Gerhard Marinell
Gabriele Steckel-Berger

Einführung in die Statistik

Anwendungsorientierte Methoden
zur Datenauswertung

Bachelor

Oldenbourg

Einführung in die Statistik

Anwendungsorientierte Methoden
zur Datenauswertung

von
Gerhard Marinell
und
Gabriele Steckel-Berger

R. Oldenbourg Verlag München Wien

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Rosenheimer Straße 145, D-81671 München
Telefon: (089) 45051-0
oldenbourg.de

Lektorat: Wirtschafts- und Sozialwissenschaften, wiso@oldenbourg.de
Herstellung: Anna Grosser
Satz: DTP-Vorlagen des Autors
Coverentwurf: Kochan & Partner, München
Gedruckt auf säure- und chlorfreiem Papier
Druck: Grafik + Druck, München
Bindung: Thomas Buchbinderei GmbH, Augsburg

ISBN 978-3-486-58200-0

VORWORT

Die Anwendung statistischer Verfahren wird heute wesentlich durch die Verfügbarkeit von statistischer Software erleichtert. Das Erstellen von Häufigkeitsverteilungen, Maßzahlen, Schätzintervallen und Tests erfolgt nach Eingabe der Daten „auf Knopfdruck“. Dieser Vorteil hat aber auch seine Schattenseiten. Oft fehlt das Verständnis für die Rechenverfahren, die der Computer anwendet. Die Folge sind fehlerhafte oder falsche Interpretationen der Computerergebnisse. Dieses Buch soll helfen, einfache univariate und multivariate statistische Verfahren zu verstehen, um solche Fehlinterpretationen zu vermeiden. Dazu wird jedes Verfahren an einem einfachen numerischen Beispiel durchgerechnet und erklärt.

Die Auswahl und Darstellung der statistischen Verfahren orientiert sich insbesondere an den Bedürfnissen von Personen wie z. B. Diplomanden oder Dissertanten, die empirische Erhebungen aufbereiten und auswerten müssen, um ihre Forschungshypothesen zu überprüfen. Die mathematische Herleitung der Formeln steht dabei nicht im Vordergrund, sondern die Erklärung des Rechenvorganges und damit die adäquate Interpretation der Auswertungsergebnisse.

Berücksichtigt werden Schätz- und Testverfahren für eine Stichprobe, für zwei und mehr Stichproben sowie Verfahren, die auf dem Zusammenhang von zwei Stichproben beruhen. Dabei wird nicht nur zwischen abhängigen und unabhängigen Stichproben unterschieden, sondern auch die Skalierung der empirischen Variablen beachtet. Neben diesen univariaten Verfahren werden auch multivariate Verfahren wie Varianz-, Regressions- und Faktorenanalyse behandelt.

Für die beschriebenen Verfahren haben die Autoren ein anwenderfreundliches Softwarepaket für Mathcad 13 erstellt, das von der Homepage

http://homepage.uibk.ac.at/~c40314/default.htm

zum Download zur Verfügung steht (Mathcad ist ein eingetragenes Warenzeichen von MathSoft, Inc. 101 Main Street Cambridge, MA 02142 USA). Diese Software besteht aus zwei Varianten. Wenn man eine Erhebung mit Rohdaten und zahlreichen Fragen auswerten will, dann wählt man die Datei „Statistik“. Will man hingegen eine einzelne Frage auswerten, die nicht unbedingt in Form von Rohdaten vorliegen muss, sondern auch als Häufigkeits- oder Kreuztabelle, dann wählt man die Datei „Stat-Rechner“.

Innsbruck, September 2006. Gerhard Marinell, Gabriele Steckel-Berger

INHALTSVERZEICHNIS

1_0 EINE STICHPROBE

Skalierung

Für die neue Spielkonsole Yoki soll der Käufermarkt untersucht werden, der circa 100000 potentielle Kunden umfasst. Jeder Teil dieser Grundgesamtheit ist eine Stichprobe. Für die Anwendung statistischer Verfahren, die in diesem Buch behandelt werden, muss vorausgesetzt werden, dass die Stichproben so genannte Zufallsstichproben sind. Der einfachste Fall einer Zufallsstichprobe liegt vor, wenn jeder potentielle Kunde der Spielkonsole Yoki die gleiche Chance hat, in die Stichprobe zu gelangen.

Bei einer Werbeveranstaltung für die neue Spielkonsole Yoki wurde den 30 (zufällig ausgewählten) potentiellen Kunden die Frage gestellt, ob sie die Spielkonsole Yoki kaufen werden. Die Zahl 30 wird als Stichprobenumfang n bezeichnet und die Gliederung der Befragten nach den Antworten "Ja", "Nein" und "Weiß nicht" als Stichprobenverteilung.

Die Antworten sind nominal skaliert. Zwischen "Ja", "Nein" und "Weiß nicht" besteht keine Rangfolge und noch weniger die Möglichkeit, eine Antwort als das Vielfache einer anderen auszudrücken. Von ordinaler Skalierung spricht man z. B. bei den Antworten sehr gut, gut, mittelmäßig, schlecht und sehr schlecht auf die Frage "Wie beurteilen Sie die Handlichkeit der neuen Spielkonsole Yoki?". "Wie alt sind Sie (in Jahren)?" ist eine metrisch skalierte Frage. Hier werden keine Antworten vorgegeben, sie sind selbsterklärend. Die Antworten sind metrisch skaliert, da zwischen ihnen sinnvolle Verhältnisse gebildet werden können. 40 Jahre alt ist doppelt so alt wie 20 Jahre.

Für die Auswertung werden die Fragen kodiert. So werden z. B. die Antworten auf die Frage "Werden Sie die Spielkonsole Yoki kaufen?" mit 1 (= Ja), 2 (= Nein) und 3 (=Weiß nicht) kodiert. Die Beurteilung der Handlichkeit wird mit 1 (= sehr gut), 2 (= gut), ..., 5 (= sehr schlecht) ausgedrückt. Nur bei einer metrisch skalierten Frage wie das Alter, ist keine Kodierung erforderlich. Hier stimmt die Antwort mit der "Kodierung" überein.

Wozu dient diese Einteilung der Stichprobenverteilungen in nominal, ordinal oder metrisch skaliert? Die Skalierung ist für die Auswahl des geeigneten Verfahrens relevant, um von der Stichprobe auf die Grundgesamtheit zu schließen. Jedes statistisches Verfahren setzt eine bestimmte Skalierung der Antworten auf eine Frage voraus.

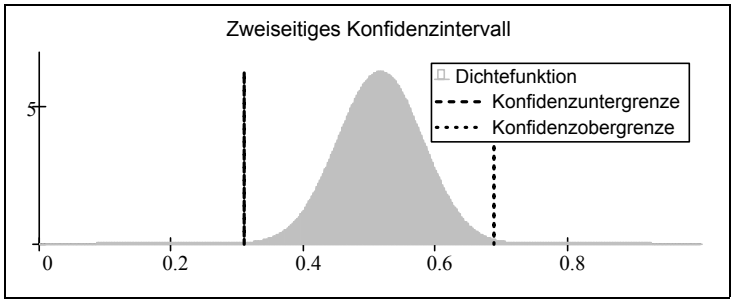
Schätzverfahren

Von den 30 zufällig ausgewählten potentiellen Käufern der Spielkonsole Yoki gaben 15 an, dass sie die Spielkonsole kaufen werden, das sind 50% der Befragten. Kann man auf Grund dieses Stichprobenergebnisses annehmen, dass auch in der Grundgesamtheit aller potentiellen Kunden 50% die Spielkonsole kaufen werden?

Fast sicher nicht. Es ist sehr unwahrscheinlich, dass in der Grundgesamtheit auch genau 50% die Spielkonsole kaufen werden. Es ist aber zu erwarten, dass der Prozentsatz der Käufer in der Nähe von 50% liegt. Was in der Nähe von 50% genau heißt, kann man mit Hilfe eines so genannten Konfidenzintervalls präzisieren. So kann man z. B. mit einem Vertrauen von 95% erwarten, dass der unbekannte Käuferanteil in der Grundgesamtheit der 100000 potentiellen Kunden zwischen 31% und 69% liegt.

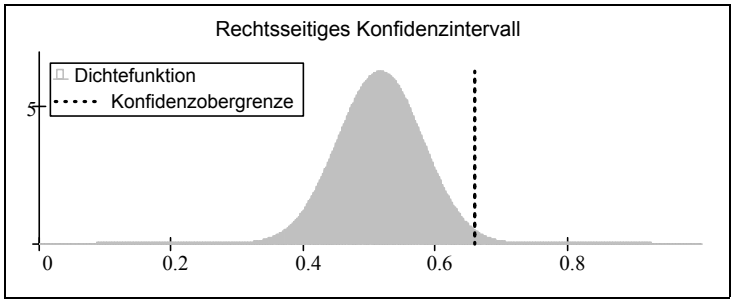
Wie kommt man zu diesem Ergebnis? Man unterscheidet zwei Arten von Verfahren: Punktschätzungen und Intervallschätzungen. Bei einer Punktschätzung wird für den zu schätzenden Anteilswert an Hand der Stichprobenergebnisse lediglich ein einziger Schätzwert (Punktschätzwert) bestimmt (hier also 0.50), oder ein Intervall, in dem der unbekannte Anteilswert der Grundgesamtheit liegt. Beide Verfahren unterscheiden sich vor allem dadurch, dass das Vertrauen fast Null ist, dass der Punktschätzwert aus der Stichprobe mit dem entsprechenden unbekannten Wert in der Grundgesamtheit übereinstimmt. Beim Konfdenzintervall kann man hingegen mit einem vorgegebenen Vertrauen (z. B. 95%) berechnen, innerhalb welcher Grenzen der unbekannte Anteil in der Grundgesamtheit zu erwarten ist.
.
Diese statistische Verfahren nennt man Schätzverfahren: An Hand einer Stichprobe wird ein Schätzwert sowie ein Konfidenzintervall für einen unbekannten Prozentsatz in der Grundgesamtheit bestimmt. Diese Verfahren werden für verschiedene Maßzahlen und Skalierungen im Abschnitt "Eine Stichprobe" behandelt.

Bei der Berechnung von Konfidenzgrenzen kann man zwischen 3 Fällen unterscheiden: Will man nur wissen, innerhalb welcher Grenzen die unbekannte Maßzahl in der Grundgesamtheit liegt, dann berechnet man ein zweiseitiges Konfidenzintervall: Der unbekannte Prozentsatz der Käufer liegt mit einem Vertrauen von 95% zwischen 31% und 69%. Hier muss man damit rechnen, dass der unbekannte Käuferprozentsatz mit 2.5% Wahrscheinlichkeit kleiner als 31% ist und mit der gleichen Wahrscheinlichkeit, dass er größer als 69% ist.



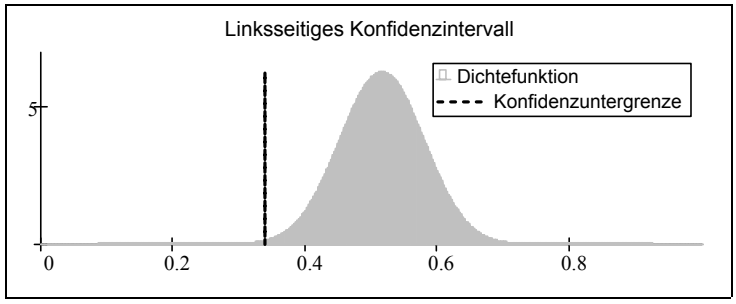
(Konfidenzuntergrenze)	(Konfidenzobergrenze)	(p – Wahrscheinlichkeit)
0.31	0.69	0.05

Will man nur wissen, mit welchem Käuferprozentsatz man beim gleichen Vertrauen von 95% höchstens rechnen kann, dann berechnet man kein zweiseitig begrenztes Konfidenzintervall, sondern ein einseitig begrenztes. Bei "höchstens" ist das Intervall rechtsseitig begrenzt durch 66%. Mit einem Vertrauen von 95% kann man den unbekannten Anteil der Käufer in der Grundgesamtheit höchstens zwischen 0% und 66% erwarten. Dass mehr als 66% Käufer in der Grundgesamtheit sind, hat nur eine Wahrscheinlichkeit von 5%.



(Konfidenzuntergrenze)	(Konfidenzobergrenze)	(p – Wahrscheinlichkeit)
$-\infty$	0.66	0.05

Neben dem rechtsseitig begrenzten Konfidenzintervall gibt es auch die linksseitig begrenzten. Sie entsprechen der Frage: Mit welchen Käuferprozentsatz kann ich in der Grundgesamtheit aller potentiellen Käufer mindestens rechnen bei einem Vertrauen von 95%? Hier ist das Konfidenzintervall linksseitig begrenzt durch 34% und rechts durch 100%. Dass weniger als 34% die Spielkonsole Yoki kaufen werden, hat nur eine Wahrscheinlichkeit von 5%.



(Konfidenzuntergrenze)	(Konfidenzobergrenze)	(p – Wahrscheinlichkeit)
0.34	∞	0.05

Testverfahren

Auf Grund der Kostenrechnung weiß der Produzent der Spielkonsole Yoki, dass mindestens 40% der potentiellen Käufer seine Spielkonsole kaufen müssen, um in die Gewinnzone zu gelangen. Von den zufällig ausgewählten 30 Besuchern einer Werbeveranstaltung haben 15 angegeben, dass sie die Spielkonsole kaufen werden. Kann der Produzent daraus schließen, dass mindestens 40% aller potentiellen Kunden seine Spielkonsole kaufen? Diese Frage wird mit Hilfe eines geeigneten Testverfahrens geklärt.

Zuerst werden die möglichen Annahmen über die Grundgesamtheit in Form von Hypothesen formuliert. Als Nullhypothese nimmt man an, dass der Anteil der Spielkonsolenkäufer in der Grundgesamtheit aller potentiellen Käufer mindestens 40% beträgt. Dies drückt man formal wie folgt aus:

$$H_0: \pi \geq 0.4.$$

π ist der unbekannte Anteil der Spielkonsolenkäufer in der Grundgesamtheit und 0.4 der oben erwähnte Break-even-point des Produzenten. Eine Alternative zu dieser Nullhypothese ist die Annahme, dass dieser unbekannte Anteil π kleiner als der Break-even-point ist:

$$H_1: \pi < 0.4.$$

An Hand der Ergebnisse obiger Stichprobe entscheidet man sich für die Annahme oder Ablehnung der Null- oder Alternativhypothese. Folgende Konsequenzen können bei dieser Entscheidung auftreten: Entscheidet man sich für die Annahme der Nullhypothese, dann kann dies die richtige Entscheidung sein. Der Käuferanteil in der Grundgesamtheit aller Käufer ist größer oder gleich 40%. Der Produzent wird in diesem Fall zumindest keinen Verlust mit der Produktion der Spielkonsole Yoki machen. Wenn es die falsche Entscheidung war, wenn also der Käuferanteil in Wirklichkeit kleiner als 40% ist, dann ist dies eine Fehlentscheidung, die als Fehler 2. Art oder β-Fehler bezeichnet wird.

Wird auf Grund der Stichprobenergebnisse die Alternativhypothese angenommen, dann kann auch dies die richtige Entscheidung sein, wenn der Käuferanteil in der Grundgesamtheit tatsächlich kleiner ist als der Break-even-point von 40%. Der Produzent wird in diesem Fall die Produktion der Spielkonsole nicht aufnehmen und keinen Verlust machen. Ist aber in Wirklichkeit der Käuferanteil über 40%, dann liegt wieder eine Fehlentscheidung vor, die als Fehler 1. Art oder α-Fehler bezeichnet wird. Der Unternehmer muss in diesem Fall mit einem entgangenen Gewinn rechnen. In folgender Übersicht sind diese vier Möglichkeiten nochmals zusammengefasst:

AKTIONEN\ZUSTAND	RICHTIGER ZUSTAND: H_0	RICHTIGER ZUSTAND: H_1
ANNAHME VON H_0	Richtige Entscheidung	Fehler 2. Art oder β–Fehler
ANNAHME VON H_1	Fehler 1. Art oder α–Fehler	Richtige Entscheidung

Beim Signifikanztest berücksichtigt man nur den α-Fehler (oder Fehler 1. Art) durch die Vorgabe eines Signifikanzniveaus. Wird z. B. ein Test mit einem 5% Signifikanzniveau durchgeführt, dann weiß man, dass die Wahrscheinlichkeit höchstens 5% beträgt, die richtige Nullhypothese abzulehnen. Wenn man also die Nullhypothese auf Grund der Stichprobenergebnisse ablehnt, dann kennt man die Obergrenze seines Fehlerrisikos.

Kann die Nullhypothese auf Grund der Stichprobenergebnisse nicht abgelehnt werden, dann kennt man beim Signifikanztest sein Fehlerrisiko nicht. Denn in diesem Fall kann nur ein β-Fehler (oder Fehler 2. Art) auftreten und über seine Höhe weiß man beim Signifikanztest nicht Bescheid.

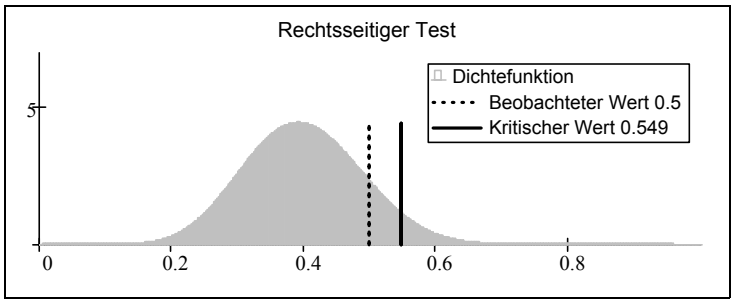
Konventionell sind zwei Signifikanzniveaus üblich, nämlich 5% und 1%. Wird die Nullhypothese auf dem Niveau von 5% abgelehnt, dann spricht man von einem signifikanten Unterschied. Bei einem 1% Signifikanzniveau sagt man, dass der Unterschied sehr signifikant ist.

So wie bei den Konfidenzintervallen kann man auch hier drei Testarten unterscheiden: Zweiseitige Tests, rechtsseitige Tests und linksseitige Tests. Da der Produzent prüfen will, ob der Käuferanteil größer als sein Break-even-point ist

$H_0: \pi \geq 0.4$,

wird er den Ablehnungsbereich seiner Nullhypothese so platzieren, dass er rechts von seinem Break-even-point liegt:

$H_1: \pi < 0.4$



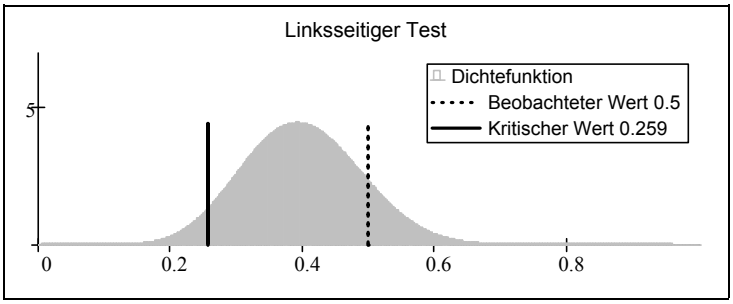
Beobachteter_Wert	Kritischer_Wert	Wahrscheinlichkeit_p
0.5	(0 0.549)	0.05

Wenn man bereit ist, in 5 von 100 Fällen eine richtige Nullhypothese abzulehnen (das Signifikanzniveau ist also 0.05), dann ist der kritische Wert c = 0.549. Die Fläche die rechts von 0.549 liegt umfasst 5% der Gesamtfläche. Wenn also der Stichprobenanteilswert größer gleich diesem kritischen Wert ist, dann kann man die Nullhypothese ablehnen. Ist er kleiner als der kritische Wert von 0.549, dann gehört er zum so genannten Nichtablehnungsbereich der Nullhypothese. Da von den 30 befragten Besuchern 15 angeben, dass sie die Spielkonsole Yoki kaufen werden, ist p der Stichprobenanteil gleich 0.50. Dieser Wert ist kleiner als 0.549. Die Nullhypothese kann daher nicht abgelehnt werden. Das Risiko, dass diese Entscheidung falsch ist, ist bei einem Signifikanztest numerisch nicht bekannt (= β-Fehler). Der Unternehmer kann jedenfalls auf Grund dieser Stichprobe noch nicht annehmen, dass sich unter allen potentiellen Käufern der Grundgesamtheit mehr als 40% tatsächliche Käufer befinden.

Von einem linksseitigen Test spricht man, wenn sich der Ablehnungsbereich der Nullhypothese links von der Nullhypothese befindet.

$H_0: \pi \geq 0.4$,

$H_1: \pi < 0.4$



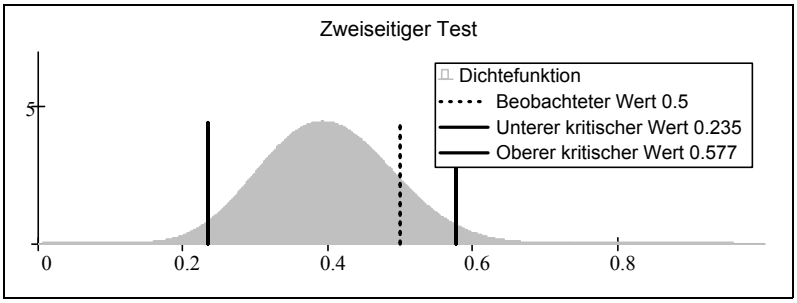
Beobachteter_Wert	Kritischer_Wert	Wahrscheinlichkeit_p
0.5	(0.259 1.000)	0.05

Zum Ablehnungsbereich der Nullhypothese zählen hier alle Werte zwischen 0 und 0.259 und zum Nichtablehnungsbereich alle Werte größer 0.259 bis 1.

Beim zweiseitigen Test zählen zum Ablehnungsbereich der Nullhypothese sowohl Werte links als auch rechts von der Nullhypothese:

H_0: $\pi = 0.4$

H_1: $\pi \neq 0.4$



Beobachteter_Wert	Kritischer_Werte	Wahrscheinlichkeit_p
0.5	(0.235 0.577)	0.05

Hier gehören alle Werte zwischen 0 und 0.235 sowie zwischen 0.577 und 1 zum Ablehnungsbereich der Nullhypothese und alle Werte zwischen 0.235 und 0.577 zum Nichtablehnungsbereich.

Vorschau

Die Frage nach dem Anteil der potentiellen Käufer der neuen Spielkonsole Yoki wird im Abschnitt "1_1 Anteilswert" untersucht. Für die Anwendung dieses Schätzverfahrens ist die Skalierung der Antworten belanglos. Selbst für metrisch skalierte Antworten wie das Alter kann beispielsweise der Anteil der 20-Jährigen in der Stichprobe bestimmt und auf den unbekannten Anteil in der entsprechenden Grundgesamtheit hochgerechnet werden.

Während im Abschnitt 1_1 Schätzverfahren für den Anteilswert gezeigt werden, sind Testverfahren für den Anteilswert einer Stichprobe der Inhalt von Abschnitt 1_2. Das oben angezeigte Problem mit dem Break-even-point wird ausführlich dargestellt und berechnet.

Im Abschnitt "1_3 Zentralwert" wird die Designbeurteilung der neuen Spielkonsole Yoki mit den Urteilen sehr gut bis sehr schlecht ausgewertet und auf die Grundgesamtheit geschlossen. Voraussetzung für die Anwendung dieses Verfahrens ist zumindest ordinale Skalierung der Antworten. Fragen mit nominal skalierten Antworten wie die Frage nach der Kaufabsicht, können mit diesen Verfahren nicht ausgewertet werden.

Testverfahren für ordinal skalierte Antworten einer Frage und Stichprobe werden im Abschnitt 1_4 behandelt. Diese Verfahren können auch für Fragen mit metrisch skalierten Antworten verwendet werden.

Wie viele Stunden verbringen die potentiellen Kunden wöchentlich im Schnitt vor dem Computer? Diese metrisch skalierte Frage wird im Abschnitt "1_5 Durchschnitt" analysiert. Fragen mit nominal oder ordinal skalierten Antworten können mit den Methoden dieses Abschnittes nicht ausgewertet werden. Die entsprechenden Testverfahren für eine Frage mit metrisch skalierten Antworten findet man im Abschnitt 1_6.

1_1 Anteilswert, Schätzverfahren

Wie viele der potentiellen Kunden werden das neue Produkt, die Spielkonsole Yoki, kaufen? Bei einer Werbeveranstaltung werden Kunden befragt. Kennzeichnet man die Kunden ohne Kaufabsicht durch "1" und die mit Kaufabsicht durch "2", dann ergibt sich folgende Stichprobe:

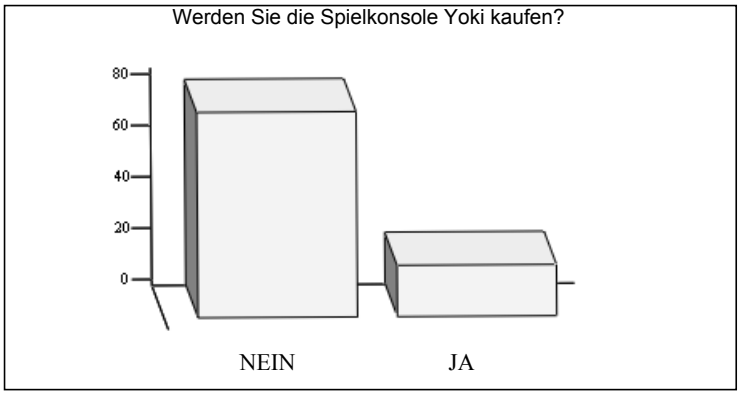
[(Kaufabsicht) 1 1 2 1 1 1 1 2 1 1]

Der Stichprobenumfang ist in diesem Fall n = 10 und das untersuchte Merkmal "Kaufabsicht" kommt x = 2 mal vor. Der Stichprobenanteil ist allgemein definiert als

$$p = \frac{x}{n}$$

und im vorliegenden Beispiel gleich p = 2/10 = 0.2. Ausgezählt nach den beiden Antworten erhält man folgende Häufigkeitstabelle:

(Kaufabsicht)	Häufigkeit	In_Prozent
Nein	8	80
Ja	2	20
SUMME	10	100



2 von 10 Kunden haben Kaufabsicht. Der Anteilswert der Kunden mit Kaufabsicht ist also 0.2 oder in Prozent ausgedrückt 20%. Kann man nun annehmen, dass 20% der 100.000 potentiellen Kunden die neue Spielkonsole "Yoki" kaufen werden?

Sicher nicht. Unter der Voraussetzung, dass die Stichprobe eine Zufallsstichprobe aus der Grundgesamtheit der 100.000 potentiellen Kunden ist, kann man nur erwarten, dass der "wahre" Anteil in der Nähe von 20% liegt. Was "in der Nähe" heißt, ist mit Hilfe der Statistik präzisierbar: Mit einem Vertrauen (Konfidenz) von 95% kann man annehmen, dass der wahre Anteil der Kunden mit Kaufabsicht zwischen 2.5% und 55.6% liegt.
Wie kommt man zu diesem Ergebnis? Da die Ergebnisse einer Stichprobenerhebung zufallsabhängig sind,

wird auch ein gefundener Punktschätzwert nur in den seltensten Fällen genau mit dem gesuchten Anteilswert der Grundgesamtheit übereinstimmen. Um wenigstens Aussagen über ein Intervall machen zu können, in dem der unbekannte Anteilswert zu erwarten ist, führt man eine Intervallschätzung durch. Ausgehend von den Ergebnissen einer Stichprobe wird ein Konfidenzintervall berechnet, das den unbekannten Anteilswert der Grundgesamtheit mit vorgegebener Wahrscheinlichkeit enthält. Folgende Wahrscheinlichkeitsverteilungen kann man dazu verwenden:

Binomial- und Betaverteilung

Für den unbekannten Anteilswert der Grundgesamtheit wird ein zweiseitiges Konfidenzintervall zum Niveau $1-\alpha$ und den Stichprobenergebnissen n und x mit Hilfe folgender Formeln berechnet: Die Untergrenze K_u und die Obergrenze K_o des Konfidenzintervalls werden so bestimmt, dass gilt

$$\sum_{k=0}^{x-1}\left[\frac{n!}{x!\,(n-x)!}\cdot p^{k}\cdot(1-p)^{n-k}\right]=1-\frac{\alpha}{2}$$

$$\sum_{k=x+1}^{n}\left[\frac{n!}{x!\,(n-x)!}\cdot p^{k}\cdot(1-p)^{n-k}\right]=1-\frac{\alpha}{2}$$

$1-\alpha$ ist das Konfidenzniveau des Schätzintervalls (meist 95% oder 99%) und p der gesuchte Anteilswert der möglichen Stichproben, der diese Gleichung erfüllt. Weiter ist n der Stichprobenumfang und n! (gelesen "n Fakultät") gleich

$$n! = n\cdot(n-1)\cdot(n-2)\,..\,3\cdot2\cdot1$$

und

$$0! = 1.$$

Der Stichprobenumfang ist im obigen Beispiel n = 10 und die Stichprobenrealisation x = 2 und $\alpha = 0.05$. K_u muss so bestimmt werden, dass gilt

$$\sum_{k=0}^{2-1}\left[\frac{10!}{k!\,(10-k)!}\cdot p^{k}\cdot(1-p)^{10-k}\right]=1-\frac{0.05}{2}=0.975$$

Von den möglichen p-Werten erfüllt p = 0.025 diese Gleichung:

$$\frac{10!}{0!\cdot(10-0)!}\cdot0.025^{0}\cdot(1-0.025)^{10-0}=0.776$$

$$\frac{10!}{1!\cdot(10-1)!}\cdot0.025^{1}\cdot(1-0.025)^{10-1}=0.199$$

$$\sum_{k=0}^{2-1}\left[\frac{10!}{k!\,(10-k)!}\cdot0.025^{k}\cdot(1-0.025)^{10-k}\right]=0.975$$

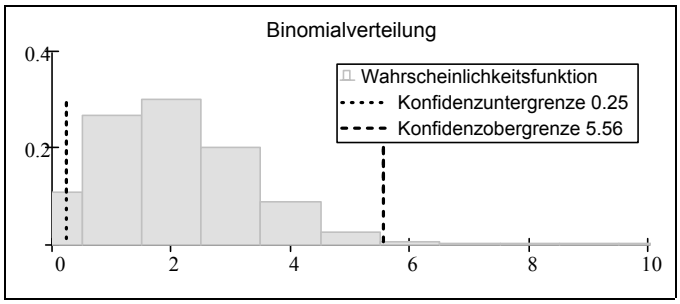
Díe Untergrenze K_u des Schätzintervalls ist daher 0.025. Die Obergrenze wird auf die gleiche Art bestimmt. Es muss gelten

$$\sum_{k=2+1}^{10}\left[\frac{10!}{k!\,(10-k)!}\cdot p^{k}\cdot(1-p)^{10-k}\right] = 1-\frac{0.05}{2} = 0.975$$

Von den möglichen p-Werten erfüllt p = 0.556 diese Gleichung:

$$\sum_{k=2+1}^{10}\left[\frac{10!}{k!\,(10-k)!}\cdot 0.556^{k}\cdot(1-0.556)^{10-k}\right] = 0.975$$

Die Obergrenze des Schätzintervalls ist daher 0.556. Mit einem Vertrauen von 95% kann man in der Grundgesamtheit aller potentiellen Kunden mit 2.5% bis 55.6% Personen mit Kaufabsicht rechnen.



(Konfidenzuntergrenze)	(Konfidenzobergrenze)	(p – Wahrscheinlichkeit)
0.25	5.56	0.05

F-Verteilung

Da die Binomialverteilung durch die F-Verteilung dargestellt werden kann, berechnet man die beiden Konfidenzgrenzen auch mit Hilfe dieser Verteilung. Die Formeln für die beiden Intervallgrenzen für ein zweiseitiges Konfidenzintervall zum Niveau von $1-\alpha$ sind

$$K_u = \frac{x}{x+(n-x+1)\cdot F_1}$$

$$K_o = \frac{(x+1)\cdot F_2}{(n-x)+(x+1)\cdot F_2}$$

mit

$$F_1 = F_{1-\frac{\alpha}{2},\,2\cdot(n-x-1),\,2\cdot x}$$

$$F_2 = F_{1-\frac{\alpha}{2},\,2\cdot(x+1),\,2\cdot(n-x)}$$

F_1 und F_2 sind die entsprechenden Quantile der F-Verteilung. Für obiges Beispiel sind die mit Hilfe dieser Formeln berechneten Konfidenzgrenzen

$$K_u = \frac{2}{2 + (10 - 2 + 1)\cdot 8.59} = 0.025$$

mit

$$F_1 = F_{1-\frac{0.05}{2},\,2\cdot(10-2+1),\,2\cdot 2} = 8.59$$

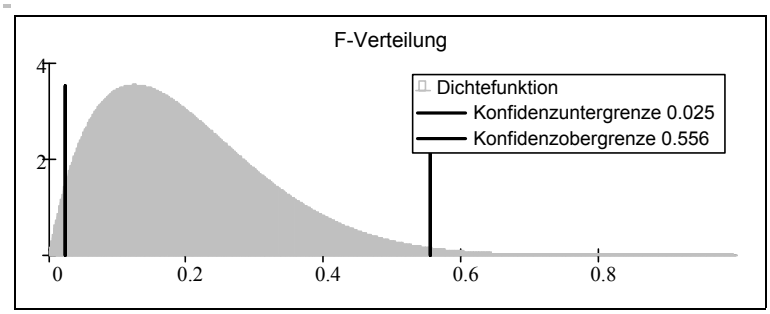
sowie

$$K_o = \frac{(2 + 1)\cdot 3.341}{(2 + 1)\cdot 3.341 + 10 - 2} = 0.556$$

mit

$$F_2 = F_{0.975,\,2\cdot(2+1),\,2\cdot(10-2)} = 3.341$$

Der Anteil der Kunden mit Kaufabsicht liegt auch hier mit 95% Vertrauen zwischen den Grenzen 0.025 und 0.556.



(Konfidenzuntergrenze)	(Konfidenzobergrenze)	(p – Wahrscheinlichkeit)
0.025	0.556	0.05

Die beiden Werte der F-Verteilung $F_{0.975,\,20,\,4} = 8.56$ und $F_{0.975,\,6,\,16} = 3.341$ bestimmt man aus geeigneten Tabellen für die F-Verteilung oder berechnet sie direkt mit Hilfe entsprechender Programme über die Dichte- und Verteilungsfunktion der F-Verteilung.

Normalverteilung

Wenn der Stichprobenumfang n so groß ist, dass die Bedingung n. (x / n). (1 – (x / n)) > 9 erfüllt ist, dann kann man die Konfidenzgrenzen mit Hilfe der Normalverteilung berechnen. Die Untergrenze des Konfidenzintervalls wird nach der Formel bestimmt

$$K_u = \frac{n}{n + z^2} \cdot \left[p + \frac{z^2}{2 \cdot n} - z \cdot \sqrt{\frac{p \cdot (1 - p)}{n} + \frac{z^2}{4 \cdot n^2}} \right]$$

mit

$$p = \frac{x}{n}.$$

z ist für ein zweiseitiges Konfidenzintervall zum Niveau von $1 - \alpha$ gleich dem $(1 - \alpha/2)$-ten Quantil der Standardnormalverteilung. Für ein Konfidenzniveau von z. B. 95% ist z gleich

$$z = 1.96$$

Die Obergrenze des Intervalls wird analog bestimmt:

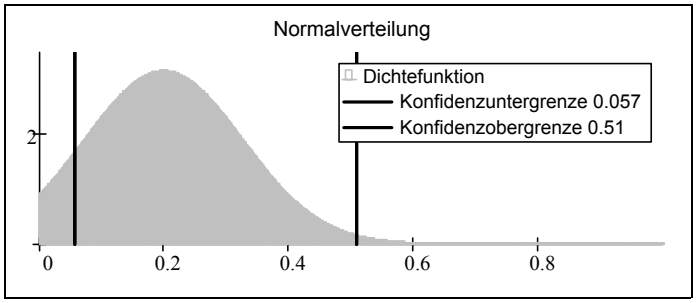
$$K_o = \frac{n}{n + z^2} \cdot \left[p + \frac{z^2}{2 \cdot n} + z \cdot \sqrt{\frac{p \cdot (1 - p)}{n} + \frac{z^2}{4 \cdot n^2}} \right]$$

Für obiges Beispiel sind die mit Hilfe dieser Näherungsformeln berechneten Konfidenzgrenzen

$$K_u = \frac{10}{10 + 1.96^2} \cdot \left[0.2 + \frac{1.96^2}{2 \cdot 10} - 1.96 \cdot \sqrt{\frac{0.2 \cdot (1 - 0.2)}{10} + \frac{1.96^2}{4 \cdot 10^2}} \right] = 0.057$$

$$K_o = \frac{10}{10 + 1.96^2} \cdot \left[0.2 + \frac{1.96^2}{2 \cdot 10} + 1.96 \cdot \sqrt{\frac{0.2 \cdot (1 - 0.2)}{10} + \frac{1.96^2}{4 \cdot 10^2}} \right] = 0.51$$

Da $10 \cdot 0.2 \cdot (1 - 0.2) = 1.6$ nicht größer als 9 ist, ist die Diskrepanz der hier berechneten Intervallgrenzen von den oben angeführten nicht überraschend. Erst wenn die Bedingung $n \cdot (x/n) \cdot (1 - (x/n)) > 9$ erfüllt ist, liefern die hier angeführten Formeln brauchbare Näherungen.



(Konfidenzuntergrenze)	(Konfidenzobergrenze)	(p – Wahrscheinlichkeit)
0.057	0.510	0.05

Zwei- und einseitige Konfidenzintervalle

Bisher lautete die Frage "Wie viele der potentiellen Kunden werden das neue Produkt, die Spielkonsole Yoki, kaufen?" An Hand vorn Stichprobenergebnissen kann man sich ein Intervall berechnen, in dem der unbekannte Anteil der Grundgesamtheit mit einem vorgegebenen Vertrauen liegt. Dieses Intervall hat eine Untergrenze und eine Obergrenze. Es ist ein zweiseitig (begrenztes) Intervall.

Man kann aber auch fragen: "Wie viele der potentiellen Kunden werden das neue Produkt, die Spielkonsole Yoki, mindestens kaufen?" oder auch "Wie viele der potentiellen Kunden werden das neue Produkt, die Spielkonsole Yoki, höchstens kaufen?" In beiden Fällen sucht man ein einseitig begrenztes Konfidenzintervall. Will man z. B. mit einem Vertrauen von 95% wissen, wie viele potentielle Kunden die Spielkonsole mindestens kaufen werden, dann bestimmt man die Untergrenze so, dass 5% (und nicht 2.5% wie beim zweiseitigen Intervall) links vor dieser Untergrenze liegen:

$$\sum_{k=0}^{x-1}\left[\frac{n!}{x!\,(n-x)!}\cdot p^{k}\cdot(1-p)^{n-k}\right] = 1-\alpha = 0.95$$

oder

$$K_u = \frac{x}{x+(n-x+1)\cdot F_1}$$

mit

$$F_1 = F_{1-\alpha,\,2\cdot(n-x-1),\,2\cdot x}$$

Für die Angaben obiger Stichprobe ist

$$F_1 = F_{0.95,\,2\cdot(10-2-1),\,2\cdot 2} = 5.821$$

und die Konfidenzuntergrenze gleich

$$K_u = \frac{2}{2 + (10 - 2 + 1) \cdot 5.821} = 0.037$$

Mit einem Vertrauen von 95% kann man erwarten, dass mindestens 3.7% der potentiellen Kunden die Spielkonsole Yoki kaufen werden. Auf die gleiche Art kann man eine Konfidenzobergrenze bestimmen, wenn die Frage nach "höchstens" geht.

Mit Hilfe von Statistiksoftware kann die umständliche Berechnung der Konfidenzintervalle vereinfacht werden. Auf Seite 180 findet man den Output der Autorensoftware für die Berechnung von Konfidenzintervallen für Anteilswerte Für diesen Output sind nur die Rohdaten (in kodierter Form) einzugeben oder die Häufigkeitsverteilung der Stichprobenergebnisse, wie sie am Beginn dieses Kapitels stehen.

1_2 Anteilswert, Testverfahren

Der Produzent der neuen Spielkonsole Yoki weiß auf Grund einer Break-even-point-Analyse, dass mindestens 35% der potentiellen Kunden die Spielkonsole Yoki kaufen müssen, um in die Gewinnzone zu gelangen. Bei einer Werbeveranstaltung werden 10 Kunden befragt. Kennzeichnet man die Kunden mit Kaufabsicht wieder durch "2" und die ohne Kaufabsicht mit "1", dann ergibt sich folgende Stichprobe:

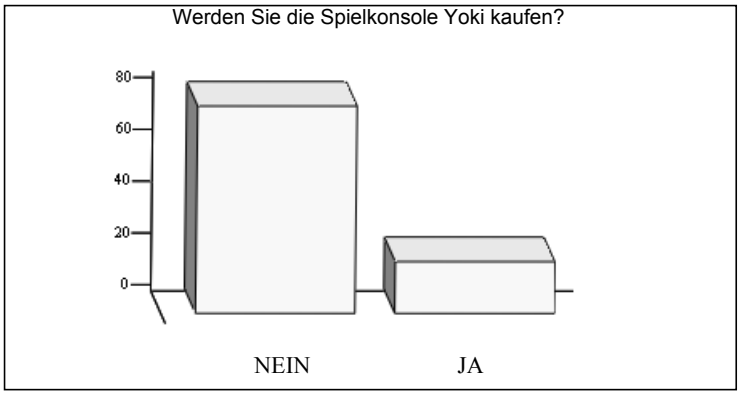
[(Kaufabsicht) 1 1 2 1 1 1 1 2 1 1]

Der Stichprobenumfang ist in diesem Fall n = 10 und das untersuchte Merkmal "Kaufabsicht" kommt x = 2 mal vor. Der Stichprobenanteil ist allgemein definiert als

$$p = \frac{x}{n}$$

und im vorliegenden Beispiel gleich p = 2/10 = 0.2. Ausgezählt nach den beiden Antworten erhält man folgende Häufigkeitstabelle und Grafik:

(Kaufabsicht)	Häufigkeit	In_Prozent
Nein	8	80
Ja	2	20
SUMME	10	100



2 von 10 Kunden haben Kaufabsicht. Der Anteilswert der Kunden mit Kaufabsicht ist also 0.2 oder in Prozent ausgedrückt 20%. Kann man auf Grund dieses Ergebnisses annehmen, dass mindestens 35% der 100000 potentiellen Kunden die neue Spielkonsole "Yoki" kaufen werden oder spricht das Stichprobenergebnis gegen diese Annahme?

Unter der Voraussetzung, dass die Stichprobe eine Zufallsstichprobe aus der Grundgesamtheit der 100000 potentiellen Kunden ist und der "wahre" Anteil in dieser Grundgesamtheit mindestens 0.35 beträgt, kann man auch erwarten, dass in einer Zufallsstichprobe von 10 potentiellen Kunden der Käuferanteil von 0.2 auftritt. Das Stichprobenergebnis von 20% spricht nicht gegen die Annahme, dass in der Grundgesamtheit der Käuferanteil mindestens 35% ausmacht. Man kann die Nullhypothese nicht ablehnen.

Wie kommt man zu diesem Ergebnis? Zuerst werden Null- und Alternativhypothese formuliert. Da der Unternehmer weiß, dass mindestens 35% der potentiellen Kunden die Spielkonsole kaufen müssen, um aus der Verlustzone zu kommen, ist die Nullhypothese

$$H_0: \pi \geq 0.35$$

Mindestens 35% der potentiellen Kunden kaufen die Spielkonsole Yoki. Die Alternativhypothese zu dieser Nullhypothese ist

$$H_1: \pi < 0.35$$

Der Käuferanteil ist kleiner als 0.35. Um die Hypothesen zu prüfen, berechnet man die Wahrscheinlichkeit für das Auftreten eines Anteilswertes von 0.2 in einer Stichprobe vom Umfang 10, wenn in der Grundgesamtheit der Anteilswert 0.35 ist. Die allgemeine Formel für die Berechnung dieser Wahrscheinlichkeit ist

$$f(x, n, \pi) = \frac{n!}{x!\,(n-x)!} \cdot \pi^k \cdot (1-\pi)^{n-k}$$

n ist der Stichprobenumfang, π der Anteil in der Grundgesamtheit und x die Anzahl der Käufer in der Stichprobe. Die Wahrscheinlichkeit für genau 2 Käufer in der Stichprobe von 10 potentiellen Kunden ist daher für einen Anteil von 0.35 in der Grundgesamtheit

$$f(2, 10, 0.35) = \left[\frac{10!}{2! \cdot (10-2)!} \cdot 0.35^2 \cdot (1-0.35)^{10-2}\right] = 0.176$$

Die Wahrscheinlichkeit für die oben ausgewertete Stichprobe ist also unter Gültigkeit der Nullhypothese 17.6%. Ist man bereit in 5 von 100 Fällen eine richtige Nullhypothese abzulehnen, dann kann die Stichprobe nicht als Beleg gegen die Gültigkeit der Nullhypothese verwendet werden, da 17.6% viel größer ist als das Signifikanzniveau von 5%. In den Ablehnungsbereich der Nullhypothese fallen nur jene Stichproben, deren Auftretenswahrscheinlichkeit kleiner gleich 5% sind. Auch bei 1 Käufer unter 10 potentiellen Kunden kann man die Nullhypothese ablehnen, da die Wahrscheinlichkeit für diese Stichprobe noch immer größer ist als das Signifikanzniveau:

$$f(1, 10, 0.35) = \left[\frac{10!}{1! \cdot (10-1)!} \cdot 0.35^1 \cdot (1-0.35)^{10-1}\right] = 0.072$$

Erst eine Stichprobe mit 0 Käufern würde gegen die Nullhypothese sprechen. Die Wahrscheinlichkeit für so eine Stichprobe ist nämlich kleiner als das Signifikanzniveau von 5%:

$$f(0, 10, 0.35) = \left[\frac{10!}{0! \cdot (10-0)!} \cdot 0.35^0 \cdot (1-0.35)^{10-0}\right] = 0.013$$

Für eine rechtsseitige Hypothese wird allgemein der kritische Wert c_u, der den Ablehnungsbereich der Nullhypothese vom Nichtablehnungsbereich trennt, nach folgender Formel berechnet:

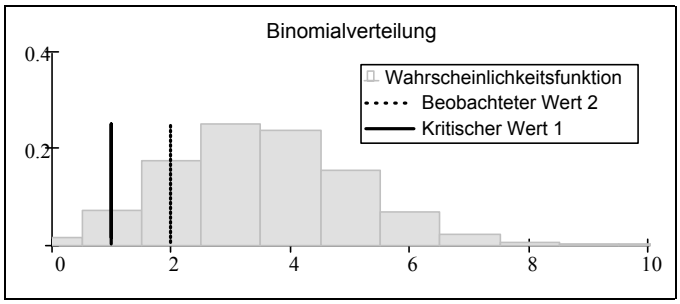
$$\sum_{k=0}^{c_u+1} \left[\frac{n!}{k!\,(n-k)!} \cdot \pi^k \cdot (1-\pi)^{n-k}\right] \leq \alpha$$

Für eine linksseitige Hypothese bestimmt man den kritischen Wert c_o nach folgender Formel

$$\sum_{k=0}^{c_o} \left[\frac{n!}{k!\,(n-k)!} \cdot \pi^k \cdot (1-\pi)^{n-k} \right] \leq 1 - \alpha$$

und für eine zweiseitige Alternativhypothese werden die beiden kritischen Werte so bestimmt, dass obige Ungleichungen für $\alpha/2$ gelten.

