	99,700	108,661	114,268	119,282	120,801	125,289	129,490	138,437
92	55,698	60,815	63,409	66,330	67,356	70,882	75,100	82,524	91,334	100,750	109,756	115,390	120,427	121,953	126,462	130,681	139,667
93	56,471	61,625	64,238	67,179	68,211	71,760	76,006	83,474	92,334	101,800	110,850	116,511	121,571	123,105	127,633	131,871	140,894
94	57,246	62,437	65,068	68,028	69,068	72,640	76,912	84,425	93,334	102,850	111,944	117,632	122,715	124,255	128,803	133,059	142,118
95	58,022	63,250	65,898	68,879	69,925	73,520	77,818	85,376	94,334	103,899	113,038	118,752	123,858	125,405	129,973	134,247	143,343
96	58,800	64,063	66,730	69,730	70,783	74,401	78,725	86,327	95,334	104,948	114,131	119,871	125,000	126,554	131,141	135,433	144,566
97	59,577	64,878	67,562	70,582	71,642	75,282	79,633	87,278	96,334	105,997	115,223	120,990	126,141	127,702	132,309	136,619	145,789
98	60,356	65,693	68,396	71,434	72,501	76,164	80,541	88,229	97,334	107,045	116,315	122,108	127,282	128,849	133,476	137,803	147,009
99	61,136	66,510	69,230	72,288	73,361	77,046	81,449	89,181	98,334	108,093	117,407	123,225	128,422	129,996	134,641	138,987	148,230
100	61,918	67,328	70,065	73,142	74,222	77,929	82,358	90,133	99,334	109,141	118,498	124,342	129,561	131,142	135,807	140,170	149,449

17.4 F -Verteilung

Definition

Y_1 und Y_2 seien χ^2 -verteilte, voneinander unabhängige Zufallsvariablen mit m bzw. n Freiheitsgraden. Die Zufallsvariable

$$X = \frac{Y_1/m}{Y_2/n}$$

ist F -verteilt.

Dichtefunktion

$$f_X(x; m, n) = \left(\frac{m}{2}\right)^{\frac{m}{2}} \cdot \left(\frac{n}{2}\right)^{\frac{n}{2}} \cdot \frac{\Gamma\left(\frac{m}{2} + \frac{n}{2}\right)}{\Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{n}{2}\right)} \cdot \frac{x^{\left(\frac{m}{2}-1\right)}}{\left(\frac{m}{2}x + \frac{n}{2}\right)^{\frac{m+n}{2}}}$$

Erwartungswert

$$E(X) = \frac{n}{n-2} \quad n > 2$$

Varianz

$$\text{Var}(X) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)} \quad n > 4$$

Parameter der F -Verteilung

m, n (Anzahl der Freiheitsgrade)

Tabelle der F -Verteilung

In Abhängigkeit von m und n sind die Werte x der F -verteilten Zufallsvariablen tabelliert, bei denen die Verteilungsfunktion den Wert 0,95 bzw. 0,99 annimmt.

<http://www.fernuni-hagen.de/ls.statistik/lehre/>

$F(x) = 0,95$

n	m																				∞						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	30	40	50	100	200	∞	
1	161	199	216	225	230	234	237	239	241	242	243	244	245	246	246	247	247	247	248	248	248	250	251	252	253	254	254
2	18,5	19,0	19,2	19,2	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5	19,5	19,5
3	10,14	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,76	8,74	8,73	8,71	8,70	8,69	8,68	8,67	8,67	8,66	8,66	8,62	8,59	8,58	8,55	8,53	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,94	5,91	5,89	5,87	5,86	5,84	5,83	5,82	5,81	5,80	5,79	5,75	5,72	5,70	5,66	5,63	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,70	4,68	4,66	4,64	4,62	4,60	4,59	4,58	4,57	4,56	4,55	4,50	4,46	4,44	4,41	4,39	4,37
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00	3,98	3,96	3,94	3,92	3,91	3,90	3,88	3,87	3,81	3,77	3,75	3,71	3,69	3,67	
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,60	3,57	3,55	3,53	3,51	3,49	3,48	3,47	3,46	3,44	3,38	3,34	3,32	3,27	3,25	3,23	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,31	3,28	3,26	3,24	3,22	3,20	3,19	3,17	3,16	3,15	3,08	3,04	3,02	2,97	2,95	2,93	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,10	3,07	3,05	3,03	3,01	2,99	2,97	2,96	2,95	2,94	2,86	2,83	2,80	2,76	2,73	2,71	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,94	2,91	2,89	2,86	2,85	2,83	2,81	2,80	2,79	2,77	2,70	2,66	2,64	2,59	2,56	2,54	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,82	2,79	2,76	2,74	2,72	2,70	2,69	2,67	2,66	2,65	2,57	2,53	2,51	2,46	2,43	2,40	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,72	2,69	2,66	2,64	2,62	2,60	2,58	2,57	2,56	2,54	2,47	2,43	2,40	2,35	2,32	2,30	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,63	2,60	2,58	2,55	2,53	2,51	2,50	2,48	2,47	2,46	2,38	2,34	2,31	2,26	2,23	2,21	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,57	2,53	2,51	2,48	2,46	2,44	2,43	2,41	2,40	2,39	2,31	2,27	2,24	2,19	2,16	2,13	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,51	2,48	2,45	2,42	2,40	2,38	2,37	2,35	2,34	2,33	2,25	2,20	2,18	2,12	2,10	2,07	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,46	2,42	2,40	2,37	2,35	2,33	2,32	2,30	2,29	2,28	2,20	2,15	2,12	2,07	2,04	2,01	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,41	2,38	2,35	2,33	2,31	2,29	2,27	2,26	2,24	2,23	2,15	2,10	2,08	2,02	1,99	1,96	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37	2,34	2,31	2,29	2,27	2,25	2,23	2,22	2,20	2,19	2,11	2,06	2,04	1,98	1,95	1,92	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,34	2,31	2,28	2,26	2,23	2,21	2,20	2,18	2,17	2,16	2,08	2,03	2,00	1,94	1,91	1,88	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,31	2,28	2,25	2,22	2,20	2,18	2,17	2,15	2,14	2,12	2,04	1,99	1,97	1,91	1,88	1,84	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,28	2,25	2,22	2,20	2,18	2,16	2,14	2,12	2,11	2,10	2,02	1,97	1,94	1,88	1,84	1,81	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,26	2,23	2,20	2,17	2,15	2,13	2,11	2,10	2,08	2,07	1,98	1,94	1,91	1,85	1,82	1,78	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,24	2,20	2,18	2,15	2,13	2,11	2,09	2,08	2,06	2,05	1,96	1,92	1,88	1,82	1,79	1,76	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,22	2,18	2,15	2,13	2,11	2,09	2,07	2,05	2,04	2,03	1,94	1,89	1,86	1,80	1,77	1,73	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,20	2,16	2,14	2,11	2,09	2,07	2,05	2,04	2,02	2,01	1,92	1,87	1,84	1,78	1,75	1,71	1,71
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,18	2,15	2,12	2,09	2,07	2,05	2,03	2,02	2,00	1,99	1,90	1,85	1,82	1,76	1,73	1,69	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,17	2,13	2,10	2,08	2,06	2,04	2,02	2,00	1,99	1,97	1,88	1,84	1,81	1,74	1,71	1,67	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,15	2,12	2,09	2,06	2,04	2,02	2,00	1,99	1,97	1,96	1,87	1,83	1,80	1,73	1,69	1,65	1,65
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,14	2,10	2,08	2,05	2,03	2,01	1,99	1,97	1,96	1,94	1,85	1,81	1,77	1,71	1,67	1,64	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,13	2,09	2,06	2,04	2,01	1,99	1,98	1,96	1,95	1,93	1,84	1,79	1,76	1,70	1,66	1,62	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,04	2,00	1,97	1,95	1,92	1,90	1,89	1,87	1,85	1,84	1,74	1,69	1,66	1,59	1,55	1,51	1,51
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,03	1,99	1,95	1,92	1,89	1,87	1,85	1,83	1,81	1,80	1,78	1,69	1,63	1,60	1,52	1,48	1,44	1,44
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,95	1,92	1,89	1,86	1,84	1,82	1,80	1,78	1,76	1,75	1,65	1,59	1,56	1,48	1,44	1,39	1,39
70	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,02	1,97	1,93	1,89	1,86	1,84	1,81	1,79	1,77	1,75	1,74	1,72	1,62	1,57	1,53	1,45	1,40	1,35	1,35
80	3,96	3,11	2,72	2,49	2,33	2,21	2,13	2,06	2,00	1,95	1,91	1,88	1,84	1,82	1,79	1,77	1,75	1,73	1,72	1,70	1,60	1,54	1,51	1,43	1,38	1,32	1,32
90	3,95	3,10	2,71	2,47	2,32	2,20	2,11	2,04	1,99	1,94	1,90	1,86	1,83	1,80	1,78	1,76	1,74	1,72	1,70	1,69	1,59	1,53	1,49	1,41	1,36	1,30	1,30
100	3,94	3,09	2,70	2,46	2,31	2,19	2,10	2,03	1,97	1,93	1,89	1,85	1,82	1,79	1,77	1,75	1,73	1,71	1,69	1,68	1,57	1,52	1,48	1,39	1,34	1,28	1,28
120	3,92	3,07	2,68	2,45	2,29	2,18	2,09	2,02	1,96	1,91	1,87	1,83	1,80	1,78	1,75	1,73	1,71	1,69	1,67	1,66	1,55	1,50	1,46	1,37	1,32	1,25	1,25
140	3,91	3,06	2,67	2,44	2,28	2,16	2,08	2,01	1,95	1,90	1,86	1,82	1,79	1,76	1,74	1,72	1,70	1,68	1,66	1,65	1,54	1,48	1,44	1,35	1,30	1,23	1,23
160	3,90	3,05	2,66	2,43	2,27	2,16	2,07	2,00	1,94	1,89	1,85	1,81	1,78	1,75	1,73	1,71	1,69	1,67	1,65	1,64	1,53	1,47	1,43	1,34	1,28	1,21	1,21
180	3,89	3,05	2,65	2,42	2,26	2,15	2,06	1,99	1,93	1,88	1,84	1,81	1,77	1,75	1,72	1,70	1,68	1,66	1,64	1,63	1,52	1,46	1,42	1,33	1,27	1,20	1,20
200	3,89	3,04	2,65	2,42	2,26	2,14	2,06	1,98	1,93	1,88	1,84	1,80	1,77	1,74	1,72	1,69	1,67	1,66	1,64	1,62	1,52	1,46	1,41	1,32	1,26	1,19	1,19
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,79	1,75	1,72	1,69	1,67	1,64	1,62	1,60	1,59	1,57	1,46	1,39	1,35	1,24	1,17	1,00	1,00