In folgender Grafik ist die Wahrscheinlichkeitsverteilung der Binomialverteilung für $\pi = 0.35$ und $n = 10$ dargestellt mit dem kritischen Wert $c_u = 1$, der den Ablehnungsbereich der Nullhypothese vom Nichtablehnungsbereich trennt und dem beobachteten Wert $x = 2$, der im Nichtablehnungsbereich der Nullhypothese liegt.



(Beobachteter_Wert)	(Kritischer_Wert)	(p – Wahrscheinlichkeit)
2	1	0.914

Da die Nullhypothese nicht abgelehnt werden kann, besteht die Gefahr eines β-Fehlers. Bei einem Signifikanztest hat man nur das Risiko eines α-Fehlers quantitativ beschränkt. Man weiß daher im konkreten Fall, dass der Käuferanteil in der Grundgesamtheit auch kleiner als 0.35 sein kann. Wie groß dieses Risiko jedoch quantitativ ist, kann man nicht allgemein angeben.

Für die Nullhypothese eines Käuferanteils von genau 40% findet man die Ergebnisse im Output der Software auf Seite 181.

1_3 Zentralwert, Schätzverfahren

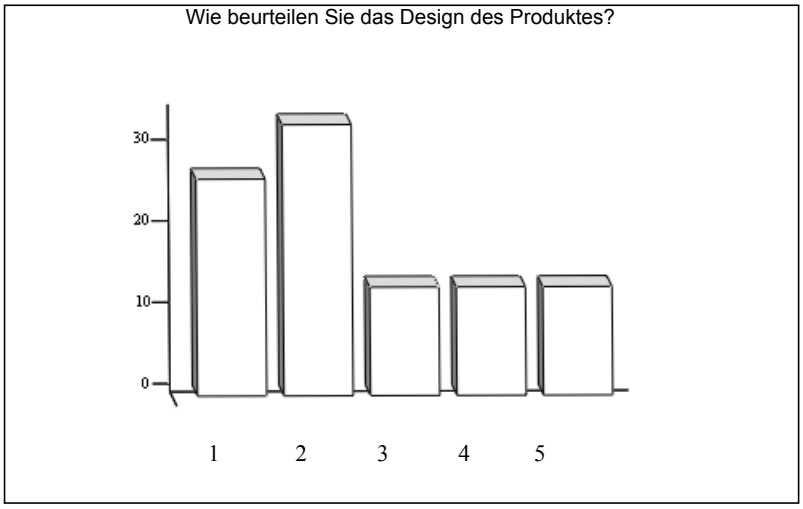
Wie wird das Design der neuen Spielkonsole von mindestens der Hälfte der potentiellen Kunden beurteilt? Bei einer Werbeveranstaltung wurde das Design von 15 Besuchern wie folgt benotet (1 = sehr gut, 2 = gut, ..., 5 = sehr schlecht):

[("Designbeurteilung") 4 1 3 3 5 2 2 2 4 1 1 5 2 1 2]

Der Stichprobenumfang ist in diesem Fall n = 15 und der Zentralwert ist "gut": Die Hälfte der Befragten beurteilt das Design der Spielkonsole mindestens mit "gut", die andere Hälfte mit höchstens "gut".

In folgender Tabelle sind die Beurteilungen nach ihren Häufigkeiten ausgezählt:

("Designbeurteilung")	Häufigkeit	Kum_Häufigkeit	in%	Kum_%
sehr_gut	4	4	27	27
gut	5	9	33	60
mittelmäßig	2	11	13	73
schlecht	2	13	13	87
sehr_schlecht	2	15	13	100
SUMME	15	"*"	100	"*"



("Nr.:")	1	2	3	4	5
"Antworten:"	sehr_gut	gut	mittelmäßig	schlecht	sehr_schlecht
"in %: "	27	33	13	13	13

Der Zentralwert (oder auch Median) einer Verteilung ist die Merkmalsausprägung jener Einheit, der 50% aller Verteilungseinheiten vorangehen und nachfolgen. Für seine Bestimmung werden die Merkmalsausprägungen zuerst der Größe nach geordnet:

$$\begin{bmatrix} (\text{Rangzahl}) & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ (\text{Beurteilung}) & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 & 2 & 3 & 3 & 4 & 4 & 5 & 5 \end{bmatrix}$$

Bei ungeradem n ist der Zentralwert jene Merkmalsausprägung, die in der Mitte dieser Reihe steht:

$$\text{median} = x_{\left(\frac{n+1}{2}\right)} \quad \text{falls n ungerade}$$

Bei geradem n erfüllt jede Merkmalsausprägung zwischen x(n/2) und x(n/2 + 1) die Bedingung des Zentralwertes. Wenn möglich wählt man den Durchschnitt aus diesen beiden Ausprägungen als Zentralwert:

$$\text{median} = \frac{1}{2} \cdot \left[x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)} \right] \quad \text{falls n gerade}$$

Da der Stichprobenumfang n = 15 ungerade ist, ist die mittlere Beurteilung die Merkmalsausprägung der (15 + 1) / 2 = 8.-ten Einheit der der Größe nach geordneten Designbeurteilungen.

Der Zentralwert dieser Verteilung ist gut (= 2), da dies die Merkmalsausprägung der 8.-ten Einheit ist. Die Hälfte der Befragten beurteilt das Design der Spielkonsole mindestens mit "gut", die andere Hälfte mit höchstens "gut". Kann man nun annehmen, dass auch die Hälfte der 100.000 potentiellen Kunden das Design der Spielkonsole mindestens mit "gut" beurteilt?

Nein. Unter der Voraussetzung, dass die Stichprobe eine Zufallsstichprobe aus der Grundgesamtheit der 100000 potentiellen Kunden ist, kann man ein Intervall für diesen unbekannten "wahren" Zentralwert berechnen.

Da die Ergebnisse einer Stichprobenerhebung zufallsabhängig sind, wird auch ein gefundener Zentralwert nur in den seltensten Fällen genau mit dem gesuchten Zentralwert der Grundgesamtheit übereinstimmen. Ausgehend von den Ergebnissen einer Stichprobe wird ein Konfidenzintervall berechnet, das den unbekannten Zentralwert der Grundgesamtheit mit vorgegebener Wahrscheinlichkeit enthält.

Wenn die Merkmalsausprägungen der Stichprobe nur ordinal skaliert sind, dann werden diese geordnet zu

$$x_{[1]} \ldots x_{[u]} \ldots x_{[i]} \ldots x_{[o]} \ldots x_{[n]}.$$

$x_{[u]}$ und $x_{[o]}$ sind die u-te und o-te Merkmalsausprägung der geordneten Stichprobe. u wird so bestimmt, dass gilt

$$\sum_{i=0}^{u} \left(\frac{n!}{i! \cdot (n-i)!} \cdot \frac{1}{2^n} \right) \leq \frac{\alpha}{2} < \sum_{i=0}^{u+1} \left(\frac{n!}{i! \cdot (n-i)!} \cdot \frac{1}{2^n} \right)$$

und o ist

$$o = n - (u + 1)$$

.

Für u = 3 ist diese Ungleichung erfüllt:

$$\left[\sum_{i=0}^{3}\left(\frac{15!}{i!\cdot(15-i)!}\cdot\frac{1}{2^{15}}\right) = 0.018\right] \le 0.025 < \left[0.059 = \sum_{i=0}^{3+1}\left(\frac{15!}{i!\cdot(15-i)!}\cdot\frac{1}{2^{15}}\right)\right]$$

Die Untergrenze des Konfidenzintervalls ist daher die Merkmalsausprägung der 3.-ten Einheit der geordneten Stichprobe. Dies ist die Beurteilung "sehr gut":

$$\begin{bmatrix} (\text{Rangzahl}) & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ (\text{Beurteilung}) & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 & 2 & 3 & 3 & 4 & 4 & 5 & 5 \end{bmatrix}$$

Die Obergrenze des Konfidenzintervalls ist die Merkmalsausprägung o = 15 – (3 + 1) = 11 der geordneten Stichprobe. Dies ist die Beurteilung "genügend". Mit einem Vertrauen von 95% kann man erwarten, dass der Zentralwert der Grundgesamtheit zwischen der Beurteilung "sehr gut" und "befriedigend“ liegt.

Bei metrischer Skalierung der Antworten einer Frage kann man ein Konfidenzintervall sowohl mit den oben angeführten Formeln berechnen, die aus der Binomialverteilung abgeleitet sind, als auch mit der Normalverteilung. Beide Methoden werden am folgenden Beispiel demonstriert.

Bei dieser Werbeveranstaltung wurde den 15 Befragten auch folgende Frage gestellt: Wie viele Stunden verbringen sie üblicherweise pro Woche vor dem Computer? Folgende Zeitangaben wurden abgegeben:

$$[\,(\text{Computerzeit})\ \ 4.4\ \ 12\ \ 9\ \ 9\ \ 5.5\ \ 8.5\ \ 20\ \ 6\ \ 4.4\ \ 6\ \ 6\ \ 5\ \ 5\ \ 10\ \ 6\,]$$

Da n = 15 ungerade ist, ist die mittlere Stundenangabe die Merkmalsausprägung der (15+1)/2 = 8.-ten Einheit der der Größe nach geordneten Stunden.

$$\begin{bmatrix} (\text{Rangzahl}) & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ (\text{Computerzeit}) & 4.4 & 4.4 & 5 & 5 & 5.5 & 6 & 6 & 6 & 6 & 8.5 & 9 & 9 & 10 & 12 & 20 \end{bmatrix}$$

Der Zentralwert dieser Verteilung ist 6 Stunden, da dies die Merkmalsausprägung der 8.-ten Einheit ist. Die Hälfte der Befragten verbringt höchstens diese Stundenanzahl wöchentlich vor dem Computer, die andere mindestens soviel.

Für ein Konfidenzniveau von 95% muss u wieder so bestimmt werden, dass gilt

$$\left[\sum_{i=0}^{3}\left(\frac{15!}{i!\cdot(15-i)!}\cdot\frac{1}{2^{15}}\right)=0.018\right]\le 0.025<\left[0.059=\sum_{i=0}^{3+1}\left(\frac{15!}{i!\cdot(15-i)!}\cdot\frac{1}{2^{15}}\right)\right]$$

Für u = 3 ist bekanntlich diese Ungleichung erfüllt. Die Untergrenze des Konfidenzintervalls ist daher die Merkmalsausprägung der 3.-ten Einheit der geordneten Stichprobe. Dies ist die Stundenanzahl 5. Die Obergrenze des Konfidenzintervalls ist die Merkmalsausprägung der o = 15 – 3 + 1 = 13. Einheit der geordneten Stichprobe. Dies ist die Merkmalsausprägung 10 Stunden. Mit einem Vertrauen von 95% kann man erwarten, dass der Zentralwert der Grundgesamtheit zwischen einer wöchentlichen Stundenzahl von 5 und 10 Stunden liegt.

Wenn die Antworten der Frage wie hier metrisch skaliert und genügend groß sind (n > 36), kann man die beiden Konfidenzgrenzen auch mit Hilfe der Normalverteilung berechnen. Die Formeln für die beiden Intervallgrenzen für ein zweiseitiges Konfidenzintervall zum Niveau von 1– α sind

$$K_u = \text{median} - 1.253\cdot z_{1-\frac{\alpha}{2}}\cdot\frac{s}{\sqrt{n}}$$

$$K_o = \text{median} + 1.253\cdot z_{1-\frac{\alpha}{2}}\cdot\frac{s}{\sqrt{n}}$$

Der "median" ist hier der Stichprobenzentralwert. s ist der aus der Stichprobe errechnete Schätzwert für die Standardabweichung der Stichprobe

$$s=\sqrt{\frac{1}{n-1}\cdot\sum_{i=1}^{n}\left(x_i-\text{xquer}\right)^2}$$

n ist der Stichprobenumfang und z ist das 1 – (α/2).-te Perzentil der Normalverteilung. Für das obiges Beispiel der wöchentlichen Stundenzahl vor dem Computer ist der Durchschnitt xquer gleich

$$\text{xquer}=\frac{1}{n}\cdot\sum_{i=1}^{n}x_i=(4.4+4.4+5+\ldots+20)/15=7.787$$

und s gleich

$$s=\sqrt{\frac{1}{15-1}\cdot\left[(4.4-7.787)^2+(4.4-7.787)^2+(5-7.787)^2\ldots(20-7.787)^2\right]}=4.068$$

z ist für ein 95% Konfidenzintervall

$$z_{1-\frac{0.05}{2}}=1.96$$

Dieser Wert kann aus einer Tabelle der Quantile der Standardnormalverteilung abgelesen werden oder mit Hilfe von Programmen berechnet werden. Die Konfidenzuntergrenze ist daher

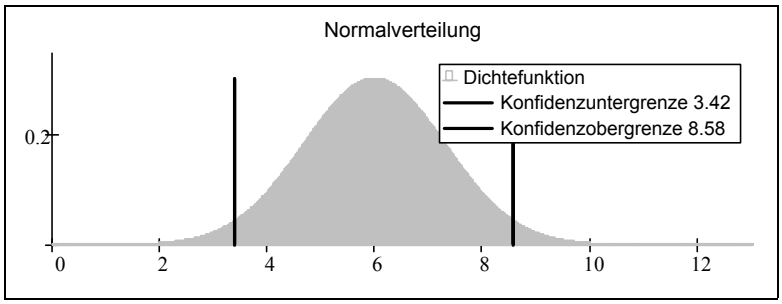
$$K_u = 6 - 1.253 \cdot 1.96 \cdot \frac{4.068}{\sqrt{15}} = 3.42$$

und die Obergrenze gleich

$$K_o = 6 + 1.253 \cdot 1.96 \cdot \frac{4.068}{\sqrt{15}} = 8.58$$

Mit einem Vertrauen von 95% kann man erwarten, dass der Zentralwert der Grundgesamtheit zwischen einer wöchentlichen Stundenzahl vor dem Computer von 3.42 Stunden und 8.58 Stunden liegt. Da n nicht größer als 36 ist, ist die Diskrepanz der hier berechneten Intervallgrenzen von den oben angeführten nicht verwunderlich. Erst wenn die Bedingung n > 36 erfüllt ist, liefern die hier angeführten Formeln brauchbare Näherungen.

Die Stichprobenverteilung des Zentralwertes samt Konfidenzgrenzen zeigt folgende Grafik:



(Konfidenzuntergrenze)	(Konfidenzobergrenze)	"Wahrscheinlichkeit p"
3.42	8.58	0.05

Der Output der Software auf Seite 182 zeigt die Ergebnisse für das Beispiel der Designbeurteilung. Den Zentralwert als auch sein Konfidenzintervall liefert das Programm, wenn man die Rohdaten oder die Häufigkeitsverteilung eingibt.

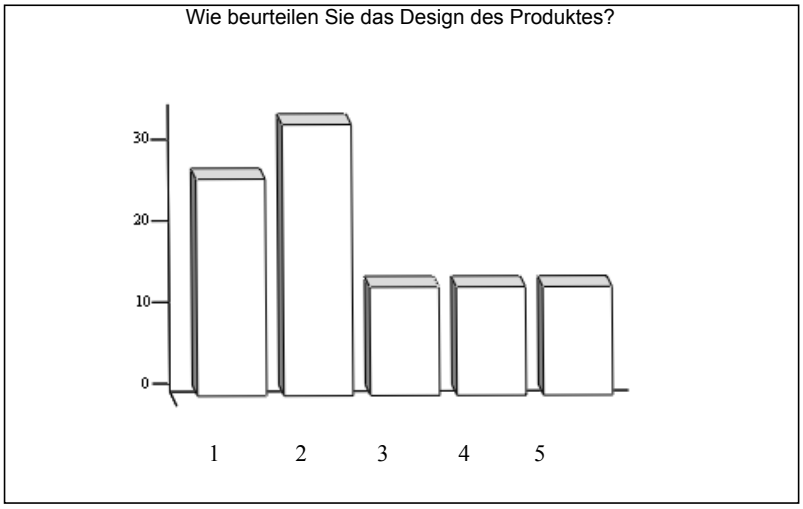
1_4 Zentralwert, Testverfahren

Kann man annehmen, dass das Design der neuen Spielkonsole Yoki von mindestens der Hälfte der potentiellen Kunden mit "mittelmäßig" beurteilt wird? Bei einer Werbeveranstaltung wurde das Design von 15 Besuchern wie folgt benotet (1 = sehr gut, 2 = gut, ..., 5 = sehr schlecht):

[(Designbeurteilung) 4 1 3 3 5 2 2 2 4 1 1 5 2 1 2]

In folgender Tabelle sind die Beurteilungen nach ihren Häufigkeiten ausgezählt:

(Designbeurteilung)	Häufigkeit	Kum_Häufigkeit	in%	Kum_%
sehr_gut	4	4	27	27
gut	5	9	33	60
mittelmäßig	2	11	13	73
schlecht	2	13	13	87
sehr_schlecht	2	15	13	100
SUMME	15	"*"	100	"*"



"Nr.:"	1	2	3	4	5
"Antworten:"	sehr_gut	gut	mittelmäßig	schlecht	sehr_schlecht
"in %: "	27	33	13	13	13

Der Stichprobenumfang ist in diesem Fall n = 15. Da der Stichprobenumfang n = 15 ungerade ist, ist die mittlere Beurteilung die Merkmalsausprägung der (15 + 1) / 2 = 8.-ten Einheit der der Größe nach geordneten Designbeurteilungen. Der Zentralwert der Stichprobe ist "gut":

$$\text{median} = x_{\left(\frac{15+1}{2}\right)} = x_{(8)} = \text{gut}$$

Eine Hälfte der Befragten beurteilt das Design der Spielkonsole mindestens mit "gut", die andere Hälfte mit höchstens "gut". Kann man nun annehmen, dass auch die Hälfte der 100000 potentiellen Kunden das Design der Spielkonsole mindestens mit "mittelmäßig" beurteilt?

Unter der Voraussetzung, dass die Stichprobe eine Zufallsstichprobe aus der Grundgesamtheit der 100000 potentiellen Kunden ist und der "wahre" Zentralwert in dieser Grundgesamtheit die Designbeurteilung "mittelmäßig" ist, kann man auch erwarten, dass in einer Zufallsstichprobe von 15 potentiellen Kunden der Zentralwert von "gut" auftritt. Das Stichprobenergebnis spricht nicht gegen die Annahme, dass in der Grundgesamtheit der Zentralwert der Designbeurteilung "mittelmäßig" ist.

Da man testen will, ob das Design der neuen Spielkonsole Yoki von mindestens der Hälfte der potentiellen Kunden mit "mittelmäßig" beurteilt wird, ist die Nullhypothese

H_0: Median = mittelmäßig

und die Alternativhypothese

H_1: Median $\neq$ mittelmäßig

.

Um die Hypothesen zu testen, berechnet man die Wahrscheinlichkeiten für das Auftreten der möglichen Zentralwerte in einer Stichprobe vom Umfang 15, wenn in der Grundgesamtheit der Zentralwertwert "mittelmäßig" ist. Unter der Voraussetzung, dass H_0 wahr ist, ist das Auftreten eines Stichprobenzentralwertes, der kleiner bzw. größer ist als der der Grundgesamtheit binomialverteilt mit $\pi = 0.5$. Stichprobenzentralwerte, die eine geringere Auftretenswahrscheinlichkeit haben als ein vorgegebenes Signifikanzniveau α führen zur Ablehnung der Nullhypothese. Die allgemeine Formel für die Berechnung dieser Grenze c_u bzw. c_o zwischen Ablehnungs- und Nichtablehnungsbereich der Nullhypothese ist

$$\sum_{i=0}^{c_u} \left(\frac{n!}{i! \cdot (n-i)!} \cdot \frac{1}{2^n} \right) \leq \alpha < \sum_{i=0}^{c_u+1} \left(\frac{n!}{i! \cdot (n-i)!} \cdot \frac{1}{2^n} \right)$$

für eine einseitige Nullhypothese und

$$\sum_{i=0}^{c_u} \left(\frac{n!}{i! \cdot (n-i)!} \cdot \frac{1}{2^n} \right) \leq \frac{\alpha}{2} < \sum_{i=0}^{c_u+1} \left(\frac{n!}{i! \cdot (n-i)!} \cdot \frac{1}{2^n} \right)$$

für eine zweiseitige Nullhypothese. Die Obergrenze c_o wird, wenn nötig, mit Hilfe der Formel

$$c_o = n - (c_u + 1)$$

berechnet. Wenn bei diesem Test Antworten auftreten, die mit dem postulierten Zentralwert übereinstimmen, dann werden diese nicht berücksichtigt, da sie nichts über die Richtung der Abweichung aussagen. In obigem Beispiel stimmen 2 Antworten mit dem postulierten Zentralwert der Grundgesamtheit "mittelmäßig" überein. Daher bestimmt man die kritischen Wert c_u und c_o nicht für eine Stichprobe mit dem Umfang 15 sondern 13. Für ein Signifikanzniveau von 5% wird c_u so bestimmt, dass gilt

$$\left[\sum_{i=0}^{2} \left(\frac{13!}{i! \cdot (13-i)!} \cdot \frac{1}{2^{13}} \right) = 0.011 \right] \leq 0.025 < \left[0.046 = \sum_{i=0}^{2+1} \left(\frac{13!}{i! \cdot (13-i)!} \cdot \frac{1}{2^{13}} \right) \right]$$

c_u ist also 2, die Merkmalsausprägung der 2.-ten Einheit der geordneten Stichprobe. Dies ist die Beurteilung "sehr gut": $c_o = 13 - (2 + 1) = 10$, die Merkmalsausprägung der 10.-ten Einheit der geordneten Stichprobe. Dies ist die Designbeurteilung "schlecht".

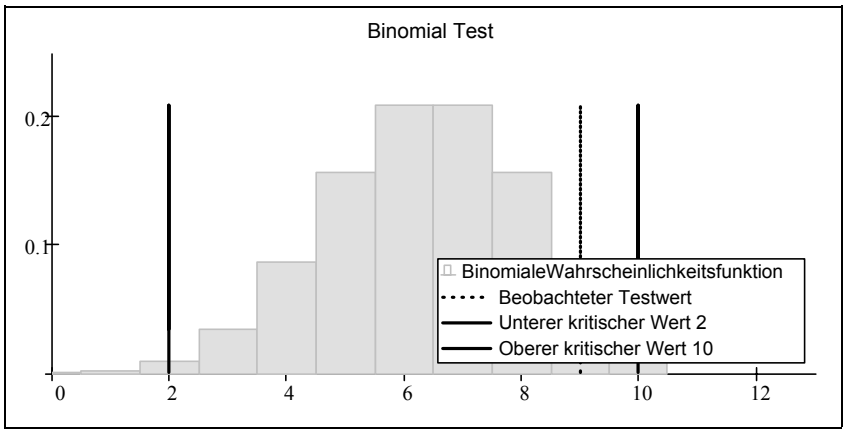
(Rangzahl)	1	2	3	4	5	6	7	8	9	10	11	12	13
(Beurteilung)	1	1	1	1	2	2	2	2	2	4	4	5	5

Die Testmaßzahl A ist die Anzahl der Designbeurteilungen, die kleiner sind als der postulierte Zentralwert der Grundgesamtheit. 9 der insgesamt 13 Beurteilungen sind besser als "mittelmäßig". Daher ist A = 9. Da A zwischen der Untergrenze c_u und der Obergrenze c_o liegt

$$c_u = 2 < A = 9 < 10 = c_o$$

kann die Nullhypothese nicht abgelehnt werden.

Da die Nullhypothese nicht abgelehnt werden kann, besteht die Gefahr eines β-Fehlers. Bei einem Signifikanztest hat man nur das Risiko eines α-Fehlers quantitativ beschränkt. Man weiß daher im konkreten Fall, dass der Zentralwert in der Grundgesamtheit auch ungleich "mittelmäßig" sein kann. Wie groß dieses Risiko jedoch quantitativ ist, kann man nicht allgemein angeben. In folgender Grafik ist die entsprechende Wahrscheinlichkeitsverteilung sowie der Ablehnungs- und Nichtablehnungsbereich dargestellt:



Beobachteter_Wert	Kritischer_Werte	p_Wahrscheinlichkeit
9	(2 10)	0.05

Den entsprechenden Softwareoutput für dieses Beispiel findet man auf Seite 183.

1_5 Durchschnitt, Schätzverfahren

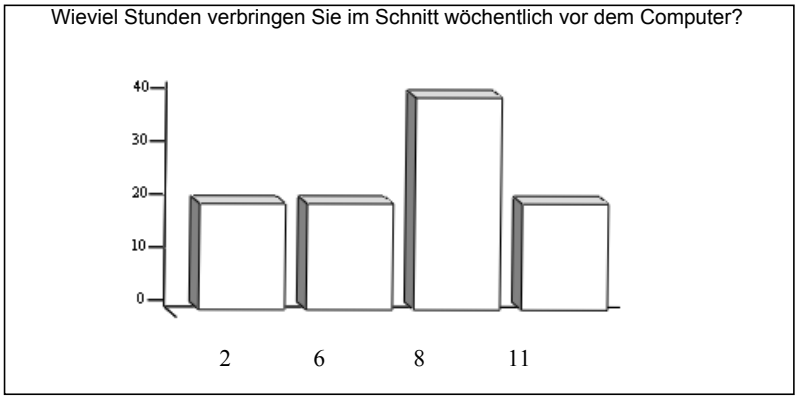
Wie viele Stunden verbringen die potentiellen Kunden der Spielkonsole Yoki im Schnitt pro Woche vor den Computer? Auf diese Frage antworteten 5 potentielle Kunden einer Werbeveranstaltung mit folgenden Angaben:

[(Computerzeit) 4.4 12 9 9 5.6]

Der Stichprobenumfang ist in diesem Fall n = 5 und der Durchschnitt ist 8 Stunden. Im Mittel sitzen die 5 Befragten wöchentlich 8 Stunden vor dem Computer.

Fasst man die einzelnen Angaben in folgenden Intervallen zusammen und zählt sie nach diesen Intervallen aus, dann erhält man folgende Tabelle:

"Unter-"	"Ober-"	"Mittel-"	"Häufig-"	"Kumulierte"	"Prozente"	"Kumulierte"
"grenze"	"grenze"	"punkt"	"keiten"	"Häufigkeit"	"%"	"%"
0	5	2	1	1	20	20
5	7	6	1	2	20	40
7	10	8	2	4	40	80
10	13	11	1	5	20	100
(Computerzeit)	"*"	"SUMME"	5	"*"	100	"*"



Der Durchschnitt (oder auch das arithmetisches Mittel) einer Verteilung ist die reale oder fiktive Merkmalsausprägung, die sich aus der Summe aller Merkmalsausprägungen dividiert durch die Anzahl der Einheiten ergibt. Für seine Bestimmung werden folgende Formeln verwendet:

$$xquer = \frac{1}{n} \cdot \left(x_1 + x_2 + ... + x_n \right) = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i$$

oder

$$xquer = \frac{1}{n} \cdot \sum_{j=1}^{m} \left(h_j \cdot x_j \right)$$

mit h_j = absolute Häufigkeit von x_j.

Wenn die Ausprägungen nur in Form einer Urliste vorliegen, dann wird die erste Formel genommen. Die zweite verwendet man bei Häufigkeitsverteilungen. Man nennt xquer meistens den Mittelwert und lässt den Zusatz arithmetisch weg. Wird das Mittel nicht aus einer Stichprobe sondern aus einer Grundgesamtheit berechnet, dann kürzt man es mit dem griechischen Buchstaben μ (sprich mü) ab.

Die Summe der Merkmalsausprägungen ist im Beispiel

$$\sum_{i=1}^{5} x_i = (4.4 + 12 + 9 + 9 + 5.6) = 40$$

und n = 5. Daher ist der Durchschnitt der Quotient aus beiden Zahlen nämlich gleich

$$xquer = \frac{40}{5} = 8$$

Fasst man die gleichen Merkmalsausprägungen zusammen, dann erhält man folgende Häufigkeitsverteilung:

(Ausprägung)	(Häufigkeit)
4.4	1
5.6	1
9.0	2
12.0	1
SUMME	5

Wenn man die Ausprägungen mit ihren Häufigkeiten multipliziert und dann diese Produkte summiert, dann erhält man wieder die Summe der Merkmalsausprägungen (letztes Element in der 3. Spalte):

(Ausprägung)	(Häufigkeit)	(Produkt)
4.4	1	4.4
5.6	1	5.6
9.0	2	18.0
12.0	1	12.0
SUMME	5	40.0

Die Division durch die Summe der Häufigkeiten liefert wieder den Mittelwert xquer = 8. Kann man nun annehmen, dass auch alle potentiellen Kunden der Spielkonsole Yoki im Mittel 8 Stunden pro Woche vor dem Computer sitzen? Sicher nicht. Man kann aber ein Konfidenzintervall für den unbekannten Durchschnitt der Grundgesamtheit von 100.000 potentiellen Kunden berechnen. Dazu benötigt man neben den Stichprobendurchschnitt noch die Stichprobenvarianz bzw. einen Schätzwert für die unbekannte Varianz der Grundgesamtheit.

Den Durchschnitt aus den Abweichungsquadraten der Merkmalsausprägungen einer Verteilung von ihrem arithmetischen Mittel bezeichnet man als Varianz. Sie wird mit s^2 abgekürzt, wenn sie aus einer Stichprobe errechnet wird und mit dem griechischen Buchstaben σ^2 (sprich sigma^2), wenn sie die Streuung einer Grundgesamtheit kennzeichnet.

$$s^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(x_i - xquer\right)^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i - n \cdot xquer^2$$

In folgender Tabelle sind die Berechnungsschritte für die Varianz der wöchentlichen Zeit vor dem Computer von 5 Befragten angeführt. In der ersten Spalte stehen die Merkmalsausprägungen x_j und ihre Häufigkeiten h_j in der zweiten Spalte. In der dritten Spalte stehen die Abweichungen der Ausprägungen vom Durchschnitt xquer = 8. Die Quadrate dieser Abweichungen vom Durchschnitt findet man in der vierten Spalte. Diese Quadrate werden mit den Häufigkeiten der zweiten Spalte gewichtet und summiert. Das Ergebnis ist die Abweichungsquadratsumme von 36.72.

x_j	h_j	$x_j - xquer$	$\left(x_j - xquer\right)^2$	$\left(x_j - xquer\right)^2 \cdot h_j$
4.4	1	–3.6	12.96	12.96
5.6	1	–2.4	5.76	5.76
9.0	2	1.0	1.00	2.00
12.0	1	4.0	16.00	16.00
SUMME	5	"*"	"*"	36.72

Wenn man die Abweichungsquadratsumme durch die Anzahl der Einheiten dividiert, dann erhält man die Varianz dieser Verteilung:

$$s^2 = \frac{36.72}{5} = 7.344$$

Die Wurzel aus der Varianz ist die Standardabweichung. Zum Unterschied zur Varianz ist die Standardabweichung eine benannte Zahl.

$$s = \sqrt{s^2}$$

Für obiges Beispiel ist die Standardabweichung gleich s = 2.71 Stunden. Im Mittel weichen die 5 wöchentlichen Computerzeiten um 2.71 Stunden von ihrem Durchschnitt 8 Stunden ab.

Wenn die Varianz der Grundgesamtheit σ^2 nicht bekannt ist, dann wird ein zweiseitiges Konfidenzintervall für den unbekannten Mittelwert μ zum Niveau $1-\alpha$ nach folgenden Formeln berechnet:

$$K_u = xquer - t_{1-\frac{\alpha}{2},n-1} \cdot \frac{s\_dach}{\sqrt{n}}$$

$$K_o = xquer + t_{1-\frac{\alpha}{2},n-1} \cdot \frac{s\_dach}{\sqrt{n}}$$

Für die unbekannte Standardabweichung σ verwendet man als besten Schätzwert nicht die Standardabweichung s wie oben definiert, sondern s_dach. s_dach wird nach folgender Formel berechnet:

$$s\_dach = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} \left(x_i - xquer\right)^2}$$

Die Abweichungsquadratsumme wird nicht durch n sondern durch (n – 1) dividiert.

$$s\_dach = \sqrt{\frac{1}{5-1} \cdot \sum_{i=1}^{5} \left(x_i - 8\right)^2} = \sqrt{\frac{36.72}{4}} = 3.03$$

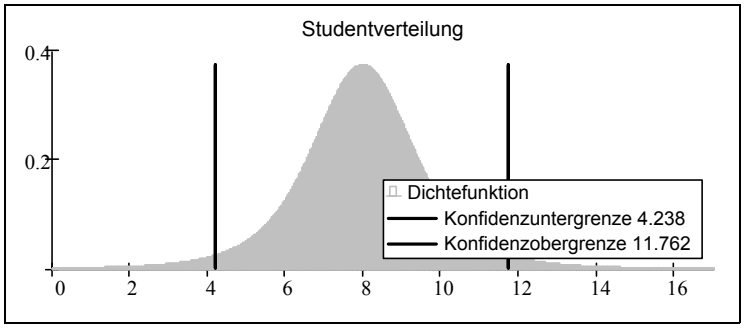
t ist das $1-(\alpha/2)$.-te Perzentil der standardisierten Studentverteilung für $\nu = n - 1$ Freiheitsgrade. Dieses Perzentil bestimmt man aus entsprechenden Tabellen der Studentverteilung oder berechnet es mit geeigneten Programmen. Für $\nu = n - 1 = 5 - 1 = 4$ und ein Konfidenzniveau von 95% ist t gleich

$$t_{1-\frac{0.05}{2}} = 2.776$$

Unter- und Obergrenze des zweiseitigen Konfidenzintervalls für den Durchschnitt sind

$$K_u = 8 - 2.776 \cdot \frac{3.03}{\sqrt{5}} = 4.238$$

$$K_o = 8 + 2.776 \cdot \frac{3.03}{\sqrt{5}} = 11.762$$



(Konfidenzuntergrenze)	(Konfidenzobergrenze)	"Wahrscheinlichkeit p"
4.238	11.762	0.05

Mit einem Vertrauen von 95% kann man den unbekannten Durchschnitt der wöchentlichen Zeit vor dem Computer im Intervall von 4.24 bis 11.76 Stunden erwarten. Dieses Ergebnis liefert auch der Output der Software der auf Seite 184 dargestellt ist.

Die Berechnung der Konfidenzintervalle mit Hilfe der oben angeführten Formeln setzt voraus, dass der Stichprobenumfang größer gleich 30 ist. In diesem Fall kann man auf Grund des zentralen Grenzwertsatzes annehmen, dass die Verteilung der möglichen Stichprobenmittel einer Normalverteilung folgen.

Wenn der Stichprobenumfang kleiner als 30 ist, dann muss man voraussetzen, dass die Stichprobe aus einer normalverteilten Grundgesamtheit stammt. Diese Voraussetzung kann man mit Hilfe des Kolmogorov-Smirnov-Tests überprüfen.

1_6 Durchschnitt, Testverfahren

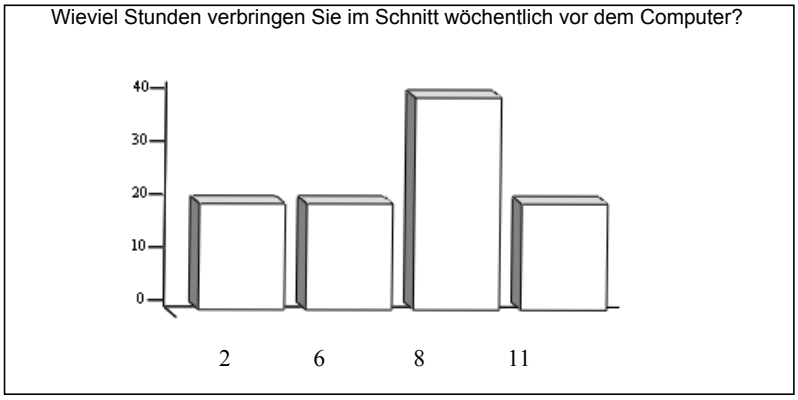
Kann man annehmen, dass die potentiellen Kunden der Spielkonsole Yoki im Schnitt pro Woche höchstens 3 Stunden vor den Computer verbringen? Auf die Frage nach der wöchentlichen Zeit vor dem Computer antworteten 5 potentielle Kunden einer Werbeveranstaltung mit folgenden Angaben:

[(Computerzeit) 4.4 12 9 9 5.6]

Der Stichprobenumfang ist in diesem Fall n = 5 und der Durchschnitt ist 8 Stunden. Im Mittel sitzen die 5 Befragten wöchentlich 8 Stunden vor dem Computer.

Fasst man die einzelnen Angaben in folgenden Intervallen zusammen und zählt sie nach diesen Intervallen aus, dann erhält man folgende Tabelle und Grafik:

"Unter-"	"Ober-"	"Mittel-"	"Häufig-"	"Kumulierte"	"Prozente"	"Kumulierte"
"grenze"	"grenze"	"punkt"	"keiten"	"Häufigkeit"	"%"	"%"
0	5	2	1	1	20	20
5	7	6	1	2	20	40
7	10	8	2	4	40	80
10	13	11	1	5	20	100
(Computerzeit)	"*"	"SUMME"	5	"*"	100	"*"



Widerlegt dieses Stichprobenergebnis die Annahme, dass alle potentiellen Kunden der Spielkonsole Yoki im Mittel höchstens 3 Stunden pro Woche vor dem Computer sitzen?

Unter der Voraussetzung, dass der "wahre" Durchschnitt in dieser Grundgesamtheit 3 Stunden pro Woche ist, kann man auch erwarten, dass in einer Zufallsstichprobe von 5 potentiellen Kunden ein Durchschnitt von 8 Stunden auftritt. Das Stichprobenergebnis spricht nicht gegen die Annahme, dass in der Grundgesamtheit aller potentiellen Kunden diese im Mittel 3 Stunden pro Woche vor dem Computer sitzen.
Die Nullhypothese ist

$$H_0: \mu \leq 3$$

Alle potentiellen Kunden der Spielkonsole Yoki verbringen im Mittel höchstens 3 Stunden pro Woche vor dem Computer. Die Alternative dazu ist

$$H_1: \mu > 3$$

Alle potentiellen Kunden der Spielkonsole Yoki verbringen im Mittel mehr als 3 Stunden pro Woche vor dem Computer.

Unter der Voraussetzung einer normalverteilten Grundgesamtheit verwendet man folgende Testmaßzahl zur Entscheidung, wenn n < 30 ist:

$$t_{beob} = \frac{xquer - \mu_0}{\frac{s\_dach}{\sqrt{n}}}$$

Diese Maßzahl ist studentverteilt mit

$$\nu = n - 1$$

Freiheitsgraden. Ist $n \geq 30$, dann muss nicht normalverteilte Grundgesamtheit vorausgesetzt werden. In diesem Fall verwendet man die Testmaßzahl

$$z_{beob} = \frac{xquer - \mu_0}{\frac{s\_dach}{\sqrt{n}}}$$

die standardnormalverteilt ist. Da im obigen Beispiel n = 5 ist, verwendet man die studentverteilte Testmaßzahl t_{beob}. xquer ist

$$xquer = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i = \frac{1}{5} \cdot (4.4 + 12 + 9 + 9 + 5.6) = 8$$

s_dach ist

$$s\_dach = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} \left(x_i - xquer\right)^2} = \sqrt{\frac{36.72}{4}} = 3.03$$

und μ_0 ist 3. t_{beob} ist daher

$$t_{beob} = \frac{8 - 3}{\frac{3.03}{\sqrt{5}}} = 3.69$$

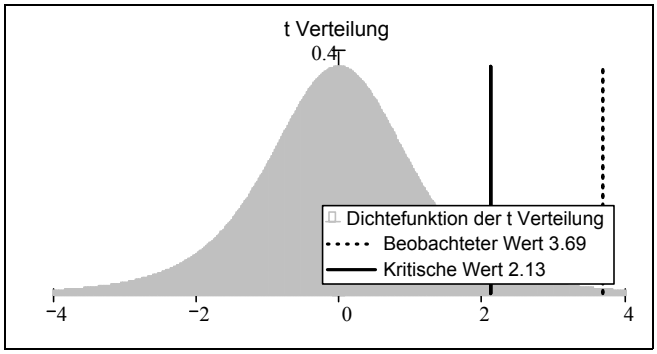
Wenn man ein Signifikanzniveau von 5% voraussetzt, dann ist der kritische Wert t_c der Studentverteilung für

$$\nu = 5 - 1 = 4$$

Freiheitsgrade gleich

$$F_t(0.95, 4) = 2.132$$

Da dieser Wert kleiner ist als der beobachtete, kann die Nullhypothese abgelehnt werden. Das Stichprobenergebnis spricht gegen die Annahme, dass die durchschnittliche Zeit pro Woche vor dem Computer der potentiellen Yokikäufer höchstens 3 Stunden beträgt. Da die Nullhypothese abgelehnt werden kann, ist mit dieser Entscheidung ein α Risiko verknüpft. Wie hoch dieses Risiko zahlenmäßig höchstens ist, weiß man bei einem Signifikanztest durch die Vorgabe des Signifikanzniveaus. Folgende Grafik zeigt den kritischen und beobachteten Wert mit der entsprechenden Studentverteilung.



Beobachteter_Wert	Kritische_Werte	Wahrscheinlichkeit_p
3.69	$(-\infty \quad 2.13)$	0.01

Auf Seite 185 steht der Computeroutput für dieses Beispiel. Er erfordert lediglich die Eingabe der Rohdaten, die Nullhypothese und die Festlegung auf eine linksseitige (Alternativ-) Hypothese.

2_0 ZWEI UNABHÄNGIGE STICHPROBEN

Wenn man den Käufermarkt für die neue Spielkonsole untersucht, dann kann es im Hinblick auf Werbemaßnahmen sinnvoll sein, zwischen männlichen und weiblichen Kunden zu unterscheiden. Wird also bei einer Werbeveranstaltung potentiellen Kunden die Frage gestellt, ob sie die Spielkonsole kaufen werden, dann werden die möglichen Antworten "Ja" und "Nein" getrennt nach Männern und Frauen ausgezählt. Man hat also nicht mehr eine Stichprobe im Hinblick auf die Frage "Kaufabsicht", sondern zwei: die Stichprobe der Männer und die Stichprobe der Frauen, jeweils ausgezählt nach den Fragen "Ja" und "Nein".

Spezielle Werbemaßnahmen für Männer und Frauen sind nur sinnvoll, wenn sich beide in der Grundgesamtheit im Hinblick auf ihre Kaufabsicht unterscheiden. Wenn man z. B. 200 Männer und 100 Frauen bei dieser Werbeveranstaltung nach ihren Kaufabsichten befragte, dann werden sich die Antwortprozentsätze in beiden Stichproben vermutlich unterscheiden. Kann man auf Grund dieser Unterschiede schon allgemein für alle potentiellen männlichen und weiblichen Kunden behaupten, dass sie sich im Hinblick auf ihre Kaufabsicht unterscheiden? Diese Frage wird mit Hilfe geeigneter Testverfahren geklärt.

Vorschau

Ist die Kaufabsicht der männlichen potentiellen Kunden für die neue Spielkonsole größer als die der weiblichen Kunden? Diese Frage wird im Abschnitt 2_1 Fisher Test analysiert. Dieser Test ist auf zwei Stichproben beschränkt, bei der die Frage die den Unterschied misst, nur zwei Antworten kennt: Werden Sie die Spielkonsole kaufen? Antwort: Ja oder Nein. Wenn es mehr als zwei Antworten gibt (z. B. noch zusätzlich die Antwort "Weiß nicht"), dann muss man auf den χ2 Test zurückgreifen, der in Abschnitt 2_2 behandelt wird.

Fisher Test und χ2 Test setzen nur nominal skalierte Antworten voraus. Wenn die Antworten zumindest ordinal skaliert sind, dann eignet sich der U Test zur Analyse eventueller Unterschiede. Der U Test wird im Abschnitt 2_3 behandelt und zwar zur Klärung der Frage, ob sich die Beurteilung des Designs der Spielkonsole von Männern und Frauen unterscheidet.

Im Abschnitt 2_4 t Test werden Verfahren gezeigt, mit denen man Unterschiede in den Durchschnitten von 2 Stichproben prüfen kann. Sitzen Männer im Durchschnitt länger vor dem Computer als Frauen?

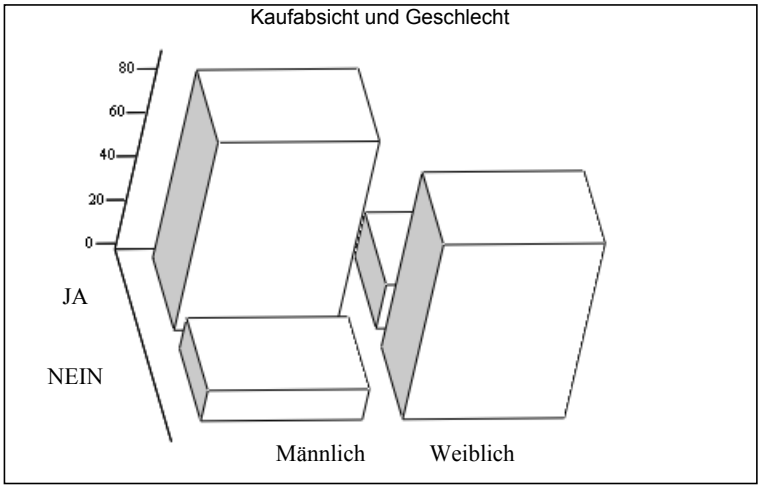
2_1 Exakter Fisher-Test

Ist die Kaufabsicht der männlichen potentiellen Kunden für die neue Spielkonsole größer als die der weiblichen Kunden? Eine Befragung bei einer Werbeveranstaltung von 5 Frauen und 7 Männern brachte folgendes Ergebnis:

(Geschlecht)	2	1	1	2	2	1	1	2	1	1	2	1
(Kaufabsicht)	2	1	1	2	2	1	1	1	2	1	2	1

In der ersten Zeile stehen die Befragten gegliedert nach dem Geschlecht: 1 = männlich und 2 = weiblich. Die Kaufabsicht steht in der zweiten Zeile: 1 = Ja und 2 = Nein. Der erste Befragte ist also weiblich (= 2) und hat keine Kaufabsicht (= 2), der zweite Befragte ist männlich (= 1) und hat Kaufabsicht (= 1), usw. In folgender Kreuztabelle sind die Befragungsergebnisse nach Geschlecht und Kaufabsicht zusammengefasst und ausgewiesen:

Geschlecht / Kaufabsicht1	Männlich	Weiblich
Ja	6	1
Nein	1	4
SUMME	7	5



6 von 7 Männern haben Kaufabsicht für die neue Spielkonsole, aber nur 1 Frau von 5. Kann man auf Grund dieser Stichprobenergebnisse auch für die Grundgesamtheit annehmen, dass sich die Kaufabsicht aller potentiellen weiblichen Kunden von allen potentiellen männlichen Kunden unterscheidet?

Ja man kann auf Grund dieser Stichprobenergebnisse annehmen, dass sich das Kaufinteresse der potentiellen weiblichen und männlichen Kunden in der Grundgesamtheit von z. B. 100000 potentiellen Kunden unterscheidet. Genauer gesagt: Das Kaufinteresse der potentiellen weiblichen Kunden ist geringer als das der männlichen. Das Risiko, dass diese Annahme falsch ist, ist höchstens 5%.

Wie kommt man zu dieser Schlussfolgerung? Zuerst werden die Hypothesen formuliert. Man nimmt als Nullhypothese an, dass der Anteil der Frauen, die Kaufabsicht für das neue Produkt haben, gleich groß ist wie der Anteil der Männer. Die Differenz der beiden Anteile sei also Null. Diese Hypothese wird formal wie folgt angeschrieben:

$$H_0: \pi_{männlich} = \pi_{weiblich} \text{ oder } H_0: \pi_{männlich} - \pi_{weiblich} = 0$$

Mit π wird der unbekannte Anteil in der Grundgesamtheit aller Frauen bzw. Männer bezeichnet, die Kaufabsicht haben. Eine Alternative zu dieser Nullhypothese ist

$$H_1: \pi_{männlich} \neq \pi_{weiblich} \text{ oder } H_1: \pi_{männlich} - \pi_{weiblich} \neq 0$$

Der Anteil der Frauen, die das neue Produkt kaufen wollen, unterscheidet sich vom Anteil der Männer. An Hand der Ergebnisse einer Stichprobe entscheidet man sich für die Annahme oder Ablehnung der Null- oder Alternativhypothese.

Um die Hypothesen zu prüfen, verwendet man die Häufigkeit $h_{11} = x$ als Testmaßzahl, wobei h_{11} aus folgender 2 X 2 Kreuztabelle stammt:

(Antwort_B / Antwort_A)	B	Nicht_B	ZEILENSUMME
A	h_{11}	h_{12}	$\sum_{j=1}^{2} h_{1j} = n_{1.}$
Nicht_A	h_{21}	h_{22}	$\sum_{j=1}^{2} h_{2j} = n_{2.}$
SPALTENSUMME	$\sum_{j=1}^{2} h_{i1} = n_{.1}$	$\sum_{j=1}^{2} h_{i2} = n_{.2}$	$\sum_{i=1}^{2}\sum_{j=1}^{2} h_{ij} = n$

$h_{11} = x$ ist die Testmaßzahl, d. h. die gemeinsame absolute Häufigkeit der Merkmalsausprägungen A und B. h_{11} ist im vorliegenden Beispiel 6. 6 Männer beabsichtigen, die neue Spielkonsole zu kaufen.

(Geschlecht / Kaufabsicht)	Männlich	Weiblich	ZEILENSUMME
Ja	6	1	7
Nein	1	4	5
SPALTENSUMME	7	5	12

Ist x kleiner als c_u oder größer als c_o, dann wird die Nullhypothese abgelehnt. c_u und c_o werden mit Hilfe der Hypergeometrischen Verteilung so bestimmt, dass gilt

$$\sum_{k=0}^{c_u-1} \frac{n_{1.}!\cdot n_{.1}!\cdot n_{2.}!\cdot n_{.2}!}{n!\cdot k!\cdot(n_{.1}-k)!\cdot(n_{1.}-k)!\cdot(n-n_{.1}-n_{2.}+k)!} \le \frac{\alpha}{2}$$

$$\frac{\alpha}{2} < \sum_{k=0}^{c_u} \frac{n_{1.}!\cdot n_{.1}!\cdot n_{2.}!\cdot n_{.2}!}{n!\cdot k!\cdot(n_{.1}-k)!\cdot(n_{1.}-k)!\cdot(n-n_{.1}-n_{2.}+k)!}$$

und für c_o

$$\sum_{k=c_0-1}^{\min(n_{.1},n_{1.})} \frac{n_{1.}!\cdot n_{.1}!\cdot n_{2.}!\cdot n_{.2}!}{n!\cdot k!\cdot(n_{.1}-k)!\cdot(n_{1.}-k)!\cdot(n-n_{.1}-n_{2.}+k)!} > \frac{\alpha}{2}$$

$$\frac{\alpha}{2} \ge \sum_{k=c_0}^{\min(n_{.1},n_{1.})} \frac{n_{1.}!\cdot n_{.1}!\cdot n_{2.}!\cdot n_{.2}!}{n!\cdot k!\cdot(n_{.1}-k)!\cdot(n_{1.}-k)!\cdot(n-n_{.1}-n_{2.}+k)!}$$

mit

$$n! = n \,.\, (n-1) \,.\, (n-2) \,.... \, 3 \,.\, 2 \,.\, 1.$$

Für ein Signifikanzniveau von $\alpha = 0.05$ ist c_u gleich 3 und c_o gleich 5. Zwischen männlichen und weiblichen potentiellen Käufern kann auf Grund dieser Stichprobe eine unterschiedliche Kaufabsicht angenommen werden. h_{11} liegt nicht zwischen c_u und c_o.

$$c_u = 3 < c_o = 5 < h_{11} = 6$$

Die Nullhypothese kann nicht abgelehnt werden, wenn h_{11} größer als c_u und kleiner als c_o ist.
c_u und c_o werden wie folgt berechnet:

$$\sum_{k=0}^{3-1} \frac{7!\cdot 7!\cdot 5!\cdot 5!}{12!\cdot k!\cdot(7-k)!\cdot(7-k)!\cdot(12-7-5+k)!} = 0 + 0.001 + 0.013 = 0.014 < 0.025$$

$$\sum_{k=0}^{3} \frac{7!\cdot 7!\cdot 5!\cdot 5!}{12!\cdot k!\cdot(7-k)!\cdot(7-k)!\cdot(12-7-5+k)!} = 0.014 + 0.037 = 0.051 > 0.025$$

c_u ist daher gleich 3:

$$0.014 < 0.025 < 0.051$$

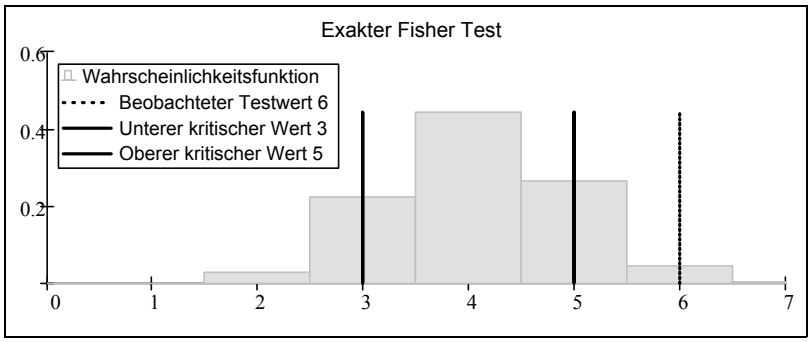
c_o wird auf die gleiche Art bestimmt:

$$\sum_{k=5-1}^{7} \frac{7!\cdot 7!\cdot 5!\cdot 5!}{12!\cdot k!\cdot(7-k)!\cdot(7-k)!\cdot(12-7-5+k)!} = 0.037 + 0.013 + 0.001 = 0.051 > 0.025$$

$$\sum_{k=5}^{7} \frac{7!\cdot 7!\cdot 5!\cdot 5!}{12!\cdot k!\cdot(7-k)!\cdot(7-k)!\cdot(12-7-5+k)!} = 0.013 + 0.001 + 0 = 0.014 < 0.025$$

c_o ist gleich 5:

$$0.00791 \le 0.025 < 0.063346$$



(Beobachteter – Wert)	(Unterer_kritischer – Wert)	(Oberer_kritischer – Wert)
6	3	5

Man kann auf Grund dieser Stichprobenergebnisse annehmen, dass sich das Kaufinteresse der potentiellen weiblichen und männlichen Kunden in der Grundgesamtheit von z. B. 100.000 potentiellen Kunden unterscheidet. Man hat einen zweiseitigen Test durchgeführt bei dem als Ergebnis herauskam, dass sich die beiden Anteilswerte unterscheiden. Will man aber genauer feststellen, ob das Kaufinteresse der potentiellen weiblichen Kunden geringer ist als das der männlichen, dann muss man einen einseitigen Test durchführen. Die entsprechende Alternativhypothese ist nicht

$$H_1: \pi_{männlich} \neq \pi_{weiblich}$$

sondern

$$H_1: \pi_{männlich} > \pi_{weiblich}$$

Diese rechtsseitige Alternativhypothese kann angenommen werden, wenn die Testmaßzahl $x = h_{11}$ größer ist als c_o. Für die Berechnung von c_o werden die gleichen Formeln verwendet wie bei einem zweiseitigen Test. Lediglich $\alpha/2$ wird durch α ersetzt:

$$\sum_{k = c_0 - 1}^{\min(n_{.1}, n_{1.})} \frac{n_{1.}! \cdot n_{.1}! \cdot n_{2.}! \cdot n_{.2}!}{n! \cdot k! \cdot (n_{.1} - k)! \cdot (n_{1.} - k)! \cdot (n - n_{.1} - n_{2.} + k)!} > \alpha$$

$$\alpha \geq \sum_{k = c_0}^{\min(n_{.1}, n_{1.})} \frac{n_{1.}! \cdot n_{.1}! \cdot n_{2.}! \cdot n_{.2}!}{n! \cdot k! \cdot (n_{.1} - k)! \cdot (n_{1.} - k)! \cdot (n - n_{.1} - n_{2.} + k)!}$$

Für das vorliegende Beispiel ist c_o ebenfalls gleich 5, da

$$\sum_{k = 5 - 1}^{7} \frac{7! \cdot 7! \cdot 5! \cdot 5!}{12! \cdot k! \cdot (7 - k)! \cdot (7 - k)! \cdot (12 - 7 - 5 + k)!} = 0.051587 > 0.05$$

$$\sum_{k = 5}^{7} \frac{7! \cdot 7! \cdot 5! \cdot 5!}{12! \cdot k! \cdot (7 - k)! \cdot (7 - k)! \cdot (12 - 7 - 5 + k)!} = 0.014761 < 0.05$$

Man kann daher nicht nur behaupten, dass sich bei allen potentiellen männlichen und weiblichen Käufern das Kaufinteresse unterscheidet, sondern noch genauer, dass das Kaufinteresse der potentiellen männlichen Kunden signifikant größer ist als das der Frauen.

Mit Hilfe der Autorensoftware findet man dieses Ergebnis durch die Eingabe der Rohdaten oder der Kreuztabelle sowie der Eingabe ob ein einseitiger oder zweiseitiger Test durchzuführen ist. Der Output für dieses Beispiel steht auf Seite 186.

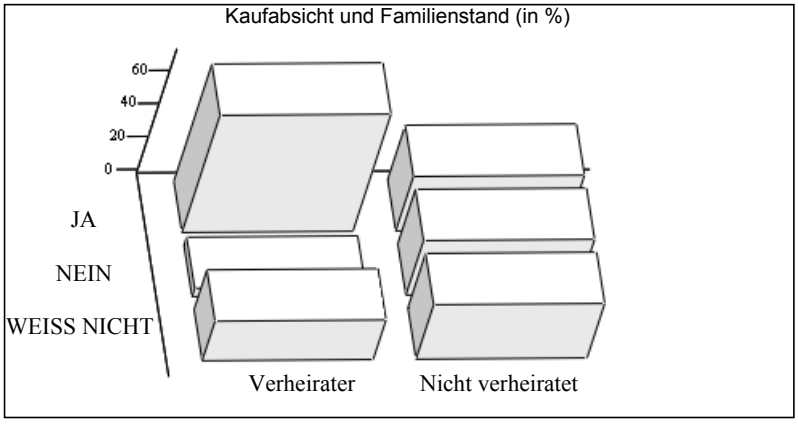
2_2 Chiquadrat-(= χ2)Test

Unterscheiden sich die verheirateten von den nicht verheirateten potentiellen Kunden für das neue Produkt "Spielkonsole Yoki" im Hinblick auf ihre Kaufabsichten? Eine Befragung von 21 Verheirateten und 15 Nicht-Verheirateten bei einer Werbeveranstaltung brachte folgendes Ergebnis:

(Familienstand / Kaufabsicht)	Verheiratet	Nichtverheiratet
Ja	15	5
Nein	1	5
Weiß_nicht	5	5
SUMME	21	15

Wenn man diese Ergebnisse auf die jeweils befragten Verheirateten und Nicht-Verheirateten bezieht, dann zeigt sich, dass 71% der verheirateten Befragten Kaufabsicht haben, aber nur 33% der Nicht-Verheirateten. Bei den "Nein" - Antworten ist es umgekehrt: 5% der befragten Verheirateten stehen hier 33% der Nicht-Verheirateten gegenüber. Bei den Unentschlossenen sind 24% verheiratet und 33% nicht verheiratet:

(Familienstand / Kaufabsicht)	Verheiratet	Nichtverheiratet
Ja	71	33
Nein	5	33
Weiß_nicht	24	33
SUMME	100	100



Kann man auf Grund dieser Stichprobenergebnisse allgemein behaupten, dass sich die Kaufabsichten aller verheirateten potentiellen Kunden von den nicht verheirateten Kunden unterscheiden?

Ja, die Unterschiede zwischen den Prozentsätzen sind groß genug, um allgemein zu behaupten, die Kaufabsichten der verheirateten potentiellen Kunden unterscheiden sich von den Kaufabsichten der nicht verheirateten Kunden. Es könnte zwar tatsächlich sein, dass die verheiraten und nicht verheirateten potentiellen Kunden keine unterschiedlichen Kaufabsichten haben. Das Risiko ist dafür mit 5% beschränkt. Wie kommt man zu diesem Ergebnis? Zuerst formuliert man Vermutungen über die Beziehungen zwischen Grundgesamtheiten in Form von Hypothesen. Man nimmt z. B. an, dass die Verteilung der Kaufabsicht der Verheirateten, gleich ist wie die entsprechende Verteilung der Nicht-Verheirateten. Diese Hypothese wird formal wie folgt angeschrieben:

$$H_0: \text{Verteilung}_{\text{Verheiratet}} = \text{Verteilung}_{\text{Nicht- Verheiratet}}$$

Mit F(x) wird die Verteilung der Kaufabsicht in der Grundgesamtheit aller Verheirateten bzw. Nicht-Verheirateten bezeichnet. Eine Alternative zu dieser Nullhypothese ist

$$H_1: \text{Verteilung}_{\text{Verheiratet}} \neq \text{Verteilung}_{\text{Nicht-Verheiratet}}$$

Die Verteilung der Kaufabsicht der Verheirateten unterscheidet sich von der entsprechenden Verteilung der Nicht-Verheirateten. An Hand der Ergebnisse einer Stichprobe entscheidet man sich für die Annahme oder Ablehnung der Null- oder Alternativhypothese. Dazu berechnet man folgende Testmaßzahl:

$$\chi 2_{beob} = n \cdot \left[\left[\sum_{j=1}^{m} \sum_{k=1}^{t} \frac{(h_{jk})^2}{h_j \cdot h_k} \right] - 1 \right]$$

n ist der Umfang der gesamten Stichprobe, hj und h_k sind die Randhäufigkeiten der Zeilen und Spalten sowie h_{jk} die Häufigkeiten der Antwortkombinationen. In einer Kreuztabelle sieht es so aus:

(Antwort_B / Antwort_A)	B_1	B_2	ZEILENSUMME
A_1	h_{11}	h_{12}	$\sum_{k=1}^{2} h_{1k} = h_{1.}$
.....	.....	.....	.....
A_j	h_{j1}	h_{j2}	$\sum_{k=1}^{2} h_{jk} = h_{j.}$
.....	.....	.....	.....
A_m	h_{m1}	h_{m2}	$\sum_{k=1}^{2} h_{mk} = h_{m.}$
SPALTENSUMME	$\sum_{j=1}^{m} h_{j1} = h_{.1}$	$\sum_{j=1}^{m} h_{j2} = h_{.2}$	$\sum_{j=1}^{m} \sum_{k=1}^{2} h_{jk} = n$

In folgender Kreuztabelle sind neben den Häufigkeiten der Antwortkombinationen auch die Zeilen und Spaltensummen des Beispiels angeführt:

(Familienstand / Kaufabsicht)	Verheiratet	Nichtverheiratet	ZEILENSUMME
Ja	15	5	20
Nein	1	5	6
Weiß_nicht	5	5	10
SPALTENSUMME	21	15	36

Nun kann man die Werte in die Formel für die Chiquadratmaßzahl einsetzen und berechnen:

$$\chi 2_{beob} = 36 \cdot \left(\frac{15^2}{20 \cdot 21} + \frac{5^2}{20 \cdot 15} + \frac{1^2}{6 \cdot 21} + \frac{5^2}{6 \cdot 15} + \frac{5^2}{10 \cdot 21} + \frac{5^2}{10 \cdot 15} - 1 \right) = 6.857$$

Die $\chi 2$-Testmaßzahl folgt einer Chiquadratverteilung mit

$$\nu = (m - 1) \cdot (t - 1)$$

Freiheitsgraden. m ist die Anzahl der Antworten in den Zeilen und t die der Spalten der Kreuztabelle. Im obigen Beispiel sind die Freiheitsgrade gleich $\nu = (3 - 1). (2 - 1) = 2$ und die Wahrscheinlichkeit für das Auftreten einer so großen Testmaßzahl von $\chi 2_{beob} = 6.857$ unter Gültigkeit der Nullhypothese gleich $F_{\chi 2}(6.857, 2) = 0.968$. (Diesen Wert liest man aus einer geeigneten Tabelle der Chiquadratverteilung ab oder berechnet ihn mit Hilfe entsprechender Programme.)

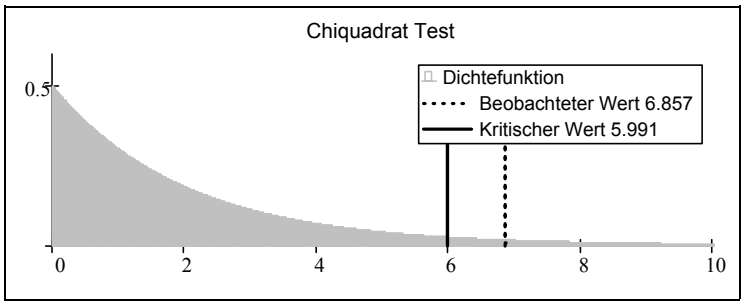
Ist man bereit, in 5 von 100 Fällen eine richtige Nullhypothese abzulehnen, dann ist das Signifikanzniveau α gleich 0.05. Unter Gültigkeit der Nullhypothese sind die 5% seltensten $\chi 2$ Werte nach unten beschränkt durch den kritischen Wert $\chi 2_c$ gleich 5.991. Da das aus der Stichprobe errechnete $\chi 2_{beob}$ von 6.857 größer ist als dieser kritische Wert, kann die Nullhypothese abgelehnt werden.

$$\chi 2_{beob} = 6.857 > 5.991 = \chi 2_c$$

Zum selben Ergebnis kommt man, wenn man die Wahrscheinlichkeiten vergleicht: Da die Wahrscheinlichkeit für einen $\chi 2$ Wert von 6.857 oder größer gleich $1 - F\chi 2(6.857, 2) = 0.032$ ist und das Signifikanzniveau mit 0.05 angenommen wurde, kann die Nullhypothese abgelehnt werden.

$$1 - F_{\chi 2}(6.857, 2) = 0.032 < 0.05 = \text{Signifikanzniveau}$$

Man kann auf Grund dieser Stichprobe behaupten, dass sich die verschiedenen Formen der Kaufabsicht unterschiedlich auf Verheiratete und Nichtverheiratete verteilen. Das Risiko, dass diese Entscheidung falsch ist, ist höchstens 5%.



(Chiquadrat_beobachtet)	(Chiquadrat_kritisch)	(p_Wahrscheinlichkeit)
6.857	5.991	0.032

Wo sind die signifikanten Unterschiede? Unterscheiden sich die Ja-Antworten der Verheirateten und Nichtverheirateten oder die Nein-Antworten oder die Weiß-nicht-Antworten? Der Chiquadrattest kam nur zum Ergebnis, dass Unterschiede in der Kaufabsicht zwischen Verheirateten und Nichtverheirateten bestehen, nicht aber bei welcher Antwort. Tatsächlich unterscheiden sich die Ja- und Nein-Antworten der Verheirateten und Nichtverheirateten signifikant, nicht aber die Weiß-nicht-Antworten.

Zu diesem Ergebnis kommt man mit Hilfe der Konfidenzintervalle. In der folgenden Tabelle sind als letzte Spalte die Prozentwerte der Antworten auf die Frage der Kaufabsicht angeführt, unabhängig vom Familienstand.

Familienstand / Kaufabsicht	Verheiratet	Nichtverheiratet	Durchschnitt
Ja	71	33	56
Nein	5	33	17
Weiß_nicht	24	33	28
SUMME	100	100	100

56% von den insgesamt 36 Befragten haben mit "Ja" geantwortet. Bei den Verheirateten sind es 71%, bei den Nichtverheirateten 33%. Innerhalb welcher Grenzen kann man mit einem Vertrauen von 95% den "wahren" Prozentsatz für "Ja" in der Grundgesamtheit aller potentiellen Kunden erwarten, wenn in der Stichprobe der Prozentsatz für "Ja" gleich 56 ist?

In der Grundgesamtheit ist der Prozentsatz für "Ja" bei der Frage nach der Kaufabsicht mit 95% Vertrauen innerhalb der Grenzen 38% und 72% zu erwarten. Da die 71% Ja-Antworten der Verheirateten unter der Obergrenze von 72% liegt, ist die Kaufabsicht der Verheirateten nicht überdurchschnittlich hoch aber die der Nichtverheirateten mit 33% unterdurchschnittlich gering (33% liegt unter der Grenze von 38%).

Im Schnitt haben 17% der Befragten die Kaufabsicht verneint. In der Grundgesamtheit aller potentiellen Kunden ist mit einem Vertrauen von 95% zu erwarten, dass dieser Prozentsatz zwischen 6% und 30% liegt. Die Verheirateten liegen hier mit 5% unter der Untergrenze, die Nichtverheirateten mit 33% oberhalb der Obergrenze des Konfidenzintervalles. Die Abweichungen sind daher signifikant (bei einem Signifikanzniveau von 5%).

Bei den Nichtentschlossenen liegt der Prozentsatz in der Stichprobe bei 28%, in der Grundgesamtheit ist er innerhalb der Grenzen 14% und 42% zu erwarten. Sowohl der Prozentsatz der Verheirateten mit 24% als auch der der Nichtverheirateten mit 33% liegt innerhalb dieser Grenzen. Eine über- oder unterdurchschnittliche signifikante Abweichung liegt hier also nicht vor. Die Konfidenzgrenzen können z. B. mit Hilfe der F-Verteilung berechnet werden (siehe Auswertung einer Frage, nominal). Für die Ja- Antworten ist n = 36 und x = 20 (= 56%). K_u ist daher

$$K_u = \frac{x}{x + (n - x + 1)\cdot F_1} = \frac{20}{20 + (36 - 20 + 1)\cdot 1.912} = 0.381$$

mit

$$F_1 = F_{1-\frac{\alpha}{2},\,2\cdot(n-x-1),\,2\cdot x} = F_{0.975,\,34,\,40} = 1.912$$

K_o ist

$$K_o = \frac{(x + 1)\cdot F_2}{(n - x) + (x + 1)\cdot F_2} = \frac{(20 + 1)\cdot 1.965}{(36 - 20) + (20 + 1)\cdot 1.965} = 0.721$$

mit

$$F_2 = F_{1-\frac{\alpha}{2},\,2\cdot(x+1),\,2\cdot(n-x)} = F_{0.975,\,42,\,32} = 1.965$$

Die Ergebnisse der signifikanten Abweichungen vom Durchschnitt sind in folgender Tabelle zusammengestellt:

(Familienstand / Kaufabsicht)	Verheiratet	Nichtverheiratet	Durchschnitt
Ja	"---"	"unter"	56
Nein	"unter"	"über"	17
Weiß_nicht	"---"	"---"	28

Im Output der Software auf Seite 187 findet man nicht nur das Gesamtergebnis für dieses Beispiel, sondern auch die Ergebnisse für die einzelnen Kaufabsichtsgruppen, so wie sie hier dargestellt sind. Als Eingabe benötigt das Programm nicht nur Rohdaten. Es kann diese Ergebnisse auch berechnen, wenn man nur mehr die Kreuztabelle besitzt.

2_3 U-Test

Unterscheidet sich die Designbeurteilung der neuen Spielkonsole Yoki durch die weiblichen Befragten wesentlich von der der männlich Befragten? Bei einer Werbeveranstaltung beurteilten 6 Frauen und 4 Männer das Design mit Hilfe folgender Attribute: 1 = sehr gut, 2 = gut, 3 = befriedigend, 4 = schlecht, 5 = sehr schlecht. Die Frauen lieferten die Beurteilungen

$$\text{weiblich} = (3 \quad 5 \quad 1 \quad 5 \quad 4 \quad 2)$$

und die Männer

$$\text{männlich} = (4 \quad 2 \quad 3 \quad 4).$$

In folgender Tabelle sind die Beurteilungen zusammengefasst:

(Geschlecht / Designbeurteilung)	Weiblich	Männlich
sehr_gut	1	0
gut	1	1
mittelmäßig	1	1
schlecht	1	2
sehr_schlecht	2	0
SUMME	6	4

Kann man allgemein behaupten, dass in der Designbeurteilung zwischen Männern und Frauen ein signifikanter Unterschied besteht? Nein, man kann nicht behaupten, dass sich die Verteilungen der Designbeurteilungen von Männern und Frauen in der Grundgesamtheit aller potentiellen Kunden, aus der die beiden Stichproben stammen, signifikant unterscheiden. Die Differenzen in den Beurteilungen zwischen den beiden Stichproben sind zu gering, um als Hinweis auf signifikante Unterschiede in der Grundgesamtheit zu dienen. (Siehe Seite 189, Softwareoutput)

Zu dieser Schlussfolgerung kommt man mit Hilfe des U-Tests, der hier dem Chiquadrattest vorzuziehen ist. Der U-Test berücksichtigt nämlich die Tatsache, dass zwischen den Beurteilungen von sehr gut bis sehr schlecht eine Ordnung besteht, nicht aber der Chiquadrattest. Zuerst formuliert man die Vermutungen über die Beziehungen zwischen Grundgesamtheiten in Form von Hypothesen. Man nimmt an, dass die Verteilungen der Designbeurteilungen des neuen Produktes bei Frauen und Männern gleich ist. Diese Hypothese wird formal wie folgt angeschrieben:

$$H_0: \text{Verteilung}_{\text{Weiblich}} = \text{Verteilung}_{\text{Männlich}}$$

Eine Alternative dazu ist

$$H_1: \text{Verteilung}_{\text{Weiblich}} \neq \text{Verteilung}_{\text{Männlich}}$$

Die Designbeurteilungen der Frauen, unterscheidet sich vom der der Männer. An Hand der Ergebnisse der Stichprobe entscheidet man sich für die Annahme oder Ablehnung der Null- oder Alternativhypothese.
Will man prüfen, ob zwei ordinale Stichproben nicht aus Grundgesamtheiten mit gleicher Verteilung

stammen, die Unterschiede zwischen ihnen also signifikant sind, dann verwendet man den U-Test von Wilcoxon-Mann-Whitney. Getestet wird die kleinere der beiden Testmaßzahlen:

$$U_1 = \left[n_1 \cdot n_2 + \frac{n_1 \cdot (n_1 + 1)}{2} \right] - S_1$$

$$U_2 = \left[n_1 \cdot n_2 + \frac{n_2 \cdot (n_2 + 1)}{2} \right] - S_2$$

n_1 und n_2 sind die Umfänge der beiden Stichproben. S_1 ist die Summe der Rangzahlen der ersten und S_2 die der zweiten Stichprobe. Zur Berechnung der Rangzahlen bringt man die Einheiten beider Stichproben in eine gemeinsame Reihenfolge und berechnet für jede Antwort die entsprechende Rangzahl. Zu jeder Rangzahl vermerkt man, aus welcher der beiden Stichproben die Einheit stammt. Haben mehrere Einheiten die gleiche Merkmalsausprägung, dann ordnet man ihnen als Rangzahlen den Durchschnitt daraus zu. In folgender Tabelle sind die Rechenschritte für die Berechnung von S_1 und S_2 zusammengefasst:

(Beurteilung)	(Geschlecht)	Rangzahlen	Rang_weiblich	Rang_männlich
1	2	1	1	_
2	2	2.5	2.5	_
2	1	2.5	_	2.5
3	1	4.5	_	4.5
3	2	4.5	4.5	_
4	1	7	_	7
4	1	7	_	7
4	2	7	7	_
5	2	9.5	9.5	_
5	2	9.5	9.5	_
SUMME	"*"	"*"	34	21

In der ersten Spalte stehen die geordneten Designbeurteilungen der weiblich und männlich Befragten. In der zweiten Spalte ist festgehalten, ob die Beurteilung von einer Frau (1) oder einem Mann (2) stammt. Anschließend findet man die zugeordneten Rangzahlen für die Gesamtstichprobe. Wenn eine Beurteilung zweimal vorkommt, dann wird beiden der Durchschnitt aus den aufeinander folgenden Rangzahlen zugeordnet. So scheint z. B. die Beurteilung "2" (= gut) zweimal auf. In der geordneten Gesamtstichprobe stehen diese beiden Beurteilungen am 2. und 3. Platz. Daher wird ihnen der Durchschnitt daraus, nämlich 2.5 zugeordnet. Die vierte Spalte enthält die Rangzahlen der ersten Stichprobe (weiblich Befragte) und die fünfte Spalte die der zweiten Stichprobe (männlich Befragten). In den entsprechenden Spalten ist jeweils die Rangsumme am Ende der Spalte angeführt. So kommt die erste Stichprobe auf eine Rangsumme von 34 und die der zweiten auf 21. Die Maßzahlen U_1 und U_2 können nun berechnet werden:

$$U_1 = 6 \cdot 4 + \frac{6(6+1)}{2} - 34 = 11 \qquad U_2 = 6 \cdot 4 + \frac{4(4+1)}{2} - 21 = 13$$

Die Testmaßzahl U ist das Minimum der beiden Zahlen:

$$U = \min(U_1, U_2) = 11$$

Diese Maßzahl ist normalverteilt mit Erwartungswert m_U und Standardabweichung s_U:

$$m_U = \frac{n_1 \cdot n_2}{2}$$

$$s_U = \sqrt{\frac{n_1 \cdot n_2}{n \cdot (n-1)} \cdot \left(\frac{n^3 - n}{12} - T \right)}$$

n ist die Summe von n_1 und n_2. T ist folgendermaßen definiert:

$$T = \sum_{i=1}^{k} \left[\frac{(f_i)^3 - f_i}{12} \right]$$

T dient der Korrektur für Durchschnittsränge, also Ränge, die mehrfach vorkommen. Da einige Rangplätze im obigen Beispiel gleich sind, muss dies über ihre Häufigkeit f berücksichtigt werden:

(Rangzahlen)	f	f^3	$f^3 - f$
1	1	1	o
2	2	8	6
3	2	8	6
4	3	27	24
5	2	8	6
SUMME	"*"	"*"	42

In der ersten Spalte stehen die Rangzahlen, in der zweiten ihre Häufigkeit, in der dritten die Häufigkeit zur dritten Potenz und in der vierten Spalte wird davon die Häufigkeit wieder abgezogen. Die Summe der vierten Spalte ergibt 42 und der Korrekturfaktor errechnet sich wie folgt:

$$T = \frac{42}{12} = 3.5$$

Die Standardabweichung der Testmaßzahl U ist daher

$$s_U = \sqrt{\frac{6 \cdot 4}{10 \cdot (10-1)} \cdot \left(\frac{10^3 - 10}{12} - 3.5 \right)} = 4.59$$

und der Erwartungswert ist

$$m_U = \frac{6 \cdot 4}{2} = 12$$

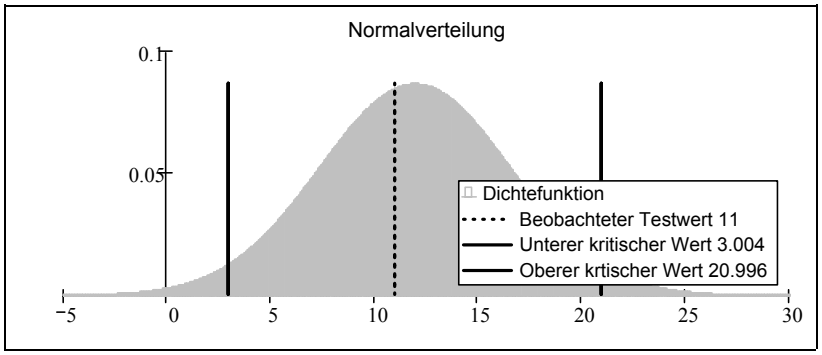
Die Wahrscheinlichkeit für das Auftreten einer so großen Testmaßzahl U = 11 unter Gültigkeit der Nullhypothese ermittelt man mit Hilfe der Normalverteilung:

$$F_N(U, m_U, s_U) = F_N(11, 12, 4.59) = 0.414$$

Wählt man $\alpha = 0.05$, dann ist diese Wahrscheinlichkeit kleiner als $1 - \alpha/2 = 0.975$ und größer als $\alpha/2 = 0.025$:

$$\frac{\alpha}{2} = 0.025 < 0.414 < 0.975 = 1 - \frac{\alpha}{2}$$

Daher wird die Gleichheit der Verteilungen beider Grundgesamtheiten nicht abgelehnt. Die Stichproben liefern keinen Hinweis auf Verteilungsunterschiede von Frauen und Männern im Hinblick auf die Designbeurteilung. In unten stehender Grafik ist dieses Testergebnis veranschaulicht. Die Testmaßzahl U (die blaue Linie) liegt zwischen c_u und c_o:



Beobachteter_U_Wert	Kritische_Werte	p_Wahrscheinlichkeit
11	(3.004 20.996)	0.586

Zum gleichen Ergebnis kommt man, wenn man nicht die Wahrscheinlichkeiten vergleicht, sondern die errechnete z-Maßzahl z_{err} und die kritische z-Maßzahl z_c:

$$\left|z_{err}\right| = \left|\frac{U - m_U}{s_U}\right| = \left|\frac{11 - 12}{4.59}\right| = 0.218$$

$$z_c = 1.96$$

1.96 ist der kritische Wert der Standardnormalverteilung für ein Signifikanzniveau von $\alpha = 0.05$. Diesen Wert findet man in entsprechenden Tabellen der Quantile der Standardnormalverteilung oder mit Hilfe von Rechenprogrammen, die den Wert mit Hilfe der Verteilungsfunktion numerisch bestimmen. Da

$$\left|z_{err}\right| = 0.218 < 1.96 = z_c$$

kommt man hier zum selben Ergebnis wie oben.

2_4 t-Test und Welch-Test

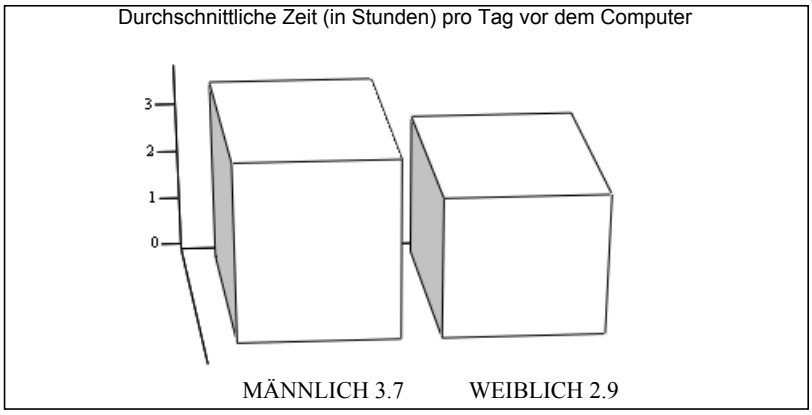
Kann man voraussetzen, dass Männer im Schnitt pro Tag länger vor dem Computer sitzen als Frauen? Bei einer Werbeveranstaltung für die neue Spielkonsole Yoki wurden 16 Männer und 9 Frauen danach gefragt, wie viele Zeit sie im Schnitt pro Tag vor dem Computer verbringen.

(Männlich 8 1 3 2.5 8 7.5 6 3.5 1 2 0.5 3 5 5 2.5 0.5)

(Weiblich 7 0 0 4.5 0 0.5 4 8 2.5)

Ausgezählt nach drei Zeitintervallen erhält man folgendes Ergebnis:

"Computerzeit" "Untergrenze"	"Computerzeit" "Obergrenze"	"Geschlecht" "Männlich"	"Geschlecht" "Weiblich"
0	2	4	4
2	4	6	1
4	9	6	4
("absolut")	("Summe")	16	9



Im Schnitt verbringen die 16 befragten Männer 3.7 Stunden täglich vor dem Computer und 2.9 Stunden die Frauen. Ist diese Zeitdifferenz schon groß genug, um allgemein zu behaupten, die Männer verbringen täglich mehr Zeit vor dem Computer als Frauen?

Nein, diese Zeitdifferenz ist noch zu gering, um allgemein anzunehmen, Männer sitzen im Schnitt pro Tag länger vor dem Computer als Frauen. Zu diesem Ergebnis kommt man wie folgt: Zuerst werden Null- und Alternativhypothese formuliert.

H_1: $\mu_{männlich} = \mu_{weiblich}$

Die Durchschnittszeit pro Tag vor dem Computer ist von Frauen und Männern der Grundgesamtheit gleich und

$$H_1: \mu_{männlich} \geqslant \mu_{weiblich}$$

Die Durchschnittszeit pro Tag vor dem Computer ist von Männern länger als die von Frauen. Um diese Hypothesen zu prüfen, berechnet man die Testmaßzahl

$$t_{beob} = \frac{xquer_{männlich} - xquer_{weiblich}}{s_d}$$

mit

$$s_d = \sqrt{\frac{n_{männlich} \cdot s_{männlich}^2 + n_{weiblich} \cdot s_{weiblich}^2}{n_{männlich} + n_{weiblich} - 2} \cdot \left(\frac{1}{n_{männlich}} + \frac{1}{n_{weiblich}} \right)}$$

Für das Beispiel errechnet sich ein Wert von

$$s_d = \sqrt{\frac{16 \cdot 6.796 + 9 \cdot 9.715}{16 + 9 - 2} \cdot \left(\frac{1}{16} + \frac{1}{9} \right)} = 1.217$$

und

$$t_{beob} = \frac{3.688 - 2.944}{1.217} = 0.611$$

mit

$$s_{männlich}^2 = \frac{1}{16} \cdot \sum_{i=1}^{16} \left(x_i - 3.688 \right)^2 = 6.796$$

$$s_{weiblich}^2 = \frac{1}{9} \cdot \sum_{i=1}^{9} \left(x_i - 2.944 \right)^2 = 9.715$$

Die Testmaßzahl t ist studentverteilt mit

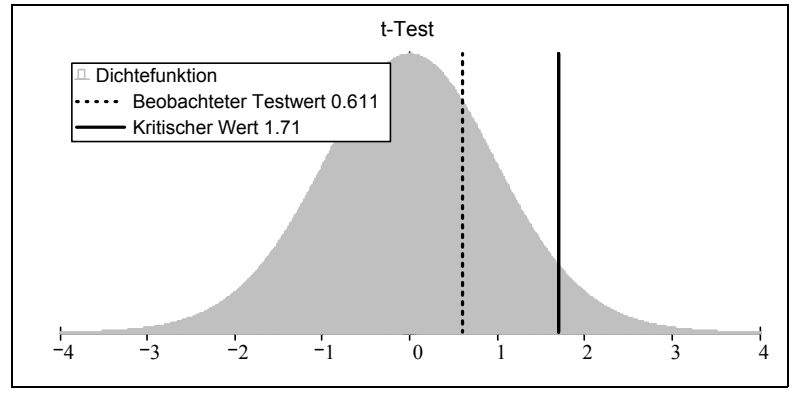
$$\nu = n_{männlich} + n_{weiblich} - 2 = 16 + 9 - 2 = 23$$

Freiheitsgraden. Die Wahrscheinlichkeit für das Auftreten einer so großen Testmaßzahl errechnet sich mit Hilfe der Studentverteilung auf

$$F_t(t_{beob}, \nu) = F_t(0.611, 23) = 0.726$$

Da diese Wahrscheinlichkeit kleiner ist als 1 – 0.05 = 0.95 wird die Gleichheit der Durchschnitte beider Grundgesamtheiten nicht abgelehnt. Zum gleichen Ergebnis kommt man, wenn man den beobachteten t-Wert mit dem kritischen t-Wert vergleicht. Für ein 5%-iges Signifikanzniveau und 23 Freiheitsgrade ist der kritische t-Wert für eine linksseitige Alternativhypothese $t_c = 1.714$. Da der beobachtete t-Wert $t_{beob} = 0.611$ kleiner ist als der kritische t-Wert, kann die Nullhypothese nicht abgelehnt werden.

In unten stehender Grafik ist dieses Testergebnis veranschaulicht. Die Testmaßzahl t_{beob} (die rote gestrichelte Linie) liegt links vor dem rechten Schwanzende (blaue Linie):



Beobachteter t-Wert	Kritische Werte	Wahrscheinlichk. p
0.61	$(-\infty \quad 1.71)$	0.265

F-Test für den Varianzvergleich

Der t-Test setzt voraus, dass die unbekannten Varianzen der beiden Grundgesamtheiten "Männlich" und "Weiblich" gleich sind. Man kann diese Voraussetzung mit Hilfe eines F-Tests prüfen, wenn man voraussetzen kann, dass beide Stichproben aus normalverteilten Grundgesamtheiten stammen. Um die beiden Hypothesen

$$H_0: \sigma_{männlich}^2 = \sigma_{weiblich}^2$$

$$H_1: \sigma_{männlich}^2 \neq \sigma_{weiblich}^2$$

zu prüfen, berechnet man die Testmaßzahl

$$F_{beob} = \frac{s_{männlich}^2}{s_{weiblich}^2} \cdot \frac{v_2}{v_1}$$

mit

$$v_1 = n_{männlich} - 1$$

$$v_2 = n_{weiblich} - 1$$

Für das Beispiel sind die entsprechenden Werte

$$F_{beob} = \frac{6.796}{9.715} \cdot \frac{8}{15} = 0.373$$

$$v_1 = 16 - 1 = 15$$

$$v_2 = 9 - 1 = 8$$

Für ein 5%-iges Signifikanzniveau und obige Freiheitsgrade ist der kritische Wert der F-Verteilung gleich

$$F\left(1 - \alpha, v_1, v_2\right) = F(0.95, 15, 8) = 3.218$$

Da der beobachtete F-Wert von 0.373 kleiner ist als der kritische F-Wert von 3.218, kann die Nullhypothese gleicher Varianzen nicht abgelehnt werden

Welch-Test für Mittelwertsunterschiede bei ungleichen Varianzen

Wenn man annehmen muss, dass die Varianzen in den beiden Grundgesamtheiten ungleich sind, dann kann man nicht obigen t-Test für die Prüfung auf Mittelwertsunterschiede durchführen, sondern muss dafür den umständlicheren Welch-Test verwenden.

Um die Hypothesen

$$H_0: \mu_{männlich} = \mu_{weiblich}$$

$$H_1: \mu_{männlich} > \mu_{weiblich}$$

zu prüfen, berechnet man hier die Testmaßzahl

$$t_{beob} = \frac{xquer_{männlich} - xquer_{weiblich}}{s_d}$$

mit

$$s_d = \sqrt{\frac{s_{männlich}^2}{n_{männlich} - 1} + \frac{s_{weiblich}^2}{n_{weiblich} - 1}}$$

Für das Beispiel errechnet sich ein Wert von

$$s_d = \sqrt{\frac{6.796}{16 - 1} + \frac{9.715}{9 - 1}} = 1.291$$

und

$$t_{beob} = \frac{3.688 - 2.944}{1.291} = 0.576$$

Die Testmaßzahl t_{beob} ist studentverteilt mit

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

Freiheitsgraden. Eingesetzt ergibt dies

$$\nu = \frac{\left(\frac{6.796}{16} + \frac{9.715}{9}\right)^2}{\frac{\left(\frac{6.796}{16}\right)^2}{16 - 1} + \frac{\left(\frac{9.715}{9}\right)^2}{9 - 1}} = 14.35$$

Freiheitsgrade. Der kritische Wert der Studentverteilung für ein 5%-iges Signifikanzniveau ist

$$F_t(1 - \alpha, \nu) = F_t(0.95, 14.35) = 1.758$$

Da der beobachtete Wert $t_{beob} = 0.576$ kleiner ist als der kritische, kann auch nach diesem Test die Nullhypothese nicht abgelehnt werden.