$F(x) = 0,99$

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	30	40	50	100	200	∞
1	4052	4999	5404	5624	5764	5859	5928	5981	6022	6056	6083	6107	6126	6143	6157	6170	6181	6191	6201	6209	6260	6286	6302	6334	6350	6366
2	98,5	99,0	99,2	99,3	99,3	99,3	99,4	99,4	99,4	99,4	99,4	99,4	99,4	99,4	99,4	99,4	99,4	99,4	99,4	99,4	99,5	99,5	99,5	99,5	99,5	99,5
3	34,1	30,8	29,5	28,7	28,2	27,9	27,7	27,5	27,3	27,2	27,1	27,1	27,0	26,9	26,9	26,8	26,8	26,8	26,7	26,7	26,7	26,5	26,4	26,4	26,2	26,1
4	21,2	18,0	16,7	16,0	15,5	15,2	15,0	14,8	14,7	14,5	14,5	14,4	14,3	14,2	14,2	14,2	14,1	14,1	14,0	14,0	13,8	13,7	13,7	13,6	13,5	13,5
5	16,3	13,3	12,1	11,4	11,0	10,7	10,5	10,3	10,2	10,1	9,96	9,89	9,82	9,77	9,72	9,68	9,64	9,61	9,58	9,55	9,38	9,29	9,24	9,13	9,08	9,02
6	13,7	10,9	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,79	7,72	7,66	7,60	7,56	7,52	7,48	7,45	7,42	7,40	7,23	7,14	7,09	6,99	6,93	6,88
7	12,2	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,54	6,47	6,41	6,36	6,31	6,28	6,24	6,21	6,18	6,16	5,99	5,91	5,86	5,75	5,70	5,65
8	11,3	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,73	5,67	5,61	5,56	5,52	5,48	5,44	5,41	5,38	5,36	5,20	5,12	5,07	4,96	4,91	4,86
9	10,6	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,18	5,11	5,05	5,01	4,96	4,92	4,89	4,86	4,83	4,81	4,65	4,57	4,52	4,41	4,36	4,31
10	10,0	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,77	4,71	4,65	4,60	4,56	4,52	4,49	4,46	4,43	4,41	4,25	4,17	4,12	4,01	3,96	3,91
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,46	4,40	4,34	4,29	4,25	4,21	4,18	4,15	4,12	4,10	3,94	3,86	3,81	3,71	3,66	3,60
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,22	4,16	4,10	4,05	4,01	3,97	3,94	3,91	3,88	3,86	3,70	3,62	3,57	3,47	3,41	3,36
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	4,02	3,96	3,91	3,86	3,82	3,78	3,75	3,72	3,69	3,66	3,51	3,43	3,38	3,27	3,22	3,17
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,86	3,80	3,75	3,70	3,66	3,62	3,59	3,56	3,53	3,51	3,35	3,27	3,22	3,11	3,06	3,00
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,73	3,67	3,61	3,56	3,52	3,49	3,45	3,42	3,40	3,37	3,21	3,13	3,08	2,98	2,92	2,87
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,62	3,55	3,50	3,45	3,41	3,37	3,34	3,31	3,28	3,26	3,10	3,02	2,97	2,86	2,81	2,75
17	8,40	6,11	5,19	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,52	3,46	3,40	3,35	3,31	3,27	3,24	3,21	3,19	3,16	3,00	2,92	2,87	2,76	2,71	2,65
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,43	3,37	3,32	3,27	3,23	3,19	3,16	3,13	3,10	3,08	2,92	2,84	2,78	2,68	2,62	2,57
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,36	3,30	3,24	3,19	3,15	3,12	3,08	3,05	3,03	3,00	2,84	2,76	2,71	2,60	2,55	2,49
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,30	3,23	3,18	3,13	3,09	3,05	3,02	2,99	2,96	2,94	2,78	2,69	2,64	2,54	2,48	2,42
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31	3,24	3,17	3,12	3,07	3,03	2,99	2,96	2,93	2,90	2,88	2,72	2,64	2,58	2,48	2,42	2,36
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,18	3,12	3,07	3,02	2,98	2,94	2,91	2,88	2,85	2,83	2,67	2,58	2,53	2,42	2,36	2,31
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,14	3,07	3,02	2,97	2,93	2,89	2,86	2,83	2,80	2,78	2,62	2,54	2,48	2,37	2,32	2,26
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	3,09	3,03	2,98	2,93	2,89	2,85	2,82	2,79	2,76	2,74	2,58	2,49	2,44	2,33	2,27	2,21
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	3,06	2,99	2,94	2,89	2,85	2,81	2,78	2,75	2,72	2,70	2,54	2,45	2,40	2,29	2,23	2,17
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09	3,02	2,96	2,90	2,86	2,81	2,78	2,75	2,72	2,69	2,66	2,50	2,42	2,36	2,25	2,19	2,13
27	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15	3,06	2,99	2,93	2,87	2,82	2,78	2,75	2,71	2,68	2,66	2,63	2,47	2,38	2,33	2,22	2,16	2,10
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03	2,96	2,90	2,84	2,79	2,75	2,72	2,68	2,65	2,63	2,60	2,44	2,35	2,30	2,19	2,13	2,06
29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09	3,00	2,93	2,87	2,81	2,77	2,73	2,69	2,66	2,63	2,60	2,57	2,41	2,33	2,27	2,16	2,10	2,03
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,91	2,84	2,79	2,74	2,70	2,66	2,63	2,60	2,57	2,55	2,39	2,30	2,25	2,13	2,07	2,01
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80	2,73	2,66	2,61	2,56	2,52	2,48	2,45	2,42	2,39	2,37	2,20	2,11	2,06	1,94	1,87	1,80
50	7,17	5,06	4,20	3,72	3,41	3,19	3,02	2,89	2,78	2,70	2,63	2,56	2,51	2,46	2,42	2,38	2,35	2,32	2,29	2,27	2,10	2,01	1,95	1,82	1,76	1,68
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,56	2,50	2,44	2,39	2,35	2,31	2,28	2,25	2,22	2,20	2,03	1,94	1,88	1,75	1,68	1,60
70	7,01	4,92	4,07	3,60	3,29	3,07	2,91	2,78	2,67	2,59	2,51	2,45	2,40	2,35	2,31	2,27	2,23	2,20	2,18	2,15	1,98	1,89	1,83	1,70	1,62	1,54
80	6,96	4,88	4,04	3,56	3,26	3,04	2,87	2,74	2,64	2,55	2,48	2,42	2,36	2,31	2,27	2,23	2,20	2,17	2,14	2,12	1,94	1,85	1,79	1,65	1,58	1,49
90	6,93	4,85	4,01	3,53	3,23	3,01	2,84	2,72	2,61	2,52	2,45	2,39	2,33	2,29	2,24	2,21	2,17	2,14	2,11	2,09	1,92	1,82	1,76	1,62	1,55	1,46
100	6,90	4,82	3,98	3,51	3,21	2,99	2,82	2,69	2,59	2,50	2,43	2,37	2,31	2,27	2,22	2,19	2,15	2,12	2,09	2,07	1,89	1,80	1,74	1,60	1,52	1,43
120	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47	2,34	2,34	2,28	2,23	2,19	2,15	2,12	2,09	2,06	2,03	1,86	1,76	1,70	1,56	1,48	1,38
140	6,82	4,76	3,92	3,46	3,15	2,93	2,77	2,64	2,54	2,45	2,38	2,31	2,26	2,21	2,17	2,13	2,10	2,07	2,04	2,01	1,84	1,74	1,67	1,53	1,45	1,35
160	6,80	4,74	3,91	3,44	3,13	2,92	2,75	2,62	2,52	2,43	2,36	2,30	2,24	2,20	2,15	2,11	2,08	2,05	2,02	1,99	1,82	1,72	1,66	1,51	1,42	1,32
180	6,78	4,73	3,89	3,43	3,12	2,90	2,74	2,61	2,51	2,42	2,35	2,28	2,23	2,18	2,14	2,10	2,07	2,04	2,01	1,98	1,81	1,71	1,64	1,49	1,41	1,30
200	6,76	4,71	3,88	3,41	3,11	2,89	2,73	2,60	2,50	2,41	2,34	2,27	2,22	2,17	2,13	2,09	2,06	2,03	2,00	1,97	1,79	1,69	1,63	1,48	1,39	1,28
∞	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,25	2,18	2,13	2,08	2,04	2,00	1,97	1,93	1,90	1,88	1,70	1,59	1,52	1,36	1,25	1,00

17.5 Student- t -Verteilung

Es seien X und Y unabhängige Zufallsvariablen und X sei $N(0, 1)$ - und Y sei $\chi^2(\nu)$ -verteilt. Dann ist

$$T = \frac{X}{\sqrt{Y/\nu}}$$

Student-verteilt mit ν Freiheitsgraden.

Die Dichtefunktion lautet

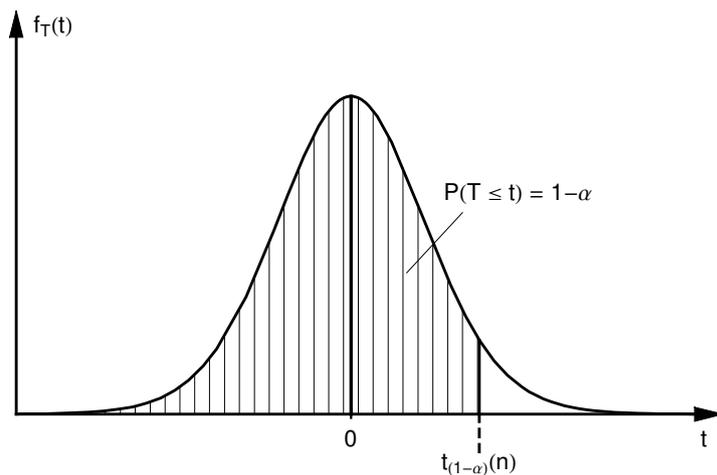
$$f_T(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi} \Gamma(\frac{\nu}{2})} (1 + t^2/\nu)^{-(\nu+1)/2}$$

$$E(T) = 0 \quad \text{für } \nu \geq 2$$

$$\text{Var}(T) = \frac{\nu}{\nu-2} \quad \text{für } \nu \geq 3$$

ν (Anzahl der Freiheitsgrade)

Für verschiedene Freiheitsgrade ν und für verschiedene Wahrscheinlichkeiten $p = P(T \leq t)$ sind die entsprechenden t -Werte $t(1-\alpha, \nu)$ tabelliert.



Definition

Dichtefunktion der Student- t -Verteilung

Erwartungswert

Varianz

Parameter der Student-Verteilung

Tabelle der Student-Verteilung

Quantile der t -Verteilung $t(1 - \alpha, \nu)$

ν	$P(T \leq t) =$							
	0,75	0,90	0,95	0,975	0,99	0,995	0,999	0,9995
1	1,000	3,078	6,314	12,706	31,821	63,657	318,300	636,619
2	0,816	1,886	2,920	4,303	6,965	9,925	22,330	31,598
3	0,765	1,638	2,353	3,182	4,541	5,841	10,210	12,941
4	0,741	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	0,727	1,476	2,015	2,571	3,365	4,032	5,893	6,869
6	0,718	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	0,711	1,415	1,895	2,365	2,998	3,499	4,785	5,408
8	0,706	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	0,703	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	0,700	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	0,697	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	0,695	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	0,694	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	0,692	1,345	1,761	2,145	2,624	2,977	3,787	4,140
15	0,691	1,341	1,753	2,131	2,602	2,947	3,733	4,073
16	0,690	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	0,689	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	0,688	1,330	1,734	2,101	2,552	2,878	3,610	3,922
19	0,688	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	0,687	1,325	1,725	2,086	2,528	2,845	3,552	3,850
21	0,686	1,323	1,721	2,080	2,518	2,831	3,527	3,819
22	0,686	1,321	1,717	2,074	2,508	2,819	3,505	3,792
23	0,685	1,319	1,714	2,069	2,500	2,807	3,485	3,768
24	0,685	1,318	1,711	2,064	2,492	2,797	3,467	3,745
25	0,684	1,316	1,708	2,060	2,485	2,787	3,450	3,725
26	0,684	1,315	1,706	2,056	2,479	2,779	3,435	3,707
27	0,684	1,314	1,703	2,052	2,473	2,771	3,421	3,690
28	0,683	1,313	1,701	2,048	2,467	2,763	3,408	3,674
29	0,683	1,311	1,699	2,045	2,462	2,756	3,396	3,659
30	0,683	1,310	1,697	2,042	2,457	2,750	3,385	3,646
40	0,681	1,303	1,684	2,021	2,423	2,704	3,307	3,551
50	0,679	1,299	1,676	2,009	2,403	2,678	3,261	3,496
100	0,677	1,290	1,660	1,984	2,364	2,626	3,174	3,390
200	0,676	1,286	1,652	1,972	2,345	2,601	3,131	3,340
∞	0,675	1,282	1,645	1,960	2,326	2,576	3,090	3,291