Sowohl der t-Test (bei Varianzengleichheit) als auch der Welch-Test setzen voraus, dass die Verteilungen der Grundgesamtheiten normalverteilt sind. Diese Voraussetzung kann mit Hilfe des Kolmogorov-Smirnov-Test überprüft werden. Wenn man die Normalverteilung ablehnen muss, dann kann man weder den t-Test noch den Welch-Test anwenden. In diesem Fall kann man mit Hilfe des U-Tests überprüfen, ob sich die beiden Verteilungen in den Grundgesamtheiten signifikant unterscheiden.

Das Softwareprogramm dessen Output auf Seite 190 zu finden ist, prüft diese Voraussetzungen und wählt automatisch das passende Programm. Es erfordert wie üblich nur die Eingabe der Daten, die Wahl ob einseitiger oder zweiseitiger Test und die Eingabe des Signifikanzniveaus. Die Voreinstellung rechnet mit einem Signifikanzniveau von 5%.

3_0 ZWEI ABHÄNGIGE STICHPROBEN

Unabhängig - Abhängig

Werden 10 Männer und 10 Frauen aus der Grundgesamtheit der potentiellen Kunden für die Spielkonsole Yoki zufällig ausgewählt und nach ihrer Kaufabsicht befragt, dann ist die Stichprobe der Männer unabhängig von der Stichprobe der Frauen. Von abhängigen Stichproben spricht man, wenn an den gleichen Personen z. B. vor und nach einer Werbeveranstaltung die gleichen Fragen gestellt werden. Fragt man 10 Besucher einer Werbeveranstaltung für die Spielkonsole Yoki vor und nach dieser Veranstaltung nach ihrer Kaufabsicht, dann liegen 2 abhängige oder verbundene Stichproben vor. Mit ihnen kann man feststellen, welchen Einfluss die Werbeveranstaltung auf das Kaufverhalten hat. Weitere Beispiele für abhängige Stichproben sind Untersuchungen des Gesundheitszustandes vor und nach einer Behandlung, Einstellungsmessungen vor und nach einer politischen Wahlveranstaltung usw.

Vorschau

Im Abschnitt 3_1 wird der Frage nachgegangen, wie die Werbeveranstaltung für die Spielkonsole Yoki die generelle Einstellung zu Konsolenspielen verändert. Den Besuchern dieser Werbeveranstaltung wurde vor und nach der Veranstaltung die Frage gestellt, ob sie Konsolenspielen gegenüber positiv oder negativ eingestellt sind. Da die Antworten auf diese Frage nominal skaliert sind, kommt für dieses Problem der Test von McNemar zur Anwendung.

Im Abschnitt 3_2 wird untersucht, ob sich durch mehrmaliges Spielen des neuen Konsolenspiels "Fantasy" die benötigte Zeit dafür verringert, ob also ein Lerneffekt auftritt. Dazu wurde die benötigte Zeit für das Spiel beim 1. Versuch und beim 3. Versuch gemessen und verglichen. Da die gemessene Zeit metrisch skaliert ist, kann man auf dieses Problem auch ein metrisches Testverfahren wie den t-Test anwenden. Dieser t-Test setzt aber voraus, dass die beiden Stichproben aus normalverteilten Grundgesamtheiten stammen. Für die Prüfung dieser Frage wurde daher der Wilcoxon Test herangezogen, der ein ordinales Testverfahren ist. Dieser Test verlangt nur, dass die Antworten ordinal skaliert sind, nicht aber, dass die Stichproben aus normalverteilten Grundgesamtheiten stammen. Jede metrisch skalierte Stichprobe kann in eine ordinal oder nominal skalierte Stichprobe transformiert werden, aber nicht umgekehrt.

Die Frage, ob sich durch mehrere Versuche die Spielzeit verringert, wurde im Abschnitt 3_2 mit Hilfe des Wilcoxon-Tests geklärt. Im Abschnitt 3_3 wird die gleiche Frage mit Hilfe des t-Tests untersucht. Für die Anwendung dieses Tests muss vorausgesetzt werden, dass beide Stichproben aus normalverteilten Grundgesamtheiten stammen.

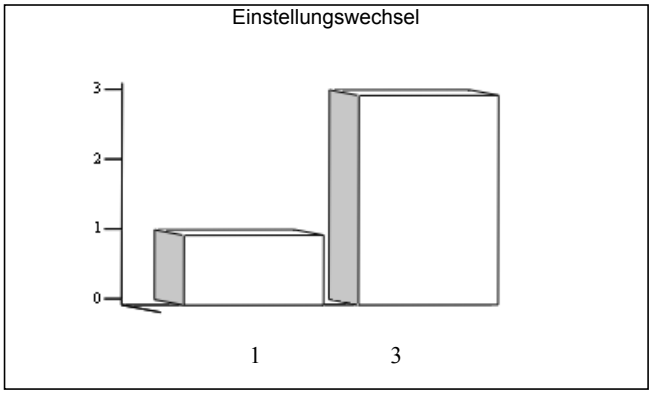
3_1 McNemar-Test

Kann man davon ausgehen, dass die Werbeveranstaltung für die neue Spielkonsole Yoki die Einstellung zu Konsolenspielen verändert? Bei einer Werbeveranstaltung für die neue Spielkonsole Yoki wurden 10 Besucher befragt, welche Einstellung zu Konsolenspielen sie vor und nach der Werbeveranstaltung hatten (1 = positiv, 2 = negativ):

(Vor_Veranstaltung)	2	1	2	1	1	1	2	1	1	2
(Nach_Veranstaltung)	1	1	1	2	1	1	1	1	1	2

Ausgezählt erhält man folgendes Ergebnis:

"*"	(Nach)	(Veranstaltung)
(Vor – Veranstaltung)	Positiv	Negativ
Positiv	5	1
Negativ	3	1
SUMME	8	2



Wechsel von positiv auf negativ (1) negativ auf positiv (3)

Durch die Werbeveranstaltung hat ein Besucher seine Einstellung von positiv auf negativ geändert und 3 von negativ auf positiv. Ist diese Einstellungsänderung schon groß genug, um allgemein zu behaupten, dass sich durch die Werbeveranstaltung die Einstellung zu Konsolenspiele signifikant veränderte? Nein, diese Einstellungsänderungen ist nicht groß genug. Man kann nicht generell behaupten, dass sich durch die Werbeveranstaltung die Einstellung zu Konsolenspiele signifikant verändert. Zu diesem Ergebnis kommt man wie folgt:

Zuerst werden Null- und Alternativhypothese formuliert. Als Nullhypothese nimmt man an, dass der Anteil an positiven und negativen Änderungen gleich sei:

$$H_0: \pi_{vor\_Veranstaltung} = \pi_{nach\_Veranstaltung}$$

Die Alternative dazu ist

$$H_1: \pi_{vor\_Veranstaltung} \neq \pi_{nach\_Veranstaltung}$$

Durch die Werbeveranstaltung treten Einstellungsänderungen im Hinblick auf Konsolenspiele auf. Um diese Hypothesen zu prüfen, berechnet man folgende Testmaßzahl:

$$\chi 2_{beob} = \frac{(b - c)^2}{b + c}$$

b und c stammen aus folgender 4-Feldertafel:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Für das Beispiel ist

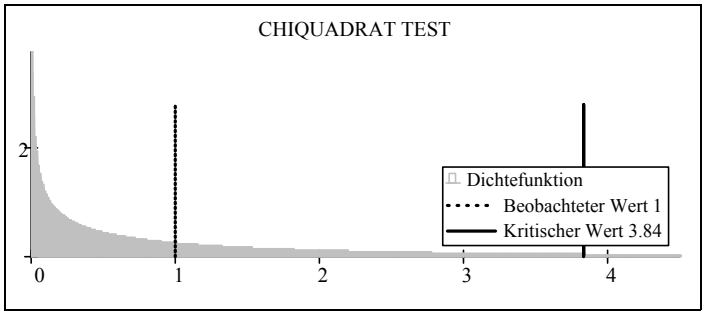
"*"	(Nach)	(Veranstaltung)
(Vor – Veranstaltung)	Positiv	Negativ
Positiv	5	1
Negativ	3	1
SUMME	8	2

und $\chi 2_{beob}$ ist

$$\chi 2_{beob} = \frac{(1 - 3)^2}{1 + 3} = 1$$

Die Testmaßzahl $\chi 2$ ist chiquadratverteilt mit 1 Freiheitsgrad. Für ein 5%-iges Signifikanzniveau ist daher der kritische Wert $\chi 2_c = 3.84$. Da der beobachtete $\chi 2$-Wert (= 1) kleiner ist als der kritische, kann die Nullhypothese nicht abgelehnt werden. Die Einstellungsänderungen der 10 Besucher sind zu gering, um daraus auf eine allgemeine Einstellungsänderung durch die Werbeveranstaltung zu schließen.

In unten stehender Grafik ist dieses Testergebnis veranschaulicht. Die Testmaßzahl $\chi 2_{beob}$ (= 1) liegt links vor dem kritischen Wert $\chi 2_c$ (= 3.84):



(Chiquadrat_beobachtet)	(Chiquadrat_kritisch)	(p – Wahrscheinlichkeit)
1	3.841	0.317

In anderer Einkleidung findet man die Interpretation für dieses Beispiel im Softwareoutput auf Seite 192. Auch für dieses Programm sind neben den Daten lediglich das Signifikanzniveaus anzugeben und die Entscheidung für einen einseitigen oder zweiseitigen Test.

Voraussetzungen und Einschränkungen:

Die Erwartungswerte der Zellen b und c in der Vierfeldertabelle müssen > 5 und alle Erwartungswerte müssen > 0 sein. Die erste Voraussetzung ist im vorliegenden Beispiel nicht gegeben. Daher ist das Testergebnis nicht gültig.

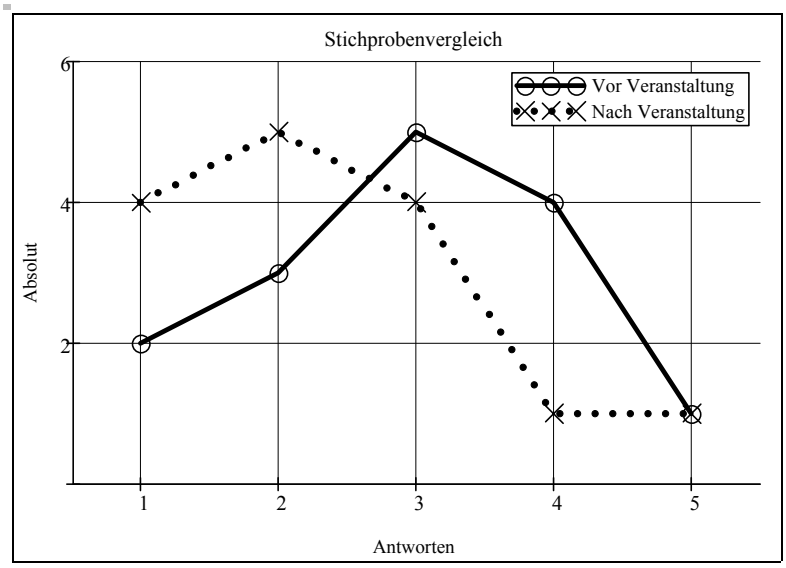
3_2 Wilcoxon-Test

Kann man allgemein annehmen, dass man das Design der neuen Spielkonsole nach der Werbeveranstaltung besser beurteilt als vor der Werbeveranstaltung? Bei einer Werbeveranstaltung für die neue Spielkonsole Yoki wurden 15 Besucher vor und nach der Veranstaltung gefragt, wie sie das Design der Spielkonsole beurteilen:

(Vor_Veranstaltung)	3	3	1	4	3	2	4	4	5	4	2	2	3	1	3
(Nach_Veranstaltung)	3	2	1	2	2	3	2	3	4	5	1	3	2	1	1

Ausgezählt nach den Beurteilungen erhält man folgendes Ergebnis:

"*"	(Vor_Veranstaltung)	(Nach_Veranstaltung)
"Beurteilung"	"absolut"	"absolut"
"Sehr gut"	2	4
"Gut"	3	5
"Befriedigend"	5	4
"Genügend"	4	1
"Nicht genügend"	1	1
"SUMME"	15	15



("Sehr gut")	"Gut"	"Befriedigend"	"Genügend"	"Nicht genügend"
1	2	3	4	5

Der mittlere Besucher beurteilt das Design der neuen Spielkonsole Yoki vor der Werbeveranstaltung mit befriedigend und nach der Veranstaltung mit gut. Ist diese Differenz schon groß genug, um allgemein zu behaupten, dass sich die Beurteilung des Designs zwischen Beginn und Ende der Werbeveranstaltung signifikant verbessert?

Ja, diese Differenz ist groß genug. Man kann generell behaupten, dass sich die Designbeurteilung zwischen Beginn und Ende der Werbeveranstaltung signifikant verbessert. Zu diesem Ergebnis kommt man wie folgt: Zuerst werden Null- und Alternativhypothese formuliert. Als Nullhypothese nimmt man an, dass die Verteilungen der Beurteilungen des Designs am Anfang und am Ende der Veranstaltung gleich sind:

$$H_0: \text{Verteilung}_{\text{Beginn}} = \text{Verteilung}_{\text{Ende}}$$

Die Alternative dazu ist

$$H_1: \text{Verteilung}_{\text{Beginn}} \neq \text{Verteilung}_{\text{Ende}}$$

Die Verteilungen unterscheiden sich zwischen Anfang und Ende der Werbeveranstaltung. Um diese Hypothesen zu prüfen, berechnet man die Maßzahlen T_1 und T_2 mit Hilfe folgender Tabelle

(Vor_Veranst)	(Nach_Veranst)	Differenz	Rang	Rang("-")	Rang("+")
3	3	0	2	_	2
3	2	1	8	_	8
1	1	0	2	_	2
4	2	2	14	_	14
3	2	1	8	_	8
2	3	−1	8	8	_
4	2	2	14	_	14
4	3	1	8	_	8
5	4	1	8	_	8
4	5	−1	8	8	_
2	1	1	8	_	8
2	3	−1	8	8	_
3	2	1	8	_	8
1	1	0	2	_	2
3	1	2	14	_	14
"*"	"*"	"*"	SUMME	24	96

Die Rangsumme der negativen Differenzen T_1 ist 24 und die Rangsumme der positiven Differenzen T_2 ist 96. Zu diesem Ergebnis kommt man durch das Bilden der Beurteilungsdifferenzen zwischen Anfang und Ende der Werbeveranstaltung. Der erste Befragte beurteilte das Design vor und nach der Veranstaltung mit befriedigend. Die Differenz ist daher 0. Diesen 15 Differenzen (3. Spalte) werden Rangplätze zugeordnet. Die kleinste Differenz ist 0. Da diese Differenz 3 Mal vorkommt, wird ihr der Durchschnitt aus den Rangzahlen 1, 2 und 3 zugeordnet. Daher erhält sie die Rangzahl 2. Wenn die Differenz negativ ist, wird sie den negativen Rängen zugeordnet (Spalte 5), sonst der Spalte 6. Die Summe der negativen Ränge

T_1 ist 24 und die der positiven 96.

Die Testmaßzahl T ist normalverteilt mit dem Durchschnitt

$$\mu_T = \frac{n \cdot (n + 1)}{4}$$

und der Standardabweichung

$$\sigma_T = \sqrt{\frac{n \cdot (n + 1) \cdot (2 \cdot n + 1) - \sum_{i=1}^{k} \frac{(t_i)^3 - t_i}{2}}{24}}$$

n ist die Anzahl der Differenzen und die Summe der t_i-Werte ist die Korrektur für Durchschnittsränge, also Ränge, die mehrfach vorkommen. Für das Beispiel ist

$$\mu_T = \frac{15 \cdot (15 + 1)}{4} = 60$$

und

$$\sigma_T = \sqrt{\frac{15 \cdot (15 + 1) \cdot (2 \cdot 15 + 1) - 384}{24}} = 17.146$$

Die Korrekturen für die Durchschnittsränge können aus folgender Tabelle entnommen werden:

(Rangzahlen)	f	f^3	$f^3 - f$
1 , 2 , 3	3	27	24
4 , 5 , 6 , 7 , 8 , 9 , 10 , 11 , 12	9	729	720
13 , 14 , 15	3	27	24
SUMME	"*"	"*"	768

In der ersten Spalte stehen die gleichen Rangzahlen, in der zweiten ihre Häufigkeit, in der dritten die Häufigkeit zur dritten Potenz und in der vierten Spalte wird davon die Häufigkeit wieder abgezogen. Die Summe der vierten Spalte ergibt 768 und der Korrekturfaktor errechnet sich wie folgt:

$$K = \frac{768}{2} = 384$$

Die Testmaßzahl z_{beob} ist

$$z_{beob} = \frac{T_1 - \mu_T}{\sigma_T} = \frac{24 - 60}{17.146} = -2.1$$

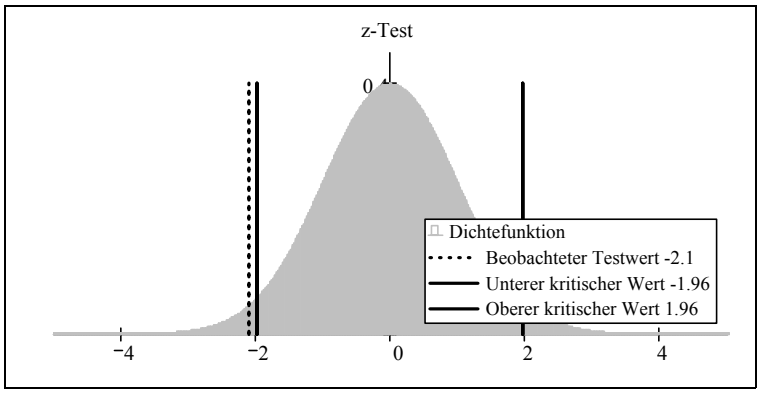
Die Wahrscheinlichkeit für das Auftreten einer so großen Testmaßzahl errechnet sich mit Hilfe der Normalverteilung auf

$$F\left(z_{beob}\right) = F_N(-2.1) = 0.018$$

Da diese Wahrscheinlichkeit kleiner ist als 0.05/2 = 0.025, kann die Nullhypothese auf einem 5% Signifikanzniveau abgelehnt werden.

Zum gleichen Ergebnis kommt man, wenn man den beobachteten z-Wert mit dem kritischen z-Wert vergleicht. Für ein 5%-iges Signifikanzniveau sind die kritischen z-Wert für eine zweiseitige Alternativhypothese $z_c = -1.96$ und $z_c = 1.96$. Da der beobachtete z-Wert $z_{beob} = -2.1$ kleiner ist als der untere kritische z-Wert −1.96, kann die Nullhypothese abgelehnt werden.

In unten stehender Grafik ist dieses Testergebnis veranschaulicht. Die Testmaßzahl z_{beob} (= −2.1) liegt links vor dem linken Schwanzende (= −1.96):



Beobachteter z-Wert	Kritische Werte	Wahrscheinlichk. p
−2.1	(−1.960 1.960)	0.018

Ohne persönlichen Rechenaufwand kommt man zu diesem Ergebnis mit Hilfe der Autorensoftware. Den Output für dieses Beispiel findet man auf Seite 193.

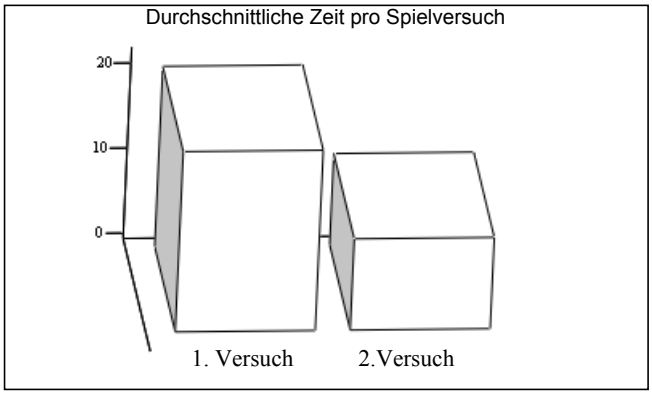
3_3 t-Test

Kann man allgemein voraussetzen, dass man für das neue Konsolenspiel "Fantasy" beim dritten Versuch im Schnitt weniger Zeit benötigt als beim ersten Versuch? Bei einer Werbeveranstaltung für die neue Spielkonsole Yoki wurden 15 Besucher getestet, wie viel Zeit sie beim 1. und 3. Spielversuch benötigten:

(Versuch_1)	21	19	8	34	21	12	34	28	35	32	12	11	22	9	20
(Versuch_3)	4	5	10	14	12	8	12	14	31	12	7	4	10	4	14

Nach drei Zeitintervallen ausgezählt erhält man folgendes Ergebnis:

Untergrenze	bis	Obergrenze	(Versuch_1)	(Versuch_3)
4	bis_unter	12	3	8
12	bis_unter	20	3	6
20	bis_unter	36	9	1
"*"	"*"	SUMME	15	15



Im Schnitt benötigten die 15 Besucher 21.1 Sekunden beim ersten Versuch und nur mehr 10.7 Sekunden beim 3. Versuch. Ist diese Zeitdifferenz schon groß genug, um allgemein zu behaupten, dass sich die Spielzeit zwischen 1. und 3. Versuch signifikant verringert?

Ja, diese Zeitdifferenz ist groß genug. Man kann generell behaupten, dass sich die Spielzeit zwischen 1. und 3. Spielversuch signifikant reduziert. Zu diesem Ergebnis kommt man wie folgt: Zuerst werden Null- und Alternativhypothese formuliert.

$$H_0: \mu_{1.\ Versuch} = \mu_{3.\ Versuch}$$

oder

$$H_0: \mu_{1.\ Versuch} - \mu_{3.\ Versuch} = \delta = 0$$

Die Durchschnittszeit beim 1. Versuch ist gleich der beim 3. Versuch und

$$H_1: \mu_{1.\ Versuch} > \mu_{3.\ Versuch}$$

oder

$$H_1: \mu_{1.\ Versuch} - \mu_{3.\ Versuch} = \delta > 0$$

Die Durchschnittszeit beim 1. Versuch ist länger als beim dritten. Um diese Hypothesen zu prüfen, berechnet man die Testmaßzahl

$$t_{beob} = \frac{xquer_{Differenz}}{\frac{s_{Differenz}}{\sqrt{n}}}$$

mit

$$s_{Differenz} = \sqrt{\frac{\sum_{i=1}^{n} \left(d_i - xquer_{Differenz}\right)^2}{n-1}}$$

und

$$d_i = x_{i,1.\ Versuch} - x_{i,\ 3.\ Versuch}.$$

d_i ist also die jeweilige Differenzzeit zwischen 1. und 3. Versuch der i-ten Person $xquer_{Differenz}$. ist der Durchschnitt aus allen Differenzen und $s_{Differenz}$ die Standardabweichung daraus.

Für das Beispiel sind die Differenzen zwischen 1. und 3. Versuch

(Versuch_1)	(Versuch_3)	Differenz
21	4	17
19	5	14
8	10	–2
34	14	20
21	12	9
12	8	4
34	12	22
28	14	14
35	31	4
32	12	20
12	7	5
11	4	7
22	10	12
9	4	5
20	14	6

Der Mittelwert aus diesen 15 Differenzen ist

$$\text{xquer}_{\text{Differenz}} = 10.467$$

und die Standardabweichung daraus

$$s_{\text{Differenz}} = \sqrt{\frac{\sum_{i=1}^{n} \left(d_i - 10.467\right)^2}{15 - 1}} = 6.917$$

Die Testmaßzahl ist

$$t_{\text{beob}} = \frac{10.467}{\frac{6.917}{\sqrt{15}}} = 5.861$$

Die Testmaßzahl t_{beob} ist studentverteilt mit

$$\nu = n - 1 = 15 + 1 = 14$$

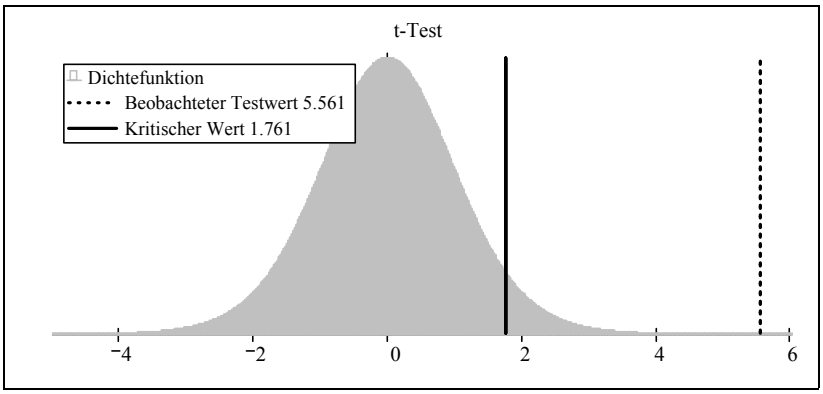
Freiheitsgraden.

Die Wahrscheinlichkeit für das Auftreten einer so großen Testmaßzahl errechnet sich mit

$$F\left(t_{beob}, \nu\right) = F(5.861, 14) = 0.999$$

Da diese Wahrscheinlichkeit größer ist als 1 – 0.05 = 0.95 kann die Nullhypothese abgelehnt werden. Zum gleichen Ergebnis kommt man, wenn man den beobachteten t-Wert mit dem kritischen t-Wert vergleicht. Für ein 5%-iges Signifikanzniveau und 14 Freiheitsgrade ist der kritische t-Wert für eine linksseitige Alternativhypothese $t_c = 1.761$. Da der beobachtete t-Wert $t_{beob} = 5.561$ größer ist als der kritische t-Wert, kann die Nullhypothese abgelehnt werden.

In unten stehender Grafik ist dieses Testergebnis veranschaulicht. Die Testmaßzahl t_{beob} (= 5.561) liegt rechts nach dem rechten Schwanzende (= 1.761):



Beobachteter t-Wert	Kritische Werte	Wahrscheinlichk. p
5.860	(–2.145 2.145)	0.000

Der t-Test setzen voraus, dass die Verteilung der Differenzen in der Grundgesamtheit normalverteilt ist. Diese Voraussetzung kann mit Hilfe des Kolmogorov-Smirnov-Test (siehe Abschnitt 9.10) überprüft werden. Wenn man die Normalverteilung ablehnen muss, dann kann man den t-Test nicht anwenden. In diesem Fall kann man mit Hilfe des Wilcoxon-Tests überprüfen, ob sich die beiden Verteilungen in den Grundgesamtheiten signifikant unterscheiden.

Der Softwareoutput für dieses Beispiel findet man auf Seite 195. Das Programm liefert nicht nur eine Maßzahl wie den p-Wert, sondern eine wörtliche Interpretation des Testergebnisses. Dadurch werden Fehlinterpretationen z. B. des p-Wertes verhindert.

4_0 ZUSAMMENHANG ZWISCHEN ZWEI STICHPROBEN

Unterschied - Zusammenhang

Wenn man 10 Männer und 10 Frauen aus der Grundgesamtheit der potentiellen Kunden für die Spielkonsole Yoki zufällig auswählt und sie nach ihrer Kaufabsicht befragt, dann wird sich als Ergebnis vermutlich herausstellen, dass sich der Anteil der Männer mit Kaufabsicht vom Anteil der Frauen mit Kaufabsicht unterscheidet. Mit Hilfe geeigneter Tests wird man prüfen, ob diese Unterschiede in den Kaufabsichtsanteilen generell für alle Männer und Frauen gelten.

Dies gilt auch für ordinal oder metrisch skalierten Antworten. Fragt man z.B. die 10 Männer und 10 Frauen nach ihren monatlichen Einkommen, dann wird sich auch beim Durchschnittseinkommen zeigen, dass vermutlich Unterschiede zwischen dem monatlichen Durchschnittseinkommen von Männern und Frauen existieren. Mit geeigneten Tests wird geprüft, ob diese Unterschiede im Durchschnittseinkommen allgemein für alle potentiellen männlichen und weiblichen Käufer gelten. Die beiden Stichproben sind in diesem Beispiel unabhängig.

Sind die beiden Stichproben nicht unabhängig sondern abhängig, dann kann die Unterschiedshypothese ebenfalls mit geeigneten Verfahren geprüft werden. Befragt man z. B. nicht 10 Männer und 10 Frauen, sondern 10 Personen nach ihrem monatlichen Einkommen vor 5 Jahren und heute, dann sind die beiden Stichproben "Personen vor 5 Jahren" und "Personen heute" abhängig. Mit geeigneten Testverfahren kann auch hier die Frage nach den Unterschieden im Durchschnittseinkommen von diesen beiden abhängigen Stichproben geklärt werden. Man kann aber auch nach der "Abhängigkeit" oder den "Zusammenhang" des Einkommens von Ehemännern und Ehefrauen fragen. Ist z. B. zu erwarten, dass Ehemänner mit hohem Einkommen auch Ehefrauen haben mit hohem Einkommen? Oder ist es umgekehrt: Ehemänner mit hohem Einkommen haben Ehefrauen mit niedrigen Einkommen?

Die Frage nach dem Zusammenhang stellt sich nicht nur bei metrisch skalierten Antworten. Auch bei ordinalen oder nominalen Stichproben ist die Frage nach dem Zusammenhang gerechtfertigt, wenn es sich um zwei abhängige Stichproben handelt. Kann man z. B. davon ausgehen, dass Ehefrauen die gleiche oder ähnliche Kaufabsicht zeigen wie ihre Ehemänner? Besteht in der Beurteilung des Designs der Spielkonole ein Zusammenhang zwischen Ehefrau und Ehemann?

Die Stärke des Zusammenhangs wird durch Korrelationskoeffizienten ausgedrückt. Diese Koeffizienten nehmen Werte zwischen 0 und 1 an. Besteht kein Zusammenhang, dann ist der Korrelationskoeffizient 0, besteht vollständiger Zusammenhang, dann nimmt der Korrelationskoeffizient den Wert 1 an. Neben der Stärke des Zusammenhangs kann bei ordinal und metrisch skalierten Antworten noch die Richtung des Zusammenhangs durch das Vorzeichen ausgedrückt werden. Ein positives Vorzeichen bedeutet, dass die Antworten gleichläufig sind. Je mehr Einkommen der Ehemann hat, umso mehr Einkommen hat auch die Ehefrau. Bei einem negativen Vorzeichen sind die Antworten gegenläufig: Je mehr Einkommen der Ehemann hat, umso weniger Einkommen hat die Ehefrau.

Mit geeigneten Testverfahren kann geprüft werden, ob die Zusammenhänge, die sich zwischen zwei Stichproben zeigen, auch allgemein gelten. Besteht z. B. zwischen dem Einkommen von Ehemännern und Ehefrauen in einer Stichprobe von 30 Ehepaaren ein positiver Zusammenhang von 0.7, dann kann geprüft werden, ob dieser Zusammenhang auch für alle Ehemänner und Ehefrauen gilt, aus denen sich die Grundgesamtheit zusammensetzt. Im Falle eines signifikanten Zusammenhanges kann man diesen bei metrisch skalierten Antworten mit Hilfe einer so genannten Regressionsfunktion zum Ausdruck bringen. Mit Hilfe dieser Funktion kann man z.B. vom gegebenen Einkommen des Ehemannes auf das seiner Ehefrau schließen.

Vorschau

Für nominal skalierte Antworten berechnet man den so genannten Kontingenzkoeffizienten, um einen eventuellen Zusammenhang zwischen zwei Stichproben auszudrücken. Da bei nominalen Antworten keine Ordnungsrelation existiert, gibt der Kontingenzkoeffizient nur über die Stärke nicht aber über die Richtung des Zusammenhanges Auskunft. Untersucht man z. B. den Zusammenhang zwischen dem gewählten Beruf und dem Geschlecht einer bestimmten Bevölkerungsgruppe, so ist zwar die Angabe der Stärke des Zusammenhangs sinnvoll nicht aber eine Richtung. Der Kontingenzkoeffizient als Maß für den Zusammenhang zwischen nominal skalierten Antworten wird im Abschnitt 4_1 beschrieben.

Im Abschnitt 4_2 wird gezeigt, wie man den Zusammenhang ordinal skalierter Fragen mit Hilfe des Rangkorrelationskoeffizienten ausdrückt und wie man prüft, ob der in den Stichproben festgestellte Zusammenhang auch für die Grundgesamtheiten gilt. Als Beispiel wird untersucht, ob zwischen der Beurteilung der Werbeveranstaltung und des Designs der Spielkonsole Yoki ein Zusammenhang besteht.

Als Beispiel für metrisch skalierte Stichproben werden Besucher einer Werbeveranstaltung verwendet, die einerseits nach ihrem Alter gefragt wurden, andererseits nach ihren monatlichen Einkommen. Die Berechnung des entsprechenden Maßkorrelationskoeffizienten wird im Abschnitt 4_3 beschrieben. Im folgenden Abschnitt 4_4 wird gezeigt, wie man den festgestellten Zusammenhang zwischen Alter und Einkommen in Form einer linearen Regressionsfunktion ausdrückt, um vom Alter auf das entsprechende Einkommen einer Person zu schließen.

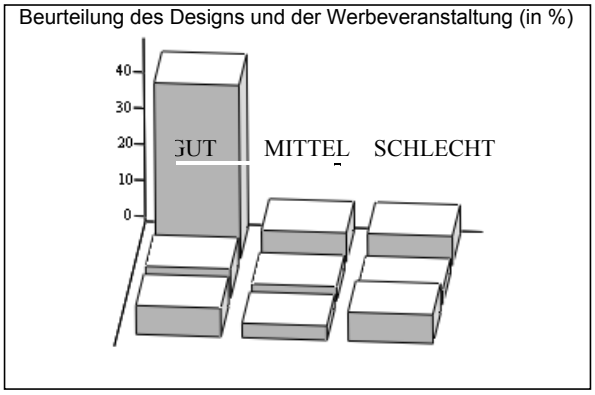
4_1 Kontingenzkoeffizient

Besteht zwischen der Beurteilung des Designs der neuen Spielkonsole und der Werbeveranstaltung ein Zusammenhang? Eine Befragung von 25 Besuchern einer Werbeveranstaltung für die Spielkonsole brachte folgendes Ergebnis:

(Designbeurteilung)	"gut"	"mittel"	"schlecht"
(Veranstaltungsbeurteilung)	absolut	absolut	absolut
"gut"	12	2	2
"mittel"	2	1	1
"schlecht"	2	1	2
SUMME	16	4	5

Wenn man diese Ergebnisse auf die 25 Befragten bezieht, dann zeigt sich, dass 48% sowohl das Design als auch die Werbeveranstaltung gut beurteilen. Nur 8% beurteilen sowohl das Design als auch die Werbeveranstaltung schlecht. Die Tabelle zeigt diese Prozentsätze:

Designbeurteilung / Veranstaltungsbeurteilung	"gut"	"mittel"	"schlecht"	ZEILENSUMME
"gut"	48	8	8	64
"mittel"	8	4	4	16
"schlecht"	8	4	8	20
SPALTENSUMME	64	16	20	100



Um den Zusammenhang zwischen der Beurteilung des Designs und der Werbeveranstaltung auszudrücken, kann man den so genannten Kontingenzkoeffizienten r_C berechnen. Er kann Werte zwischen 0 und 1 annehmen, wobei ein Wert von 0 keinen Zusammenhang bedeutet und 1 vollständiger Zusammenhang. Im vorliegenden Beispiel ist der Kontingenzkoeffizient r_C gleich 0.39, d. h. zwischen der Beurteilung des Designs und der Werbeveranstaltung besteht bei den 25 Befragten zwar ein Zusammenhang. Dieser ist jedoch nicht sehr stark.

Kann man auf Grund dieses Stichprobenergebnisses allgemein behaupten, dass zwischen der Beurteilung des Designs der Spielkonsole Yoki und der Werbeveranstaltung ein Zusammenhang besteht?

Nein, der Zusammenhang zwischen den Beurteilungen von Design und Werbeveranstaltung ist nicht groß genug, um allgemein zu behaupten, dass ein Zusammenhang besteht. Es könnte zwar tatsächlich sein, dass zwischen der Beurteilung von Design und Werbeveranstaltung in der Grundgesamtheit aller potentiellen Kunden ein Zusammenhang besteht. Die vorliegende Stichprobe liefert dafür aber keinen ausreichenden Hinweis. Dies ist das Ergebnis der Berechnungen des Computerprogrammes, dessen Output man auf Seite 196 findet.

Wie kommt man zu diesem Ergebnis? Zuerst formuliert man die Null- und Alternativhypothese. Die Nullhypothese besagt, dass zwischen der Beurteilung von Design und Werbeveranstaltung kein (= 0) Zusammenhang besteht. Diese Hypothese wird formal wie folgt angeschrieben:

$$H_0: \rho_C = 0.$$

Eine Alternative zu dieser Nullhypothese ist die Vermutung, dass zwischen der Beurteilung von Design und Werbeveranstaltung in der Grundgesamtheit ein Zusammenhang besteht. Formal

$$H_1: \rho_C > 0.$$

An Hand der Ergebnisse der Stichprobe entscheidet man sich für die Annahme oder Ablehnung der Null- oder Alternativhypothese. Dazu berechnet man folgende Testmaßzahl:

$$\chi 2_{beob} = n \cdot \left[\left[\sum_{j=1}^{m} \sum_{k=1}^{t} \frac{(h_{jk})^2}{h_j \cdot h_k} \right] - 1 \right]$$

n ist der Umfang der gesamten Stichprobe, h_j und h_k sind die Randhäufigkeiten der Zeilen und Spalten sowie h_{jk} die Häufigkeiten der Antwortkombinationen. In einer Kreuztabelle sind die Häufigkeiten der Antworten und Antwortkombinationen wie folgt angeordnet:

$$\begin{bmatrix} \begin{pmatrix}\text{Antwort\_B}\\ \text{Antwort\_A}\end{pmatrix} & B_1 & .. & B_j & .. & B_m & \begin{pmatrix}\text{Zeilen}\\ \text{summe}\end{pmatrix} \\ A_1 & h_{11} & .. & h_{1j} & .. & h_{1m} & \sum_{j=1}^{m} h_{1j} = h_{1.} \\ ..... & ..... & .. & ..... & .. & ..... & ..... \\ A_i & h_{i1} & .. & h_{ij} & .. & h_{im} & \sum_{j=1}^{m} h_{ij} = h_{i.} \\ ..... & ..... & . & ..... & .. & ..... & ..... \\ A_k & h_{k1} & .. & h_{kj} & .. & h_{km} & \sum_{j=1}^{m} h_{kj} = h_{k.} \\ \begin{pmatrix}\text{Spalten}\\ \text{summe}\end{pmatrix} & \sum_{i=1}^{k} h_{i1} = h_{.1} & .. & \sum_{i=1}^{k} h_{ij} = h_{.j} & .. & \sum_{i=1}^{k} h_{im} = h_{.m} & \sum_{i=1}^{k}\sum_{j=1}^{m} h_{ij} = n \end{bmatrix}$$

In folgender Kreuztabelle sind neben den Häufigkeiten der Beurteilungskombinationen auch die Zeilen und Spaltensummen des Beispiels nochmals angeführt:

Designbeurteilung / Veranstaltungsbeurteilung	"gut"	"mittel"	"schlecht"	ZEILENSUMME
"gut"	12	2	2	16
"mittel"	2	1	1	4
"schlecht"	2	1	2	5
SPALTENSUMME	16	4	5	25

Nun kann man die Werte in die Formel für die Chiquadratmaßzahl einsetzen und berechnen:

$$\chi 2_{beob} = 25 \cdot \left(\frac{12^2}{16 \cdot 16} + \frac{2^2}{16 \cdot 4} + \frac{2^2}{16 \cdot 5} + \frac{2^2}{4 \cdot 16} + \frac{1^2}{4 \cdot 4} + \frac{1^2}{4 \cdot 5} + \frac{2^2}{5 \cdot 16} + \frac{1^2}{5 \cdot 4} + \frac{2^2}{5 \cdot 5} - 1 \right) = 2.75$$

Der Kontingenzkoeffizient wird mit Hilfe folgender Formel bestimmt:

$$r_c = \sqrt{\frac{\chi^2 \cdot k}{\left(\chi^2 + n\right) \cdot (k-1)}}$$

Mit

$$k = \min(1, m)$$

l ist die Zahl der Antwortmöglichkeiten der ersten Verteilung und m die entsprechende Zahl der zweiten Verteilung. Für das Beispiel ist der Kontingenzkoeffizient

$$r_c = \sqrt{\frac{2.75 \cdot 3}{(2.75 + 25) \cdot (3 - 1)}} = 0.386$$

Die $\chi 2$-Testmaßzahl folgt einer Chiquadratverteilung mit

$$\nu = (l - 1) \cdot (m - 1)$$

Freiheitsgraden. Im obigen Beispiel sind die Freiheitsgrade gleich

$$\nu = (3 - 1) \cdot (3 - 1) = 4$$

und die Wahrscheinlichkeit für das Auftreten einer so großen Testmaßzahl von $\chi 2_{beob} = 2.75$ unter Gültigkeit der Nullhypothese gleich

$$F_{\chi 2}(2.75, 4) = 0.40.$$

(Diesen Wert liest man aus einer geeigneten Tabelle der Chiquadratverteilung ab oder berechnet ihn mit Hilfe entsprechender Programme.)

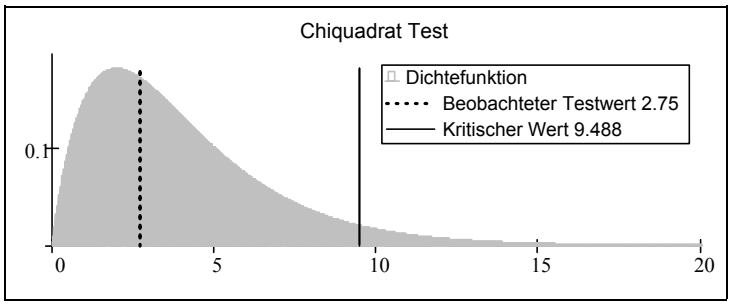
Ist man bereit, in 5 von 100 Fällen eine richtige Nullhypothese abzulehnen, dann ist das Signifikanzniveau α gleich 0.05. Unter Gültigkeit der Nullhypothese sind die 5% seltensten $\chi 2$ Werte nach unten beschränkt durch den kritischen Wert $\chi 2_c$ gleich 9.488. Da das aus der Stichprobe errechnete $\chi 2_{beob}$ von 2.75 kleiner ist als dieser kritische Wert, kann die Nullhypothese nicht abgelehnt werden.

$$\chi 2_{beob} = 2.75 < 9.488 = \chi 2_c$$

Zum selben Ergebnis kommt man, wenn man die Wahrscheinlichkeiten vergleicht: Da die Wahrscheinlichkeit für einen $\chi 2$ Wert von 2.75 oder größer, gleich $1 - F_{\chi 2}(2.75, 4) = 0.60$ ist, und das Signifikanzniveau mit $\alpha = 0.05$ angenommen wurde, kann die Nullhypothese nicht abgelehnt werden.

$$1 - F_{\chi 2}(2.75, 4) = 0.60 > 0.05 = \text{Signifikanzniveau}$$

Man kann auf Grund dieser Stichprobe nicht behaupten, dass zwischen der Beurteilung des Designs und der Werbeveranstaltung allgemein ein Zusammenhang besteht.



(Chiquadrat_beobachtet)	(Chiquadrat_kritisch)	(p – Wahrscheinlichkeit)
2.75	9.488	0.6

Voraussetzungen für den Chiquadrattest

Der Chiquadrattest ist nur dann zulässig, wenn keine Antwortkombination 0-Mal vorkommt (also keine Zelle der Kreuztabelle den Wert 0 hat) und wenn höchstens 20% aller Antwortkombinationen eine Häufigkeit kleiner gleich 5 aufweisen (höchstens 20% aller Zellen in der Kreuztabelle mit einer Häufigkeit $\leq$ 5). Ist dies nicht der Fall, dann sind die Schlussfolgerungen aus dem Chiquadrattest ungültig. Durch Zusammenfassen der Antwortmöglichkeiten oder durch Erhöhung des Stichprobenumfanges kann eventuell diese Voraussetzung des Chiquadrattests erfüllt werden. Ist dies nicht möglich, dann ist der exakte Test von Fisher zu verwenden.

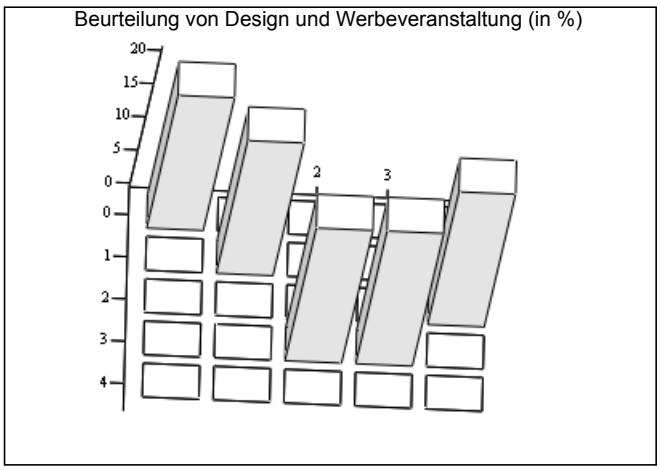
Im oben angeführten Beispiel sind zwar alle Häufigkeiten größer 0 aber mehr als 20% aller Zellen sind kleiner als 5. Die Voraussetzungen des Chiquadrattests sind daher nicht erfüllt. Wenn man daher anstelle des Chiquadrattests den exakten Test von Fisher auf diese Daten anwendet, dann kommt man jedoch zum selben Ergebnis.