Bonferroni- t -Statistik: $t(1 - \alpha/(2k), \nu), \alpha = 0.05$

		$\alpha = 0.05$															
		k															
ν	2	3	4	5	6	7	8	9	10	15	20	25	30	35	40	45	50
5	3.16	3.53	3.81	4.03	4.22	4.38	4.53	4.66	4.77	5.25	5.60	5.89	6.14	6.35	6.54	6.71	6.87
6	2.97	3.29	3.52	3.71	3.86	4.00	4.12	4.22	4.32	4.70	4.98	5.21	5.40	5.56	5.71	5.84	5.96
7	2.84	3.13	3.34	3.50	3.64	3.75	3.86	3.95	4.03	4.36	4.59	4.79	4.94	5.08	5.20	5.31	5.41
8	2.75	3.02	3.21	3.36	3.48	3.58	3.68	3.76	3.83	4.12	4.33	4.50	4.64	4.76	4.86	4.96	5.04
9	2.69	2.93	3.11	3.25	3.36	3.46	3.55	3.62	3.69	3.95	4.15	4.30	4.42	4.53	4.62	4.71	4.78
10	2.63	2.87	3.04	3.17	3.28	3.37	3.45	3.52	3.58	3.83	4.00	4.14	4.26	4.36	4.44	4.52	4.59
15	2.49	2.69	2.84	2.95	3.04	3.11	3.18	3.23	3.29	3.48	3.62	3.73	3.82	3.90	3.96	4.02	4.07
20	2.42	2.61	2.74	2.85	2.93	3.00	3.06	3.11	3.15	3.33	3.46	3.55	3.63	3.70	3.75	3.80	3.85
25	2.38	2.57	2.69	2.79	2.86	2.93	2.99	3.03	3.08	3.24	3.36	3.45	3.52	3.58	3.64	3.68	3.73
30	2.36	2.54	2.66	2.75	2.82	2.89	2.94	2.99	3.03	3.19	3.30	3.39	3.45	3.51	3.56	3.61	3.65
40	2.33	2.50	2.62	2.70	2.78	2.84	2.89	2.93	2.97	3.12	3.23	3.31	3.37	3.43	3.47	3.51	3.55
50	2.31	2.48	2.59	2.68	2.75	2.81	2.85	2.90	2.94	3.08	3.18	3.26	3.32	3.38	3.42	3.46	3.50
60	2.30	2.46	2.58	2.66	2.73	2.79	2.83	2.88	2.91	3.06	3.16	3.23	3.29	3.34	3.39	3.43	3.46
100	2.28	2.43	2.54	2.63	2.69	2.75	2.79	2.83	2.87	3.01	3.10	3.17	3.23	3.28	3.32	3.36	3.39
∞	2.24	2.39	2.50	2.58	2.64	2.69	2.73	2.77	2.81	2.94	3.02	3.09	3.14	3.19	3.23	3.26	3.29

Bonferroni- t -Statistik: $t(1 - \alpha/(2k), \nu), \alpha = 0.01$

		$\alpha = 0.01$															
		k															
ν	2	3	4	5	6	7	8	9	10	15	20	25	30	35	40	45	50
5	4.77	5.25	5.60	5.89	6.14	6.35	6.54	6.71	6.87	7.50	7.98	8.36	8.69	8.98	9.24	9.47	9.68
6	4.32	4.70	4.98	5.21	5.40	5.56	5.71	5.84	5.96	6.43	6.79	7.07	7.31	7.52	7.71	7.87	8.02
7	4.03	4.36	4.59	4.79	4.94	5.08	5.20	5.31	5.41	5.80	6.08	6.31	6.50	6.67	6.81	6.94	7.06
8	3.83	4.12	4.33	4.50	4.64	4.76	4.86	4.96	5.04	5.37	5.62	5.81	5.97	6.11	6.23	6.34	6.44
9	3.69	3.95	4.15	4.30	4.42	4.53	4.62	4.71	4.78	5.08	5.29	5.46	5.60	5.72	5.83	5.92	6.01
10	3.58	3.83	4.00	4.14	4.26	4.36	4.44	4.52	4.59	4.85	5.05	5.20	5.33	5.44	5.53	5.62	5.69
15	3.29	3.48	3.62	3.73	3.82	3.90	3.96	4.02	4.07	4.27	4.42	4.53	4.62	4.70	4.77	4.83	4.88
20	3.15	3.33	3.46	3.55	3.63	3.70	3.75	3.80	3.85	4.02	4.15	4.24	4.32	4.39	4.44	4.49	4.54
25	3.08	3.24	3.36	3.45	3.52	3.58	3.64	3.68	3.73	3.88	4.00	4.08	4.15	4.21	4.27	4.31	4.35
30	3.03	3.19	3.30	3.39	3.45	3.51	3.56	3.61	3.65	3.80	3.90	3.98	4.05	4.11	4.15	4.20	4.23
40	2.97	3.12	3.23	3.31	3.37	3.43	3.47	3.51	3.55	3.69	3.79	3.86	3.92	3.98	4.02	4.06	4.09
50	2.94	3.08	3.18	3.26	3.32	3.38	3.42	3.46	3.50	3.63	3.72	3.79	3.85	3.90	3.94	3.98	4.01
60	2.91	3.06	3.16	3.23	3.29	3.34	3.39	3.43	3.46	3.59	3.68	3.75	3.81	3.85	3.89	3.93	3.96
100	2.87	3.01	3.10	3.17	3.23	3.28	3.32	3.36	3.39	3.51	3.60	3.66	3.72	3.76	3.80	3.83	3.86
∞	2.81	2.94	3.02	3.09	3.14	3.19	3.23	3.26	3.29	3.40	3.48	3.54	3.59	3.63	3.66	3.69	3.72

17.6 Wilcoxon-Vorzeichen-Rangtest

$$H_0: f(x) \text{ ist symmetrisch zum Median } x(0.5) = \delta_0 \quad (17.1)$$

Testgröße nach Wilcoxon:

$$W^+ = \sum_n R_n Z_n \quad (17.2)$$

$$R_n = \text{rg}|X_n - \delta_0| \quad (17.3)$$

$$Z_n = \begin{cases} 1 & \text{wenn } X_n > \delta_0 \\ 0 & \text{sonst} \end{cases} \quad (17.4)$$

Erwartungswert und Varianz der Teststatistik unter H_0 :

$$E[W^+] = \frac{N(N+1)}{4} \quad (17.5)$$

$$\text{Var}[W^+] = \frac{N(N+1)(2N+1)}{24}. \quad (17.6)$$

Ablehnungsbedingung (asymptotischer Test)

$$|W^+| > E[W^+] + z(1 - \alpha/2)\sqrt{\text{Var}[W^+]} \quad (17.7)$$

Exakte Verteilung für W^+

Quantile (siehe Tabelle):

$$w^+(\alpha/2) \quad (17.8)$$

$$w^+(1 - \alpha/2) = \frac{N(N+1)}{2} - w^+(\alpha/2) \quad (17.9)$$

N	$w^+(0.01)$	$w^+(0.025)$	$w^+(0.05)$	$w^+(0.1)$	$w^+(0.9)$	$w^+(0.95)$	$w^+(0.975)$	$w^+(0.99)$
4	0	0	0	1	9	10	10	10
5	0	0	1	3	12	14	15	15
6	0	1	3	4	17	18	20	21
7	1	3	4	6	22	24	25	27
8	2	4	6	9	27	30	32	34
9	4	6	9	11	34	36	39	41
10	6	9	11	15	40	44	46	49
11	8	11	14	18	48	52	55	58
12	10	14	18	22	56	60	64	68
13	13	18	22	27	64	69	73	78
14	16	22	26	32	73	79	83	89
15	20	26	31	37	83	89	94	100
16	24	30	36	43	93	100	106	112
17	28	35	42	49	104	111	118	125
18	33	41	48	56	115	123	130	138
19	38	47	54	63	127	136	143	152
20	44	53	61	70	140	149	157	166

**Test des
Medians
(Wilcoxon-
Vorzeichen-
Rangtest)**

17.7 Wilcoxon-Rangsummentest

**Wilcoxon-
Rangsummen-Test**

2-seitiger Test:

$$H_0 : \mu_x = \mu_y \quad (17.10)$$

$$H_1 : \mu_x \neq \mu_y \quad (17.11)$$

oder einseitige Alternativen.

Daten:

X_1, \dots, X_N (1. Stichprobe)

Y_1, \dots, Y_M (2. Stichprobe)

Ränge in der zusammengefaßten (gepoolten) Stichprobe.

Teststatistik

$$T_W = \sum_n \text{rg}(X_i) \quad (17.12)$$

Ränge der 1. Stichprobe (in der gemeinsamen Stichprobe vergeben)

Erwartungswert und Varianz von T_W unter H_0

$$E[T_W] = N(N + M + 1)/2 \quad (17.13)$$

$$\text{Var}[T_W] = NM(N + M + 1)/12. \quad (17.14)$$

Kritische Werte:

$$w(\alpha), w(1 - \alpha) = N(N + M + 1) - w(\alpha)$$

$\alpha = 0.025$

Umfang größere Stichprobe	Umfang kleinere Stichprobe																								
	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25		
3	-																								
4	-	11																							
5	7	12	18																						
6	8	13	19	27																					
7	8	14	21	28	37																				
8	9	15	22	30	39	50																			
9	9	15	23	32	41	52	63																		
10	10	16	24	33	43	54	66	79																	
11	10	17	25	35	45	56	69	82	97																
12	11	18	27	36	47	59	72	85	100	116															
13	11	19	28	38	49	61	74	89	104	120	137														
14	12	20	29	39	51	63	77	92	107	124	142	161													
15	12	21	30	41	53	66	80	95	111	128	146	165	185												
16	13	22	31	43	55	68	83	98	114	132	151	170	191	212											
17	13	22	33	44	57	71	85	101	118	136	155	175	196	218	241										
18	14	23	34	46	59	73	88	104	122	140	159	180	201	223	247	271									
19	14	24	35	47	61	75	91	108	125	144	164	184	206	229	253	278	304								
20	15	25	36	49	63	78	94	111	129	148	168	189	211	235	259	284	310	338							
21	15	26	38	51	65	80	96	114	132	152	172	194	217	240	265	291	317	345	374						
22	16	27	39	52	67	82	99	117	136	156	177	199	222	246	271	297	324	352	382	412					
23	16	28	40	54	69	85	102	120	140	160	181	204	227	252	277	304	331	360	389	420	452				
24	17	28	41	55	71	87	105	123	143	164	186	208	232	257	283	310	338	367	397	428	460	493			
25	17	29	43	57	73	90	108	127	147	168	190	213	238	263	289	317	345	374	405	436	469	502	537		

$\alpha = 0.05$

Umfang größere Stichprobe	Umfang kleinere Stichprobe																								
	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25		
3	7																								
4	7	12																							
5	8	13	20																						
6	9	14	21	29																					
7	9	15	22	30	40																				
8	10	16	24	32	42	52																			
9	11	17	25	34	44	55	67																		
10	11	18	27	36	46	57	70	83																	
11	12	19	28	38	48	60	73	87	101																
12	12	20	29	39	50	63	76	90	105	121															
13	13	21	31	41	53	65	79	93	109	126	143														
14	14	22	32	43	55	68	82	97	113	130	148	167													
15	14	23	34	45	57	70	85	100	117	134	153	172	193												
16	15	25	35	47	59	73	88	104	121	139	157	177	198	220											
17	16	26	36	48	62	76	91	107	124	143	162	183	204	226	250										
18	16	27	38	50	64	78	94	111	128	147	167	188	209	232	256	281									
19	17	28	39	52	66	81	97	114	132	151	172	193	215	238	263	288	314								
20	18	29	41	54	68	84	100	118	136	156	176	198	221	244	269	295	321	349							
21	18	30	42	56	70	86	103	121	140	160	181	203	226	250	275	302	329	357	386						
22	19	31	44	58	73	89	106	124	144	164	186	208	232	256	282	308	336	365	394	425					
23	20	32	45	59	75	91	109	128	148	169	190	213	237	262	288	315	343	372	402	433	466				
24	20	33	46	61	77	94	112	131	152	173	195	219	243	268	295	322	351	380	411	442	475	508			
25	21	34	48	63	79	97	115	135	156	177	200	224	249	274	301	329	358	388	419	451	484	518	553		

$\alpha = 0.1$

Umfang kleinere Stichprobe

Umfang größere Stichprobe

	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
3	8																							
4	8	14																						
5	9	15	21																					
6	10	16	23	31																				
7	11	17	24	33	42																			
8	12	18	26	35	45	56																		
9	12	20	28	37	47	59	71																	
10	13	21	29	39	50	61	74	88																
11	14	22	31	41	52	64	77	92	107															
12	15	23	33	43	55	67	81	95	111	128														
13	16	24	34	45	57	70	84	99	115	132	150													
14	17	26	36	47	60	73	87	103	119	137	155	175												
15	17	27	38	49	62	76	91	107	124	142	160	180	201											
16	18	28	39	51	65	79	94	110	128	146	166	186	207	230										
17	19	29	41	53	67	82	98	114	132	151	171	191	213	236	260									
18	20	31	43	56	70	85	101	118	136	156	176	197	219	243	267	292								
19	21	32	44	58	72	88	104	122	140	160	181	203	225	249	274	300	326							
20	22	33	46	60	75	91	108	126	145	165	186	208	231	256	281	307	334	362						
21	22	34	48	62	77	93	111	129	149	170	191	214	237	262	288	314	342	371	400					
22	23	36	49	64	80	96	114	133	153	174	196	219	243	268	295	322	350	379	409	440				
23	24	37	51	66	82	99	118	137	157	179	201	225	249	275	301	329	358	387	418	449	482			
24	25	39	52	68	85	102	121	141	162	184	206	230	255	281	308	336	365	395	426	458	492	526		
25	26	39	54	70	87	105	124	145	166	188	212	236	261	288	315	344	373	404	435	468	501	536	571	

17.8 Zufallszahlentabelle

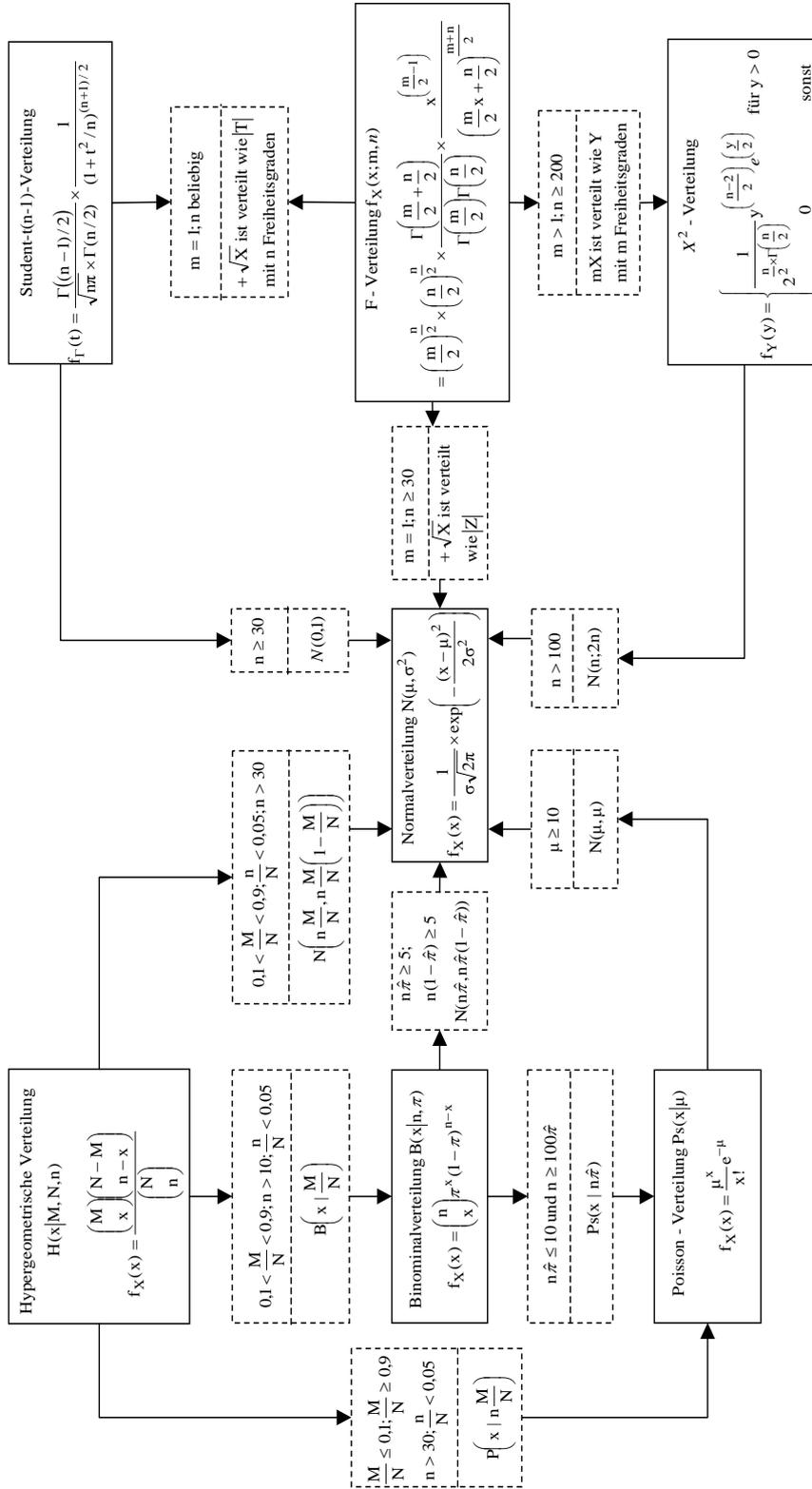
Zufallszahlentabelle

Die nachfolgende Tabelle enthält gleichverteilte Pseudozufallszahlen. Um die Handhabung der Tabelle zu erleichtern, sind die Zufallszahlen waagrecht in 10 Kolonnen zu je 4 Ziffern und senkrecht in 5 Blöcke zu je 10 Zeilen eingeteilt. Beim Ablesen von Zufallszahlen wird im Allgemeinen der Startpunkt d.h. die Startkolonne und die Richtung des Ablesens, bezeichnet durch die Nummer von Block, Zeile, Spalte, vorgegeben oder zufällig ausgewählt. Die Zufallszahlen können zeilenweise von links nach rechts oder von rechts nach links oder spaltenweise von oben nach unten bzw. unten nach oben abgelesen werden.

Je nach Stellenzahl der gewünschten Zufallszahlen werden die abgelesenen Ziffern einzeln oder in 2er-, 3er-, 4er- usw. Blöcken verwendet.

		PSEUDO-ZUFALLSZAHLEN (näherungsweise gleichverteilt)									
0	0	1761	2305	1332	6659	3052	8865	6920	8923	2937	6832
	1	1102	2592	1416	6204	3878	7668	8489	6983	5248	4150
	2	4155	9169	5211	2039	8415	2760	3768	1333	1269	7758
	3	0919	2034	2717	4162	6663	4141	2758	1972	1001	7654
	4	1393	1188	3933	2574	8622	1811	5459	8900	4445	3840
	5	5579	6632	8861	7275	4292	5770	1871	2117	1599	6315
	6	3475	8364	7500	8265	3673	6018	1994	1623	2464	5078
	7	5082	6386	9849	5545	6764	2555	5828	7418	7040	0131
	8	0401	0697	5910	9113	3567	5382	3373	9502	5327	1473
	9	9430	1296	5681	8971	4081	4497	4629	7875	7325	9104
1	0	2170	8185	9163	5117	8305	9901	9595	2537	3034	3024
	1	8621	1363	6357	7553	6240	1595	8273	3488	2454	3233
	2	8783	0830	7261	6278	7887	9577	0661	0729	5584	9731
	3	2656	6586	1876	1292	3244	3849	6761	4258	2426	2519
	4	0240	8631	0202	2594	2312	4410	6571	4076	2979	1595
	5	1535	6964	2240	0185	5093	1258	0095	0182	7245	6958
	6	6501	1625	7909	4143	1464	4513	7168	2735	5021	8808
	7	5022	2733	7171	4509	1469	4137	7916	1617	6510	6948
	8	7254	0131	0145	1165	5185	0050	2374	6788	1711	1377
	9	3196	3817	6830	4109	2612	2252	0543	8248	0622	2094
2	0	2850	3792	7226	3342	3750	0744	2423	5996	3244	9101
	1	6214	0057	1332	8865	8599	5524	8014	0034	9577	2397
	2	3289	2611	9150	0676	7158	6593	7316	0361	9621	1982
	3	4075	1453	0678	8776	9429	3951	0328	6978	3376	7855
	4	8573	6585	5917	3166	5411	7599	7052	9883	0842	0018
	5	6781	8006	4868	3844	5104	7534	7487	9076	2020	8470
	6	8701	5715	7530	0811	8508	3759	1634	4559	6908	3210
	7	4332	9713	3903	4067	5623	6273	9491	6330	5508	4239
	8	3763	0000	3987	3612	6449	5076	1059	4391	7818	1558
	9	6726	6576	7782	9446	0986	0168	6339	8741	3840	5165
3	0	3489	9442	8287	1569	9234	1549	5329	9380	3572	5062
	1	3964	8596	6504	9982	1196	9219	8030	6307	7015	1248
	2	8149	4039	1432	4683	6863	3178	4442	9524	4170	3722
	3	6046	5772	0070	5673	6243	3426	4565	9030	5035	2486
	4	7653	3794	2420	2953	9335	9963	8399	4825	9611	7539
	5	2049	8655	8349	6281	6797	1759	7394	5088	0139	6268
	6	5032	5301	2074	2184	1265	6920	2604	9507	6091	9945
	7	1727	8235	9511	4377	9444	8370	1525	0214	5754	9911
	8	2132	7459	0659	2858	1334	6109	4157	7211	9129	6166
	9	6249	2971	5517	7628	6934	0137	0500	0497	6214	8709
4	0	4076	4773	4087	8688	6246	0454	0554	0073	7010	7542
	1	5614	2863	6367	6036	9269	7060	4319	5936	1517	2663
	2	0864	7242	2359	9673	6003	9955	1795	8089	9736	4074
	3	9825	7910	2062	9599	6448	9139	2983	6530	1666	1773
	4	2497	4868	5476	5814	0605	4611	7881	1261	7306	5761
	5	8879	8114	2601	8318	8472	6373	6490	2280	6658	6039
	6	8973	7649	3437	7111	0050	4424	8811	9589	9720	2605
	7	2778	3473	7984	2193	5339	8764	4842	3187	6493	5461
	8	0293	5587	6242	3564	4339	9393	4584	3073	6978	4606
	9	1520	3989	8210	1224	7049	6311	8037	9249	1173	0039
	0	1	2	3	4	5	6	7	8	9	

17.9 Übersicht zur Approximation von Verteilungen



Kapitel 18

Notation und Rechenregeln

In diesem Kapitel werden die im Kurs benutzte Notation sowie wichtige Rechenregeln für Ereignisse und Wahrscheinlichkeiten alphabetisch zusammengestellt. Wenn Ihnen diese nicht bekannt sind, konsultieren Sie bitte Ihren Bachelor-Statistikurs bzw. entsprechende Lehrbücher (z.B. Fahrmeir et al., 2007).