4_2 Rangkorrelation

Besteht zwischen der Beurteilung des Designs der neuen Spielkonsole und der Werbeveranstaltung ein Zusammenhang? Eine Befragung von 5 Besuchern einer Werbeveranstaltung für die Spielkonsole brachte folgendes Ergebnis:

(Besucher)	Design	Werbeveranstaltung
1	gut	gut
2	genügend	schlecht
3	befriedigend	sehr_schlecht
4	genügend	mittelmäßig
5	sehr_gut	sehr_gut

Wenn man diese Ergebnisse auf die 5 Befragten bezieht, dann zeigt sich, dass 20% sowohl das Design als auch die Werbeveranstaltung sehr gut beurteilen. 20% beurteilen sowohl das Design als auch die Werbeveranstaltung schlecht bzw. nicht genügend. Die Grafik zeigt diese Prozente:



("sehr gut")	"gut"	"mittelmäßig"	"schlecht"	("sehr schlecht")
1	2	3	4	5

Um den Zusammenhang zwischen der Beurteilung des Designs und der Werbeveranstaltung auszudrücken, kann man den so genannten Rangkorrelationskoeffizienten berechnen. Er kann Werte zwischen −1 und +1 annehmen, wobei ein Wert von 0 keinen Zusammenhang bedeutet und −1 einen negativen und +1 einen positiven Zusammenhang. Im vorliegenden Beispiel ist der Rangkorrelationskoeffizient r_S gleich 0.667, d. h. zwischen der Beurteilung des Designs und der Werbeveranstaltung besteht bei den 5 Befragten ein positiver Zusammenhang. Je besser ein Besucher das Design beurteilt, desto besser beurteilt er

auch die Werbeveranstaltung. (Wäre der Rangkorrelationskoeffizient negativ, dann wäre die Interpretation: Je besser ein Besucher das Design beurteilt, umso schlechter beurteilt er die Werbeveranstaltung).

Kann man auf Grund dieses Stichprobenergebnisses allgemein behaupten, dass zwischen der Beurteilung des Designs der Spielkonsole Yoki und der Werbeveranstaltung ein positiver Zusammenhang besteht?

Nein, der Zusammenhang zwischen den Beurteilungen von Design und Werbeveranstaltung ist nicht groß genug, um allgemein zu behaupten, dass ein Zusammenhang vorliegt. Es könnte zwar tatsächlich sein, dass zwischen der Beurteilung von Design und Werbeveranstaltung in der Grundgesamtheit aller potentiellen Kunden ein Zusammenhang besteht. Die Stichprobe liefert dafür aber keine hinreichende Begründung.

Wie kommt man zu diesem Ergebnis? Zuerst formuliert man die Null- und Alternativhypothese. Die Nullhypothese besagt, dass zwischen der Beurteilung von Design und Werbeveranstaltung kein (= 0) Zusammenhang besteht. Diese Hypothese wird formal wie folgt angeschrieben:

$$H_0: \rho_S = 0.$$

Eine Alternative zu dieser Nullhypothese ist die Vermutung, dass zwischen der Beurteilung von Design und Werbeveranstaltung in der Grundgesamtheit ein positiver Zusammenhang besteht. Formal

$$H_1: \rho_S > 0.$$

An Hand der Ergebnisse der Stichprobe entscheidet man sich für die Annahme oder Ablehnung der Null- oder Alternativhypothese. Dazu berechnet man folgende Testmaßzahl:

$$t_{beob} = r_S \cdot \sqrt{\frac{n-2}{1-r_S^2}}$$

n ist der Umfang der gesamten Stichprobe und r_S ist der Rangkorrelationskoeffizient der Stichprobe. Diese Testmaßzahl folgt einer Studentverteilung mit

$$\nu = n - 2$$

Freiheitsgraden. In folgender Tabelle sind die Zwischenschritte zur Berechnung des Rangkorrelationskoeffizienten angegeben. Zuerst werden die Beurteilungen in Rangzahlen transformiert:

(Besucher)	Design	$R(x_i)$	Werbeveranstaltung	$R(y_i)$
1	gut	2	gut	2
2	genügend	4.5	schlecht	4
3	befriedigend	3	sehr_schlecht	5
4	genügend	4.5	mittelmäßig	3
5	sehr_gut	1	sehr_gut	1

Die Rangzahlen für die Beurteilung des Designs (= x) erhält man durch die Zuordnung der Rangzahlen zu den Beurteilungen. Der Besucher 5 beurteilte das Design mit "sehr gut" und bekommt die Rangzahl 1 (= $R(x_5)$), der Besucher 1 beurteilte das Design mit "gut". Daher bekommt er die Rangzahl 2 (= $R(x_1)$). Mit befriedigend“ beurteilte der Besucher 3 das Design und bekommt daher die Rangzahl 3. Die Besucher 2 und 4 gaben die gleiche Beurteilung nämlich "genügend". Sie erhalten daher den Durchschnitt aus den letzten Rangzahlen 4 und 5, nämlich 4.5 zugeordnet. Auf die gleiche Art bestimmt man die Rangzahlen $R(y_i)$ für die Beurteilungen der Werbeveranstaltung (= y). Nun kann der Rangkorrelationskoeffizient berechnet werden. In folgender Tabelle sind die Zwischenschritte angegeben:

(Besucher)	$R(x_i)$	$R(y_i)$	$R(x_i)^2$	$R(y_i)^2$	$R(x_i)\cdot R(y_i)$
1	2	2	4	4	4
2	4.5	4	20.25	16	18
3	3	5	9	25	15
4	4.5	3	20.25	9	13.5
5	1	1	1	1	1
SUMMEN	15	15	54.5	55	51.5

Den Rangkorrelationskoeffizienten kann man nach folgender einfachen Formel berechnen:

$$r_S = \frac{\sum_{i=1}^{n} \left(R(x_i)\cdot R(y_i) - n\cdot R_{xquer}\cdot R_{yquer}\right)}{\sqrt{\left[\sum_{i=1}^{n} \left[R(x_i)^2 - n\cdot(R_{xquer})^2\right]\right]\cdot\left[\sum_{i=1}^{n} \left[R(y_i)^2 - n\cdot(R_{yquer})^2\right]\right]}}.$$

Wenn man die Zahlen des Beispiels einsetzt, erhält man folgendes Ergebnis:

$$r_S = \frac{51.5 - 5\cdot\frac{15}{5}\cdot\frac{15}{5}}{\sqrt{\left[54.50 - 5\cdot\left(\frac{15}{5}\right)^2\right]\cdot\left[55 - 5\cdot\left(\frac{15}{5}\right)^2\right]}} = 0.667$$

Für die Testmaßzahl erhält man

$$t_{beob} = 0.667 \cdot \sqrt{\frac{5-2}{1-0.667^2}} = 1.551$$

mit 5 – 2 = 3 Freiheitsgraden. Die Wahrscheinlichkeit für das Auftreten einer so großen Testmaßzahl von $t_{beob} = 1.551$ unter Gültigkeit der Nullhypothese ist gleich $F_t(1.551, 3) = 0.891$. (Diesen Wert liest man aus einer geeigneten Tabelle der t-Verteilung ab oder berechnet ihn mit Hilfe entsprechender Programme.)

Ist man bereit, in 5 von 100 Fällen eine richtige Nullhypothese abzulehnen, dann ist das Signifikanzniveau α gleich 0.05. Unter Gültigkeit der Nullhypothese sind die 5% seltensten t Werte nach unten beschränkt durch den kritischen Wert t_c gleich 2.353. Da das aus der Stichprobe errechnete t_{beob} von 1.551 kleiner ist als dieser kritische Wert, kann die Nullhypothese nicht abgelehnt werden.

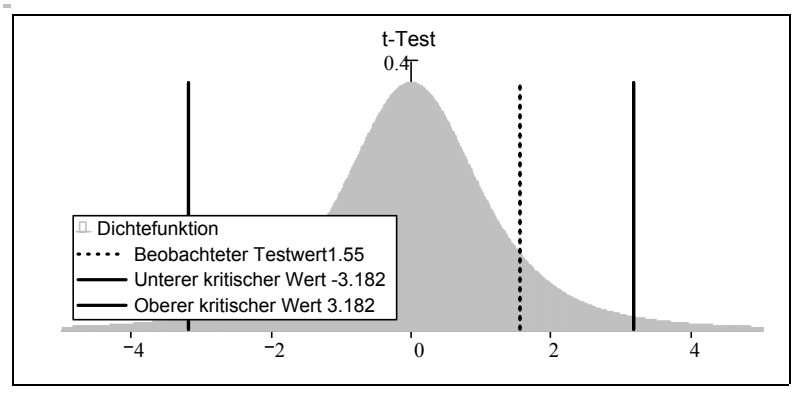
$$t_{beob} = 1.551 < 2.353 = t_c$$

Zum selben Ergebnis kommt man, wenn man die Wahrscheinlichkeiten vergleicht: Da die Wahrscheinlichkeit für einen t-Wert von 1.551 oder größer gleich $1 - F_t(1.551, 3) = 0.109$ ist und das Signifikanzniveau mit 0.05 angenommen wurde, kann die Nullhypothese nicht abgelehnt werden.

$$1 - F_t(1.551, 3) = 0.109 > 0.05 = \text{Signifikanzniveau}$$

Man kann auf Grund dieser Stichprobe nicht behaupten, dass zwischen der Beurteilung des Designs und der Werbeveranstaltung allgemein ein Zusammenhang besteht.

Die Grafik zeigt die Studentverteilung, die kritischen Werte t_u und t_o sowie den beobachteten Wert t_{beob}:



Beobachteter t-Wert	Kritische Werte	Wahrscheinlichk. p
1.55	(−3.182 3.182)	0.109

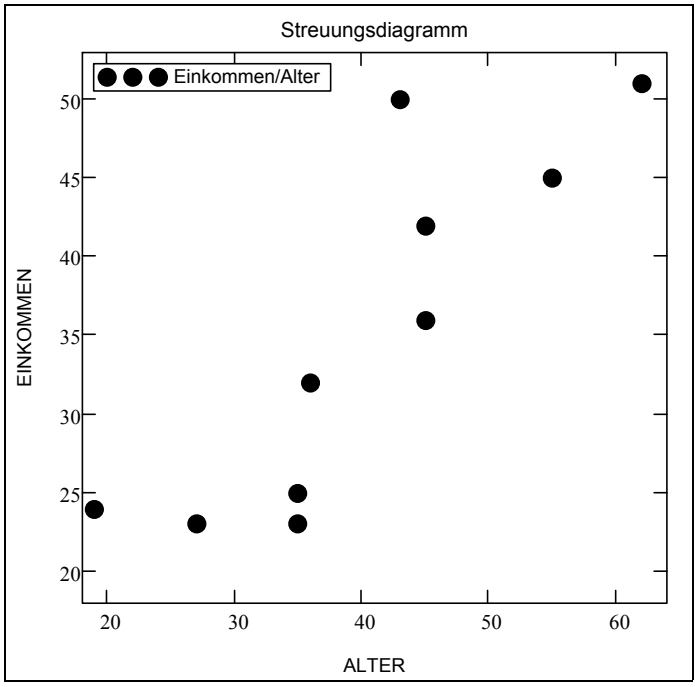
Der Output des Softwareprogrammes für dieses Beispiel steht auf Seite 197 und erfordert nur die Dateneingabe und die Wahl eines Signifikanzniveaus und eines ein- oder zweiseitigen Tests.

4_3 Maßkorrelation

Besteht zwischen dem Alter und dem Einkommen der potentiellen Kunden für die neue Spielkonsole Yoki ein Zusammenhang? Eine Befragung von 10 Besuchern einer Werbeveranstaltung für die Spielkonsole brachte folgendes Ergebnis:

36	45	19	62	27	35	45	55	43	35
32	42	24	51	23	25	36	45	50	23
1	2	3	4	5	6	7	8	9	10

In der ersten Zeile steht das Alter der befragten Personen und in der zweiten das jeweilige Monatseinkommen (in 100 €). Wenn man das Alter der Befragten auf der horizontalen Achse aufträgt und das entsprechende Einkommen auf der vertikalen, dann sieht man, wie stark die Wertepaare streuen. Liegen die Punkte eng um eine gedachte Linie, die von links unten nach rechts oben ansteigt, dann besteht ein starker positiver Zusammenhang zwischen Alter und Einkommen: Je älter die befragte Person umso höher das Einkommen. Die Grafik zeigt dies:



Wenn die Punkte eng um eine gedachte Linie liegen, die von links oben nach rechts unten abfällt, dann würde dies einen negativen Zusammenhang zum Ausdruck bringen: Je älter die befragte Person, umso geringer ist sein Einkommen. Verteilen sich die Punkte zufällig im gezeigten Quadranten, dann würde kein Zusammenhang zwischen Alter und Einkommen bestehen.

Um den Zusammenhang zwischen dem Alter und dem Einkommen durch eine Zahl auszudrücken, kann man den so genannten Maßkorrelationskoeffizienten berechnen. Er kann Werte zwischen −1 und +1 annehmen, wobei ein Wert von 0 keinen Zusammenhang bedeutet, −1 einen negativen und +1 einen positiven Zusammenhang. Im vorliegenden Beispiel ist der Maßkorrelationskoeffizient r_M gleich 0.851, d. h. zwischen dem Alter und dem Einkommen des Befragten besteht ein positiver Zusammenhang. Je älter ein Besucher ist, desto mehr monatliches Einkommen bezieht er.

Kann man auf Grund dieses Stichprobenergebnisses allgemein behaupten, dass zwischen dem Alter der potentiellen Kunden und deren Einkommen ein positiver Zusammenhang besteht?

Ja, der Zusammenhang zwischen dem Alter und dem Einkommen ist groß genug, um allgemein zu behaupten, dass ein Zusammenhang vorliegt. Es könnte zwar tatsächlich sein, dass zwischen dem Alter und dem Einkommen in der Grundgesamtheit aller potentiellen Kunden kein Zusammenhang besteht. Das Risiko dafür ist höchstens 5%.

Wie kommt man zu diesem Ergebnis? Zuerst formuliert man die Null- und Alternativhypothese. Die Nullhypothese besagt, dass zwischen dem Alter und dem Einkommen der potentiellen Kunden kein (= 0) Zusammenhang besteht. Diese Hypothese wird formal wie folgt angeschrieben:

$$H_0: \rho_m = 0.$$

Eine Alternative zu dieser Nullhypothese ist die Vermutung, dass zwischen dem Alter und dem Einkommen in der Grundgesamtheit ein positiver Zusammenhang besteht. Formal

$$H_1: \rho_m > 0.$$

An Hand der Ergebnisse der Stichprobe entscheidet man sich für die Annahme oder Ablehnung der Null- oder Alternativhypothese. Dazu berechnet man folgende Testmaßzahl:

$$t_{beob} = r_m \cdot \sqrt{\frac{n-2}{1-r_m^2}}$$

n ist der Umfang der gesamten Stichprobe und r_m ist der Maßkorrelationskoeffizient der Stichprobe. Diese Testmaßzahl folgt einer Studentverteilung mit

$$\nu = n - 2$$

Freiheitsgraden.

r_m ist der Maßkorrelationskoeffizient der Stichprobe und wie folgt definiert:

$$r_m = \frac{\sum_{i=1}^{n} \left(x_i \cdot y_i - n \cdot xquer \cdot yquer\right)}{\sqrt{\left[\sum_{i=1}^{n} \left[\left(x_i\right)^2 - n \cdot (xquer)^2\right]\right] \cdot \left[\sum_{i=1}^{n} \left[\left(y_i\right)^2 - n \cdot (yquer)^2\right]\right]}} = \frac{{s_{xy}}^2}{\sqrt{{s_x}^2 \cdot {s_y}^2}}$$

s_x^2 und s_y^2 sind die Varianzen von Alter und Einkommen, xquer und yquer die entsprechenden Durchschnitte. ${s_{xy}}^2$ ist die Kovarianz zwischen Alter und Einkommen. In folgender Tabelle sind die Zwischenschritte zur Berechnung des Maßkorrelationskoeffizienten angegeben.

(Besucher)	x_i	y_i	$(x_i)^2$	$(y_i)^2$	$x_i \cdot y_i$
1	36	32	1296	1024	1152
2	45	42	2025	1764	1890
3	19	24	361	576	456
4	62	51	3844	2601	3162
5	27	23	729	529	621
6	35	25	1225	625	875
7	45	36	2025	1296	1620
8	55	45	3025	2025	2475
9	43	50	1849	2500	2150
10	35	23	1225	529	805
Summen	402	351	17604	13469	15206

Der Maßkorrelationskoeffizienten kann nun wie folgt berechnen werden:

$$r_m = \frac{15206 - 10 \cdot \frac{402}{10} \cdot \frac{351}{10}}{\sqrt{\left[17604 - 10 \cdot \left(\frac{402}{10}\right)^2\right] \cdot \left[13469 - 10 \cdot \left(\frac{351}{10}\right)^2\right]}} = 0.851$$

Für die Testmaßzahl erhält man

$$t_{beob} = 0.851 \cdot \sqrt{\frac{10 - 2}{1 - 0.851^2}} = 4.581$$

mit 10 – 2 = 8 Freiheitsgraden.

Die Wahrscheinlichkeit für das Auftreten einer so großen Testmaßzahl von $t_{beob} = 4.581$ unter Gültigkeit der Nullhypothese ist gleich $F_t(4.581, 8) = 0.999$. (Diesen Wert liest man aus einer geeigneten Tabelle der t-Verteilung ab oder berechnet ihn mit Hilfe entsprechender Programme.)

Wenn man bereit ist, in 5 von 100 Fällen eine richtige Nullhypothese abzulehnen, dann ist das Signifikanzniveau α gleich 0.05. Unter Gültigkeit der Nullhypothese sind die 5% seltensten t Werte nach unten beschränkt durch den kritischen Wert t_c gleich 1.86. Da das aus der Stichprobe errechnete t_{beob} von 4.583 größer ist als dieser kritische Wert, kann die Nullhypothese abgelehnt werden.

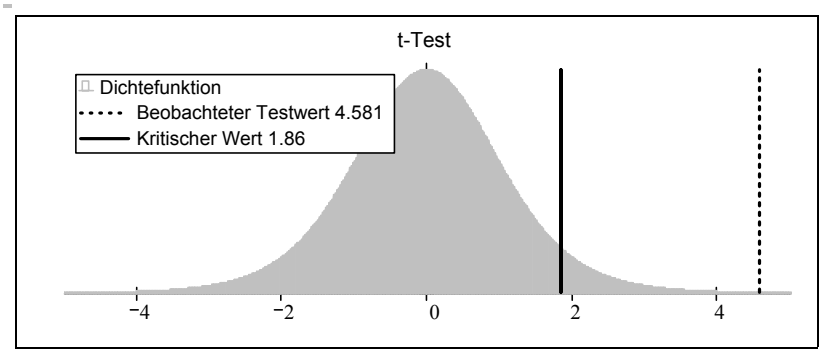
$$t_{beob} = 4.583 > 1.86 = t_c$$

Zum selben Ergebnis kommt man, wenn man die Wahrscheinlichkeiten vergleicht: Da die Wahrscheinlichkeit für einen t-Wert von 4.583 oder größer gleich $1 - F_t(4.581, 8) = 0.001$ ist und das Signifikanzniveau mit 0.05 angenommen wurde, kann die Nullhypothese nicht abgelehnt werden.

$$1 - F_t(4.581, 8) = 0.001 < 0.05 = \text{Signifikanzniveau}$$

Man kann auf Grund dieser Stichprobe behaupten, dass zwischen dem Alter und dem Einkommen aller potentiellen Kunden allgemein ein positiver Zusammenhang besteht.

Die Grafik zeigt die Studentverteilung, die kritischen Wert t_c sowie den beobachteten Wert t_{beob}:



Beobachteter_t_Wert	Kritische_Wert	p_Wahrscheinlichkeit
4.581	$(-\infty \quad 1.86)$	0.001

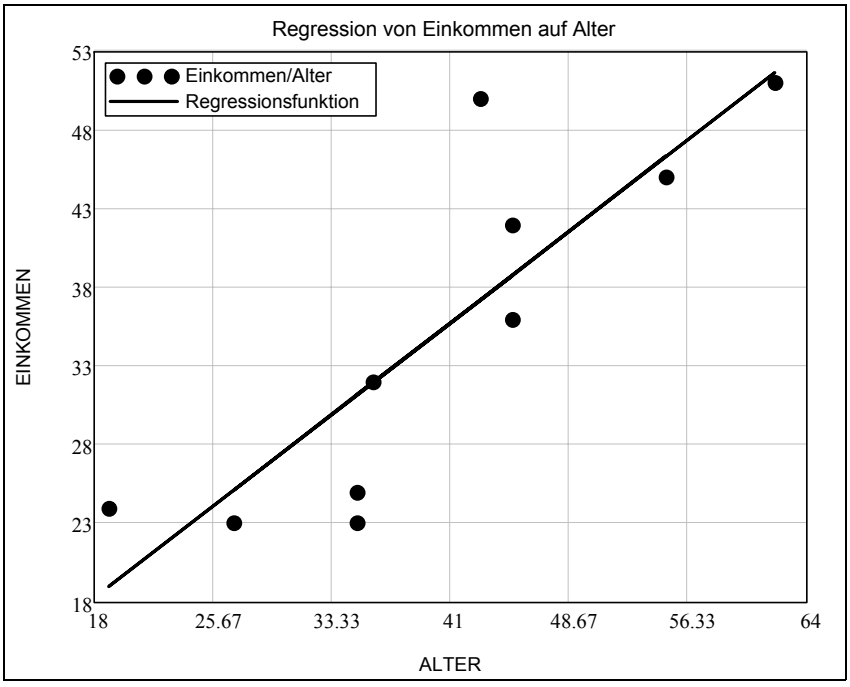
Einfacher kommt man zu dieser Interpretation des Zusammenhanges zwischen Alter und Einkommen mit Hilfe des Softwareprogrammes. Der Output für dieses Beispiel findet sich auf Seite 198.

4_4 Einfach-Regression

Kann man vom Alter der potentiellen Kunden für die neue Spielkonsole Yoki auf ihr monatliches Einkommen schließen? Eine Befragung von 10 Besuchern einer Werbeveranstaltung für die Spielkonsole brachte folgendes Ergebnis:

36	45	19	62	27	35	45	55	43	35
32	42	24	51	23	25	36	45	50	23
1	2	3	4	5	6	7	8	9	10

Das Alter der befragten Personen steht in der ersten Zeile und in der zweiten das jeweilige Monatseinkommen (in 100 €). Wenn man das Alter der Befragten auf der horizontalen Achse aufträgt und das entsprechende Einkommen auf der vertikalen, dann sieht man, wie stark die Wertepaare streuen. Liegen die Punkte eng um eine Gerade (hier blau eingetragen), der so genannten Regressionsgerade, dann ist es möglich, vom Alter der potentiellen Kunden auf ihr monatliches Einkommen zu schließen.



Um das Ergebnis vorwegzunehmen: Die beobachteten Wertepaare liegen eng genug um die Regressionsgerade. Man kann allgemein behaupten, dass die aus den Stichproben errechnete Regressionsfunktion zwischen Alter und Einkommen für alle potentiellen Kunden der Spielkonsole Yoki gilt. Diese lineare Regressionsfunktion ist

$$\text{Einkommen}_i = 4.585 + 0.759 \cdot \text{Alter}_i.$$

Mit dieser Funktion kann man für jedes Alter der potentiellen Kunden einen Schätzwert für das monatliches Einkommen bestimmen. Für einen 30-Jährigen ergibt sich z. B. durch Einsetzen ein monatliches Einkommen von 2735.5 € (= 27.357 · 100)

$$27.357 = 4.585 + 0.759 \cdot 30$$

Ein 30-jähriger Kunde wird sicher nicht genau 2735.5 € verdienen. Die Wahrscheinlichkeit für diese Punktschätzung ist Null. Man kann jedoch mit einem Vertrauen von 95% erwarten, dass sein monatliches Einkommen zwischen 2133.5 € und 3338,0 € liegt.

Der Regressionskoeffizient von 0.759 (der der Steigung der Geraden entspricht) besagt, dass pro Altersjahr das Einkommen um 75.90 € (= 0.759 · 100) zunimmt.

Wie kommt man zu diesem Ergebnis? Zuerst berechnet man aus den Stichprobenwerten die lineare Regressionsfunktion und prüft dann, ob die Beobachtungen eng genug um die Regressionsgerade liegen. Anders ausgedrückt: Man prüft, ob die Streuung der beobachteten Alters- und Einkommensangaben um die Regressionsfunktion klein genug ist, um allgemein zu behaupten, dass zwischen Alter und Einkommen ein linearer Zusammenhang besteht.

Für die lineare Regressionsfunktion der Stichprobe

$$y_i = b_0 + b_1 \cdot x_i$$

werden die beiden Regressionskoeffizienten b_0 und b_1 mit Hilfe folgender Formeln bestimmt:

$$b_1 = \frac{s_{xy}^2}{s_x^2} = r_m \cdot \frac{s_y}{s_x} = \frac{\sum_{i=1}^{n} \left[\left(x_i - \text{xquer}\right) \cdot \left(y_i - \text{yquer}\right)\right]}{\sum_{i=1}^{n} \left(x_i - \text{xquer}\right)^2} = \frac{\sum_{i=1}^{n} \left[\left(x_i - \text{xquer}\right) \cdot y_i\right]}{\sum_{i=1}^{n} \left(x_i - \text{xquer}\right)^2}$$

$$b_0 = \text{yquer} - b_1 \cdot \text{xquer}$$

s_{xy}^2 ist die Kovarianz zwischen x_i und y_i, s_x^2 ist die Varianz von x_i und yquer, xquer sind die beiden Durchschnitte. Folgende Tabelle zeigt die Berechnung für obiges Beispiel mit x = Alter (die so genannte unabhängige Variable) und y = Einkommen (die so genannte abhängige Variable):

x_i	y_i	$x_i - xquer$	$(x_i - xquer)\cdot y_i$	$(x_i - xquer)^2$
36	32	–4.2	–134.4	17.64
45	42	4.8	201.6	23.04
19	24	–21.2	–508.8	449.44
62	51	21.8	1111.8	475.24
27	23	–13.2	–303.6	174.24
35	25	–5.2	–130	27.04
45	36	4.8	172.8	23.04
55	45	14.8	666	219.04
43	50	2.8	140	7.84
35	23	–5.2	–119.6	27.04
402	351	"*"	1095.8	1443.6

Mit Hilfe dieser Zwischenergebnisse kann man nun die beiden Regressionskoeffizienten berechnen

$$b_1 = \frac{\sum_{i=1}^{n} \left[(x_i - xquer)\cdot y_i\right]}{\sum_{i=1}^{n} (x_i - xquer)^2} = \frac{1095.8}{1443.6} = 0.759$$

$$b_0 = yquer - b_1 \cdot xquer = \frac{351}{10} - 0.759 \cdot \frac{402}{10} = 4.585$$

und hat als Ergebnis die Stichprobenregressionsfunktion

$$ydach_i = 4.585 + 0.759 \cdot x_i$$

ydach ist der Schätzwert für das beobachtete y. Wenn diese Regressionsfunktion allgemein gültig ist, dann muss der Regressionskoeffizient β_1 in der Grundgesamtheit aller potentiellen Käufer der Spielkonsole von Null verschieden sein. Ein Regressionskoeffizient von Null bedeutet, dass die Steigung der linearen Funktion Null ist und daher die unabhängige Variable keinen Einfluss auf die abhängige Variable ausübt. Diese Hypothese wird formal wie folgt angeschrieben:

H_0: β_0 beliebig und $\beta_1 = 0$.

Eine Alternative zu dieser Nullhypothese ist die Vermutung, dass das Alter einen signifikanten Einfluss auf das Einkommen in der Grundgesamtheit hat.

H_1: β_0 beliebig und $\beta_1 \neq 0$.

An Hand der Ergebnisse der Stichprobe entscheidet man sich für die Annahme oder Ablehnung der Null- oder Alternativhypothese. Dazu berechnet man folgende Testmaßzahl:

$$F_{beob} = \frac{\sum_{i=1}^{n} \left(ydach_i - yquer\right)^2}{\sum_{i=1}^{n} \left(y_i - ydach_i\right)^2} \cdot \frac{\nu_2}{\nu_1}$$

n ist der Umfang der gesamten Stichprobe, y_i und $ydach_i$ sind die beobachteten und geschätzten Werte der abhängigen Variablen und yquer ist der Durchschnitt aus den beobachteten Werten der Stichprobe. Diese Testmaßzahl folgt einer F-Verteilung mit

$$\nu_1 = 1$$

$$\nu_2 = n - 2$$

Freiheitsgraden. In folgender Tabelle sind die Zwischenschritte zur Berechnung der Testmaßzahl angegeben.

(Pers)	y_i	$ydach_i$	$y_i - ydach_i$	$\left(y_i - ydach_i\right)^2$	$ydach_i - yquer$	$\left(ydach_i - yquer\right)^2$
1	32	31.912	0.088	0.008	–3.188	10.164
2	42	38.744	3.256	10.604	3.644	13.276
3	24	19.008	4.992	24.924	–16.092	258.965
4	51	51.648	–0.648	0.42	16.548	273.831
5	23	25.08	–2.08	4.327	–10.02	100.396
6	25	31.153	–6.153	37.857	–3.947	15.58
7	36	38.744	–2.744	7.527	3.644	13.276
8	45	46.334	–1.334	1.78	11.234	126.21
9	50	37.225	12.775	163.19	2.125	4.517
10	23	31.153	–8.153	66.468	–3.947	15.58
SUMME	351	"*"	"*"	317.105	"*"	831.795

Die Testmaßzahl kann nun wie folgt berechnen werden:

$$F_{beob} = \frac{\sum_{i=1}^{n} \left(ydach_i - yquer\right)^2}{\sum_{i=1}^{n} \left(y_i - ydach_i\right)^2} \cdot \frac{\nu_2}{\nu_1} = \frac{831.795}{317.105} \cdot \frac{8}{1} = 20.985$$

mit

$$\nu_1 = 1 \qquad \nu_2 = n - 2 = 10 - 2 = 8$$

Freiheitsgraden. Die Wahrscheinlichkeit für das Auftreten einer so großen Testmaßzahl von $F_{boeb} = 20.985$ unter Gültigkeit der Nullhypothese ist gleich $F_F(20.985, 1, 8) = 0.998$. (Diesen Wert liest man aus einer geeigneten Tabelle der F-Verteilung ab oder berechnet ihn mit Hilfe entsprechender Programme.) Ist man bereit, in 5 von 100 Fällen eine richtige Nullhypothese abzulehnen, dann ist das Signifikanzniveau α gleich 0.05. Unter Gültigkeit der Nullhypothese sind die 5% seltenstem F-Werte nach unten beschränkt durch den kritischen Wert F_c gleich 5.318. Da das aus der Stichprobe errechnete F_{beob} von 20.985 größer ist als dieser kritische Wert, kann die Nullhypothese abgelehnt werden.

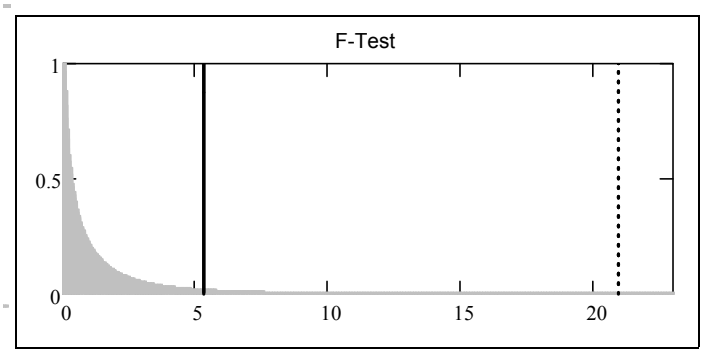
$$F_{beob} = 20.985 > 5.318 = F_c$$

Zum selben Ergebnis kommt man, wenn man die Wahrscheinlichkeiten vergleicht: Da die Wahrscheinlichkeit für einen F-Wert von 20.985 oder größer gleich $1 - F_F(20.985, 1, 8) = 0.002$ ist und das Signifikanzniveau mit 0.05 angenommen wurde, kann die Nullhypothese abgelehnt werden.

$$1 - F_F(20.985, 1, 8) = 0.002 < 0.05 = \text{Signifikanzniveau}$$

Man kann auf Grund dieser Stichprobe behaupten, dass zwischen dem Alter als unabhängige Variable und dem Einkommen aller potentiellen Kunden als abhängige Variable allgemein ein linearer Zusammenhang besteht.

Die Grafik zeigt die F-Verteilung, den kritischen Wert F_c sowie den beobachteten Wert F_{boeb}:



(Beobachteter t-Wert)	(Kritischer Wert)	(Wahrscheinlichkeit p)
20.985	5.328	0.002

Um für einen individuellen Punktschätzwert ein Konfidenzintervall zu bestimmen (wird auch als Prognoseintervall bezeichnet), benötigt man die Streuung der Residuen, das sind die Differenzen zwischen beobachteten und geschätzten Werten der abhängigen Variablen. Diese Varianz ist gegeben durch

$$s_{res}^{\,2} = \frac{1}{n-2}\cdot\sum_{i=1}^{n}\left(y_i - ydach_i\right)^2 = \frac{1}{8}\cdot 317.105 = 39.638$$

Unter der Voraussetzung, dass die Residuen normalverteilt sind, bestimmt man die Konfidenzgrenzen mit Hilfe der Studentverteilung.

$$K_u = b_0 + b_1\cdot x_0 - t\left(1-\frac{\alpha}{2}, n-2\right)\cdot s_{res}\cdot\sqrt{1+\frac{1}{n}+\frac{\left(x_0 - xquer\right)^2}{\sum_{i=1}^{n}\left(x_i - xquer\right)^2}}$$

$$K_0 = b_0 + b_1\cdot x_0 + t\left(1-\frac{\alpha}{2}, n-2\right)\cdot s_{res}\cdot\sqrt{1+\frac{1}{n}+\frac{\left(x_0 - xquer\right)^2}{\sum_{i=1}^{n}\left(x_i - xquer\right)^2}}$$

Setzt man 30 Jahre ein, dann erhält man folgende Konfidenzgrenzen für $\alpha = 0.05$:

$$K_u = \left[27.357 - 2.306\cdot\sqrt{39.638}\cdot\sqrt{1+\frac{1}{10}+\frac{\left(30-\frac{402}{10}\right)^2}{1443.6}}\right] = 11.63$$

$$K_o = \left[27.357 + 2.306\cdot\sqrt{39.638}\cdot\sqrt{1+\frac{1}{10}+\frac{\left(30-\frac{402}{10}\right)^2}{1443.6}}\right] = 43.08$$

Mit einem Vertrauen von 95% kann man das monatliche Einkommen eines 30-Jährigen zwischen 1163 € und 4308 € erwarten.

Dass man auch ein so komplexes Ergebnis wie die lineare Einfachregression nur an Hand der Dateneingabe wörtlich sinnvoll interpretieren kann, zeigt der Output der Software auf Seite 199.

5_0 DREI UND MEHR UNABHÄNGIGE STICHPROBEN

5_1 χ2-Test
5_2 H-Test
5_3 F-Test

Zwei und Drei Stichproben

Bei zwei Stichproben kann man ein- oder zweiseitig testen. Bei einem einseitigen Test ist generell das Fehlerrisiko geringer als beim zweiseitigen Test. Die Alternativhypothese lautet beim einseitigen Test üblicherweise "Der Parameter ist kleiner (oder größer) als der Parameterwert der Nullhypothese". Beim zweiseitigen Test ist die Alternativhypothese hingegen: "Der Parameter ist ungleich (also größer oder kleiner) dem der Nullhypothese.

Bei drei oder mehr Stichproben kann man nur zweiseitig testen. Eine einseitige Alternativhypothese macht keinen Sinn. Dafür ist das Testen bei drei oder mehr Stichproben noch nicht beendet, wenn man die Nullhypothese ablehnen kann. In diesem Fall will man wissen, zwischen welchen Paaren signifikanten Unterschieden existieren. Bei drei Stichproben will man z. B. wissen, ob die gefundenen signifikanten Unterschiede sich auch in den Grundgesamtheiten der beiden ersten Stichproben, der ersten und dritten Stichprobe oder der zweiten und dritten Stichprobe auch manifestieren. Dazu müssen drei Tests für 2 Stichproben durchgeführt werden.

Vorschau

Ob sich die Käufer der Spielkonsole Yoki von den Nichtkäufern und Unentschlossenen im Hinblick auf die Beurteilung der Handlichkeit dieser Konsole unterscheiden, wird im Abschnitt 5_1 mit Hilfe des χ2 Tests untersucht. Der χ2 Test setzt nur nominal skalierte Antworten voraus. Bei der Beurteilung der Handlichkeit wurden ordinal skalierte Antworten vorgegeben: sehr gut, gut, mittelmäßig und schlecht. Diese ordinal skalierten Antworten können mit einem Test für nominal skalierte Antworten analysiert werden. Man verzichtet dabei auf einen Teil der Informationen, die in den Antworten stecken. Will man darauf nicht verzichtet, dann muss man für dieses Problem den H Test verwenden. Dieser wird im Abschnitt 5_2 dargestellt. In diesem Abschnitt wird wieder untersucht, ob sich die Unterschiede in der Beurteilung der Handlichkeit der Spielkonsole Yoki zwischen Käufern, Nichtkäufern und Unentschlossenen verallgemeinern lassen.

Wenn man beim χ2-Test und H-Test zu unterschiedlichen Schlussfolgerungen kommt, dann ist für den Vergleich der Handlichkeit zwischen Käufern, Nichtkäufern und Unentschlossenen der H-Test relevant. Denn er berücksichtigt alle Informationen, die in den Antworten stecken und nicht nur einen Teil wie der χ2-Test.

Stellt sich heraus, dass sich die 3 Stichproben Käufer, Nichtkäufer und Unentschlossene signifikant unterscheiden, dann kann man mit Hilfe des U-Tests für zwei unabhängige Stichproben feststellen, wo die signifikanten Unterschiede vorliegen. Dazu prüft man die Unterschiede zwischen Käufern und Nichtkäufern,

Käufern und Unentschlossenen sowie Nichtkäufern und Unentschlossenen. Wenn der H-Test zu keinen signifikanten Ergebnissen führt, erübrigen sich natürlich die U-Tests für zwei unabhängige Stichproben.

Auch im dritten Abschnitt 5_3 werden Unterschiede zwischen den drei Stichproben Käufer, Nichtkäufer und Unentschlossene analysiert. Da der in diesem Abschnitt beschriebene F-Test metrisch skalierte Antworten voraussetzt, werden diese 3 Stichproben nicht im Hinblick auf die Beurteilung der Handlichkeit der Spielkonsole Yoki verglichen, sondern im Hinblick auf das Alter der Befragten. Auch hier gilt: Wenn sich die Grundgesamtheiten der drei Stichproben im Hinblick auf ihr Durchschnittsalter signifikant unterscheiden, dann kann man mit Hilfe des t-Tests für zwei unabhängige Stichproben prüfen, wo diese signifikanten Unterschiede auftreten.

Bonferroni Korrektur

Wenn man die Nullhypothese mit einem der drei Tests ablehnt und man mit Hilfe der Tests für zwei unabhängige Stichproben feststellen will, wo die signifikanten Unterschiede sind, dann muss man die Bonferroni Korrektur des Signifikanzniveaus berücksichtigen. Plant man z. B. im Vorhinein drei Einzeltests und möchte das gemeinsame Signifikanzniveau kontrollieren, dann muss man das Signifikanzniveau der drei Einzeltests verkleinern. Will man z. B. ein gemeinsames Signifikanzniveau von 5% (α = 0.05), dann muss jeder der drei Einzeltests auf einem Signifikanzniveau von

$$\alpha' = 0.05 / 3 = 0.0166$$

durchgeführt werden. Allgemein gilt

$$\alpha' = \alpha / n$$

mit

α = globale Signifikanzniveau,

α' = Signifikanzniveau des Einzeltests und

n = Anzahl der geplanten Einzeltests.

5_1 χ2-Test

Unterscheiden sich die Käufer der Spielkonsole Yoki von den Nichtkäufern und den Unentschlossenen im Hinblick auf die Beurteilung der Handlichkeit? Eine Befragung von 93 Besuchern einer Werbeveranstaltung brachte folgendes Ergebnis:

(Kaufabsicht / Handlichkeit)	Ja	Nein	Weiß_nicht
Sehr_gut	30	8	1
Gut	18	5	5
Mittelmäßig	5	5	5
Schlecht	5	5	1
SUMME	58	23	12

Wenn man diese Ergebnisse jeweils auf die Summe der Käufer, Nichtkäufer und Unentschlossenen bezieht, dann zeigt sich, dass 52% der Käufer die Handlichkeit der Spielkonsole sehr gut beurteilen, aber nur 35% der Nichtkäufer und nur 8% der Unentschlossenen. Die folgende Tabelle zeigt die Antwortkombinationen in Prozenten:

(Kaufabsicht / Handlichkeit)	Ja	Nein	Weiß_nicht
Sehr_gut	52	35	8
Gut	31	22	42
Mittelmäßig	9	22	42
Schlecht	9	22	8
SUMME	100	100	100

Kann man auf Grund dieser Stichprobenergebnisse allgemein behaupten, dass sich Käufer, Nichtkäufer und Unentschlossene im Hinblick auf die Beurteilung der Handlichkeit der neuen Spielkonsole Yoki unterscheiden?

Ja, die Unterschiede zwischen den Prozentsätzen sind groß genug. Es könnte zwar tatsächlich sein, dass die Käufer, Nichtkäufer und Unentschlossenen die Handlichkeit nicht unterschiedlich beurteilen. Das Risiko ist dafür mit 5% beschränkt. (Vergleiche Softwareoutput auf Seite 200)

Wie kommt man zu diesem Ergebnis? Zuerst formuliert man zuerst die Hypothesen. Man nimmt als Nullhypothese an, dass die Verteilungen der Beurteilungen der Handlichkeit für die Käufer, Nichtkäufer und Unentschlossenen gleich sind. Diese Hypothese wird formal wie folgt angeschrieben:

$$H_0: \text{Verteilung}_{\text{Käufer}} = \text{Verteilung}_{\text{Nichtkäufer}} = \text{Verteilung}_{\text{Unentschlossen}}$$

Mit F(x) wird die Verteilung der Beurteilung der Handlichkeit in der Grundgesamtheit aller Käufer, Nichtkäufer und Unentschlossenen bezeichnet. Eine Alternative zu dieser Nullhypothese ist

$$H_1: \text{Verteilung}_{\text{Käufer}} \neq \text{Verteilung}_{\text{Nichtkäufer}} \neq \text{Verteilung}_{\text{Unentschlossen}}$$

Die Verteilungen der Handlichkeitsbeurteilung unterscheiden sich. An Hand der Ergebnisse der Stichprobe entscheidet man sich für die Annahme oder Ablehnung der Null- oder Alternativhypothese. Dazu berechnet man folgende Testmaßzahl:

$$\chi 2_{beob} = n \cdot \left[\left[\sum_{j=1}^{m} \sum_{k=1}^{t} \frac{(h_{jk})^2}{h_j \cdot h_k} \right] - 1 \right]$$

n ist der Umfang der gesamten Stichprobe, h_j und h_k sind die Randhäufigkeiten der Zeilen und Spalten sowie h_{jk} die Häufigkeiten der Antwortkombinationen. In folgender Kreuztabelle sind neben den Häufigkeiten der Antwortkombinationen auch die Zeilen und Spaltensummen des Beispiels angeführt:

(Kaufabsicht / Handlichkeit)	Ja	Nein	Weiß_nicht	SUMME
Sehr_gut	30	8	1	39
Gut	18	5	5	28
Mittelmäßig	5	5	5	15
Schlecht	5	5	1	11
SUMME	58	23	12	93

Nun kann man die Werte in die Formel für die Chiquadratmaßzahl einsetzen und berechnen:

$$\chi 2_{beob} = 93 \cdot \left(\frac{30^2}{39 \cdot 58} + \frac{8^2}{39 \cdot 23} + ... + \frac{5^2}{11 \cdot 58} + \frac{5^2}{11 \cdot 23} + \frac{1^2}{11 \cdot 12} - 1 \right) = 15.788$$

Die $\chi 2$-Testmaßzahl folgt einer Chiquadratverteilung mit

$$\nu = (m - 1) \cdot (t - 1)$$

Freiheitsgraden. m ist die Anzahl der Antworten in den Zeilen und t die der Spalten der Kreuztabelle. Im obigen Beispiel sind die Freiheitsgrade gleich $\nu = (4 - 1) \cdot (3 - 1) = 6$ und die Wahrscheinlichkeit für das Auftreten einer so großen Testmaßzahl von $\chi 2_{beob} = 15.788$ unter Gültigkeit der Nullhypothese gleich. $F\chi 2(15.788, 6) = 0.985$. (Diesen Wert liest man aus einer geeigneten Tabelle der Chiquadratverteilung ab oder berechnet ihn mit Hilfe entsprechender Programme.)

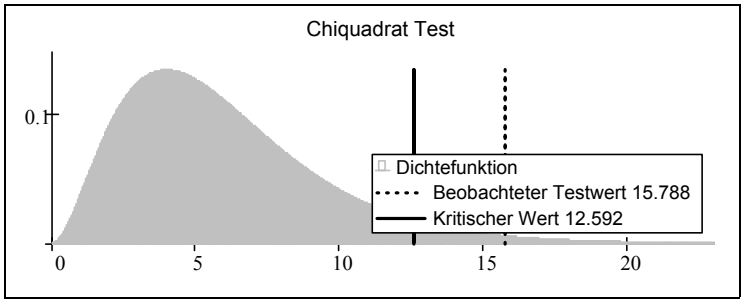
Ist man bereit, in 5 von 100 Fällen eine richtige Nullhypothese abzulehnen, dann ist das Signifikanzniveau α gleich 0.05. Unter Gültigkeit der Nullhypothese sind die 5% seltensten $\chi 2$ Werte nach unten beschränkt durch den kritischen Wert $\chi 2_c$ gleich.12.592 Da das aus der Stichprobe errechnete $\chi 2_{beob}$ von 15.788 größer ist als dieser kritische Wert, kann die Nullhypothese abgelehnt werden.

$$\chi 2_{beob} = 15.788 > 12.592 = \chi 2_c$$

Zum selben Ergebnis kommt man, wenn man die Wahrscheinlichkeiten vergleicht: Da die Wahrscheinlichkeit für einen $\chi 2$ Wert von $\chi 2_{beob}$ 15.788 oder größer gleich $1 - F\chi 2(15.788, 6) = 0.015$ ist und das Signifikanzniveau mit 0.05 angenommen wurde, kann die Nullhypothese abgelehnt werden.

$$1 - F_{\chi 2}(15.788, 6) = 0.015 < 0.05 = \text{Signifikanzniveau}$$

Man kann auf Grund dieser Stichproben behaupten, dass sich die Beurteilungen der Handlichkeit unterschiedlich auf Käufer, Nichtkäufer und Unentschlossene verteilen. Das Risiko, dass diese Entscheidung falsch ist, ist höchstens 5%.