18.1 Formeln (alphabetisch)

$$P(A|B) := \frac{P(A \cap B)}{P(B)} \quad (\text{bedingte Wahrscheinlichkeit}) \quad (18.1)$$

Bayes-Formel

$$= P(A) \quad \text{bei Unabhängigkeit} \quad (18.2)$$

**bedingte
Wahrscheinlichkeit**

$$\delta_{jj'} = \begin{cases} 1 & j = j' \\ 0 & \text{sonst} \end{cases} \quad (18.3)$$

**Kronecker-Delta-
Symbol**

$$f(x) = P(x \leq X \leq x + dx)/dx \quad (\text{stetige Zufallsvariablen}) \quad (18.4)$$

Dichtefunktion

$$= P(X = x) \quad (\text{diskrete Zufallsvariablen}) \quad (18.5)$$

**Dichtefunktion
(bivariat)**

$$f(x, y) = P(x \leq X \leq x + dx, y \leq Y \leq y + dy) / dx dy$$

(stetige Zufallsvariablen) (18.6)

$$= P(X = x, Y = y)$$

(diskrete Zufallsvariablen) (18.7)

$$f(x, y) = f_x(x) \cdot f_y(y) \text{ bei Unabhängigkeit} \quad (18.8)$$

$$f_x(x), f_y(y) : \text{ Randverteilungen} \quad (18.9)$$

$$f_x(x) = \int f(x, y) dy$$

(stetige Zufallsvariablen) (18.10)

$$= \sum_y f(x, y)$$

(diskrete Zufallsvariablen) (18.11)

Ereignisse

$$A = \{\omega | \omega \in \Omega\} \quad (18.12)$$

$$\Omega = \text{Ergebnismenge} \quad (18.13)$$

$$P(\Omega) = 1 \text{ (sicheres Ereignis)} \quad (18.14)$$

$$\bar{A} : A \text{ tritt nicht ein} \quad (18.15)$$

$$P(\bar{A}) = 1 - P(A) \quad (18.16)$$

$$\emptyset = \bar{\Omega} \text{ (unmögliches Ereignis)} \quad (18.17)$$

$$P(\emptyset) = 1 - P(\Omega) = 0 \quad (18.18)$$

$$A \cap B : A \text{ und } B \text{ tritt ein} \quad (18.19)$$

$$P(A \cap B) = P(A) \cdot P(B) \text{ bei Unabhängigkeit} \quad (18.20)$$

$$A \cup B : A \text{ oder } B \text{ tritt ein} \quad (18.21)$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (18.22)$$

$$= P(A) + P(B), \text{ falls } A \cap B = \emptyset \quad (18.23)$$

$$A \subset B : \text{ Wenn } A, \text{ dann auch } B \quad (18.24)$$

$$A \setminus B = A \cap \bar{B} \text{ (} A \text{ ohne } B) \quad (18.25)$$

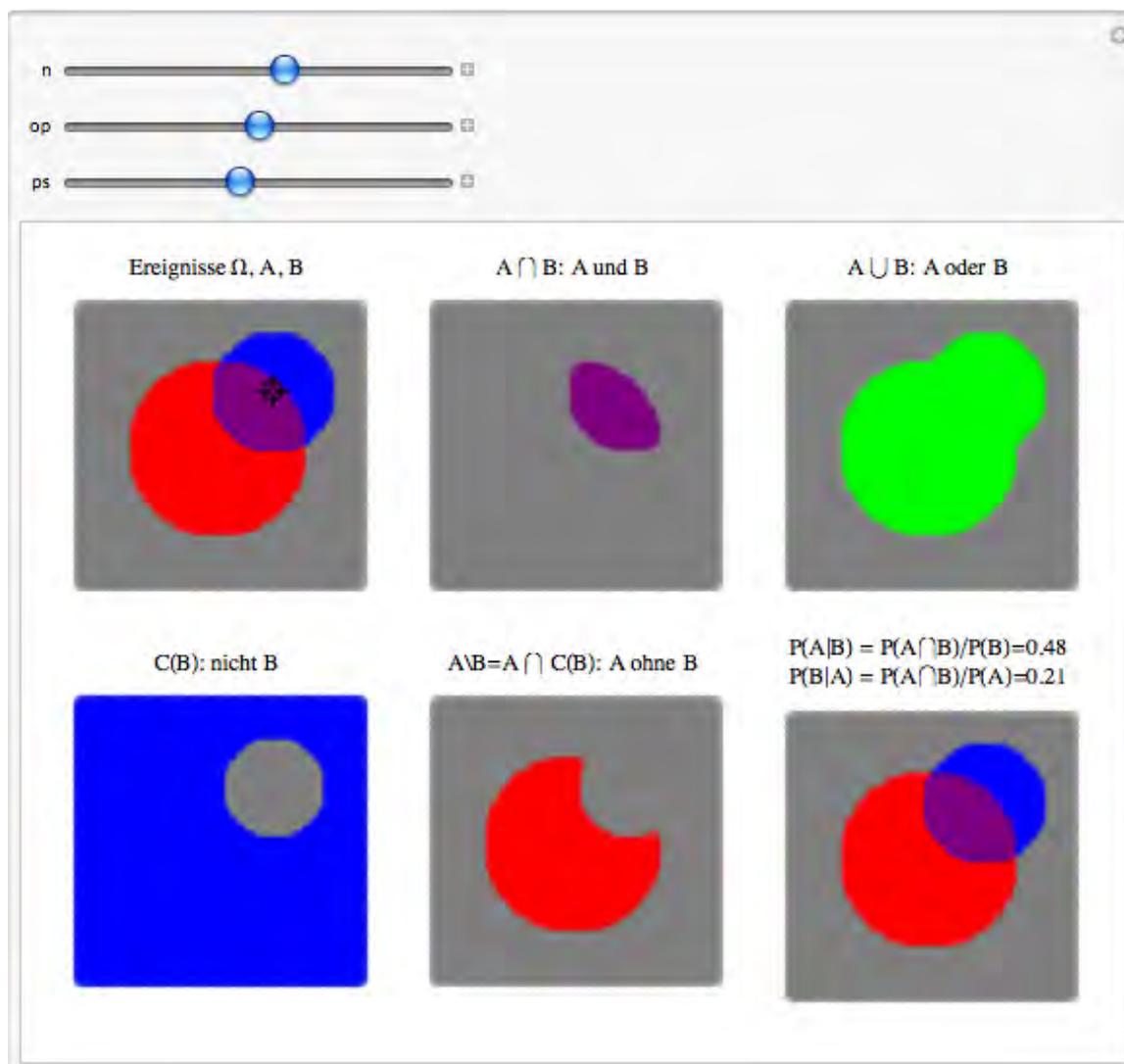


Abbildung 18.1: Ereignisse.

http://www.fernuni-hagen.de/ls_statistik/lehre/).

Erwartungswert

$$E[X] = \mu \quad (18.26)$$

$$= \int xf(x)dx$$

(stetige Zufallsvariablen) (18.27)

$$= \sum_i x_i f(x_i)$$

(diskrete Zufallsvariablen) (18.28)

$$E[h(X)] = \int h(x)f(x)dx$$

(stetige Zufallsvariablen) (18.29)

$$= \sum_i h(x_i)f(x_i)$$

(diskrete Zufallsvariablen) (18.30)

Kovarianz

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] \quad (18.31)$$

$$E(X) = \mu_X \quad (18.32)$$

$$E(Y) = \mu_Y \quad (18.33)$$

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] \quad (18.34)$$

$$= \int (x - \mu_X)(y - \mu_Y)f(x, y)dxdy$$

(stetige Zufallsvariablen) (18.35)

$$= \sum (x_i - \mu_X)(y_j - \mu_Y)f(x_i, y_j)$$

(diskrete Zufallsvariablen) (18.36)

$$\text{Var}(X) = \text{Cov}(X, X) = E[X^2] - \mu^2 \quad (18.37)$$

$$\text{Var}(X + Y) = \text{Cov}(X + Y, X + Y) \quad (18.38)$$

$$= \text{Cov}(X, X) + 2 \text{Cov}(X, Y) + \text{Cov}(Y, Y) \quad (18.39)$$

$$= \text{Var}(X) + 2 \text{Cov}(X, Y) + \text{Var}(Y) \quad (18.40)$$

$$\text{Cov}(aX, bY) = a \cdot b \text{Cov}(X, Y) \quad (18.41)$$

$$\begin{aligned} \text{Cov}(X + Y, Z + W) &= \text{Cov}(X, Z + W) + \text{Cov}(Y, Z + W) \\ &= \text{Cov}(X, Z) + \text{Cov}(X, W) \\ &\quad + \text{Cov}(Y, Z) + \text{Cov}(Y, W) \end{aligned} \quad (18.42)$$

$$\sigma_{ij} = \text{Cov}(X_i, X_j); \quad i, j = 1, \dots, p \quad (18.43)$$

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} \quad (p = 3) \quad (18.44) \quad \textbf{Kovarianzmatrix}$$

$$\bar{X} = \frac{1}{N} \sum_{n=1}^N X_n \quad (18.45) \quad \textbf{Mittelwert}$$

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{n=1}^N \mathbf{x}_n \quad (18.46)$$

$$X \sim N(\mu, \sigma^2) \quad (18.47) \quad \textbf{Normalverteilung}$$

$$\mu = E(X) \quad (18.48) \quad \textbf{(univariat)}$$

$$\sigma^2 = \text{Var}(X) \quad (18.49)$$

Dichtefunktion:

$$f(x) = (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (18.50)$$

**Normalverteilung
(bivariat)**

$$\mathbf{x} = \begin{bmatrix} X \\ Y \end{bmatrix} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (18.51)$$

$$\boldsymbol{\mu} = E[\mathbf{x}] = \begin{bmatrix} E(X) \\ E(Y) \end{bmatrix} = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix} \quad (18.52)$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \text{Var}(X) & \text{Cov}(X, Y) \\ \text{Cov}(X, Y) & \text{Var}(Y) \end{bmatrix} \quad (18.53)$$

$$= \begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} \end{bmatrix} \quad (18.54)$$

$$= \begin{bmatrix} \sigma_x^2 & \sigma_x \sigma_y \rho_{xy} \\ \sigma_x \sigma_y \rho_{xy} & \sigma_y^2 \end{bmatrix} \quad (18.55)$$

$$(18.56)$$

Dichtefunktion:

$$\begin{aligned} f(x, y) &= \det(2\pi\boldsymbol{\Sigma})^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right) \\ &= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho_{xy}^2}} \\ &\times \exp\left\{-\frac{1}{2(1-\rho_{xy}^2)} \left[\frac{(x-\mu_x)^2}{\sigma_x^2} - 2\rho_{xy} \frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2} \right]\right\} \end{aligned}$$

Unabhängigkeit

Bei $\rho_{xy} = 0$ gilt $f(x, y) = f_x(x)f_y(y)$ (Unabhängigkeit).

Spur

$$\text{tr}(\mathbf{A}) = \sum_{i=1}^p A_{ii} \quad (\text{Spur} = \text{engl. trace}) \quad (18.57)$$

Stichprobenvarianz

$$s^2 = \frac{1}{N-1} \sum_{n=1}^N (x_n - \bar{x})^2 \quad (18.58)$$

$$\mathbf{S} = \frac{1}{N-1} \sum_{n=1}^N (\mathbf{x}_n - \bar{\mathbf{x}})(\mathbf{x}_n - \bar{\mathbf{x}})' \quad (18.59)$$

$$\text{Var}(X) = \sigma^2 \quad (18.60) \quad \text{Varianz}$$

$$= \int (x - \mu)^2 f(x) dx$$

(stetige Zufallsvariablen) (18.61)

$$= \sum_i (x_i - \mu) f(x_i)$$

(diskrete Zufallsvariablen) (18.62)

$$P(A \cap B) = P(A) \cdot P(B) \quad (18.63) \quad \text{Unabhängigkeit}$$

$$P(A|B) := \frac{P(A \cap B)}{P(B)} = P(A) \quad (18.64)$$

$$F(x) = P(X \leq x) \quad (18.65) \quad \text{Verteilungsfunktion}$$

$$= \int_{-\infty}^x f(y) dy$$

(stetige Zufallsvariablen) (18.66)

$$= \sum_{y_i \leq x} f(y_i)$$

(diskrete Zufallsvariablen) (18.67)

$$P(A) = \text{Wahrscheinlichkeit des Ereignisses } A \quad (18.68) \quad \text{Wahrscheinlichkeit}$$

$$[0 \leq P(A) \leq 1] \quad (18.69)$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (18.70)$$

$$P(A \cap B) = P(A) \cdot P(B) \text{ bei Unabhängigkeit} \quad (18.71)$$

$$P(A|B) := \frac{P(A \cap B)}{P(B)} \text{ (bedingte Wahrscheinlichkeit)} \quad (18.72)$$

$$= P(A) \text{ bei Unabhängigkeit} \quad (18.73)$$

Zufallsvariable $X = X(\omega); \omega \in \Omega$ (18.74)

$$X : \Omega \rightarrow \mathbb{R} \quad (18.75)$$

Zufallsvektor $\mathbf{x} = \mathbf{x}(\omega) = [X_1(\omega), \dots, X_p(\omega)]'; \omega \in \Omega$ (18.76)

$$\mathbf{x} : \Omega \rightarrow \mathbb{R}^p \quad (18.77)$$

18.2 Beispiel: symmetrischer Würfel

18.2.1 Einmaliges Würfeln

$$\Omega = \{1, 2, 3, 4, 5, 6\} = \{\omega_1, \dots, \omega_6\}$$

(Ergebnismenge)

$$P(\Omega) = 1 \text{ (sicheres Ereignis)}$$

$$P(\omega_i) = 1/6$$

$$\omega_i = i \text{ (Elementarereignisse)}$$

$$X(\omega_i) = i \text{ (Zufallsvariable)}$$

$$f(x_i) = P(X = x_i) = 1/6; x_i = 1, 2, 3, 4, 5, 6$$

(Dichtefunktion)

$$F(x_i) = \sum_{x_j \leq x_i} f(x_j)$$

(Verteilungsfunktion)

$$E[X] = \sum x_i f(x_i) = (1 + 2 + 3 + 4 + 5 + 6)/6 = 3.5$$

$$A = \{1, 3, 5\} \text{ (ungerade Zahlen)}$$

$$P(A) = 1/2$$

$$\bar{A} = \{2, 4, 6\} \text{ (A tritt nicht ein, gerade Zahlen)}$$

$$P(\bar{A}) = 1 - P(A) = 1 - 1/2 = 1/2$$

$$\emptyset = \bar{\Omega} \text{ (unmögliches Ereignis)}$$

$$P(\emptyset) = 1 - P(\Omega) = 0$$

$$B = \{2, 4, 6\} \text{ (gerade Zahlen)}$$

$$A \cap B = \emptyset \text{ (gerade und ungerade Zahl unmöglich)}$$

$$P(A \cap B) = P(\emptyset) = 0$$

$$A \cup B = A \text{ oder } B \text{ tritt ein}$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= 1/2 + 1/2 - 0 = 1$$

18.2.2 Zweimaliges Würfeln

$$\Omega = \{(1, 1), (1, 2), (1, 3), \dots, (6, 6)\} = \{\omega_{11}, \dots, \omega_{66}\}$$

(Ergebnismenge)

$$P(\Omega) = 1 \text{ (sicheres Ereignis)}$$

$$P(\omega_{ij}) = 1/36 = 1/6 \cdot 1/6 \text{ (Würfe unabhängig)}$$

$$\omega_{ij} = (i, j) \text{ (Elementarereignisse, Tupel)}$$

$$X(\omega_{ij}) = i + j = X_1(\omega_{ij}) + X_2(\omega_{ij})$$

= (Zufallsvariable: Würfelsumme)

$$f(x) = P(X = x), x = 2, \dots, 12 \text{ (Dichtefunktion)}$$

$$f(2) = P(X = 2) = 1/36$$

$$f(3) = P(X = 3) = 2/36$$

$$f(4) = P(X = 4) = 3/36$$

$$f(5) = P(X = 5) = 4/36$$

$$f(6) = P(X = 6) = 5/36$$

$$f(7) = P(X = 7) = 6/36$$

$$f(8) = P(X = 8) = 5/36$$

$$f(9) = P(X = 9) = 4/36$$

$$f(10) = P(X = 10) = 3/36$$

$$f(11) = P(X = 11) = 2/36$$

$$f(12) = P(X = 12) = 1/36$$

$$F(x_i) = \sum_{x_j \leq x_i} f(x_j)$$

(Verteilungsfunktion)

$$A = \{(6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)\}$$

(6 im ersten Wurf, 2. Wurf beliebig = 1. Zeile in Omega)

$$P(A) = 6/36 = 1/6$$

$$B = \{(1, 6), (2, 6), (3, 6), (4, 6), (5, 6), (6, 6)\}$$

(6 im zweiten Wurf, 1. Wurf beliebig = 1. Spalte in Omega)

$$A \cap B = \{(6, 6)\}$$

$$P(A \cap B) = P(\{(6, 6)\}) = 1/36 = P(A) \cdot P(B)$$

(Ereignisse sind unabhängig)

In übersichtlicher Tabellenform kann man die Elementarereignisse als Tupel $\omega_{ij} = (i, j)$

ω_{ij}	1	2	3	4	5	6
1	(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
2	(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
3	(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
4	(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
5	(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
6	(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

und die Summenwerte als

$X(\omega_{ij})$	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

anschreiben. Daraus ergeben sich unmittelbar die Wahrscheinlichkeiten $f(x) = P(X = x)$, z.B. $P(X = 7) = 6/36$.

In Abb. 18.2 sind die Ergebnismenge Ω sowie die Ereignisse $A = \{(3, 2), \dots, (5, 4)\}$ (rot), $B = \{(3, 4), \dots, (4, 5)\}$ (blau) sowie die Mengen $A \cap B = \{(3, 4), (4, 4)\}$ (lila), $A \cup B$ (grün), \bar{B} (blau), $A \setminus B$ (rot) sowie die bedingten Wahrscheinlichkeiten $P(A|B) = P(A \cap B)/P(B) = (2/36)/(4/36) = 2/4 = .5$ und $P(B|A) = P(A \cap B)/P(A) = (2/36)/(9/36) = 2/9 = .22$ dargestellt.

Die Ereignisse A und B sind voneinander abhängig, da $P(A \cap B) = 2/36 \neq P(A) \cdot P(B) = 9/36 \cdot 4/36 = 1/36$.

Entsprechend gilt für die bedingten Wahrscheinlichkeiten: $P(A|B) = 2/4 \neq P(A) = 1/4$ und $P(B|A) = 2/9 \neq P(B) = 4/36$.

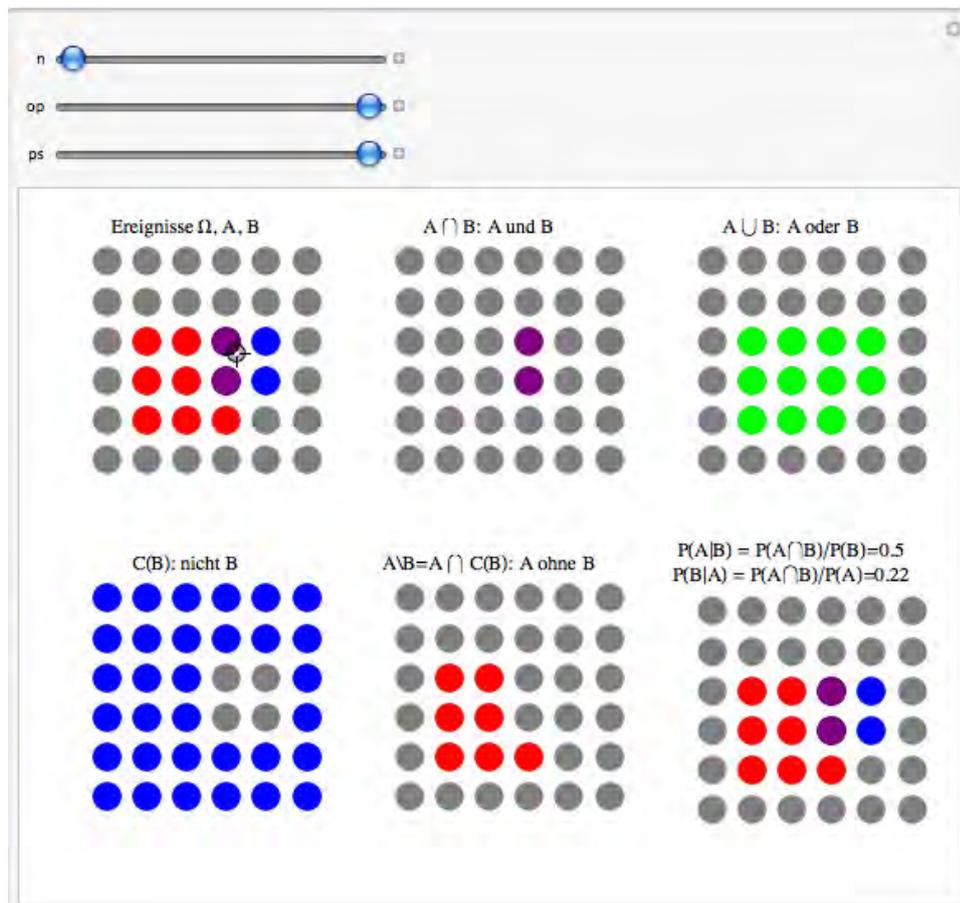


Abbildung 18.2: Ereignisse beim zweimaligen Würfeln.