(Chiquadrat – beobachtet)	(Chiquadrat – kritisch)	(Wahrscheinlichkeit – p)
15.788	12.592	0.015

Wo sind die signifikanten Unterschiede? Unterscheiden sich die Käufer, Nichtkäufer und Unentschlossenen im Hinblick auf die sehr gute, gute, mittelmäßige oder schlechte Beurteilung der Handlichkeit? Der Chiquadrattest kam nur zum Ergebnis, dass Unterschiede bestehen, nicht aber bei welcher Antwort. Tatsächlich unterscheiden sich die Käufer und Unentschlossenen im Hinblick auf die Beurteilung der Handlichkeit der Spielkonsole Yoki.

Zu diesem Ergebnis kommt man mit Hilfe der Konfidenzintervalle. In der folgenden Tabelle sind als letzte Spalte die Prozentwerte der Beurteilung der Handlichkeit angeführt, unabhängig von der Kaufabsicht.

(Kaufabsicht)	"Ja"	"Nein"	"Weiß nicht"	DURCHSCHNITT
(Handlichkeit)	in_%	in_%	in_%	in_%
"Sehr gut"	52	35	8	42
"Gut"	31	22	42	30
"Mittelmäßig"	9	22	42	16
"Schlecht"	9	22	8	12
SUMME	100	100	100	100

42% von den insgesamt 93 Befragten haben die Handlichkeit der Spielkonsole mit "sehr gut" beurteilt. Bei den Käufern sind es 52%, bei den Nichtkäufern 35% und bei den Unentschlossenen 8%. Innerhalb welcher Grenzen kann man mit einem Vertrauen von 95% den "wahren" Prozentsatz für die Beurteilung "sehr gut" in der Grundgesamtheit aller potentiellen Kunden erwarten, wenn in der Stichprobe der Prozentsatz für "sehr gut" gleich 42 ist?

In der Grundgesamtheit ist der Prozentsatz für "sehr gut" mit 95% Vertrauen innerhalb der Grenzen 32% und 52.5% zu erwarten. Da die 51.7% der Käufer über der Obergrenze von 52.5% liegt, ist die Beurteilung "sehr gut" bei den Käufern überdurchschnittlich stark vertreten und den Unentschlossenen mit 8% unterdurchschnittlich. Die folgende Tabelle zeigt die gesammelten Ergebnisse:

(Kaufabsicht / Handlichkeit)	"Ja"	"Nein"	"Weiß nicht"	DURCHSCHNITT
"Sehr gut"	"über"	"----"	"unter"	41.9
"Gut"	"----"	"----"	"über"	30.1
"Mittelmäßig"	"unter"	"----"	"über"	16.1
"Schlecht"	"----"	"über"	"----"	11.8

Die Konfidenzgrenzen können mit Hilfe der F-Verteilung berechnet werden (siehe Auswertung einer Frage, nominal). Für die "sehr gut"- Antworten ist n = 93 und x = 39 (= 41.9%). K_u ist daher

$$K_u = \frac{78}{78 + 110 \cdot 1.522} = 0.318$$

mit

$$\nu_1 = 2 \cdot (93 - 39 + 1) = 110$$

$$\nu_2 = 2 \cdot 39 = 78$$

und

$$F_{0.975,110,78} = 1.522$$

K_o ist

$$K_o = \frac{78 \cdot 1.5}{110 + 78 \cdot 1.5} = 0.515$$

mit

$$F_{0.975,78,110} = 1.5$$

Der Chiquadrattest ist nur dann zulässig, wenn keine Antwortkombination Null-Mal vorkommt (also keine Zelle der Kreuztabelle den Wert 0 hat) und wenn höchstens 20% aller Antwortkombinationen eine Häufigkeit kleiner als 5 aufweisen (höchstens 20% aller Zellen in der Kreuztabelle mit einer Häufigkeit < 5). Ist dies nicht der Fall, dann sind die Schlussfolgerungen aus dem Chiquadrattest ungültig. Durch Zusammenfassen der Antwortmöglichkeiten oder durch Erhöhung des Stichprobenumfanges kann eventuell diese Voraussetzung des Chiquadrattest erfüllt werden. Im oben angeführten Beispiel sind die Voraussetzungen des Chiquadrattest erfüllt.

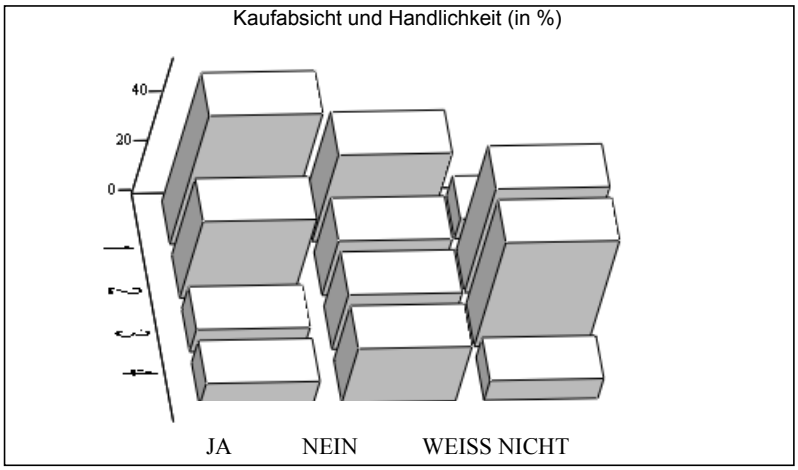
5_2 H-Test

Unterscheiden sich die Käufer der Spielkonsole Yoki von den Nichtkäufern und den Unentschlossenen im Hinblick auf die Beurteilung der Handlichkeit? Eine Befragung von 93 Besuchern einer Werbeveranstaltung brachte folgendes Ergebnis:

(Kaufabsicht / Handlichkeit)	Ja	Nein	Weiß_nicht
Sehr_gut	30	8	1
Gut	18	5	5
Mittelmäßig	5	5	5
Schlecht	5	5	1
SUMME	58	23	12

Wenn man diese Ergebnisse jeweils auf die Käufer, Nichtkäufer und Unentschlossenen bezieht, dann zeigt sich, dass 52% der Käufer die Handlichkeit der Spielkonsole sehr gut beurteilen, aber nur 35% der Nichtkäufer und nur 8% der Unentschlossenen. Die folgende Tabelle zeigt die Antwortkombinationen in Prozent:

(Kaufabsicht / Handlichkeit)	Ja	Nein	Weiß_nicht
Sehr_gut	52	35	8
Gut	31	22	42
Mittelmäßig	9	22	42
Schlecht	9	22	8
SUMME	100	100	100



Kann man auf Grund dieser Stichprobenergebnisse allgemein behaupten, dass sich Käufer, Nichtkäufer und Unentschlossene im Hinblick auf die Beurteilung der Handlichkeit der neuen Spielkonsole Yoki unterscheiden?

Ja, die Unterschiede zwischen den Beurteilungen sind groß genug. Es könnte zwar tatsächlich sein, dass die Käufer, Nichtkäufer und Unentschlossenen die Handlichkeit nicht unterschiedlich beurteilen. Das Risiko ist dafür mit 5% beschränkt.

Wie kommt man zu diesem Ergebnis? Zuerst formuliert man die Hypothesen. Man nimmt als Nullhypothese an, dass die Verteilungen der Beurteilungen der Handlichkeit für die Käufer, Nichtkäufer und Unentschlossenen gleich sind. Diese Hypothese wird formal wie folgt angeschrieben:

$$H_0: \text{Verteilung}_{\text{Käufer}} = \text{Verteilung}_{\text{Nichtkäufer}} = \text{Verteilung}_{\text{Unentschlossen}}$$

Mit F(x) wird die Verteilung der Beurteilung der Handlichkeit in der Grundgesamtheit aller Käufer, Nichtkäufer und Unentschlossenen bezeichnet. Eine Alternative zu dieser Nullhypothese ist

$$H_1: \text{Verteilung}_{\text{Käufer}} \neq \text{Verteilung}_{\text{Nichtkäufer}} \neq \text{Verteilung}_{\text{Unentschlossen}}$$

Die Verteilungen der Handlichkeitsbeurteilung unterscheiden sich. An Hand der Ergebnisse der Stichprobe entscheidet man sich für die Annahme oder Ablehnung der Null- oder Alternativhypothese. Dazu berechnet man folgende Testmaßzahl:

$$H = \frac{\frac{12}{n\cdot(n+1)}\cdot\left(\sum_{i=1}^{k}\frac{T_i^2}{n_i}\right) - 3\cdot(n+1)}{1 - \frac{\sum_{i=1}^{k}\left(f_i^3 - f_i\right)}{n^3 - n}}$$

n ist die Summe der einzelnen Stichprobenumfänge $n = n_1 + n_2 + \ldots + n_l$. T_i ist die Summe der Rangzahlen der i-ten Stichprobe. Zur Berechnung der Rangzahlen bringt man die Einheiten der Stichproben in eine gemeinsame Reihenfolge und berechnet für jede Beurteilung die entsprechende Rangzahl. Zu jeder Rangzahl vermerkt man, aus welcher der Stichproben die Einheit stammt. Haben mehrere Einheiten die gleiche Merkmalsausprägung, dann ordnet man ihnen als Rangzahlen den Durchschnitt daraus zu. In folgender Tabelle sind die einzelnen Schritte dargestellt:

(Beurteilung)	h_{i1}	h_{i2}	h_{i3}	f_i	cf_i	R_i	$h_{i1}\cdot R_i$	$h_{i2}\cdot R_i$	$h_{i3}\cdot R_i$	f_i^3	$f_i^3 - f_i$
sehr_gut	30	8	1	39	39	20	600	160	20	59319	59280
gut	18	5	5	28	67	53.5	963	267.5	267.5	21952	21924
mittelmäßig	5	5	5	15	82	75	375	375	375	3375	3360
schlecht	5	5	1	11	93	88	440	440	88	1331	1320
SUMME	58	23	12	93	"*"	"*"	2378	1242.5	750.5	"*"	85884

Die Rangzahlen für die vier Beurteilungen ergeben sich aus den kumulierten Häufigkeiten cf_i. 1 (= sehr gut) wird insgesamt 39 Mal genannt (= $f_1 = cf_1$). Die ersten 39 Rangzahlen (= 1, 2, ..., 39) werden daher zusammengezählt und durch 39 dividiert. Das Ergebnis ist der Durchschnittsrang R_1 von 20. Man kann ihn auch mit Hilfe folgender Formel berechnen:

$$R_1 = \frac{1}{2}\cdot(0 + 39 + 1) = 20$$

Die Beurteilung "gut" (= 2) wird insgesamt 28 Mal genannt (= f_2). Einschließlich der "sehr guten" Beurteilung sind dies insgesamt 67 Nennungen (= cf_2). Der Durchschnitt aus den weiteren 28 Rangzahlen (= 40, 41, ..., 67) wird mit folgender Formel bestimmt:

$$R_2 = \frac{1}{2}\cdot(39 + 67 + 1) = 53.5$$

"Mittelmäßig" (= 3) wird 15 Mal genannt (=f_3). Mit "sehr guter" und "guter" Beurteilung sind dies nun insgesamt 82 Nennungen (= cf_3). Der Durchschnitt aus den weiteren 15 Rangzahlen (= 68, 69, ..., 82) wird mit der folgenden Formel berechnet:

$$R_3 = \frac{1}{2}\cdot(67 + 82 + 1) = 75$$

"Schlecht" (= 4) wird 11 Mal genannt (= f_4). Mit "sehr guter“, "guter" und "mittelmäßiger" Beurteilung sind dies nun insgesamt 93 Nennungen (= cf_4). Der Durchschnitt aus den weiteren 11 Rangzahlen (= 83, 84, ..., 93) wird mit der folgenden Formel bestimmt:

$$R_4 = \frac{1}{2}\cdot(82 + 93 + 1) = 88.$$

Mit Hilfe dieser Zwischenergebnisse kann man nun den Zähler von H berechnen. Für jede der drei Stichproben wird die mit der Häufigkeit gewichtete Rangzahl bestimmt. Die Beurteilung "sehr gut" kommt bei den Käufern 30 Mal vor. Daher wird dort $30 \cdot 20 = 600$ eingetragen (= $h_{11}.R_1$). Bei den Nichtkäufern wird

"sehr gut" 8 Mal genannt. Daher ist dort der gewichtete Durchschnittsrang gleich 160 = 8 · 20 (= $h_{12}.R_1$). Schließlich wird "sehr gut" bei den Unentschlossenen 1 Mal genannt. Der gewichtete Durchschnittsrang ist daher gleich 20 (= $h_{13}.R_1$). Die weiteren Beurteilungsränge werden analog auf die drei Stichproben aufgeteilt und dann jeweils für die einzelnen Stichproben summiert. Für die erste Stichprobe ergibt sich so eine Rangsumme von 2378 (=T_1), für die zweite von 1242.5 (=T_2) und für die dritte von 750.5 (=T_3). Berücksichtigt man noch die Stichprobenumfänge von 58, 23 und 12 Beurteilungen, dann kann man in H einsetzen:

$$H = \frac{\frac{12}{93\cdot(93+1)}\cdot\left(\frac{2378^2}{58}+\frac{1242.5^2}{23}+\frac{750.5^2}{12}\right)-3\cdot(93+1)}{1-\frac{85884}{93^3-93}} = 9.406$$

Für den Nenner von H müssen die Häufigkeiten der Rangzahlen wie folgt berücksichtigt werden: $f^3_i - f_i$ Die Beurteilung "sehr gut" wurde z. B. 39 Mal abgegeben. Daher ist $39^3 - 39 = 59280$. Die weiteren Häufigkeiten werden analog berücksichtigt. Die Summe aus diesen Beträgen ergibt 85884.

Die Testmaßzahl H folgt einer Chiquadratverteilung mit

$$\nu = (k - 1)$$

Freiheitsgraden. k ist die Anzahl der Stichproben. Im obigen Beispiel sind die Freiheitsgrade gleich $\nu = (3 - 1) = 2$ und die Wahrscheinlichkeit für das Auftreten einer so großen Testmaßzahl von H = 9.406 unter Gültigkeit der Nullhypothese ist gleich. $F\chi2(9.406, 2) = 0.991$ (Diesen Wert liest man aus einer geeigneten Tabelle der Chiquadratverteilung ab oder berechnet ihn mit Hilfe entsprechender Programme.)

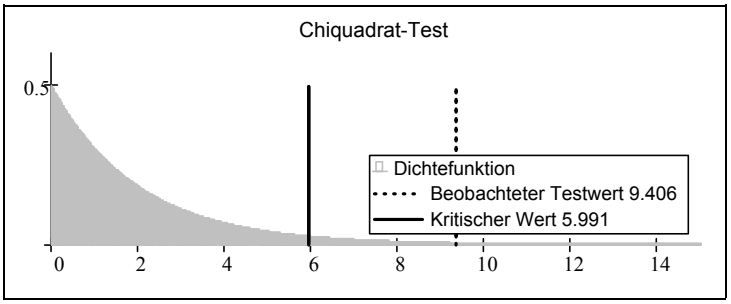
Ist man bereit, in 5 von 100 Fällen eine richtige Nullhypothese abzulehnen, dann ist das Signifikanzniveau α gleich 0.05. Unter Gültigkeit der Nullhypothese sind die 5% seltensten $\chi2$ Werte nach unten beschränkt durch den kritischen Wert $\chi2_{beob}$ gleich 5.991. Da das aus der Stichprobe errechnete $\chi2_{beob}$ von 9.406 größer ist als dieser kritische Wert, kann die Nullhypothese abgelehnt werden.

$$\chi2_{beob} = 9.406 > 5.991 = \chi2_c$$

Zum selben Ergebnis kommt man, wenn man die Wahrscheinlichkeiten vergleicht: Da die Wahrscheinlichkeit für einen $\chi2$ Wert von 9.406 oder größer gleich $1 - F\chi^2(9.406, 2) = 0.009$ ist und das Signifikanzniveau mit 0.05 angenommen wurde, kann die Nullhypothese abgelehnt werden.

$$1 - F_{\chi2}(9.406, 2) = 0.009 < 0.05 = \text{Signifikanzniveau}$$

Man kann auf Grund dieser Stichprobe behaupten, dass sich die verschiedenen Formen der Kaufabsicht hinsichtlich der Beurteilung der Handlichkeit der Spielkonsole Yoki signifikant unterscheiden. Das Risiko, dass diese Entscheidung falsch ist, ist höchstens 5%.



(Chiquadrat – beobachtet)	(Chiquadrat – kritisch)	(p – Wahrscheinlichkeit)
9.406	5.991	0.009

Wo sind die signifikanten Unterschiede? Unterscheiden sich die Käufer, Nichtkäufer und Unentschlossenen im Hinblick auf die sehr gute, gute, mittelmäßige oder schlechte Beurteilung der Handlichkeit? Der H-Test kam nur zum Ergebnis, dass Unterschiede bestehen, nicht aber bei welcher Antwort. Um dies zu eruieren, muss man mit je zwei Stichproben U-Tests durchführen, also für das Beispiel mit drei Stichproben insgesamt drei Tests. Die Software führt diese Tests automatisch durch, auch wenn der H-Test zu keinem signifikanten Ergebnis führt. Auf Seite 202 findet man den Output für dieses Beispiel samt den drei U-Tests.

5_3 F-Test

Unterscheiden sich die Käufer der Spielkonsole Yoki von den Nichtkäufern und den Unentschlossenen im Hinblick auf ihr Alter? Eine Befragung von 17 Besuchern einer Werbeveranstaltung brachte folgendes Ergebnis:

(Käufer_Ja 17 25 18 22 30 18)

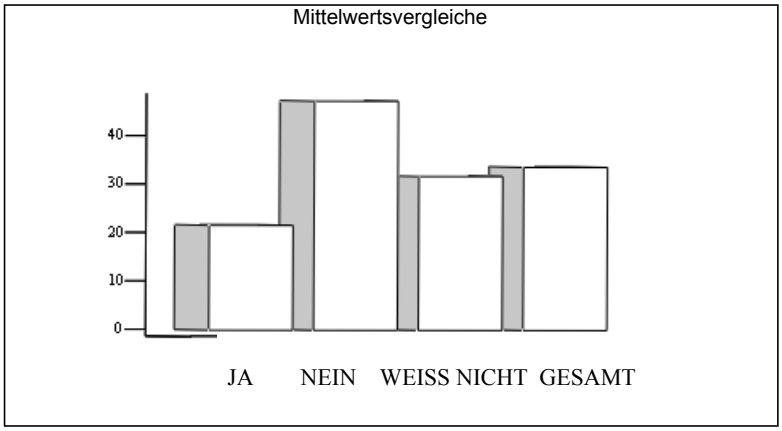
(Käufer_Nein 23 35 61 44 56 65)

(Weiß_nicht 33 27 18 54 27)

Wenn man diese Ergebnisse jeweils auf die Käufer, Nichtkäufer und Unentschlossenen bezieht, dann zeigt sich, dass 100% der Käufer zwischen 15 und 30 Jahre alt sind, aber nur 17% der Nichtkäufer und 60.8% der Unentschlossenen. Die folgende Tabelle zeigt die Antwortkombinationen mit drei Altersklassen:

(U_grenze)	(O_grenze)	Kaufabs_Ja	Kaufabs_Nein	Kaufabs_Weiß_nicht
unter	30	6	1	3
30	60	0	3	2
60	und_mehr	0	2	0
(Alter)	SUMME	6	6	5

Die Durchschnitte aus den Altersangaben für die drei Stichproben zeigt folgende Grafik:



"Grafikbalken"	1	2	3	4
("Kaufabsicht")	"Ja"	"Nein"	"Weiß nicht"	"Insgesamt"
(Alter)	21.667	47.333	31.8	33.706

Kann man auf Grund dieser Stichprobenergebnisse allgemein behaupten, dass sich Käufer, Nichtkäufer und Unentschlossene im Hinblick auf ihr Durchschnittsalter unterscheiden?

Ja, die Unterschiede zwischen den Durchschnitten sind groß genug. Es könnte zwar tatsächlich sein, dass das Durchschnittsalter die Käufer, Nichtkäufer und Unentschlossenen unterschiedlich ist. Das Risiko ist dafür mit 5% beschränkt.

Wie kommt man zu diesem Ergebnis? Zuerst formuliert man die Hypothesen. Man nimmt als Nullhypothese an, dass das Durchschnittsalter der Käufer, Nichtkäufer und Unentschlossenen in der Grundgesamtheit aller potentiellen Kunden gleich ist. Diese Hypothese wird formal wie folgt angeschrieben:

$$H_0: \mu_{Käufer} = \mu_{Nichtkäufer} = \mu_{Unentschlossen}$$

Mit μ wird der jeweilige Altersdurchschnitt in der Grundgesamtheit aller Käufer, Nichtkäufer und Unentschlossenen bezeichnet. Eine Alternative zu dieser Nullhypothese ist

$$H_1: \mu_{Käufer} \neq \mu_{Nichtkäufer} \neq \mu_{Unentschlossen}$$

Das Durchschnittsalter der Käufer, Nichtkäufer und Unentschlossenen unterscheidet sich. An Hand der Ergebnisse der Stichprobe entscheidet man sich für die Annahme oder Ablehnung der Null- oder Alternativhypothese. Dazu berechnet man folgende Testmaßzahl:

$$F_{beob} = \frac{\sum_{j=1}^{k} \left[\left(xquer_j - xquer_{gesamt} \right)^2 \cdot n_j \right]}{\sum_{j=1}^{k} \sum_{i=1}^{n_j} \left(x_{ij} - xquer_j \right)^2} \cdot \frac{\nu_2}{\nu_1}$$

Im Nenner dieser Maßzahl stehen die Abweichungsquadrate innerhalb der Gesamtstichprobe und im Zähler die Abweichungsquadrate zwischen den Stichproben. Die Freiheitsgrade ν_1 und ν_2 sind wie folgt definiert:

$$\nu_1 = k - 1$$

$$\nu_2 = n - k$$

k ist die Anzahl der Stichproben und n ist der Umfang der Gesamtstichprobe. Die Berechnung des Nenners ist in folgender Tabelle für die erste Stichprobe dargestellt:

x_{i1}	$x_{i1} - 21.667$	$(x_{i1} - 21.667)^2$
17	−4.667	21.778
25	3.333	11.111
18	−3.667	13.444
22	0.333	0.111
30	8.333	69.444
18	−3.667	13.444
Käufer_Ja	SUMME	129.333

In der 1. Spalte stehen die Altersangaben der Käufer und in der 2. Spalte die Abweichungen des Alters vom Durchschnitt 21.667. In der 3. Spalte sind die Quadrate von Spalte 2 ausgewiesen. Die Summe dieser letzten Spalte ist 129.333, die Abweichungsquadratsumme der ersten Stichprobe von ihrem Durchschnitt.

Die Abweichungsquadratsummen der 2. Strichprobe sind 1329.333 und die der 3. Stichprobe ist 730.8. Der Nenner des Quotienten der Testmaßzahl ist die Summe aus diesen drei Werten, die sogenannte Abweichungsquadratsumme innerhalb der Stichproben:

$$\sum_{j=1}^{3} \sum_{i=1}^{n_j} (x_{i,j} - xquer_j)^2 = 129.333 + 1329.333 + 730.8 = 2189.466$$

Die Berechnung des Zählers des ersten Quotienten der Testmaßzahl F_{beob} ergibt sich wie folgt:

$$(21.667 - 33.706)^2 \cdot 6 + (47.333 - 33.706)^2 \cdot 6 + (31.8 - 33.706)^2 \cdot 5 = 2001.96$$

Die beiden Freiheitsgrade sind

$$\nu_1 = k - 1 = 3 - 1 = 2$$

$$\nu_2 = n - k = (6 + 6 + 5) - 3 = 14$$

Die Testmaßzahl F_{beob} ist daher

$$F_{beob} = \frac{2001.96}{2189.466} \cdot \frac{14}{2} = 6.401$$

Die Testmaßzahl F_{beob} ist F-verteilt mit ν_1 und ν_2 Freiheitsgraden. Die Wahrscheinlichkeit für das Auftreten einer so großen Testmaßzahl F_{beob} wie im vorliegenden Beispiel ist bei Gültigkeit der Nullhypothese gleich

$$F_F(3.269, 2, 14) = 0.989$$

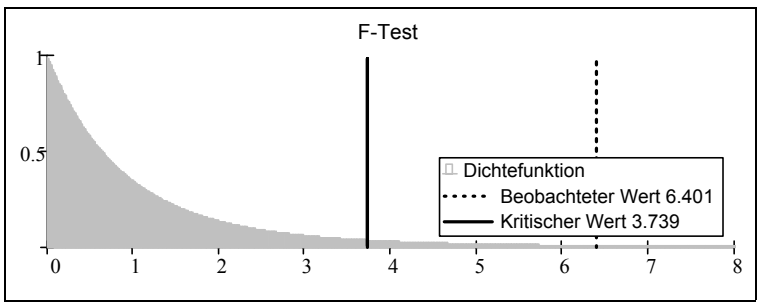
(Diesen Wert liest man aus einer geeigneten Tabelle der Chiquadratverteilung ab oder berechnet ihn mit Hilfe entsprechender Programme.) Die Wahrscheinlichkeit für eine so große oder größere Testmaßzahl ist dann gleich

$$1 - F_F(3.269, 2, 14) = 0.011$$

Da diese Wahrscheinlichkeit kleiner ist, als das Signifikanzniveau von 5% ($\alpha = 0.05$), kann die Nullhypothese abgelehnt werden. Käufer, Nichtkäufer und Unentschlossene unterscheiden sich signifikant im Hinblick auf ihr Durchschnittsalter. Das Risiko, dass diese Entscheidung falsch ist, ist höchstens 5%.

Zum selben Ergebnis kommt man, wenn man F_{beob} mit dem kritischen Wert der F-Verteilung vergleicht: Unter Gültigkeit der Nullhypothese sind die 5% seltensten F Werte nach unten beschränkt durch den kritischen Wert F_c gleich 3.739. Da das aus der Stichprobe errechnete F_{beob} von 6.401 größer ist als dieser kritische Wert, kann die Nullhypothese abgelehnt werden.

$$F_{beob} = 6.401 > 3.739 = F_c$$



(F – beobachtet)	(F – kritisch)	(p – Wahrscheinlichkeit)
6.401	3.739	0.011

Wo sind die signifikanten Unterschiede? Unterscheiden sich die Käufer, Nichtkäufer und Unentschlossenen im Hinblick auf die sehr gute, gute, mittelmäßige oder schlechte Beurteilung der Handlichkeit? Der F-Test kam nur zum Ergebnis, dass Unterschiede bestehen, nicht aber bei welcher Antwort. Um dies zu eruieren, muss man mit je zwei Stichproben Tests durchführen, also für das Beispiel mit drei Stichproben insgesamt drei t-Tests. Dabei ist die Bonferroni Korrektur zu beachten.

Das Computerprogramm führt nicht nur den F-Test durch, sondern auch die möglichen t-Tests für zwei Stichproben. Die Ergebnisse für dieses Beispiel findet man auf Seite 204.

6_0 DREI UND MEHR ABHÄNGIGE STICHPROBEN

6_1 Cochran-Test
6_2 Friedman-Test
6_3 F-Test

Vorschau

Im Abschnitt 6_1 wird untersucht, ob sich durch wiederholte Spielversuche die Quote der Spieler erhöht, die die Sollzeit des Spieles erreichen. Die ausgewerteten Antworten sind bei diesem Test nominal, nämlich "Sollzeit erreicht" oder "Sollzeit nicht erreicht". Die Abhängigkeit der Stichproben ist dadurch gegeben, dass die Auswertung nach den beiden Antworten für jeden Spieler bei seinen ersten drei Spielversuchen erfolgte.

Mit Hilfe des Friedman-Tests von Abschnitt 6_2 wurde die gleiche Frage analysiert. Die Antworten sind in diesem Fall nicht "Sollzeit erreicht" und "Sollzeit nicht erreicht", sondern die benötigte Zeit zum Beenden des Spiels in Minuten. Die Antworten der drei Versuche sind daher metrisch skaliert. Für den Friedman-Test wurde jedoch nur die Ordnungsrelation dieser Antworten berücksichtigt. Welche Spielzeit der drei gemessenen Versuche war die kürzeste, die mittlere und die längste? Nur diese drei Rangzahlen werden beim Friedman-Test ausgewertet.

Wenn man aus den gemessenen Zeiten für die Beendigung des Spiels jeweils den Durchschnitt berechnet und diese für die drei Versuche miteinander vergleicht, dann kann dieser Vergleich mit Hilfe eines F-Tests durchgeführt werden. Mit ihm wird festgestellt, ob sich die Durchschnitte der drei Versuche auch in der Grundgesamtheit aller Spieler dieses Spiels unterscheiden oder nicht. Dieser Test wird im Abschnitt 6_3 dargestellt.

Sowohl für den Cochran-, Friedman- als auch den F-Test ist die Analyse der Daten abgeschlossen, wenn die Nullhypothese nicht abgelehnt werden kann. Kann man jedoch die Nullhypothese ablehnen, dann will man wissen, wo die signifikanten Unterschiede auftreten: Zwischen dem 1. und 2. Versuch, dem 1. und 3. Versuch oder zwischen dem 2. und 3. Versuch? Diese Fragen können mit Hilfe der entsprechenden Tests für zwei abhängige Stichproben geklärt werden. Dabei ist wieder die im Kapitel 6 erwähnte Bonferroni Korrektur des Signifikanzniveaus zu berücksichtigen.

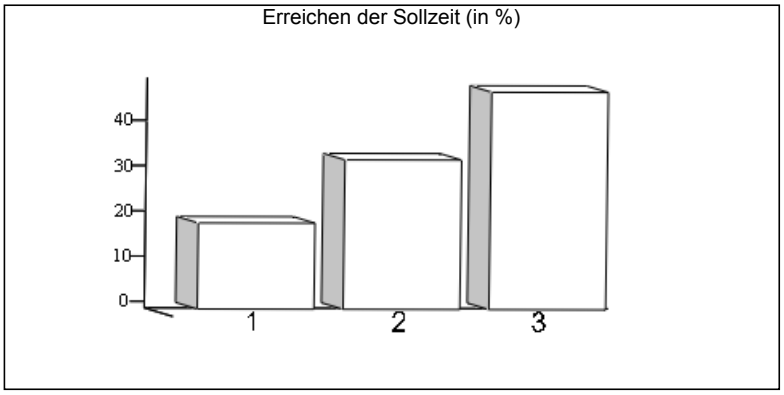
6_1 Cochran-Test

Kann man davon ausgehen, dass wiederholtes Probieren zum Erreichen der Sollzeit des neuen Konsolenspiels "Fantasy" führt? Bei einer Werbeveranstaltung für die neue Spielkonsole Yoki wurden 15 Besucher bei 3 Versuchen getestet, ob sie die Sollzeit des Spieles "Fantasy" erreichten (1 = Sollzeit erreicht, 0 = Sollzeit nicht erreicht):

1	0	0	1	0	0	1	0	1	1	0	0	0	0	0	"1.V"
0	1	1	1	0	0	1	1	0	1	1	1	0	0	1	"2.V"
1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	"3.V"

Ausgezählt erhält man folgendes Ergebnis:

(Sollzeit)	Häufigkeit	in_%
"1. Versuch"	5	19
"2. Versuch"	9	33
"3. Versuch"	13	48
SUMME	27	100



(Sollzeit)	1	2	3
"Versuche"	"1. Versuch"	"2. Versuch"	"3. Versuch"
"in %"	19	33	48

Beim 1. Versuch hatten erst 19% der 15 Besucher die Sollzeit erreicht, beim 2. Versuch schon 33% und beim 3. Versuch 48%. Ist diese Steigerung schon groß genug, um allgemein zu behaupten, dass mehrfaches Probieren zum erreichen der Sollzeit für das Spiel "Fantasy" führt?

Ja, die Steigerungsquote ist groß genug. Man kann generell behaupten, dass sich durch mehrfaches Probieren das Erreichen der Sollzeit signifikant verändert. Zu diesem Ergebnis kommt man wie folgt: Zuerst werden Null- und Alternativhypothese formuliert. Als Nullhypothese nimmt man an, dass die Anteile der Personen, die die Sollzeit erreichten, in jedem Versuch gleich hoch sind:

$$H_0: \pi_{1\_Versuch} = \pi_{2\_Versuch} = \pi_{3\_Versuch}$$

Die Alternative dazu ist

$$H_1: \pi_{1\_Versuch} \neq \pi_{2\_Versuch} \neq \pi_{3\_Versuch}$$

Durch mehrmaliges Probieren ändern sich die Erreichungsquoten der Sollzeit des Spiels "Fantasy". Um diese Hypothesen zu prüfen, berechnet man folgende Testmaßzahl:

$$Q_{beob} = \frac{(k-1)\cdot\left[k\cdot\sum_{j=1}^{k}(T_j)^2 - \left(\sum_{j=1}^{k} T_j\right)^2\right]}{k\cdot\sum_{i=1}^{n} L_i - \sum_{i=1}^{n}(L_i)^2}$$

Diese Testmaßzahl ist quiquadratverteilt mit $\nu = k - 1$ Freiheitsgraden. k ist die Anzahl der Stichproben und n die Anzahl der Befragten. T_j ist die Summe der positiven Ergebnisse (= 1) über die Versuche und L_i die positiven Ergebnisse über die Befragten. Die Berechnung ist in folgender Tabelle angeführt:

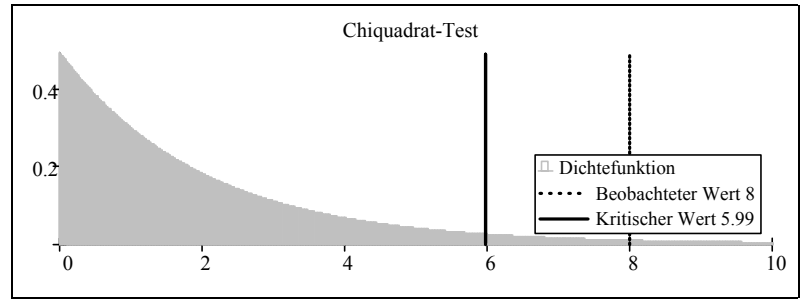
("1.Versuch")	("2.Versuch")	("3.Versuch")	L_i	L_i^2
1	0	1	2	4
0	1	1	2	4
0	1	1	2	4
1	1	1	3	9
0	0	1	1	1
0	0	1	1	1
1	1	0	2	4
0	1	1	2	4
1	0	1	2	4
1	1	1	3	9
0	1	1	2	4
0	1	1	2	4
0	0	0	0	0
0	0	1	1	1
0	1	1	2	4
$T_1 = 5$	$T_2 = 9$	$T_3 = 13$	27	57
25	81	169	"*"	"*"

Eingesetzt erhält man für die Testmaßzahl

$$Q_{beob} = \frac{(3-1)\cdot\left[k\cdot\sum_{j=1}^{3}(T_j)^2-\left(\sum_{j=1}^{3}T_j\right)^2\right]}{3\cdot\sum_{i=1}^{15}L_i-\sum_{i=1}^{15}(L_i)^2} = \frac{(3-1)\cdot(3\cdot 275-729)}{3\cdot 27-57} = 8$$

Die Testmaßzahl Q_{beob} ist chiquadratverteilt mit 2 Freiheitsgraden. Für ein 5%-iges Signifikanzniveau ist daher der kritische Wert $\chi 2_c = 5.991$. Da der beobachtete Wert (= 8) größer ist als der kritische, kann die Nullhypothese abgelehnt werden. Der Trainingseffekt bei den 15 Besuchern ist groß genug, um daraus auf einen allgemeinen Trainingseffekt zu schließen.

In unten stehender Grafik ist dieses Testergebnis veranschaulicht. Die Testmaßzahl Q_{beob} (= 8) liegt rechts nach dem kritischen Wert $\chi 2_c$ (= 5.991):



(Chiquadrat_beobachtet)	(Chiquadrat_kritisch)	(p – Wahrscheinlichkeit)
8	5.99	0.317

Den Output der Software für dieses Beispiel findet man auf Seite 205.

Voraussetzungen und Einschränkungen:

Der Cochran Test sollte nur angewendet werden, wenn n. k > 30 ist. Im vorliegenden Beispiel ist 15 · 3 = 45 und daher die Voraussetzung erfüllt.

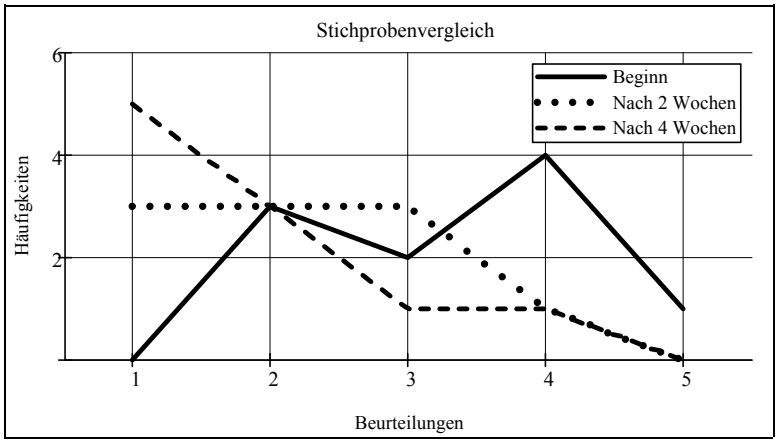
6_2 Friedman-Test

Kann man davon ausgehen, dass wiederholtes Sehen eines Fernsehspots einen Lerneffekt hat? 10 Personen wurden am Beginn einer für 4 Wochen geplanten Werbekampagne im Fernsehen danach befragt, wie sie die Handlichkeit einer neuen Spielkonsole beurteilen. Die gleichen 10 Personen wurden nach 2 Wochen und noch einmal nach 4 Wochen wieder befragt, wie sie die Handlichkeit beurteilen. Das Ergebnis zeigt folgende Tabelle (1 = sehr gut, 5 = sehr schlecht):

2	4	3	4	5	2	3	4	4	2	"0 W."
2	3	2	1	4	1	3	3	2	1	"2 W."
1	2	2	2	3	1	1	4	1	1	"4 W."

Die erste Person beurteilte die Handlichkeit am Beginn der Werbekampagne mit gut (= 2), nach 2 Wochen noch immer mit gut (= 2) und nach 4 Wochen mit sehr gut (= 1). Die Verteilung der Beurteilungen zu den drei Befragungspunkten zeigt folgende Tabelle und Grafik:

("Handlichkeit")	"Beginn"	"Nach 2 Wochen"	"Nach 4 Wochen"
("Beurteilung")	"absolut"	"absolut"	"absolut"
"sehr gut"	0	3	5
"gut"	3	3	3
"neutral"	2	3	1
"schlecht"	4	1	1
"sehr schlecht"	1	0	0
"SUMME "	10	10	10



("Beurteilungen")	"sehr gut"	"gut"	"neutral"	"schlecht"	"sehr schlecht"
"*"	1	2	3	4	5

Am Beginn der Werbekampagne beurteilte keine der Personen die Handlichkeit mit sehr gut. Nach 2 Wochen waren es schon drei Personen und nach 4 Wochen 5 Personen, die die Handlichkeit mit sehr gut bezeichneten. Ist diese Verbesserung schon groß genug, um allgemein zu behaupten, dass die Werbekampagne die Beurteilung der Handlichkeit verbesserte? (Vergleiche den Output von Seite 207)

Ja, die Verbesserung der Beurteilung ist groß genug. Man kann generell behaupten, dass die 4-wöchige Werbekampagne im Fernsehen die Beurteilung der Handlichkeit signifikant veränderte. Zu diesem Ergebnis kommt man wie folgt: Zuerst werden Null- und Alternativhypothese formuliert. Als Nullhypothese nimmt man an, dass die Verteilungen der Grundgesamtheiten, aus der die drei Stichproben stammen, gleich sind:

$$H_0: \text{Verteilung}_{\text{Beginn}} = \text{Verteilung}_{\text{Nach 2 Wochen}} = \text{Verteilung}_{\text{Nach 4 Wochen}}$$

Die Alternative dazu ist

$$H_1: \text{Verteilung}_{\text{Beginn}} \neq \text{Verteilung}_{\text{Nach 2 Wochen}} \neq \text{Verteilung}_{\text{Nach 4 Wochen}}$$

Um diese Hypothesen zu prüfen, berechnet man folgende Testmaßzahl:

$$FQ_{beob} = \left[\frac{12}{n \cdot k \cdot (k+1) - B} \cdot \sum_{j=1}^{k} \left(\sum_{i=1}^{n} R_{i,j} \right)^2 \right] - [3 \cdot n \cdot (k+1)]$$

n ist die Anzahl der Befragten, k die der Stichproben, R_i sind die Rangzahlen und B eine Korrektur für Bindungen:

$$B = \frac{1}{k-1} \cdot \sum_{j=1}^{n} \left[\left[\sum_{i=1}^{g_j} \left(b_{i,j} \right)^3 \right] - k \right]$$

Diese Testmaßzahl ist chiquadratverteilt mit n = k – 1 Freiheitsgraden. Die Berechnung der Zwischensummen zeigt folgende Tabelle:

Besucher	Beginn	Nach_2_Wo	Nach_4_Wo	$R_{i,1}$	$R_{i,2}$	$R_{i,3}$	b_i	$(b_i)^3 - 3$
1	2	2	1	2.5	2.5	1	2	5
2	4	3	2	3	2	1	0	0
3	3	2	2	3	1.5	1.5	2	5
4	4	1	2	3	1	2	0	0
5	5	4	3	3	2	1	0	0
6	2	1	1	3	1.5	1.5	2	5
7	3	3	1	2.5	2.5	1	2	5
8	4	3	4	2.5	1	2.5	2	5
9	4	2	1	3	2	1	0	0
10	2	1	1	3	1.5	1.5	2	0
SUMME	"*"	"*"	"*"	28.5	17.5	14	"*"	30

Der erste Befragte beurteilte am Beginn der Werbekampagne die Handlichkeit der Spielkonsole mit gut (= 2), nach 2 Wochen immer noch mit gut und nach 4 Wochen mit sehr gut (= 1). Da 1 die kleinste Zahl ist, wird ihr der Rang 1 zugeordnet. Die Beurteilung gut kommt zweimal vor. Daher wird beiden Beurteilungen (am Beginn und nach 2 Wochen) der Durchschnitt aus den beiden Rangzahlen 2 und 3 nämlich 2.5 zugeordnet. Die Rangsummen für die 3 Stichproben sind 28.5, 17.5 und 14.

Mit diesen Zwischenergebnissen kann man nun FQ_{beob} berechnen:

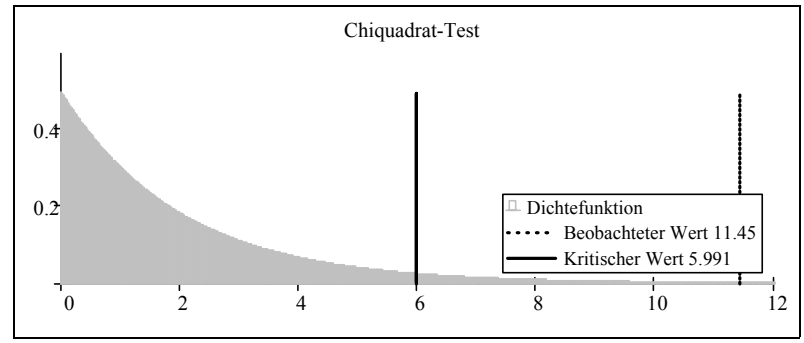
$$FQ_{beob} = \frac{12}{10 \cdot 3 \cdot (3+1) - 15} \cdot \left(28.5^2 + 17.5^2 + 14^2\right) - 3 \cdot 10 \cdot (3+1) = 30.229$$

Die Bindungskorrektur B für gleiche Rangzahlen bei den einzelnen Befragten errechnet man mit Hilfe der letzten Spalte der Tabelle:

$$B = \frac{1}{k-1} \cdot \sum_{j=1}^{n} \left[\left[\sum_{i=1}^{g_j} \left(b_{i,j}\right)^3 \right] - k \right] = \frac{1}{(3-1)} \cdot 30 = 15$$

Die Testmaßzahl FQ_{beob} ist chiquadratverteilt mit n = 3 – 1 = 2 Freiheitsgraden. Für ein 5%-iges Signifikanzniveau ist daher der kritische Wert der $\chi 2$-Verteilung $\chi 2_c = 5.991$. Da der beobachtete Wert (= 30.229) größer ist als der kritische, kann die Nullhypothese abgelehnt werden. Das Werbefernsehen verändert die Beurteilung der Handlichkeit der Spielkonsole. Das Risiko, das diese Behauptung falsch ist, ist höchstens 5%

In unten stehender Grafik ist dieses Testergebnis veranschaulicht. Die Testmaßzahl FQ_{beob} (= 11.45) liegt rechts vor dem kritischen Wert $\chi 2_c$ (= 5.991):



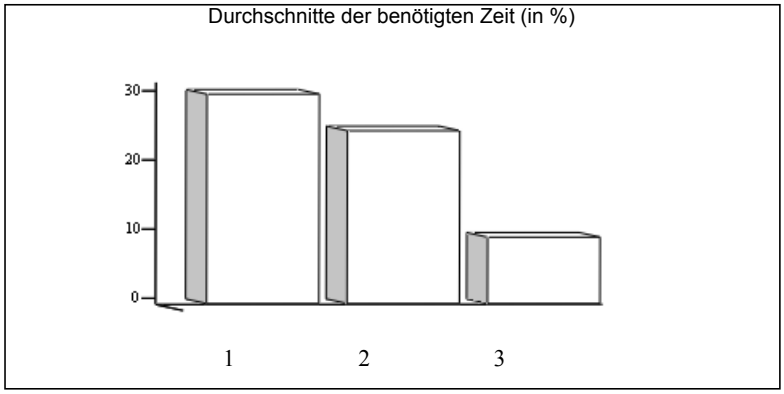
(Beobachteter_Wert)	(Kritischer_Wert)	(p – Wahrscheinlichkeit)
11.45	5.991	0.000

6_3 F-Test

Kann man davon ausgehen, dass wiederholtes Spielen einen Lerneffekt aufweist? Verringert mehrfaches Spielen des neuen Konsolenspiels "Fantasy" die benötigte Zeit zum Erreichen des Spielzieles? Bei einer Werbeveranstaltung für die neue Spielkonsole Yoki wurde die Spielzeit von 10 Besuchern bei 3 Versuchen gestoppt (in Minuten):

34	38	45	32	12	11	22	9	20	80	"1.V."
32	41	31	38	21	13	17	22	24	11	"2.V."
12	14	15	12	7	4	10	4	14	4	"3.V."

Der erste Besucher benötigte 34 Minuten beim ersten Versuch, 32 Minuten beim zweiten und nur mehr 12 beim dritten Versuch. Die Durchschnitte aus den benötigten Zeiten bei den 3 Versuchen zeigt folgende Grafik



"Grafikbalken"	1	2	3
("Versuche")	"1. Versuch"	"2. Versuch"	"3. Versuch"
"Durchschnitte"	30.3	25	9.6

Beim 1. Versuch benötigten die Besucher im Schnitt 30.3 Minuten, beim 2. Versuch 25 Minuten und beim 3. Versuch nur mehr 9.6 Minuten. Ist diese Verringerung schon groß genug, um allgemein zu behaupten, dass mehrfaches Spielen die Zeit zum Erreichen des Spielzieles des Computerspieles "Fantasy" reduziert?

Ja, die Zeitverringerung ist groß genug. Man kann generell behaupten, dass sich durch mehrfaches Spielen die Durchschnittszeit zum Erreichen des Spielziels signifikant verändert. Zu diesem Ergebnis kommt man wie folgt: Zuerst werden Null- und Alternativhypothese formuliert. Als Nullhypothese nimmt man an, dass die Durchschnitte der Grundgesamtheiten, aus der die drei Stichproben stammen, gleich sind:

$$H_0: \mu_{Versuch\_1} = \mu_{Versuch\_2} = \mu_{Versuch\_3}$$

Die Alternative dazu ist

$$H_1: \mu_{Versuch\_1} \neq \mu_{Versuch\_2} \neq \mu_{Versuch\_3}$$

Durch mehrmaliges Spielen ändert sich die Durchschnittszeit zum Erreichen des Spielziels von "Fantasy". Um diese Hypothesen zu prüfen, berechnet man folgende Testmaßzahl:

$$F_{beob} = \frac{\left[\frac{\sum_{j=1}^{k} (Q_j)^2}{n}\right] - \frac{\left(\sum_{i=1}^{n} P_i\right)^2}{n \cdot k}}{\left[\frac{\left(\sum_{i=1}^{n} P_i\right)^2}{n \cdot k} + \sum_{i=1}^{n}\sum_{j=1}^{k} (x_{i,j})^2\right] - \left[\frac{\sum_{j=1}^{k} (Q_j)^2}{n} + \frac{\sum_{i=1}^{n} (P_i)^2}{k}\right]} \cdot \frac{\nu_2}{\nu_1}$$

Diese Testmaßzahl ist F-verteilt mit

$$\nu_1 = k - 1 \qquad \nu_2 = (n - 1) \cdot (k - 1)$$

Freiheitsgraden. k ist die Anzahl der Stichproben und n die Anzahl der Befragten. P_i ist die i-te Zeilensumme

$$P_i = \sum_{j=1}^{k} x_{ij}$$

und Q_j die j-te Spaltensumme

$$Q_j = \sum_{i=1}^{n} x_{ij}$$

Die Berechnung der Zwischensummen P_i und Q_j zeigt folgende Tabelle:

(Besucher)	Versuch_1	Versuch_2	Versuch_3	P_i	$\frac{P_i}{3}$	P_i^2
1	34	32	12	78	26	6084
2	38	41	14	93	31	8649
3	45	31	15	91	30.33	8281
4	32	38	12	82	27.33	6724
5	12	21	7	40	13.33	1600
6	11	13	4	28	9.33	784
7	22	17	10	49	16.33	2401
8	9	22	4	35	11.67	1225
9	20	24	14	58	19.33	3364
10	80	11	4	95	31.67	9025
SUMME	303	250	96	649	216.33	48137

P_i ist die Summe aus den drei Versuchen und $P_i/3$ der Durchschnitt daraus. Die Summe der einzelnen Zeiten zum Quadrat ist 21571:

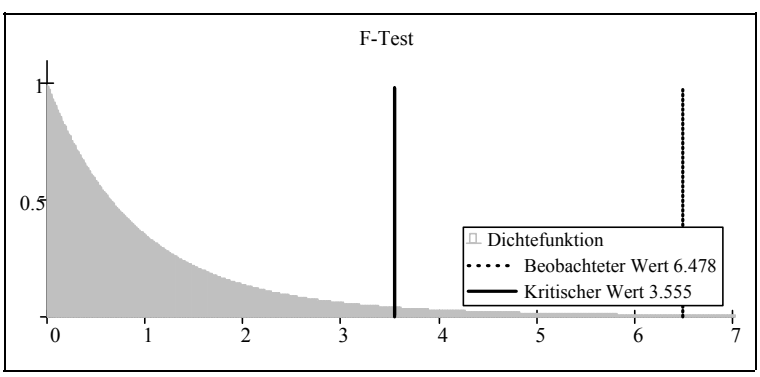
$$\sum_{i=1}^{n}\sum_{j=1}^{k}\left(x_{ij}\right)^2 = 21571$$

Mit diesen Zwischenergebnissen kann man nun F_{beob} berechnen:

$$F_{beob} = \frac{\left(\frac{303^2 + 250^2 + 96^2}{10}\right) - \frac{649^2}{30}}{\left(\frac{649^2}{30} + 21571\right) - \left(\frac{303^2 + 250^2 + 96^2}{10} + \frac{48137}{3}\right)} \cdot \frac{(10-1)\cdot(3-1)}{3-1} = 6.478$$

Die Testmaßzahl F_{beob} ist F-verteilt mit $\nu_1 = 3 - 1 = 2$ und $\nu_2 = (10 - 1).(3 - 1) = 18$ Freiheitsgraden. Für ein 5%-iges Signifikanzniveau ist daher der kritische Wert der F-Verteilung $F_c = 3.555$. Da der beobachtete Wert (= 6.478) größer ist als der kritische, kann die Nullhypothese abgelehnt werden. Das mehrfache Spielen verringert die benötigte Zeit zum Erreichen des Spielziels. Das Risiko, das diese Behauptung falsch ist, ist höchstens 5% (genau 0.8%).

In unten stehender Grafik ist dieses Testergebnis veranschaulicht. Die Testmaßzahl F_{beob} (= 6.478) liegt rechts vor dem kritischen Wert F_c (= 3.555):



(Beobachteter_Wert)	(Kritischer_Wert)	(p – Wahrscheinlichkeit)
6.478	3.555	0.008

Mit Hilfe der Software kann man dieses Ergebnis bekommen, wenn man die Daten und das gewünschte Signifikanzniveau eingibt. Auch der Paarvergleich wird automatisch durchgeführt. Für dieses Beispiel findet man den Output auf Seite 208.

7_0 MULTIVARIATE STICHPROBEN, NOMINAL

7_1 Loglineare Modelle
7_2 Kategoriale Regression
7_3 Korrespondenzanalyse

Univariat - Multivariat

Wenn man die Stichproben der Käufer der Spielkonsole Yoki, der Nichtkäufer und der Unentschlossenen im Hinblick auf ihr Durchschnittsalter analysiert, dann handelt es sich um eine univariate Analyse. Die einzige Variable die den eventuellen Unterschied zwischen den drei Grundgesamtheiten der Käufer, Nichtkäufer und Unentschlossenen misst, ist das Alter. Wenn man noch zusätzlich das Einkommen der Befragten hinzunimmt, dann liegt eine multivariate Analyse vor. Der eventuelle Unterschied wird durch die zwei Variablen Alter und Einkommen gemessen.

Diese Unterscheidung kann man nicht nur bei metrisch skalierten Antworten wie Alter und Einkommen machen, sondern auch bei nominal skalierten Antworten. Werden die Käufer der Spielkonsole, die Nichtkäufer und die Unentschlossenen z. B. im Hinblick auf die Beurteilung der Handlichkeit der Spielkonsole analysiert, dann handelt es sich um eine univariate Analyse. Wird neben der Handlichkeit auch noch die Beurteilung des Designs herangezogen, dann ist die Analyse multivariat.

Im multivariaten Fall werden also mindestens zwei Variable bei der Analyse berücksichtigt, die den Unterschied bzw. den Zusammenhang messen. Dies gilt sowohl für nominal skalierte Antworten wie ordinal und metrisch skalierte Antworten. Einige multivariate Verfahren für nominal skalierte Antworten werden in diesem Kapitel dargestellt und einige multivariate Verfahren für metrisch skalierte Antworten im folgenden Kapitel.

Vorschau

Im Abschnitt 7_1, Loglineare Modelle, wird der Frage nachgegangen, welche Wechselwirkungen zwischen der Kaufabsicht, dem Geschlecht und dem Familienbestand bestehen. Kann man z. B. allgemein annehmen, dass zwischen Kaufabsicht und Geschlecht keine Wechselwirkung besteht, aber zwischen Kaufabsicht und Familienstand?

Interessiert man sich für die Frage, wie die Kaufabsicht für die Spielkonsole Yoki durch Geschlecht und Familienstand beeinflusst wird, dann findet man die Antwort im Abschnitt 7_2, Kategoriale Regression. In diesem Abschnitt wird untersucht, ob und wie es die beiden unabhängigen Variablen Geschlecht und Familienstand erlauben, die Kaufabsicht vorher zusagen.

Im Abschnitt 7_3 wird die Korrespondenzanalyse beschrieben. Sie dient dazu, Antworten verschiedener Fragen auf Grund ihrer so genannten Euklidischen Abstände zusammen zu fassen. Dies soll eine einfache Interpretation der komplexen Zusammenhänge zwischen den Antworten erlauben. Durch welche Merkmale sind z. B. die Käufer der Spielkonsole Yoki, die Nichtkäufer und die Unentschlossenen gekennzeichnet? Ist der typische Käufer männlich, ledig und unter 25 Jahre alt?

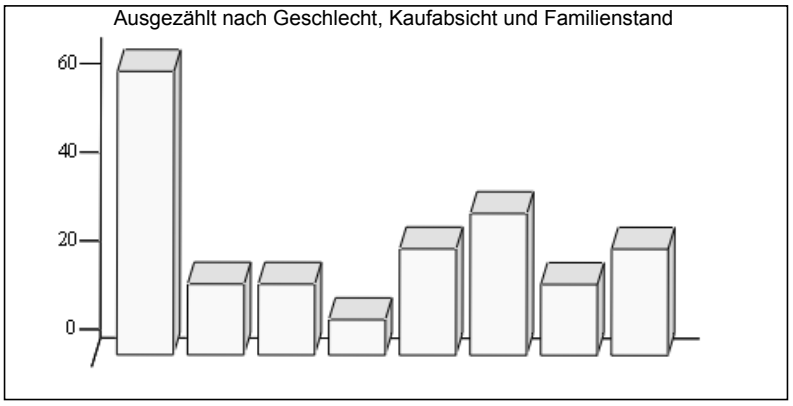
7_1 Loglineare Modelle

Welche Wechselwirkung besteht zwischen den Fragen nach dem Geschlecht, der Kaufabsicht und dem Familienstand? Welches Modell beschreibt diese Abhängigkeiten am besten? Kann man z. B. annehmen, dass männlich Ledige eine stärkere Kaufabsicht haben als weiblich Ledige oder umgekehrt? Bei einer Werbeveranstaltung für die neue Spielkonsole Yoki wurden 200 Besuchern folgende Fragen gestellt:

1. Geschlecht? mit den Antworten "Weiblich" oder "Männlich".
2. Werden Sie die Spielkonsole Yoki kaufen? mit den Antworten "Ja" oder "Nein".
3. Familienstand? mit den Antwortmöglichkeiten "Ledig" und "Verheiratet".

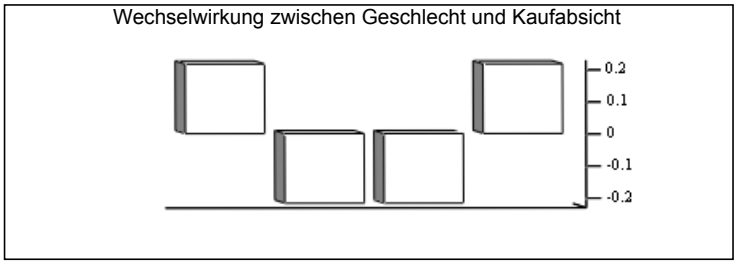
Das Ergebnis dieser Befragung zeigt folgende Tabelle:

(Geschlecht)	(Kaufabsicht)	(Familienstand)	Beobachtete_Häufigkeiten
"Männlich"	"Ja"	"Ledig"	64
"Männlich"	"Ja"	"Verheiratet"	16
"Männlich"	"Nein"	"Ledig"	16
"Männlich"	"Nein"	"Verheiratet"	8
"Weiblich"	"Ja"	"Ledig"	24
"Weiblich"	"Ja"	"Verheiratet"	32
"Weiblich"	"Nein"	"Ledig"	16
"Weiblich"	"Nein"	"Verheiratet"	24



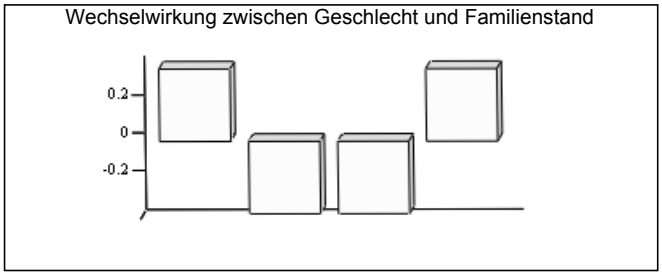
(Grafikbalken)	1	2	3	4	5	6	7	8
Häufigkeiten	64	16	16	8	24	32	15	24
"Geschlecht: Männlich oder Weiblich"	M	M	M	M	W	W	W	W
"Kaufabsicht: Ja oder Nein"	J	J	N	N	J	J	N	N
"Familienstand: Ledig oder Verheiratet"	L	V	L	V	L	V	L	V

Von den möglichen 9 Modellen zur Erklärung der Abhängigkeiten zwischen diesen drei Fragen ist das Modell am besten geeignet, das einerseits zwischen Geschlecht und Kaufabsicht Wechselwirkung annimmt und andererseits zwischen Geschlecht und Familienstand. Mit Hilfe dieses Modells kann man zeigen, dass Männer überdurchschnittlich stark eine Kaufabsicht an der Spielkonsole Yoki bekunden, Frauen hingegen unterdurchschnittlich stark.



("Geschlecht")	("Kaufabsicht")	"Koeffizienten"	"Sig.-Niveau 5 %"	"Wahrsch."
"Männlich"	"Ja"	0.217	"signifikant"	0.005
"Männlich"	"Nein"	–0.217	"signifikant"	0.005
"Weiblich"	"Ja"	–0.217	"signifikant"	0.005
"Weiblich"	"Nein"	0.217	"signifikant"	0.005

Weiteres kann man zeigen, dass ledige Männer und verheiratete Frauen überdurchschnittlich oft vorkommen und verheiratete Männer und ledige Frauen unterdurchschnittlich oft.



(Geschlecht)	(Familienstand)	"Koeffizienten"	"Sig.-Niveau 5 %"	"Wahrsch."
"Männlich"	"Ledig"	0.385	"signifikant"	0
"Männlich"	"Verheiratet"	–0.385	"signifikant"	0
"Weiblich"	"Ledig"	–0.385	"signifikant"	0
"Weiblich"	"Verheiratet"	0.385	"signifikant"	0

Wie kommt man zu diesen interessanten Ergebnissen? Zuerst muss man feststellen, welches der neun Modelle die möglichen Wechselwirkungen am besten erklärt. Hier nochmals die Ausgangstabelle:

(Geschlecht)	(Kaufabsicht)	(Familienstand)	Beobachtete_Häufigkeiten
"Männlich"	"Ja"	"Ledig"	64
"Männlich"	"Ja"	"Verheiratet"	16
"Männlich"	"Nein"	"Ledig"	16
"Männlich"	"Nein"	"Verheiratet"	8
"Weiblich"	"Ja"	"Ledig"	24
"Weiblich"	"Ja"	"Verheiratet"	32
"Weiblich"	"Nein"	"Ledig"	16
"Weiblich"	"Nein"	"Verheiratet"	24

Die beobachteten Häufigkeiten in obiger Tabelle können durch folgende Komponenten beeinflusst werden:

μ_{x1} der Haupteffekt der Variablen x1 (= Geschlecht)
μ_{x2} der Haupteffekt der Variablen x2 (= Kaufabsicht)
μ_{x3} der Haupteffekt der Variablen x3 (= Familienstand)
$\mu_{x1,x2}$ die Wechselwirkung der Variablen x1 und x2 (= Geschlecht und Kaufabsicht)
$\mu_{x1,x3}$ die Wechselwirkung der Variablen x1 und x3 (= Geschlecht und Familienstand)
$\mu_{x2,x3}$ die Wechselwirkung der Variablen x2 und x3 (= Kaufabsicht und Familienstand)
$\mu_{x1,x2,x3}$ die Wechselwirkung der Variablen x1, x2 und x3 (= Geschlecht, Kaufabsicht und Familienstand)

Wenn man sich für ein Modell ohne Wechselwirkungen entscheidet, dann hat das Modell zur Erklärung der Antworthäufigkeiten folgende Struktur:

$$\ln(m_{ijk}) = \mu + \mu_{x1_i} + \mu_{x2_j} + \mu_{x3_k}$$

bzw.

$$m_{ijk} = e^{\mu+\mu_{x1_i}+\mu_{x2_j}+\mu_{x3_k}}$$

m_{ijk} sind die erwarteten Zellhäufigkeiten, $\ln(m_{ijk})$ sind die Logarithmen der erwarteten Zellhäufigkeiten für die Zellen in der Kontingenztabelle und e ist die Euler'sche Zahl. μ ist der Gesamtdurchschnitt aus den natürlichen Logarithmen der erwarteten Häufigkeiten. μ_Z-Ausdrücke repräsentieren "Effekte", die die Variablen auf die Zellhäufigkeiten haben. x1, x2, und x3 sind Variable und i, j und k beziehen sich auf die Ausprägungen der Variablen. Für dieses Erklärungsmodell sind die berechneten Koeffizienten gleich

$$\mathbf{b} = (\,3.128 \quad 0.040 \quad -0.377 \quad -0.203\,)$$

und eingesetzt

$$m_{ijk} = e^{\mathbf{X}\,\mathbf{b}} = e^{3.128+0.04\cdot x1_i-0.377\cdot x2_j-0.203\cdot x3_k}$$

Um mit dieser Formel die Schätzwerte für die Häufigkeiten zu bestimmen, müssen die Antworten auf die drei Fragen geeignet kodiert werden, um sie in **X** einsetzen zu können.

Bei der Effektkodierung erfolgt die Darstellung der Antworten auf eine Frage mit Dummy-Variablen.

$$X_i = \begin{pmatrix} \text{"-1 falls Kategorie i der Variablen A vorliegt"} \\ \text{"1 falls Kategorie k der Variable A vorliegt"} \\ \text{"0 sonst"} \end{pmatrix}$$

So wird z. B. "Weiblich" mit 1 und "Männlich" mit –1 kodiert, "Nein" mit 1 und "Ja" mit –1, "Verheiratet" mit 1 und "Ledig" mit –1. Die möglichen Antwortkombinationen von obigen Beispiel werden daher wie folgt kodiert: (für μ, die Modellkonstante wird ein 1 Vektor eingeführt!)

(Geschlecht)	(Kaufabsicht)	(Stand)	Kon	Geschlecht	Kaufabsicht	Stand
"Männlich"	"Ja"	"Ledig"	1	–1	–1	–1
"Männlich"	"Ja"	"Verheiratet"	1	–1	–1	1
"Männlich"	"Nein"	"Ledig"	1	–1	1	–1
"Männlich"	"Nein"	"Verheiratet"	1	–1	1	1
"Weiblich"	"Ja"	"Ledig"	1	1	–1	–1
"Weiblich"	"Ja"	"Verheiratet"	1	1	–1	1
"Weiblich"	"Nein"	"Ledig"	1	1	1	–1
"Weiblich"	"Nein"	"Verheiratet"	1	1	1	1

Die Matrix **X** ist daher

$$\mathbf{X} = \begin{pmatrix} 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Zu dem Koeffizientenvektor **b** kommt man mit Hilfe der Newton-Raphson-Methode. Zuerst berechnet man einen Ausgangsvektor für $\mathbf{b}_0$ mit Hilfe folgender Formel:

$$\mathbf{b}_0 = ((\mathbf{X}^T \cdot \mathbf{X})^{-1} \,.\, \mathbf{X}^T \cdot \ln(\mathbf{n})) = \begin{pmatrix} 3.047 \\ 0.101 \\ -0.347 \\ -0.173 \end{pmatrix}$$

X ist die oben definierte Designmatrix und **n** der Vektor der beobachteten Häufigkeiten.

$$\mathbf{n}^T = (\,64 \quad 16 \quad 16 \quad 8 \quad 24 \quad 32 \quad 16 \quad 24\,)$$

Mit Hilfe dieses Ausgangsvektors $\mathbf{b}_0$ kann man folgende Schätzwerte für die beobachteten Häufigkeiten bestimmen:

$$\mathbf{m}_0 = e^{\mathbf{V}\mathbf{b}_0} = \begin{pmatrix} 32 \\ 22.627 \\ 16 \\ 11.314 \\ 39.192 \\ 27.713 \\ 19.596 \\ 13.856 \end{pmatrix}$$

Die so genannte Hesse'sche Matrix $\mathbf{H}_0$ bestimmt man nun mit Hilfe von $\mathbf{m}_0$:

$$\mathbf{H}_0 = (\mathbf{X}^T \,.\, \mathrm{diag}(\mathbf{m}_0) \,.\, \mathbf{X}) = \begin{pmatrix} 182.298 & 18.416 & -60.766 & -31.277 \\ 18.416 & 182.298 & -6.139 & -3.16 \\ -60.766 & -6.139 & 182.298 & 10.426 \\ -31.277 & -3.16 & 10.426 & 182.298 \end{pmatrix}$$

diag($\mathbf{m}_0$) ist eine Diagonalmatrix mit den Werten von $\mathbf{m}_0$ als Diagonalelemente. Die Differenz zwischen den geschätzten und beobachteten Häufigkeiten gewichtet man mit der Designmatrix $\mathbf{X}$:

$$\mathbf{q}_0 = \mathbf{X}^T \,.\, (\mathbf{n} - \mathbf{m}_0) = \begin{pmatrix} 17.702 \\ -26.416 \\ -11.234 \\ -8.723 \end{pmatrix}$$

Man multipliziert dieses Ergebnis mit der inversen Hesse'schen Matrix H_0 und addiert dieses Produkt zu den Ausgangsvektor $\mathbf{b}_0$ und erhält $\mathbf{b}_1$:

$$\mathbf{b}_1 = \mathbf{b}_0 + \mathbf{H}_0^{-1} \cdot \mathbf{q}_0 = \begin{pmatrix} 3.144 \\ -0.055 \\ -0.379 \\ -0.205 \end{pmatrix}$$

Nun berechnet man den Vektor $\mathbf{b}_2$ auf die gleiche Art indem man an Stelle von $\mathbf{b}_0$ $\mathbf{b}_1$ verwendet. $\mathbf{b}_2$ ist

$$\mathbf{b}_2 = \mathbf{b}_1 + \mathbf{H}_1^{-1} \cdot \mathbf{q}_1 = \begin{pmatrix} 3.128 \\ -0.040 \\ -0.377 \\ -0.203 \end{pmatrix}$$

Da sich $\mathbf{b}_3$ von $\mathbf{b}_2$ auf 3 Dezimalstellen nicht unterscheidet, kann man den Iterationszyklus nach $\mathbf{b}_3$ beenden, wenn man mit dieser Genauigkeit zufrieden ist. Der Gewichtungsvektor **b** ist auf drei Dezimalstellen genau

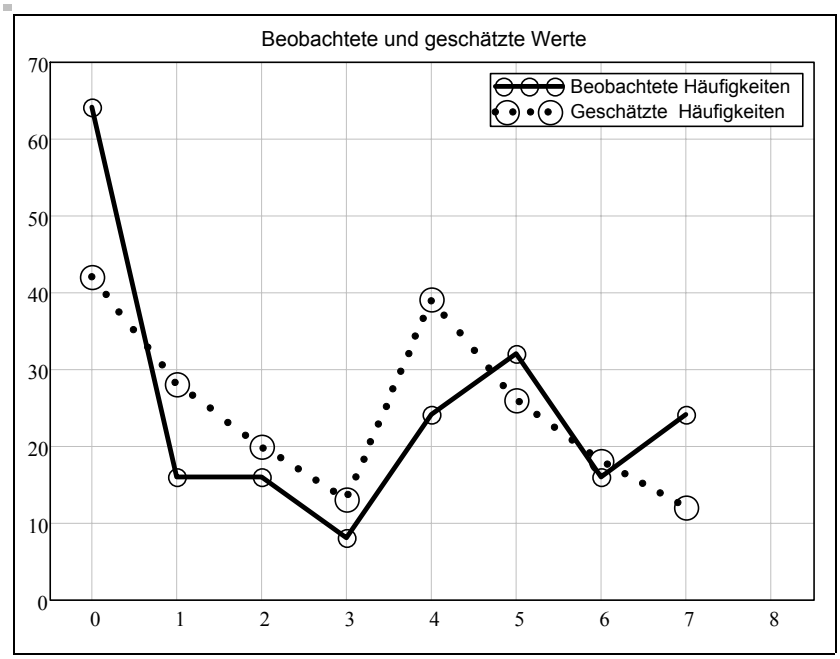
$$\begin{pmatrix} 3.128 \\ -0.040 \\ -0.377 \\ -0.203 \end{pmatrix}$$

Die beobachteten Häufigkeiten **n** können nun mit Hilfe der kodierten Antwortmatrix **X** und den ermittelten **b**-Koeffizienten geschätzt werden

$$\mathbf{m} = e^{\mathbf{Xb}} = e^{\begin{pmatrix} 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3.128 \\ -0.04 \\ -0.377 \\ -0.203 \end{pmatrix}} = \begin{pmatrix} 42 \\ 28 \\ 20 \\ 13 \\ 39 \\ 26 \\ 18 \\ 12 \end{pmatrix}$$

In folgender Tabelle sind die beobachteten Häufigkeiten **n** den geschätzten **m** gegenüber gestellt.

("Geschlecht")	("Kaufabs.")	("Stand")	"Beob."	"in %"	"Geschätzt"	"in %"
"Männlich"	"Ja"	"Ledig"	64	32	42	21
"Männlich"	"Ja"	"Verheiratet"	16	8	28	14
"Männlich"	"Nein"	"Ledig"	16	8	20	10
"Männlich"	"Nein"	"Verheiratet"	8	4	13	7
"Weiblich"	"Ja"	"Ledig"	24	12	39	20
"Weiblich"	"Ja"	"Verheiratet"	32	16	26	13
"Weiblich"	"Nein"	"Ledig"	16	8	18	9
"Weiblich"	"Nein"	"Verheiratet"	24	12	12	6



Aus der Grafik erkennt man, dass die geschätzten Werte mehr oder weniger stark von den beobachteten Häufigkeiten abweichen. Man muss daher als nächstes prüfen, ob diese Abweichungen signifikant sind. In diesem Fall wäre das Modell mit drei Haupteffekten nicht geeignet, die Beziehung zwischen Geschlecht, Kaufabsicht und Familienstand passend zu erklären. Ob die gefundenen Schätzwerte von den beobachteten signifikant abweichen, kann man mit einem Chiquadrattest oder einem Likelihood-Quotienten-Test prüfen. Die Likelihood-Quotienten-Testmaßzahl G wird nach folgender Formel berechnet:

$$G = 2 \cdot \sum_{i=1}^{q} \left(n_i \cdot \ln\left(\frac{n_i}{m_i} \right) \right)$$

n_i sind die beobachteten Häufigkeiten und m_i die geschätzten. q ist die Anzahl der Zellen. Für das Beispiel zeigt folgende Tabelle die Berechnung der Zwischensumme

n_i	m_i	$\frac{n_i}{m_i}$	$\ln\left(\frac{n_i}{m_i}\right)$	$n_i \cdot \ln\left(\frac{n_i}{m_i}\right)$
64	42.43	1.51	0.41	26.3
16	28.29	0.57	–0.57	–9.12
16	19.97	0.8	–0.22	–3.54
8	13.31	0.6	–0.51	–4.07
24	39.17	0.61	–0.49	–11.76
32	26.11	1.23	0.2	6.51
16	18.43	0.87	–0.14	–2.26
24	12.29	1.95	0.67	16.07
(SUMME)	"*"	"*"	"*"	18.13

Die Summe aus 18.13 mal 2 ist G:

$$G = 18.13 \cdot 2 = 36.26$$

(Das nicht gerundete Ergebnis ist 36.241). Die Maßzahl G ist chiquadratverteilt mit 4 Freiheitsgraden für dieses Modell. Für ein Signifikanzniveau von 5% ist der kritische Wert der Chiquadratverteilung $\chi 2 = 9.488$. Da der beobachtete G-Wert größer ist, als der kritische, muss man die Nullhypothese ablehnen. Die Abweichungen zwischen beobachteten und geschätzten Antworthäufigkeiten sind signifikant. Das Modell mit den drei Haupteffekten Geschlecht, Kaufabsicht und Familienstand eignet sich nicht für die Erklärung der Beziehungen dieser drei Fragen.

Welche weiteren Modelle stehen zur Verfügung, um die Beziehungen zwischen Geschlecht, Kaufabsicht und Familienstand zu analysieren? Wenn man alle Haupteffekte und alle Wechselwirkungen berücksichtigt, dann kommt man zum saturierten Modell.

Da in diesem Modell alle Komponenten (x1 = Geschlecht, x2 = Yoki kaufen, x3 = Familienstand, [x1,x2] = Wechselwirkung "Geschlecht+Yoki kaufen", [x1,x3] = Wechselwirkung "Geschlecht+Familienstand", [x2,x3] = Wechselwirkung "Yoki kaufen+ Familienstand", [x1,x2,x3] = Wechselwirkung "Geschlecht+Yoki kaufen+Familienstand") enthalten sind, spricht man vom saturierten Modell. Die erwarteten Zellhäufigkeiten m fallen in diesem Modell immer mit den beobachteten Häufigkeiten n zusammen. Das Modell hat 0 Freiheitsgrade. Man sucht daher ein nichtsaturiertes Modell, das nicht nur eine gute Abbildung der Abhängigkeitsstruktur bietet, sondern auch die geringste Anzahl an Komponenten aufweist. Ein Modell das die Abhängigkeit zwischen den Merkmalen ausreichend beschreibt, aber mit weniger Parametern auskommt als ein zweites Modell, wird bevorzugt.

Folgende nichtsaturierte Modelle stehen zur Wahl für die Modellierung der Beziehungsstruktur:

I) Totale Unabhängigkeit:

1. Modell [x1][x2][x3]: Es wird keinerlei Abhängigkeit zwischen den Faktoren "Geschlecht", "Yoki kaufen" und "Familienstand" angenommen. Die drei Variablen sind voneinander unabhängig.

II) Gemeinsame Unabhängigkeit:

2. Modell [x1,x2][x3]: Es wird angenommen, dass der Familienstand unabhängig ist von der Kaufabsicht und dem Geschlecht. Die Abhängigkeiten können durch Assoziationen zwischen Kaufabsicht und Geschlecht erklärt werden: Weibliche Befragte geben überzufällig an Yoki zu kaufen und/oder männliche Befragte Yoki nicht zu kaufen, und/oder Frauen geben überzufällig häufig an, Yoki nicht zu kaufen und/oder umgekehrt.

3. Modell [x1,x3][x2]: Es wird angenommen, dass die Kaufabsicht unabhängig ist vom Familienstand des Befragten und dem Geschlecht. Die Abhängigkeiten können durch Zusammenhänge zwischen Familienstand und Geschlecht erklärt werden: Weibliche Befragte sind überzufällig verheiratet und/oder männliche Befragte sind ledig, und/oder Frauen sind überzufällig häufig ledig und/oder umgekehrt.

4. Modell [x2,x3][x1]: Hier wird eine Abhängigkeit zwischen Familienstand und Kaufabsicht angenommen. Das Geschlecht spielt keine Rolle. Ledige geben überzufällig an Yoki zu kaufen und/oder Verheiratete Yoki nicht zu kaufen, und/oder Ledige geben überzufällig häufig an, Yoki nicht zu kaufen und/oder umgekehrt.

III) Bedingte Unabhängigkeit:

5. Modell [x1,x2][x2,x3]: Es wird angenommen, dass es (a) eine Abhängigkeit zwischen Geschlecht und Familienstand gibt und andererseits (b) eine Abhängigkeit zwischen Kaufabsicht und Familienstand. So kann die Tatsache, dass ein Befragter männlich ist, die Wahrscheinlichkeit dafür erhöhen, dass er auch ledig ist, unabhängig von der Kaufabsicht. Andererseits kann die Tatsache, dass ein Befragter ledig ist, auch die Kaufwahrscheinlichkeit erhöhen, unabhängig vom Geschlecht. Kaufabsicht und Geschlecht sind daher bedingt unabhängig vom Familienstand.

6. Modell [x1,x2][x1,x3]: Hier wird angenommen, dass es (a) eine Abhängigkeit zwischen Geschlecht und Kaufabsicht gibt und andererseits (b) eine Abhängigkeit zwischen Geschlecht und Familienstand. So kann die Tatsache, dass ein Befragter männlich ist, die Wahrscheinlichkeit dafür erhöhen, dass er auch ledig ist, unabhängig von der Kaufabsicht. Andererseits kann die Tatsache, dass ein Befragter männlich ist, auch die Kaufwahrscheinlichkeit erhöhen, unabhängig vom Familienstand. Kaufabsicht und Familienstand sind daher bedingt unabhängig vom Geschlecht.

7. Modell [x1,x2][x2,x3]: Es wird angenommen, dass es eine Abhängigkeit zwischen der Kaufabsicht und dem Geschlecht einerseits und zwischen Kaufabsicht und Familienstand andererseits gibt. Die Abhängigkeiten können durch Assoziationen zwischen Kaufabsicht und Geschlecht geklärt werden: Weibliche Befragte geben z. B. überzufällig an, Yoki zu kaufen, unabhängig vom Familienstand und ledige Befragte haben die Kaufabsicht, unabhängig vom Geschlecht. Geschlecht und Familienstand sind daher bedingt unabhängig von der Kaufabsicht.

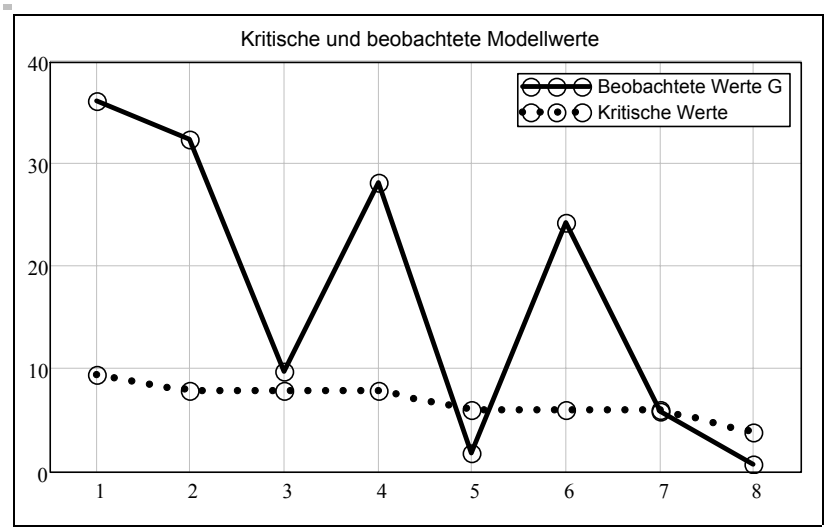
IV) Bedingte Abhängigkeit:

8. Modell [x1,x2][x1,x3][x2,x3]: Hier wird paarweise Unabhängigkeit zwischen dem Variablenpaaren Kaufabsicht und Geschlecht, zwischen Familienstand des Befragten und Geschlecht und zwischen dem Variablenpaar Kaufabsicht und Familienstand angenommen.

Die folgende Tabelle zeigt für jedes der 8 Modelle die Testmaßzahl "G", den kritischen Wert der Chiquadratverteilung "c" für ein Signifikanzniveau von 5% und die entsprechenden Freiheitsgrade "df". Als letzte Spalte wird die Wahrscheinlichkeiten für das Auftreten einer so großen oder größeren Testmaßzahl G angeführt, wenn in Wahrheit diese in der Grundgesamtheit Null ist.

("Nr")	"Modelle"	"G-Werte"	"c"	"df"	"W für G"
1	"[x1][x2][x3]"	36.241	9.488	4	0
2	"[x1][x2,x3]"	32.355	7.815	3	0
3	"[x2][x1,x3]"	9.804	7.815	3	0.02
4	"[x3][x1,x2]"	28.261	7.815	3	0
5	"[x1,x2][x1,x3]"	1.824	5.991	2	0.402
6	"[x1,x2][x2,x3]"	24.375	5.991	2	0
7	"[x1,x3][x2,x3]"	5.918	5.991	2	0.052
8	"[x1,x2][x1,x3][x2,x3]"	0.739	3.841	1	0.39

In der Grafik wird der beobachtete Likelihoodquotient G dem theoretischen Wert c der Chiquadratverteilung gegenübergestellt, wobei c unter der Voraussetzung bestimmt wird, dass in der Grundgesamtheit G Null ist.



("Modell Nr")	1	2	3	4	5	6	7	8
"W für G"	0.000	0.000	0.020	0.000	0.402	0.000	0.052	0.390

Man kann der Grafik (oder der Tabelle) entnehmen, dass die beobachteten G-Werte der ersten 4 Modelle alle größer sind als die entsprechenden kritischen Werte c. D.h. keines dieser Modelle ist zur Modellierung der Abhängigkeit zwischen den Variablen "Geschlecht", "Kaufabsicht" und "Familienstand" geeignet. Beim fünften und siebten Modell sind die G-Werte kleiner als die c-Werte. Dies gilt auch für das achte Modell. Welches Modell soll nun herangezogen werden für die Erklärung der Abhängigkeitsstruktur von Geschlecht, Kaufabsicht und Familienstand? Folgende Grafik zeigt die Modellhierarchie für eine dreidimensionale Kontingenztabelle:

$$\begin{bmatrix} \text{"Modell 1: [x1][x2][x3]"} \\ \text{("Modell 2: [x3][x1,x2]" "Modell 3: [x2][x1,x3]" "Modell 4: [x1][x2][x3]")} \\ \text{("Modell 7: [x1,x3][x2,x3]" "Modell 6: [x1,x2][x2,x3]" "Modell 5: [x1,x2][x1,x3]")} \\ \text{"Modell 8: [x1,x2][x1,x3][x2,x3]"} \end{bmatrix}$$

Nach Goodman geht man vom einfachsten Modell aus, das ist hier das Modell 1 (= [x1],[x2],[x3]). Lehnt man dieses Modell nicht ab, dann beendet man die Suche und verwendet dieses Modell zur Erklärung der Abhängigkeitsstruktur. Da der beobachtete Wert (G = 36.241) dieses Modells größer ist als der kritische Wert (c = 9.488), lehnt man dieses Modell ab. Die Anpassung ist nicht ausreichend. Nun vergleicht man dieses Modell 1 mit den drei nachgelagerten Modellen 2 (= [x1,x2],[x3]), 3 (= [x1,x3],[x2]) und 4 (= [x2,x3],[x1]). Man untersucht, ob die Anpassung durch den Übergang von Modell 1 zu Modell 2, 3 oder 4 signifikant verbessert wird. Dazu berechnet man die Differenzen aus den G-Werten.

$$G(1) - G(2) = 36.241 - 32.355 = 3.886$$
$$G(1) - G(3) = 36.241 - 9.804 = 26.437$$
$$G(1) - G(4) = 36.241 - 28.261 = 7.98$$

Die größte Verbesserung tritt beim Modell 3 ([x1,x3],[x2]) auf. Da diese Differenzen chiquadratverteilt sind mit Freiheitsgraden, die sich aus der Differenz der beiden Modelle ergeben, kann man prüfen, ob der Übergang die Anpassung signifikant verbessert. Da die Wahrscheinlichkeit einer so großen oder noch größeren Differenz (= 26.437) kleiner ist als das Signifikanzniveau von 5%, wird die Nullhypothese abgelehnt.

$$1 - F_{\chi 2}(26.437,\ 1) = 0.000$$

Die Verbesserung ist also signifikant. Nun prüft man, ob das Modell 3 ([x1,x3],[x2]) gut an die beobachteten Daten angepasst ist. Aus obiger Tabelle (oder Grafik) kann man entnehmen, dass dies nicht der Fall ist. Man testet daher, ob die Anpassung durch den Übergang von Modell 3 zu Modell 5 oder 7 signifikant verbessert wird.

$$G(3) - G(5) = 9.804 - 1.824 = 7.98$$

$$G(3) - G(7) = 9.804 - 5.918 = 3.886$$

Beim Übergang von Modell 3 auf Modell 5 tritt die größte Verbesserung auf. Diese Verbesserung ist signifikant. Da dieses Modell nicht signifikant von den beobachteten Werten abweicht (siehe Tabelle oder Grafik), ist das am besten geeignete Modell für die Beschreibung der Abhängigkeitsstruktur gefunden: Die Variablen "Geschlecht" und "Kaufabsicht" sind bedingt unabhängig von den Variablen "Geschlecht" und "Familienstand".

$$\text{Bestes\_Modell} = [\ 5 \quad \text{"[x1,x2][x1,x3]"} \quad (\ \text{"Geschlecht, Kaufabsicht"} \quad \text{"Geschlecht, Familienstand"}\ )\ ]$$

Um zu testen, ob die ermittelten Schätzwerte $\mathbf{b}$ für $\boldsymbol{\beta}$ signifikant von Null abweichen, benötigt man die Varianz-Kovarianzmatrix dieser Schätzungen. Dies ist die Inverse der schon bekannten Hesse'schen Matrix $\mathbf{H}$:

$$\text{Var}(\mathbf{b}) = \mathbf{H}^{-1} = (\mathbf{X}^{T} \cdot \text{diag}(\mathbf{m}) \cdot \mathbf{X})^{-1} =$$

$$= \begin{pmatrix} 200 & -8 & -72 & -40 & 40 & 72 & 40 \\ -8 & 200 & 40 & 72 & -72 & -40 & -29.781 \\ -72 & 40 & 200 & 40 & -8 & -29.781 & -40 \\ -40 & 72 & 40 & 200 & -29.781 & -8 & -72 \\ 40 & -72 & -8 & -29.781 & 200 & 40 & 72 \\ 72 & -40 & -29.781 & -8 & 40 & 200 & 40 \\ 40 & -29.781 & -40 & -72 & 72 & 40 & 200 \end{pmatrix}$$

Mit ihrer Hilfe kann man prüfen, ob die Koeffizienten b_i signifikant von Null abweichen. Dazu berechnet man die Wald-Testmaßzahl

$$W_i = \frac{(b_i)^2}{Var(b_i)}$$

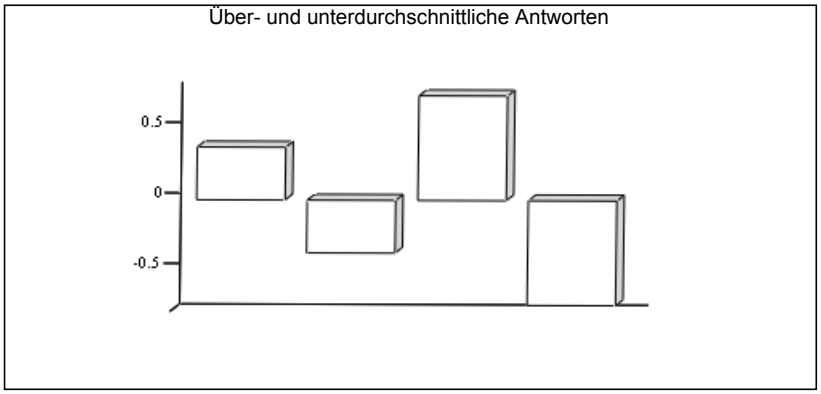
mit den Diagonalelementen obiger Varianz- Kovarianzmatrix als Var(bi). Diese ist chiquadratverteilt mit einem Freiheitsgrad. In folgender Tabelle ist einerseits festgehalten, ob dieser Test zur Ablehnung der Nullhypothese bei einem 5%-igem Signifikanzniveau führt. Andererseits wird auch die Wahrscheinlichkeit angegeben, dass so eine große oder größere Testmaßzahl auftritt, wenn in der Grundgesamtheit diese Maßzahl Null ist.