Diese Seite bleibt aus technischen Gründen frei

Index

- 1 – α -Quantil, 324
- F -Test, 324
- F -Verteilung, 38
 - Dichtefunktion, 508
 - Erwartungswert, 508
 - Parameter, 508
 - Varianz, 508
- α -Fehler, 74
- β -Fehler, 74, 78
- χ^2 -verteilt, 32
- χ^2 -Verteilung
 - Approximation, 504
 - Definition, 504
 - Dichtefunktion, 504
 - Erwartungswert, 504
 - Parameter, 504
 - Tabelle, 504
 - Varianz, 504
- k -tes
 - Stichprobenmoment, 54
- p -Wert, 92, 95, 120, 125, 215, 253, 264, 269, 283, 290
- t -Konfidenzintervalle für Regressionsparameter, 152
- t -Test für gepaarte Stichproben, 253
- t -Tests für Regressionsparameter, 153
- t -Verteilung, 35
 - Dichtefunktion, 511
 - Erwartungswert, 511
 - Parameter, 511
 - Tabelle, 511
 - Varianz, 511
- (un)abhängige Variablen, 203
- a posteriori-Fehlermaße, 232
- Abschwächungs-Korrektur , 369
- adjustiertes Signifikanzniveau, 325
- Äquivalenz, 226, 309
- Analytisch-Nomologische Wissenschaft, 238
- Annahmen beim klassischen Regressionsmodell, 149
- ANOVA-Tafel, 177, 324
- Anpassungstest, 100
- Antezedenz-Bedingungen, 202, 237
- apriori-Fehlermaße, 228
- Arten von Grundgesamtheiten, 20
- Assoziationen, 296
- Asymmetrie beim Test, 83
- asymmetrisch, 297
- Aufbau eines Hypothesentests, 71
- Aussagen
 - Je-Desto-, 236
 - Wenn-Dann-, 236
- Bayes-Formel, 305, 311, 525
- Begründungszusammenhang, 238
- bedingte Häufigkeiten(Spalte), 282
- bedingte Häufigkeiten(Zeile), 282
- bedingte Wahrscheinlichkeit, 525
- Behrens-Fisher-Problem, 105
- Beobachtungs-Studie, 206
- Bernoulli-Variablen, 25
- Bernoulli-Verteilung, 280
- Bestimmtheitsmaß
 - (Regression), 163
- Beurteiler, 211
- Bevölkerungspyramide, 414
- Bindestrich-Metriken, 187
- Binomialverteilung, 25
 - Additionseigenschaft, 492
 - Erwartungswert, 492
 - Tabelle, 492
 - Varianz, 492
 - Verteilungsfunktion, 492
 - Wahrscheinlichkeitsfunktion, 492
- bivariate Analyse, 248
- bivariate Normalverteilung, 389

- Bonferroni-Methode, 268
 Bonferroni-Ungleichung, 171
 Bonferroni-Verfahren, 325
 Boolesche Ungleichung, 174
 Box-Plot, 248

 casewise deletion, 264
 Chi-Quadrat, 283
 Cluster der Korrelationen, 422
 Cohen κ , 213, 228, 315
 Cramérs v , 283
 Credit-Scoring, 246
 Cronbach's Alpha, 363

 Daten-Matrix, 258, 271
 Datenerhebung, 187, 210
 Datenerzeugung, 188
 Datenmatrix, 212
 deduktiv-nomologische Erklärung, 237
 Definition, 508, 511
 Design, 205
 Determinationskoeffizient, 324
 Diagnose, 167
 Dichotome Variablen (Indikatoren), 25
 Dichtefunktion, 508, 525
 Dichtefunktion (bivariat), 526
 Dichteschätzungen, 258
 Dimensionalität, 371
 Diskrepanzfunktion
 gewichtete kleinste Quadrate, 397
 ungewichtete kleinste Quadrate, 397
 diskret, 296
 Drittvariable, 296
 Duhem-Quine-These, 218, 240

 Eigenwert (Spektral)-Zerlegung, 381
 Eigenwertzerlegung, 385
 einfache Stichprobe, 19
 einfache Zufallsstichprobe, 207
 Einheitswissenschaft, 238
 Einzelvarianz, 380
 empirische Kovarianzmatrix, 272
 empirische Sozialforschung als übergreifende Disziplin, 190
 empirische Sozialforschung, 235
 Entdeckungszusammenhang, 238
 Ereignisse, 305, 526
 Erklärung, 237
 erklärte Streuung MQE , 342
 Erwartungstreue, 49
 Erwartungswert, 508, 511, 528
 ex-post-facto-Anordnung, 236
 exaktes Signifikanzniveau, 171
 Exhaustion, 202, 217, 240
 Experiment, 236
 Feldexperiment, 237
 Experiment vs. Beobachtung, 205
 experimentelles Design, 322
 Explanandum, 237
 Explanans, 237
 Exploration, 195
 explorative Statistik, 214
 explorative Studien, 197

 Faktor-Ladungs-Matrix, 375
 Faktoren, 375
 Faktoren-Schätzung (Bartlett), 376
 Faktoren-Schätzung (Thompson), 375
 Faktorenanalyse, 209
 Faktorenanalyse (Maximum-Likelihood), 390
 Faktorenanalyse-Modell, 375
 Faktorladungen (Hauptkomponenten-Analyse), 382
 Falsifikation, 218, 239
 Falsifikatoren, 198
 fehlende Daten, 264
 Fehlende Werte, 212
 Fehler 1. Art (α -Fehler), 216, 223
 Fehler 2. Art (β -Fehler), 217, 223
 Fehlerzellen, 225, 233
 Fisher- Z -Statistik, 130
 Freiheitsgrade, 324
 Freiheitsgrade, 32
 Fundamentaltheorem, 376

 Güte-Funktion, 78
 Gamma-Koeffizient (Goodman-Kruskal), 304
 Gauß-Glocke, 31

- Gegenbeispiel, 221
 gemeinsame Häufigkeiten, 281
 geometrische Interpretation, 339
 geschichtete Stichproben, 208
 Gesetz, 236
 Globaler F-Test, 165
 Goodman-Kruskal λ , 285
 Goodman-Kruskal $\lambda(x \rightarrow y)$, 316
 Goodman-Kruskal $\lambda(y \rightarrow x)$, 316
 graphische Datenanalyse, 214
 Grundgesamtheit, 18, 207, 267
- Hat-Matrix \mathbf{P} , 354
 Hauptkomponenten, 383
 Hauptkomponenten-Analyse, 381
 Histogramme, 247
 Homogenität, 371
 Hypothese, 236
 Quanten-Hypothese, 236
 Hypothesen, 70
- Identifikation, 393
 Implikation, 226, 305
 Indikator-Variablen, 280, 293
 Indikatoren, 197, 207, 300, 360
 Induktionslogik, 239
 induktiv-statistische Erklärung, 237
 induktive Statistik, 17
 Interaktionseffekt, 217
 Interpretation von Statistiken, 299
 Intersubjektivität, 238
 Intervall-Skala, 295
 Item-Analyse, 207, 370
 Item-Schwierigkeit, 370
- kategorial, 296
 Kausalanalyse, 236
 Kausalitätsprinzip, 238
 Kerntheorie, 218
 klassische Testtheorie, 213, 360
 kleinste Quadrate-Schätzer
 (KQ, OLS), 144
 Klumpenstichproben, 207
 Kodierung, 210
 Kointegration, 251
 Kommunalität, 380
 Konditionalsatz, 198
 Konfidenzintervall
- Anteilswert, 65
 Erwartungswert, 58
 Erwartungswert, σ unbekannt, 62
 Korrelation, 69
 Prognose, 158
 Regressionsgerade, 156
 Varianz, 66
- Konkrete Stichproben, 20
 Konstanthaltung von Störvariablen,
 206
 Kontingenzkoeffizient, 283
 kontinuierlich, 296
 Kontrapositionsgesetz, 307
 Kontrast, 450
 Kontrollvariablen, Moderator-
 variablen, 203
 Kontrollvariablen, 329
 Konvergenz in Wahrscheinlichkeit, 52
 Konversationsanalyse, 240
 Kooperationsbereitschaft, 201
 Korrelation, 248, 263, 302
 Korrelations-Matrix, 263
 Korrelationskoeffizient (Population),
 127
 Korrelationskoeffizient (Stichprobe),
 128
 Korrelationskoeffizient, 44
 Korrelationsmatrix, 419
 Korrelierte Stichprobenvariablen, 24
 Kovarianz, 528
 Kovarianzmatrix, 529
 KQ-Schätzer, 320, 353
 KQ-Schätzwerte, 288
 Kronecker Delta-Symbol, 46
 Kronecker-Delta-Symbol, 525
 kumulierte
 Verteilungsfunktion, 31
- Labor vs. Feld, 205
 latent, 298
 Likelihood-Funktion, 54
 Likert-Skala, 197, 369
 linearer Prädiktor, 320, 341
 lineares Modell, 251
 lineares Regressionsmodell, 141
 listwise deletion, 265
 logische Widersprüche, 199

- manifest, 298
 MANOVA, 330
 Matrix-Streudiagramm, 256
 Maximum-a posteriori(MAP)-Schätzer, 54
 Maximum-Likelihood(ML)-Schätzer, 54, 56, 392
 McNemar-Test (Approximation), 123
 Meßniveau, 203
 Meßwiederholungen, 116, 206
 Median, 44
 Messung
 Meßprozeß, 235
 Methodenkanon der empirischen Sozialforschung, 191
 metrisch, 296
 metrische Variable, 247
 Meßprozeß als Interaktion, 189
 Mittelwert, 18, 22, 272, 529
 Moderatorvariablen, 401
 Modus, 300
 modus tollens, 307
 Mosaik-Plot, 419
 Multinomialverteilung, 44
 multiple Korrelation, 343, 359
 multiples Regressionsmodell, 253
 multivariate Analyse, 297
 multivariates Regressionsmodell, 352

 nichtparametrische Regressionsmodelle, 338
 nichtparametrische Verfahren, 93
 Nominal-Skala, 295
 nominale Variable, 247
 nomologisch, 238
 Normal (Gauß)-Verteilung, 29
 Normal-Quantil-Plot, 167
 Normalverteilung, 31, 248
 bivariat, 530
 Dichtefunktion, 500
 Erwartungswert, 500
 Parameter, 500
 Tabelle, 501
 Unabhängigkeit, 530
 univariat, 529
 Varianz, 500
 Verteilungsfunktion, 500

 Normalverteilung der KQ-Schätzer, 150
 Notation und Rechenregeln, 525
 Nullhypothese (ANOVA), 323
 Nullhypothese der Varianzanalyse, 171, 174
 numerisches Relativ, 213

 objektive Realität/
 Realismus, 235
 Objektivität, 213
 Operationalisierung, 197
 Optimale lineare Prognose, 297
 optimale lineare Prognose, 341
 Optimale Prognose, 297
 optimale Prognose, 338
 Ordinal-Skala, 295
 ordinale Variablen, 226
 ordinale Zusammenhangsmaße, 303

 Panel-Studie, 236
 Paneldaten, 253
 Paradigma
 interpretatives, 240
 Parallel-Test
 (Re-Test)-
 Reliabilität, 361
 Parameter der F -Verteilung, 508
 partielle Korrelation, 297, 334, 344
 Phasen des Forschungsprozesses, 195
 Phi-Koeffizient, 283
 Positivismusstreit, 238
 Positivismus, 238
 Prädiktor, 251
 Prädiktionsregeln, 315
 PRE, 213
 PRE-Maß, 315
 PRE-Maß
 (Regression), 162
 pre-test/post-test, 116
 Produkt-Moment-Korrelation, 68, 289, 321
 Profilplot, 442
 Prognose-Fehler, 338, 342
 Prognose-Fehler-Varianz, 339
 Prognose-Regeln, 181, 323, 330
 Prognosefehler, Residuen, 355
 prognostizierte Werte, 354

- Projektionsmatrizen, 353, 355
 proportionale Fehlerreduktion, 315
 prozentuale Fehlerreduktion PRE, 182
 psychometrische Qualität, 213
 Punktschätzung, 41

 qualitativ, 296
 qualitative Methoden, 192, 240
 Quantil, 31
 quantitativ, 296
 quantitative Auswertung qualitativer Variablen, 235
 quantitative Methoden, 191
 quantitative Methoden/
 quantitatives Paradigma, 235
 quasi-experimentelle Anordnung, 236
 Querschnittsanalyse, 236
 Quoten-Auswahl, 208

 Randbedingungen, 218
 Randhäufigkeiten, 282
 Randomisierung, 205
 Rater, 213
 Rating-Skala, 212
 Reaktivität, 189
 Realisation, 19
 Regressionsmodell, 236
 Regressionsmodelle, 251
 Reihenfolge der Messungen, 188
 relative Häufigkeit, 18, 298
 Reliabilität, 213, 360
 Repräsentativität, 207
 residuale Kovarianzen, 358
 Residualstreuung, 343
 Residuen, 145

 Säkulargleichung, 385
 SAS/JMP, 269
 Satz von der
 Normalkorrelation, 341
 Schätzer, 42
 Schätzfehler, 49
 Schätzwert, 42
 Scheinkorrelation, 131, 204, 251, 296
 Schereneffekt, 436
 scree plot, 382

 Selektionsprozesse, 206
 Signifikanz und Relevanz, 208
 Signifikanztest, 222
 Signifikanzniveau, 216, 267
 simultane Wahrscheinlichkeitsverteilung, 189
 Sinnzusammenhang, 240
 Skalen-Niveau, 295
 Skalenniveau, 235, 278, 296
 Spearmans Rangkorrelation, 133
 Sphärizitäts-Test, 268
 Split-Half-Reliabilität, Spearman-Brown-Formel, 362
 SPSS, 269
 Spur, 385, 530
 SQE, 176, 323, 334
 SQR, 176, 323, 334
 SQT, 176, 323, 334
 Störvariablen, 204
 Standard-
 normalverteilung, 500
 Standardabweichung, 28
 Standardisierung, 34, 238
 Statistik, 22, 265
 Statistik als übergreifende Disziplin, 187
 Statistiken, 214, 245
 statistische Einheiten, 212
 Stem and Leaf-Diagramm, 414
 Stichprobe, 19
 Stichproben-Kovarianz-Matrix, 356
 Stichprobenverteilungen, 29
 Stichprobengröße, 208
 Stichprobenvariablen, 19
 Stichprobenvarianz, 26, 530
 Stichprobenvarianz (ML), 55
 stochastische Regressoren, 150, 203
 Streudiagramm, 255
 Streuungszersetzung, 176, 321, 324, 334, 336, 340, 342, 357
 Strukturgleichungsmodell, 236, 449
 Student-*t*-Verteilung, 37
 Survey-Studie, 203, 236
 Symmetrie, Asymmetrie, 251
 symmetrische, asymmetrische Fragestellung, 214
 Systematik, 295

- Tabelle
t-Verteilung, 511
Tabelle der *F*-Verteilung, 508
Tabelle der Binomialverteilung, 492
Tabelle der Normalverteilung, 501
Tau-Koeffizient (Kendall), 304
Tautologien, Kontradiktionen, 198
Test \Leftrightarrow
 Konfidenzintervall, 125
Test der Varianz, 88
Test des
 Medians
 (Wilcoxon-Vorzeichen-Rangtest), 515
Test des Anteilswerts, 91
Test des Erwartungswerts, 86
Test des Erwartungswerts
 (2-Stichprobenfall), 104, 114
Test des Erwartungswerts
 (abhängige Stichproben),
 117
Test des Korrelationskoeffizienten,
 130
Test des Medians (Vorzeichentest),
 93
Test des Medians (Wilcoxon-
 Vorzeichen-Rangtest), 96
Test von Anteilswerten (2-
 Stichprobenfall), 109
Test von Varianzen
 (2-Stichprobenfall), 112
theoretische Struktur, 197
Theorie, 236
 Instrumententheorie, 240
 Kerntheorie, 240
transponierte Matrix, 272
Trennschärfe, 370
true score, 360
Tschebyscheff-Ungleichung, 52
Überschreitungswahrscheinlichkeit
 (*p*-Wert), 125, 215
Übereinstimmung von Beobachtern,
 226
Unabhängigkeit, 19, 137, 531
Unabhängigkeit und Korrelation, 138
Unabhängigkeitstest, 137
Union-Intersection-Prinzip, 269
univariate Signifikanztests, 267
Validität
 externe, 205, 237
 interne, 205, 237
Variablen, 202
 abhängige, 236
 unabhängige, 236
Variablen-Ansicht, 275
Variablenansicht, 247
Varianz, 508, 511, 531
Varianz-Analyse (ANOVA), 322
Varianz-Zerlegung, 160, 363
Varianzanalyse, 209, 215
Varianzanalyse (Effekt-Darstellung),
 180
Varianzanalyse (Grundmodell), 176
Verhältnis-Skala, 295
Verifikation, 223, 239
Verifikation, Falsifikation, 200, 239
Verkürzung der Realität, 210
Verschiebungssatz, 27
Versuchsleiter-Effekte, 210
Verteilungsfunktion, 492
Verteilungsfunktion, 500, 531
Verwertungszusammenhang, 238
Verzerrung, 49
wahre Gerade, 144
wahrer Wert, 298
Wahrheitstafel, 225
Wahrscheinlichkeit, 531
Wahrscheinlichkeitsfunktion, 492
Wahrscheinlichkeit der Ursache, 308
Wertfreiheit, 238
Wertneutralität, Wertfreiheit, 238
Wilcoxon-Rangsummen-Test, 114,
 516
Wilcoxon-Vorzeichen-Rangtest, 97,
 515
zentraler Grenzwertsatz, 63
Zentrierungsmatrix, 353
Zufallsvariable, 532
Zufallsvektor, 532
Zufallszahlen, 20
Zufallszahlentabelle, 520

zusammengesetzte Hypothese, 83
Zusammenhang, 263
Zusatzhypothesen, 218

Literaturverzeichnis

- Adorno, T. (Hrsg.) (1979). *Der Positivismusstreit in der deutschen Soziologie*, 9. Aufl., Luchterhand, Darmstadt, Neuwied.
- Bamberg, G., Baur, F. und Krapp, M. (2008). *Statistik*, 14. Aufl., Oldenbourg, München.
- Bergmann, J. (1987). *Klatsch*, de Gruyter, Berlin.
- Bergmann, J. (1988). *Ethnomethodologie und Konversationsanalyse*, Fernuniversität in Hagen, Hagen. Kurs Nr. 03772.
- Bortz, J. (2005). *Statistik für Sozialwissenschaftler*, 6. Aufl., Springer, Berlin.
- Bortz, J. und Döring, N. (2006). *Forschungsmethoden und Evaluation*, 4. Aufl., Springer, Berlin.
- Box, G. und Tiao, G. (1973). *Bayesian Inference in Statistical Analysis*, Addison Wesley, Reading, MA.
- Davidson, R., MacKinnon, J. G. et al. (2004). *Econometric theory and methods*, Vol. 5, Oxford University Press New York.
- Duhem, P. (1998 (1906)). *Ziel und Struktur der physikalischen Theorien*, Meiner, Hamburg.
- Fahrmeir, L., Hamerle, A. und Tutz, G. (Hrsg.) (1996). *Multivariate Statistische Methoden*, 2. Aufl., de Gruyter, Berlin, New York.
- Fahrmeir, L., Künstler, R., Pigeot, I. und Tutz, G. (2007). *Statistik*, 6. Aufl., Springer, Berlin, Heidelberg, New York.
- Friedrichs, J. (1982). *Methoden empirischer Sozialforschung*, 10. Aufl., Westdeutscher Verlag, Opladen.
- Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N. und Altman, D. G. (2016). Statistical tests, p values, confidence intervals, and power: a guide to misinterpretations, *European Journal of Epidemiology* **31**(4): 337–350.
- Grünbaum, A. (1976). Ad hoc auxiliary hypotheses and falsificationism, *British Journal for the Philosophy of Science* **27**: 329–362.

- Gudder, S. (1979). *Stochastic Methods in Quantum Mechanics*, Elsevier, New York.
- Haken, H. (1977). *Synergetics*, Springer, Berlin.
- Harding, S. (Hrsg.) (1976). *Can Theories be Refuted*, Dordrecht, Boston.
- Heinze, T. und Kleining, G. (1996). *Qualitative Sozialforschung*, Fernuniversität in Hagen, Hagen. Kurs Nr. 3655, 3656.
- Holzkamp, K. (1972). *Kritische Psychologie*, Fischer, Frankfurt a.M.
- Jaynes, E. (2003). *Probability Theory – The Logic of Science*, Cambridge University Press, Cambridge, UK.
- Jordan, P. (1951). *Verdrängung und Komplementarität*, Stromverlag, Hamburg.
- Kromrey, H. (2001). *Empirische Sozialforschung*, Fernuniversität in Hagen, Hagen. Kurs Nr. 03607.
- Kromrey, H. (2002). *Empirische Sozialforschung*, UTB, Leske + Budrich, Opladen.
- Kuffner, T. A. und Walker, S. G. (2018). Why are p-values controversial?, *The American Statistician* pp. 1–3.
- Luhmann, N. (1988). *Soziale Systeme*, Suhrkamp, Frankfurt am Main.
- Mardia, K., Kent, J. und Bibby, J. (1979). *Multivariate Analysis*, Academic Press, London.
- McGrayne, S. B. (2014). *Die Theorie, die nicht sterben wollte*, Springer, Berlin.
- Meehl, P. (1978). Theoretical Risks and Tabular Asterisks: Sir Karl, Sir Ronald and the Slow Progress of Soft Psychology, *J. of Consulting and Clinical Psychology* **46,4**: 806–834.
- Möbus, C. und Nagl, W. (1983). Messung, Analyse und Prognose von Veränderungen (Measurement, analysis and prediction of change; in german), *Hypothesenprüfung, Band 5 der Serie Forschungsmethoden der Psychologie der Enzyklopädie der Psychologie*, Hogrefe, Göttingen, pp. 239–470.
- Nagl, W. (1992). *Statistische Datenanalyse mit SAS*, Campus, Frankfurt/New York.
- Nelson, E. (1967). *Dynamical Theories of Brownian Motion*, Princeton University Press, Princeton.
- Pispers, V. (2001). *Gefühlte Wirklichkeiten, con anima*, Düsseldorf.

- Popper, K. (1984, 1935). *Logik der Forschung*, 8. Aufl., J.C.B Mohr, Tübingen.
- Rao, C. (1973). *Linear Statistical Inference and Its Applications*, 2 Aufl., Wiley, New York.
- Rothenberg, T. (1971). Identification in parametric models, *Econometrica* **39,3**: 577–591.
- Schnell, R., Hill, P. und Esser, E. (1999). *Methoden der empirischen Sozialforschung*, Oldenbourg, München, Wien.
- Shiryayev, A. N. (1995). *Probability*, 2. Aufl., Springer-Verlag, Berlin, Heidelberg.
- Strauss, A. (1984). *Qualitative Analysis in Social Research: Grounded Theory Methodology*, Fernuniversität in Hagen, Hagen. Kurs Nr. 03758.
- van Dijkum, C., Blasius, J., Kleijer, H. und Van Hilten, B. (Hrsg.) (2004). *Sixth International Conference on Social Science Methodology*, Netherlands Institute for the Social Sciences (SISWO), CD ROM, ISBN 90-6706-176-x.
- Wasserstein, R. L., Lazar, N. A. et al. (2016). The ASA's statement on p-values: context, process, and purpose, *The American Statistician* **70**(2): 129–133.
- Weidlich, W. und Haag, G. (1983). *Quantitative Sociology*, Springer, Berlin.