"Geschlecht [x1]"	"Signifikanz"	("Wahrscheinlichkeit")
0.104	"nicht signifikant"	0.220
"Yoki kaufen [x2]"	"*"	"*"
−0.371	"signifikant"	0.000
"Familienstand [x3]"	"*"	"*"
−0.187	"signifikant"	0.024
"Wechselwirkung [x1,x2]"	"*"	"*"
0.187	"signifikant"	0.024
"Wechselwirkung [x1,x3]"	"*"	"*"
0.371	"signifikant"	0.000
"Wechselwirkung [x2,x3]"	"*"	"*"
0.086	"nicht signifikant"	0.297

Von den Faktoren hat lediglich der Haupteffekt "Geschlecht" keinen signifikanten Einfluss auf die Häufigkeitsschätzung, von den Wechselwirkungen nur die Abhängigkeit x2 (= Kaufabsicht) und x3 (= Familienstand) nicht. (Vergleiche den Output der Software für dieses Beispiel auf Seite 209)

7_2 Kategoriale (Logistische) Regression

Kann man vom Geschlecht und Familienstand potentieller Kunden der neuen Spielkonsole Yoki auf die Kaufabsicht schließen? (Vergleiche den Output für dieses Beispiel auf Seite 211). Folgende Grafik zeigt das Analyseergebnis von 200 Befragten:



(Nummer)	1	2	3	4
Antwort	"Männlich"	"Weiblich"	"Ledig"	"Verheiratet"
Wahrscheinlichkeit_p	0.024	0.024	0.000	0.000

Sowohl das Geschlecht des Befragten als auch sein Familienstand haben auf die Kaufabsicht einen signifikanten Einfluss: Männliche Befragte und Ledige weisen eine überdurchschnittliche Kaufabsicht auf, weibliche Befragte und Verheiratete eine unterdurchschnittliche Kaufabsicht.

Bei einer Werbeveranstaltung für die neue Spielkonsole Yoki wurden 200 Besuchern folgende Fragen gestellt:

1. Werden Sie die Spielkonsole Yoki kaufen? mit den Antworten "Ja" oder "Nein".

2. Geschlecht? mit den Antworten "Männlich" oder "Weiblich".

3. Familienstand? mit den Antwortmöglichkeiten "Ledig" und "Verheiratet".

Das Ergebnis dieser Befragung zeigt folgende Tabelle:

Kaufabsicht	Geschlecht	Familienstand	(Häufigkeit)
"Ja"	"Männlich"	"Ledig"	64
"Ja"	"Männlich"	"Verheiratet"	16
"Ja"	"Weiblich"	"Ledig"	16
"Ja"	"Weiblich"	"Verheiratet"	8
"Nein"	"Männlich"	"Ledig"	24
"Nein"	"Männlich"	"Verheiratet"	32
"Nein"	"Weiblich"	"Ledig"	16
"Nein"	"Weiblich"	"Verheiratet"	24

Man kann diese Ergebnisse auch folgendermaßen zusammenfassen:

Geschlecht	Familienstand	Nichtkaufabsicht	Kaufabsicht	Anteile	Odds	(Logit)
"Männlich"	"Ledig"	24	64	0.727	2.667	0.981
"Männlich"	"Verheiratet"	32	16	0.333	0.500	–0.693
"Weiblich"	"Ledig"	16	16	0.500	1.000	0.000
"Weiblich"	"Verheiratet"	24	8	0.250	0.333	–1.099

Von den befragten ledigen Männern beabsichtigen 64 die Spielkonsole Yoki zu kaufen und 24 nicht zu kaufen. 72.7% der ledigen Männer haben also Kaufabsicht (64 / (64 +24) = 0.727, siehe 5. Spalte und 1. Zeile) und 27.3% keine. Bezieht man die ledigen Männer mit Kaufabsicht auf jene ohne Kaufabsicht, dann erhält man die so genannten Odds:

$$\text{Odds}_1 = \frac{p_1}{1 - p_1} = \frac{0.727}{1 - 0.727} = 2.667$$

Die Kaufabsicht ist also bei den ledigen Männern 2.667 Mal größer als die Nichtkaufabsicht (siehe Spalte 6 und Zeile 1). Wenn man die Odds logarithmiert, dann erhält man die Logits:

$$\text{Logit}_1 = \ln\left(\frac{p_1}{1 - p_1}\right) = \ln(2.667) = 0.981$$

Wozu dienen die Logits? Man benötigt sie, um vom Geschlecht und Familienstand der Befragten auf die Kaufabsicht zu schließen. Geschlecht und Familienstand als die unabhängigen Variablen sind linear mit der abhängigen Variable der Logits der Kaufabsicht verbunden. Diese so genannte Regressionsfunktion ermöglicht die Berechnung von Schätzwerten für die Logits der Kaufabsicht.

$$\ln\left(\frac{p}{1 - p}\right) = \mathbf{X} \,.\, \mathbf{b} = b_0 + b_1 \cdot x_{i1} + b_2 \cdot x_{i2} = -0.208 - 0.374 \cdot x_{i1} - 0.742 \cdot x_{i2}$$

Um die unabhängigen Variablen „Geschlecht“ und „Familienstand“ in der Analyse verwenden zu können, müssen die einzelnen Antworten durch Dummy-Variable kodiert werden. Bei der Effekt-Kodierung erfolgt die Darstellung der Antworten auf eine Frage mit Dummy-Variablen.

$$X_i = \begin{pmatrix} \text{"-1 falls Kategorie i der Variablen A vorliegt"} \\ \text{"1 falls Kategorie k der Variable A vorliegt"} \\ \text{"0 sonst"} \end{pmatrix}$$

So wird z. B. "Männlich" mit −1 und "Weiblich" mit 1 kodiert, "Ledig" mit −1 und "Verheiratet" mit 1. Die möglichen Antwortkombinationen von obigem Beispiel werden daher wie folgt kodiert

$$\mathbf{X} = \begin{pmatrix} 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} \text{"Männlich"} & \text{"Ledig"} \\ \text{"Männlich"} & \text{"Verheiratet"} \\ \text{"Weiblich"} & \text{"Ledig"} \\ \text{"Weiblich"} & \text{"Verheiratet"} \end{pmatrix}$$

Mit Hilfe dieser Kodierung und der oben angeführten linearen Regressionsfunktion kann man nun z. B. einen Schätzwert für die Logits von ledigen Männern berechnen.

$$\ln\left(\frac{p_1}{1-p_1}\right) = -0.208 - 0.374\cdot(-1) - 0.742\cdot(-1) = 0.908$$

Der tatsächlich beobachtete Wert war 0.981. Der Schätzwert von 0.908 weicht nur wenig von diesem ab. Exponentiert man beide Seiten und löst nach p_1 auf, dann erhält man einen Schätzwert für den Anteil lediger Männer mit Kaufabsicht:

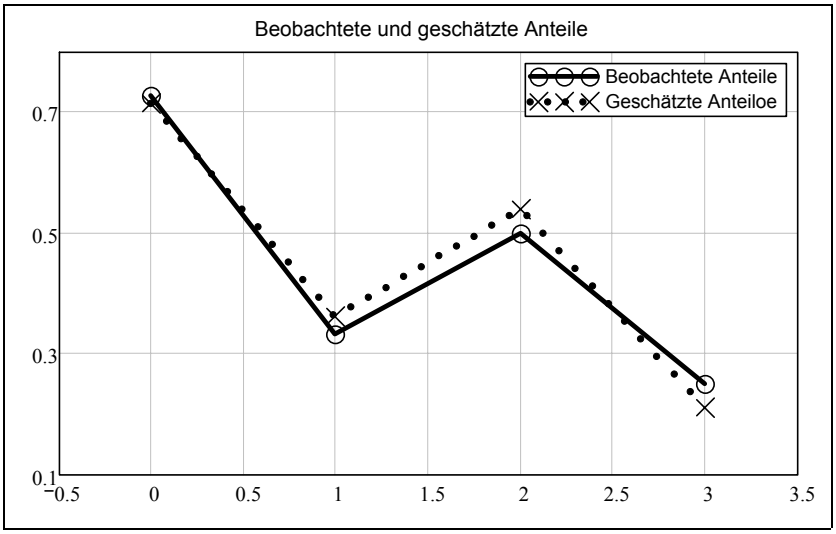
$$\ln\left(\frac{p_1}{1-p_1}\right) = 0.908$$

$$\frac{p_1}{1-p_1} = e^{0.908}$$

$$p_1 = 0.713$$

Der beobachtete Anteil der ledigen Männer mit Kaufabsicht war 0.727. Auch hier zeigt sich nur eine geringe Abweichung des Schätzwertes vom beobachteten Wert. Die folgende Tabelle zeigt alle beobachteten und geschätzten Werten der Stichprobe ebenso die Grafik.

"Geschlecht"	"Familienstand"	"Beobachtet in %"	"Geschätzt in %"
"Männlich"	"Ledig"	72.7	71.3
"Männlich"	"Verheiratet"	33.3	36.0
"Weiblich"	"Ledig"	50.0	54.0
"Weiblich"	"Verheiratet"	25.0	21.0



("Nr.")	0	1	2	3
"Geschlecht"	"Männlich"	"Männlich"	"Weiblich"	"Weiblich"
"Familienstand"	"Ledig"	"Verheiratet"	"Ledig"	"Verheiratet"

Man kann die Transformation der geschätzten Logits in Anteilsschätzwerte für die gesamte Regressionsfunktion durchführen. Aus der linearen Regressionsfunktion für Logits wird dadurch eine (nichtlineare) logistische Regressionsfunktion für Anteilswerte.

$$\ln\left(\frac{p_S}{1 - p_S}\right) = \mathbf{X} \,.\, \mathbf{b}$$

$$p_S = e^{\mathbf{Xb}}/(1 + e^{\mathbf{Xb}}) = 1 / (1 + e^{-\mathbf{Xb}})$$

Die Regressionskoeffizienten **b** werden mit Hilfe der Maximum-Likelihood-Methode bestimmt. Da dafür keine Lösung in geschlossener Form angeb bar ist. wird die Lösung iterativ bestimmt. Das Iterationsschema für das binäre Logit-Modell lautet:

$$\mathbf{b}^{(k)} = \mathbf{b}^{(k-1)} + (\mathbf{H}^{(k-1)})^{-1} \,.\, \mathbf{s}^{(k-1)}$$

$\mathbf{b}^{(k)}$ ist die k-te Approximation des Regressionkoeffizientenvektors, **H** ist die Hesse'sche Matrix, die als Elemente die zweiten partiellen Ableitungen der Maximum-Likelihood-Funktion enthält und **s** ist der Score-Vektor, der die erste Ableitung der Log-Likelihoodfunktion nach **b** enthält:

$$s(\mathbf{b}) = \frac{d}{d\beta} l(\beta, y)$$

mit der Log-Likelihoodfunktion

$$l(\mathbf{b}, \mathbf{y}) = \ln(L(\mathbf{b}, \mathbf{y})) = \sum_i \left[y_i \cdot \ln(p_i) + (1 - y_i) \ln(1 - p_i) \right]$$

Als Startwert für die Berechnung des Regressionkoeffizientenvektors wird der Nullvektor verwendet. Für obiges Beispiel ist also

$$\mathbf{b}^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

und $\mathbf{p}_S$ gleich

$$\mathbf{p}_S = e^{\mathbf{Xb}}/(1 + e^{\mathbf{Xb}}) = 1 / (1 + e^{-\mathbf{Xb}}) = \begin{pmatrix} 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \end{pmatrix}$$

mit

$$\mathbf{X} = \begin{pmatrix} 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

p ist der Anteilsvektor, berechnet aus den Beobachtungen

$$\mathbf{p} = \begin{pmatrix} 0.727 \\ 0.333 \\ 0.500 \\ 0.250 \end{pmatrix}$$

und **n** ist die Anzahl an Beobachtungen für die Antwortkombinationen

$$\mathbf{n} = \begin{pmatrix} 88 \\ 48 \\ 32 \\ 32 \end{pmatrix}$$

$\mathbf{s}^{(0)}$ ergibt sich daraus als

$$\mathbf{s}^{(0)} = \mathbf{X} \,.\, (\mathbf{n} \,.\, (\mathbf{p} - \mathbf{ps})) = \begin{pmatrix} 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}^T \cdot \begin{pmatrix} 19.976 \\ -8.016 \\ 0 \\ -8 \end{pmatrix} = \begin{pmatrix} 4 \\ -20 \\ -36 \end{pmatrix}$$

Die Hesse'sche Matrix **H** berechnet man nach der Formel

$$\mathbf{H}^{(0)} = \mathbf{X}^{T} \cdot \mathbf{W} \cdot \mathbf{X} =$$

$$= \begin{pmatrix} 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}^{T} \cdot \begin{pmatrix} 22 & 0 & 0 & 0 \\ 0 & 12 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 50 & -18 & -10 \\ -18 & 50 & 10 \\ -10 & 10 & 50 \end{pmatrix}$$

Die Diagonalelemente der Matrix **W** sind $n_i \cdot p_S \cdot (1 - p_S)$. 22 ist z. B. gleich

$$22 = 88 \,.\, 0.5 \,.\, (1 - 0.5)$$

Die erste Approximation des Regressionskoeffizientenvektors ist nun

$$\mathbf{b}^{(1)} = \mathbf{b}^{(0)} + (\mathbf{H}^{(0)})^{-1} \,.\, \mathbf{s}^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 50 & -18 & -10 \\ -18 & 50 & 10 \\ -10 & 10 & 50 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 4 \\ -20 \\ -36 \end{pmatrix} = \begin{pmatrix} -0.175 \\ -0.325 \\ -0.69 \end{pmatrix}$$

Nun wird dieser erste Regressionsvektor als Ausgangspunkt für die nächste Iteration verwendet. In der Tabelle sind die ersten 6 Approximationen ausgewiesen. Da sich zwischen der 5. und 6. Approximation die Regressionskoeffizienten auf drei Dezimalstellen nicht mehr ändern, kann man nach der 5. Iteration abbrechen.

Appr	App_0	App_1	App_2	App_3	App_4	App_5
b_1	−0.175	−0.199	−0.205	−0.207	−0.208	−0.208
b_2	−0.325	−0.362	−0.371	−0.374	−0.374	−0.374
b_3	−0.69	−0.73	−0.739	−0.741	−0.742	−0.742

Um zu prüfen, ob die Abweichungen zwischen beobachteten und geschätzten Anteilen nur zufällig oder schon signifikant sind, verwendet man die Pearson´sche Chiquadratstatistik: Man bildet die Differenz zwischen der relativen Häufigkeit p_i und der aus dem Modell resultierenden Schätzung p_{Si}, quadriert diese und dividiert sie durch die Varianz $p_{Si} \cdot (1 - p_{Si})$. Die Summe daraus liefert die Testmaßzahl.

$$\chi 2_{beob} = \sum_i \left[n_i \cdot \frac{(p_i - ps_i)^2}{ps_i \cdot (1 - ps_i)} \right]$$

Diese Maßzahl ist chiquadratverteilt mit

$$\nu = n - g$$

Freiheitsgraden. g ist die Anzahl der Parameter und n die Anzahl der Ausprägungskombinationen. Im vorliegenden Beispiel ist n = 3 – 2 = 1 und $\chi 2$ wird in folgender Tabelle berechnet:

p_i	p_{s_i}	$(p_i - p_{s_i})^2$	$p_{s_i} \cdot (1 - p_{s_i})$	n_i	$n_i \cdot \frac{(p_i - ps_i)^2}{ps_i \cdot (1 - ps_i)}$
0.727	0.713	0.000	0.205	88	0.091
0.333	0.36	0.001	0.230	48	0.148
0.500	0.54	0.002	0.248	32	0.205
0.250	0.21	0.002	0.166	32	0.307
"*"	"*"	"*"	SUMME	200	0.751

$\chi 2_{beob} = 0.751$

Für ein Signifikanzniveau von 5% ist der kritische Wert der Chiquadratverteilung $\chi 2 = 3.841$. Da der beobachtete $\chi 2$-Wert kleiner ist, als der kritische, kann man die Nullhypothese nicht ablehnen. Die Abweichungen zwischen beobachteten und geschätzten Anteilswerten für die Kaufabsicht sind nicht signifikant. Das Regressionsmodell mit den 2 unabhängigen Variablen Geschlecht und Familienstand ist geeignet, die abhängige Variable Kaufabsichtsanteile vorher zusagen. Mit Hilfe des Wald-Tests prüft man, ob die einzelnen Regressionskoeffizienten b_1 und b_2 einen signifikanten Einfluss auf die abhängige Variable haben. Die Testmaßzahl für den Wald-Test ist

$$W_i = \frac{(b_i)^2}{(H^{-1})_{i,i}}$$

b_i sind die Regressionskoeffizienten und $H^{-1}{}_{i,i}$ sind die Diagonalelemente der inversen Hesse'sche Matrix, der Varianz- Kovarianzmatrix dieser Regressionskoeffizienten. Die Hesse'sche Matrix **H** berechnet man nach folgender Formel:

$$\mathbf{H} = \mathbf{X}^T \cdot \mathbf{W} \cdot \mathbf{X}$$

W ist eine Diagonalmatrix mit

$$n_i \cdot \left[p_{s_i} \cdot (1 - p_{s_i}) \right]$$

als Diagonalelemente.

Für das Beispiel ist **W** gleich

$$\mathbf{W} = \begin{pmatrix} 18.017 & 0 & 0 & 0 \\ 0 & 11.058 & 0 & 0 \\ 0 & 0 & 7.949 & 0 \\ 0 & 0 & 0 & 5.31 \end{pmatrix}$$

(z. B. 18.017 = 88 . 0.713 . (1 – 0.713)) und **H** ist

$$\mathbf{H} = \begin{pmatrix} 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}^{T} \cdot \begin{pmatrix} 18.017 & 0 & 0 & 0 \\ 0 & 11.058 & 0 & 0 \\ 0 & 0 & 7.949 & 0 \\ 0 & 0 & 0 & 5.31 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 42.334 & -15.816 & -9.598 \\ -15.816 & 42.334 & 4.32 \\ -9.598 & 4.32 & 42.334 \end{pmatrix}$$

Mit Hilfe der Diagonalelemente der inversen Hesse'schen Matrix **H** kann man die Wald-Testmaßzahlen bestimmen:

$$\mathbf{H}^{-1} = \begin{pmatrix} 0.029 & 0.01 & 0.005 \\ 0.01 & 0.027 & -0.001 \\ 0.005 & -0.001 & 0.025 \end{pmatrix}$$

$$W_1 = \frac{(b_1)^2}{\left(H^{-1}\right)_{1,1}} = \frac{(-0.208)^2}{0.029} = 1.507$$

$$W_2 = 5.104$$

$$W_3 = 22.115$$

W_i ist chiquadratverteilt mit 1 Freiheitsgrad. Für ein 5%-iges Signifikanzniveau ist $\chi 2_c$ gleich 3.841. Daher ist sowohl der Regressionskoeffizient des Geschlechts als auch des Familienstandes signifikant von 0 verschieden.

Wenn die unabhängige Variable Regressionskoefizienten hat, die signifikant von 0 verschieden sind, dann ist es sinnvoll festzustellen, welchen Einfluss die einzelnen Antworten der unabhängigen Variablen auf die abhängige haben. Dazu multipliziert man die jeweilige Designmatrix **X** mit den entsprechenden Regressionskoeffizienten. Für die unabhängige Variable "Geschlecht" und "Familienstand" erhält man folgende Koeffizienten:

$$-0.374 \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.374 \\ -0.374 \end{pmatrix} = \begin{pmatrix} \text{"Männlich"} \\ \text{"Weiblich"} \end{pmatrix}$$

$$-0.742 \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.742 \\ -0.742 \end{pmatrix} = \begin{pmatrix} \text{"Ledig"} \\ \text{"Verheiratet"} \end{pmatrix}$$

An diesen Koeffizienten ist ersichtlich, dass männliche potentielle Käufer, die ledig sind, überdurchschnittlich starke Kaufabsicht für die Spielkonsole Yoki zeigen und weibliche, verheiratete Kunden unterdurchschnittliche Kaufabsicht.

7_3 Korrespondenzanalyse

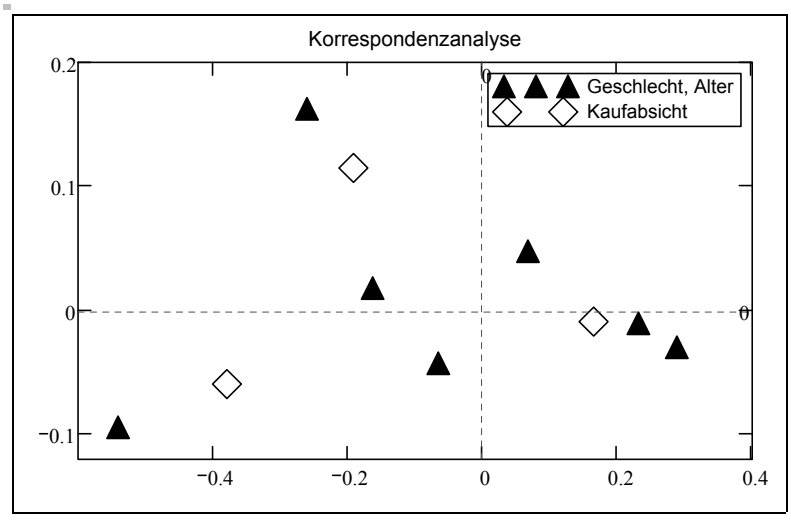
Durch welche Merkmale ist der Käufer bzw. Nichtkäufer der neuen Spielkonsole Yoki gekennzeichnet? Kann man behaupten, dass der typische Käufer männlich und unter 25 Jahre alt ist? Um das Ergebnis vorwegzunehmen, ja man kann behaupten, der typische Yoki-Käufer ist männlich und unter 25 Jahre alt. Der Nichtkäufer ist weiblich und 25 bis 50 Jahre alt.

Um das Kaufverhalten potentieller Kunden für die neue Spielkonsole zu eruieren, wurden bei einer Werbeveranstaltung den Besuchern folgende vier Fragen gestellt:
Werden Sie die Spielkonsole kaufen? (1 = Ja, 2 = Nein, 3 = Weiß nicht)
Geschlecht? (1 = Weiblich, 2 = Männlich)
Wie alt sind Sie? (1 = bis 25 Jahre, 2 = 25 bis 50 Jahre, 3 = 50 und mehr Jahre)
Familienstand? (1 = Nicht verheiratet, 2 = verheiratet)

Die ausgezählten Antworten zeigt folgende Tabelle:

(Antworten)	Ja	Nein	Weiss_nicht
"Weiblich"	9	3	4
"Männlich"	7	1	1
"bis 25 Jahre"	12	2	2
"25 bis 50 Jahre"	2	1	1
"50 Jahre und mehr"	2	1	2
"Nicht verheiratet"	8	2	2
"Verheiratet"	8	2	3

In folgender Grafik sind die Ergebnisse der Korrespondenzanalyse dargestellt:



In den 4 Quadranten sind folgende Antworten zu finden:

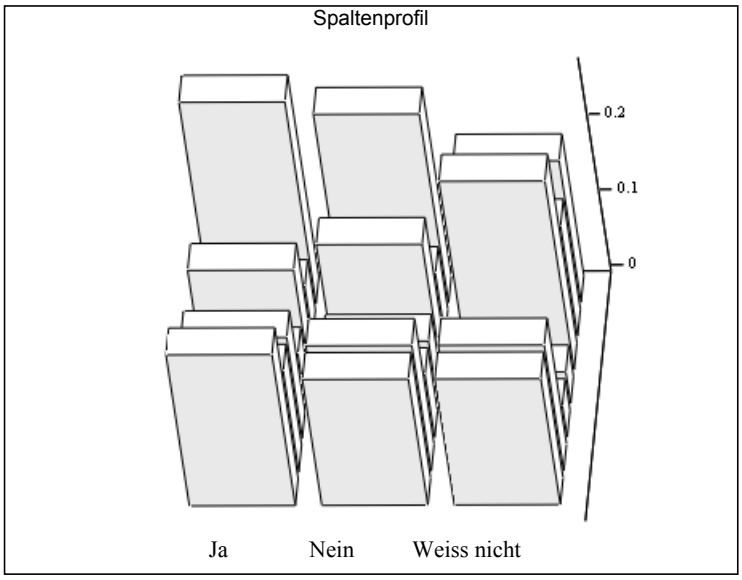
"Geschlecht - Weiblich: 0.99" "Alter - 25 bis 50 Jahre: 0.72" "Kaufabsicht - Nein: 0.74"	("Familienstand - Nicht verheiratet: 0.69")
"Alter - 50 Jahre und mehr: 0.97" "Familienstand - Verheiratet: 0.69" "Kaufabsicht - Weiss nicht: 0.98"	"Geschlecht - Männlich: 0.99" "Alter - bis 25 Jahre: 0.99" "Kaufabsicht - Ja: 0.99"

Berücksichtigt man nur jene Antworten, die den Faktor hoch laden, dann sieht man, dass dies die Antworten "Kaufabsicht-Nein" mit "Geschlecht-weiblich" und "Alter-25 bis 50 Jahre" im 2. Quadranten sind, im 3. Quadranten "Alter-50 Jahre und mehr" sowie "Kaufabsicht-Weiß nicht" und im 4. Quadranten die Antworten "Kaufabsicht-Ja" mit "Geschlecht-männlich" und Alter-bis 25 Jahre". Alle anderen Antwortkombinationen laden diesen so genannten Kauffaktor nicht hoch. Hoch laden heißt, die Antworten korrelieren mit dem Faktor mit mindestens 0.7. Bei einem Korrelationskoeffizienten von 0.7 werden rund 50% ($0.7^2 = 0.49$) der gemeinsamen Varianz erklärt.

Die potentiellen Käufer der Spielkonsole Yoki findet man also besonders häufig unter den Männern bis 25 Jahre und die Nichtkäufer unter den Frauen zwischen 25 und 50 Jahren.

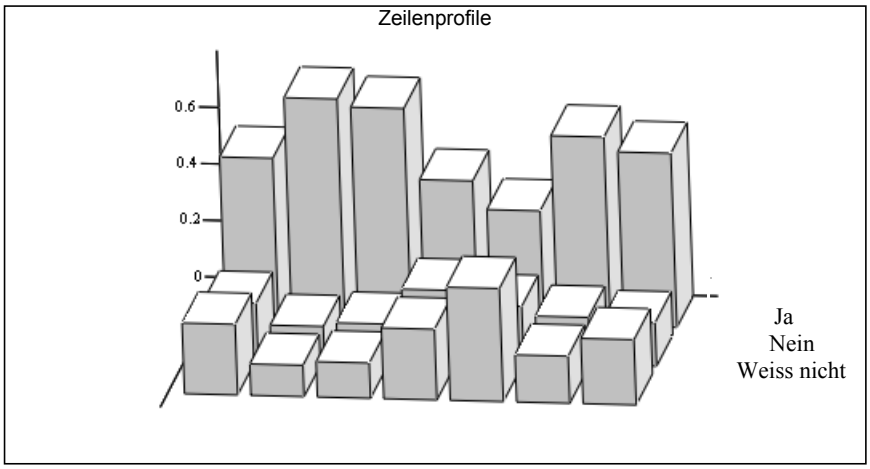
Wie kommt man zu diesen Ergebnissen? Zuerst bezieht man die Häufigkeiten der Antwortkombinationen auf die jeweiligen Spaltensummen:

("SPALTENPROFIL")	"Ja"	"Nein"	"Weiss nicht"
"Zeilen"	"Anteil"	"Anteil"	"Anteil"
"Geschlecht - Weiblich"	0.188	0.250	0.267
"Geschlecht - Männlich"	0.146	0.083	0.067
"Alter - bis 25 Jahre"	0.250	0.167	0.133
"Alter - 25 bis 50 Jahre"	0.042	0.083	0.067
"Alter - 50 Jahre und mehr"	0.042	0.083	0.133
"Familienstand - Nicht verheiratet"	0.167	0.167	0.133
"Familienstand - verheiratet"	0.167	0.167	0.200
"SUMME"	1.000	1.000	1.000



Zwischen Befragten mit Kaufabsicht und jenen ohne besteht eine gewisse Ähnlichkeit. Will man die Zeilen der Kontingenztabelle vergleichen, dann geht man analog vor. Man dividiert die Elemente einer Zeile jeweils durch ihre Zeilensumme und erhält dadurch das entsprechende Zeilenprofil.

("ZEILENPROFIL")	"Ja"	"Nein"	"Weiss nicht"	"SUMME"
"Zeilen"	"Anteil"	"Anteil"	"Anteil"	"Anteil"
"Geschlecht - Weiblich"	0.563	0.188	0.250	1.000
"Geschlecht - Männlich"	0.778	0.111	0.111	1.000
"Alter - bis 25 Jahre"	0.750	0.125	0.125	1.000
"Alter - 25 bis 50 Jahre"	0.500	0.250	0.250	1.000
"Alter - 50 Jahre und mehr"	0.400	0.200	0.400	1.000
"Familienstand - Nicht verheiratet"	0.667	0.167	0.167	1.000
"Familienstand - verheiratet"	0.615	0.154	0.231	1.000



(1)	2	3	4	5	6	7
"Weibl."	"Männl"	"- 25"	"25 - 50"	"50 u. m"	"Nicht verh."	"Verh."

Auch hier kann man Ähnlichkeiten zwischen den Zeilenprofilen entdecken. So haben z. B. männliche Befragte und Personen bis 25 Jahre ähnliche Kaufabsichten. Bezieht man die absoluten Häufigkeiten der Kontingenztabelle auf die Summe der Häufigkeiten insgesamt, dann erhält man die Korrespondenzmatrix. Die Zeilen- und Spaltensummen bezogen auf die Gesamtsumme der Häufigkeiten werden in diesem Zusammenhang als Zeilen- und Spaltenmassen bezeichnet.

(KORRESPONDENZTABELLE)	"Ja"	"Nein"	"Weiss nicht"	"ZEILENMASSEN "
"Geschlecht - Weiblich"	0.120	0.040	0.053	0.213
"Geschlecht - Männlich"	0.093	0.013	0.013	0.120
"Alter - bis 25 Jahre"	0.160	0.027	0.027	0.213
"Alter - 25 bis 50 Jahre"	0.027	0.013	0.013	0.053
"Alter - 50 Jahre und mehr"	0.027	0.013	0.027	0.067
"Familienstand - Nicht verheiratet"	0.107	0.027	0.027	0.160
"Familienstand - verheiratet"	0.107	0.027	0.040	0.173
"SPALTENMASSEN"	0.640	0.160	0.200	1.000

Ziel der Korrespondenzanalyse ist die Darstellung der Profile durch Punkte in einen möglichst niedrig dimensionalen (meist 2-dimensionalen) Euklidischen Raum. Die einfachen Euklidischen Distanzen zwischen zwei Profilen sollen in guter Näherung ihren Chiquadratabständen entsprechen. Ausgangspunkt dafür sind nicht die Chiquadratabstände, sondern die so genannten Residuen **P**:

$$\mathbf{P} = \mathbf{D}(\mathbf{z}^{(-1/2)}) \cdot (\mathbf{K} - \mathbf{z} \cdot \mathbf{s}^{T}) \cdot \mathbf{D}(\mathbf{s}^{(-1/2)})$$

K ist der Teil der Korrespondenzmatrix, der die relativen Häufigkeiten enthält

$$\chi 2 = \sum_{i=1}^{z} \sum_{j=1}^{s} \frac{\left(h_{ij} - e_{ij}\right)^2}{e_{ij}} = \sum_{i=1}^{z} n_i \cdot \sum_{j=1}^{s} \frac{\left(\frac{h_{ij}}{n_i} - \frac{e_{ij}}{n_i}\right)^2}{\frac{e_{ij}}{n_i}} = n \cdot \sum_{i=1}^{z} \frac{n_i}{n} \cdot \sum_{j=1}^{s} \frac{\left(p_{ij} - c_j\right)^2}{c_j}$$

mit

$$n_i = \sum_{i=1}^{z} h_{ij}$$

z ist die Anzahl der Zeilen,

$$n_j = \sum_{j=1}^{s} h_{ij}$$

s ist die Anzahl der Spalten und e_{ij}

$$e_{ij} = \frac{n_i \cdot n_j}{n}$$

sind die erwarteten Häufigkeiten.

Der i-te Summand der äußeren Summe ist der quadrierte und mit der Zeilenmasse $r_i = n_i/n$ gewichtete $\chi 2$-Abstand des i-ten Zeilenprofils p_i vom Zeilenrandprofil c.

$$\left|p_i - c\right| = \sqrt{\sum_{j=1}^{s} \frac{\left(p_{ij} - c_j\right)^2}{c_j}}$$

Mit dieser Vereinbarung kann die Chiquadratstatistik als gewichtetes Euklidisches Distanzmaß geschrieben werden:

$$\chi 2 = n \cdot \sum_{i=1}^{z} r_i \cdot \left(\left|p_i - c\right|\right)^2$$

Dieser Ausdruck wird umso größer, je stärker sich die Zeilenprofile unterscheiden. Er quantifiziert also die gesamte Variation die in einer Kontingenztabelle steckt. Der Quotient aus diesem Ausdruck und der Gesamtanzahl n bezeichnet man als Inertia:

$$\text{Inertia}(K) = \frac{\chi 2}{n}$$

Im Unterschied zur Chiquadratstatistik ist die Inertia normiert und damit leichter zu interpretieren:

$$0 \leq \text{Inertia}(K) \leq \min(z, s) - 1$$

Da die Anzahl der Zeilen gleich 7 und die der Spalten gleich 3 ist, ist die Obergrenze für die Inertia durch 2 beschränkt:

$$\min(7,3) - 1 = 2$$

$\chi 2$ kann wie üblich aus der Kontingenztabelle errechnet werden. Es ist

$$\chi 2 = 4.093$$

und die Inertia daher gleich

$$\text{Inertia} = \frac{\chi 2}{n} = \frac{4.092}{75} = 0.055$$

$$\mathbf{K} = \begin{pmatrix} 0.120 & 0.040 & 0.053 \\ 0.093 & 0.013 & 0.013 \\ 0.160 & 0.027 & 0.027 \\ 0.027 & 0.013 & 0.013 \\ 0.027 & 0.013 & 0.027 \\ 0.107 & 0.027 & 0.027 \\ 0.107 & 0.027 & 0.040 \end{pmatrix}$$

z ist die letzte Zeile der Korrespondensmatrix, die relativen Häufigkeiten der Zeilensummen

$$\mathbf{z} = \begin{pmatrix} 0.2133 \\ 0.1200 \\ 0.2133 \\ 0.0533 \\ 0.0667 \\ 0.1600 \\ 0.1733 \end{pmatrix}$$

und **s** ist der Vektor der relativen Häufigkeiten der Spaltensummen

$$\mathbf{s} = \begin{pmatrix} 0.64 \\ 0.16 \\ 0.20 \end{pmatrix}$$

$\mathbf{D}(\mathbf{z}^{(-1/2)})$ und $\mathbf{D}(\mathbf{s}^{(-1/2)})$ sind Diagonalmatrizen, die die reziproken Wurzeln der Elemente von **z** und **s** als Diagonalelemente enthalten. Eingesetzt und ausgerechnet erhält man die Residualmatrix **P**

$$\mathbf{P} = \begin{pmatrix} -0.0447 & 0.0318 & 0.0501 \\ 0.0585 & -0.0447 & -0.071 \\ 0.0636 & -0.0386 & -0.0758 \\ -0.0385 & 0.0484 & 0.0227 \\ -0.0759 & 0.0225 & 0.1183 \\ 0.0144 & 0.0087 & -0.028 \\ -0.0117 & -0.0044 & 0.0287 \end{pmatrix}$$

Auf diese Matrix **P** wird nun die Singulärwertzerlegung angewandt, um die Koordinaten sowohl der Zeilenprofile als auch der Spaltenprofile zu erhalten. Die Singulärwertzerlegung einer Matrix beruht auf der Tatsache, dass man für jede Datenmatrix folgende zwei Eigenwertgleichungen anschreiben kann:

$$\mathbf{P}^T \cdot \mathbf{P} = \mathbf{V} \cdot \mathbf{D}_l \cdot \mathbf{V}^T$$

mit der Diagonalmatrix $\mathbf{D}_l$ der Eigenwerte und den Eigenvektoren **V** mit $\mathbf{V}^T \cdot \mathbf{V} = \mathbf{I}$ (**I** = Einheitsmatrix), und

$$\mathbf{P}^T \cdot \mathbf{P} = \mathbf{U} \cdot \mathbf{D}_l \cdot \mathbf{U}^T$$

mit der Diagonalmatrix $\mathbf{D}_l$ der Eigenwerte und den Eigenvektoren **U** mit $\mathbf{U}^T \cdot \mathbf{U} = \mathbf{I}$.
Diese beiden Gleichungen implizieren, dass es die folgende Zerlegung der Matrix **P** gibt, die Singulärwertzerlegung genannt wird:

$$\mathbf{P} = \mathbf{U} \,.\, \mathbf{D}_q \,.\, \mathbf{V}^T$$

wobei

$$\mathbf{D}_q = \mathbf{D}_l^{(1/2)}$$

eine Diagonalmatrix ist, die als Diagonalelemente die Singulärwerte q enthält. Diese Singulärwerte ergeben sich als Wurzel der Eigenwerte.

Die Singulärwerte der Residualmatrix **P** sind

$$\mathbf{sw} = \begin{pmatrix} 0.2274 \\ 0.0534 \\ 0.0000 \end{pmatrix}$$

die entsprechende Eigenvektormatrix **U** ist

$$\mathbf{U} = \begin{pmatrix} -0.3295 & 0.1456 & 0.5145 \\ 0.4394 & -0.1942 & -0.0552 \\ 0.4744 & -0.0933 & 0.431 \\ -0.2641 & 0.6985 & 0.0358 \\ -0.6125 & -0.4579 & -0.0513 \\ 0.1214 & 0.3494 & 0.3526 \\ -0.1167 & -0.3357 & 0.6467 \end{pmatrix}$$

und die Eigenvektormatrix **V** ist

$$\begin{pmatrix} 0.5818 & -0.1465 & -0.8 \\ -0.3352 & 0.853 & -0.4 \\ -0.741 & -0.5009 & -0.4472 \end{pmatrix}$$

Werden nun die ersten beiden Spalten von **U** bzw. die ersten beiden Zeilen von **V** benutzt, dann lässt sich **P** approximativ darstellen als

$$\mathbf{P} = (\mathbf{U} \,.\, \mathbf{D}_q^{1/2}) \,.\, (\mathbf{D}_q^{1/2} \,.\, \mathbf{V}^T)$$

Die Koordinaten für die Zeilenprofile ergeben sich daraus als

$$\begin{pmatrix} u_{11}\cdot\sqrt{\theta_1} & u_{12}\cdot\sqrt{\theta_2} \\ u_{21}\cdot\sqrt{\theta_1} & u_{22}\cdot\sqrt{\theta_2} \\ \ldots\ldots\ldots & \ldots\ldots \\ u_{n1}\cdot\sqrt{\theta_1} & u_{n2}\cdot\sqrt{\theta_2} \end{pmatrix}$$

und die Spaltenprofile durch die Koordinaten

$$\begin{pmatrix} v_{11}\cdot\sqrt{\theta_1} & v_{12}\cdot\sqrt{\theta_2} \\ v_{21}\cdot\sqrt{\theta_1} & v_{22}\cdot\sqrt{\theta_2} \\ \ldots\ldots\ldots & \ldots\ldots \\ v_{n1}\cdot\sqrt{\theta_1} & v_{n2}\cdot\sqrt{\theta_2} \end{pmatrix}$$

Für das Beispiel sind die entsprechenden Zeilen-Koordinaten

(ZEILENKOORDINATEN)	"1. Achse"	"2. Achse"
"Geschlecht - Weiblich"	−0.162	0.017
"Geschlecht - Männlich"	0.288	−0.030
"Alter - bis 25 Jahre"	0.234	−0.011
"Alter - 25 bis 50 Jahre"	−0.260	0.162
"Alter - 50 Jahre und mehr"	−0.539	−0.095
"Familienstand - Nicht verheiratet"	0.069	0.047
"Familienstand - verheiratet"	−0.064	−0.043

und Spalten-Koordinaten

(SPALTENKOORDINATEN)	"1. Achse"	"2. Achse"
"Ja"	0.165	−0.010
"Nein"	−0.191	0.114
"Weiss nicht"	−0.377	−0.060

Die Kommunalitäten (die Quadrate der Faktorladungen) können nun aus diesen Zeilen- und Spaltenkoordinaten bestimmt werden. Für die Spaltenkoordinaten zeigt folgende Tabelle die Berechnung:

(s_i)	x_{i1}	x_{i2}	x_{i1}^2	x_{i2}^2	$s_i \cdot x_{i1}^2$	$s_i \cdot x_{i2}^2$
0.64	0.1654	−0.0098	0.0274	0.0001	0.0175	0.0001
0.16	−0.1906	0.1139	0.0363	0.013	0.0058	0.0021
0.2	−0.3768	−0.0598	0.142	0.0036	0.0284	0.0007

In der ersten Spalte steht der Vektor der relativen Häufigkeiten der Spaltensummen. Die 2. und 3. Spalte enthält die Spalten-Koordinaten. Das Quadrat dieser Spalten-Koordinaten steht in der 4. und 5. Spalte und die mit den relativen Häufigkeiten gewichteten Quadrate in den letzten beiden Spalten. Diese letzten beiden Spalten werden in der folgenden Tabelle weiter verarbeitet.

$s_i \cdot x_{i1}^2$	$s_i \cdot x_{i2}^2$	$\sum_{j=1}^{2} \left(s_i \cdot x_{ij}^2\right) = S$	$\frac{s_i \cdot x_{i1}^2}{S}$	$\frac{s_i \cdot x_{i2}^2}{S}$
0.0175	0.0001	0.0176	0.9965	0.0035
0.0058	0.0021	0.0079	0.7367	0.2633
0.0284	0.0007	0.0291	0.9754	0.0246

In der 3. Spalte steht die Summe aus den jeweiligen Elementen der 1. und 2. Spalte. Dividiert man die ersten beiden Spalten durch diese Summe, dann erhält man die Kommunalitäten, die in der 4. und 5. Spalte stehen. Diese Kommunalitäten sind die quadrierten Korrelationskoeffizienten der Ausprägung mit den Achsen. Man sieht, dass die Antworten "Ja", "Nein" und "Weiß nicht" stark mit der ersten Achse korrelieren, nicht aber mit der zweiten Achse.

Dividiert man die Elemente der 1. und 2. Spalte nicht durch die jeweilige Zeilensumme, sondern durch die Spaltensummen, dann erhält man die Inertia:

$\left(s_i \cdot x_{i1}^2\right)$	$\left(s_i \cdot x_{i2}^2\right)$	$\frac{s_i \cdot x_{i1}^2}{0.0517}$	$\frac{s_i \cdot x_{i2}^2}{0.0029}$
0.0175	0.0001	0.3385	0.0345
0.0058	0.0021	0.1122	0.7241
0.0284	0.0007	0.5493	0.2414
"------"	"------"	"------"	"------"
0.0517	0.0029	"*"	"*"

Anhand der Inertia kann abgelesen werden, wie stark die einzelnen Antworten die Ausrichtung der Achsen determinieren. Die Antwort "Weiß nicht" hat mit 0.5493 den stärksten Einfluss auf die Ausrichtung der ersten Achse. Sie erklärt rund 55% der Streuung dieser Achse. Die Antwort "Nein" hat mit 0.7241 den stärksten Einfluss auf die zweite Achse. 72% der Variation dieser Achse wird durch die Antwort "Nein" erklärt.

Auf die gleiche Art berechnet man die Kommunalitäten für die Zeilenkoordinaten:

"1. Achse"	"2. Achse"	(KOMMUNALITÄTEN_DER_ZK)
0.989	0.011	"Geschlecht - Weiblich"
0.989	0.011	"Geschlecht - Männlich"
0.998	0.002	"Alter - bis 25 Jahre"
0.721	0.279	"Alter - 25 bis 50 Jahre"
0.970	0.030	"Alter - 50 Jahre und mehr"
0.686	0.314	"Familienstand - Nicht verheiratet"
0.686	0.314	"Familienstand - verheiratet"

Auch hier zeigt sich, dass die Antworten stark mit der ersten Achse korrelieren, nicht aber mit der zweiten. Die größten Korrelationen findet man beim Geschlecht und Alter. Etwas schwächer korreliert der Familienstand mit der ersten Achse.

"1. Achse"	"2. Achse"	(INERTIA_DER_ZK)
0.109	0.021	"Geschlecht - Weiblich"
0.193	0.038	"Geschlecht - Männlich"
0.225	0.009	"Alter - bis 25 Jahre"
0.070	0.488	"Alter - 25 bis 50 Jahre"
0.375	0.210	"Alter - 50 Jahre und mehr"
0.015	0.122	"Familienstand - Nicht verheiratet"
0.014	0.113	"Familienstand - verheiratet"

Den stärksten Einfluss auf die Ausrichtung der 1. Achse hat mit 0.375 die Antwort "Alter-50 Jahre und mehr". Auf die Ausrichtung der 2. Achse hat die Antwort "Alter-25 bis 50 Jahre" den stärksten Einfluss. Sie erklärt 48.8% der Variation dieser Achse.

Die Berechnung einer Korrespondenzanalyse ist auch für kleine Tatbestände nur mit einer geeigneten Software sinnvoll durchführbar. Der ausführlich dargestellte Rechengang kann aber zum besseren Verständnis dienen, wenn man mit den mathematischen Hilfsmitteln vertraut ist. Den Output für dieses Beispiel findet man auf Seite 213. Dort sind jedoch nur die Hauptergebnisse angeführt. Detailergebnisse findet man im Programm unter dem Abschnitt “DETAILS“.

8_0 MULTIVARIATE STICHPROBEN METRISCH

8_1 Multivariate Varianzanalyse
8_2 Multivariate Regressionsanalyse
8_3 Faktorenanalyse

Vorschau

Im Abschnitt 8_1, Multivariate Varianzanalyse, wird untersucht, ob sich Käufer der Spielkonsole Yoki von den Nichtkäufern im Hinblick auf Alter und Einkommen unterscheiden. Da nicht nur das Durchschnittsalter von Käufern und Nichtkäufern verglichen wird, sondern auch das Durchschnittseinkommen, ist es eine multivariate Analyse. Voraussetzung für die Unterschiedsprüfung ist die multivariate Normalverteilung der Grundgesamtheiten und ihre Homoskedastizität. Beide Voraussetzungen können mit den im Abschnitt 9 angeführten Verfahren überprüft werden.

Mit Hilfe der multivariaten Regressionsanalyse von Abschnitt 8_2 wird versucht, vom Alter der Befragten und seiner täglichen Fernsehzeit auf sein monatliches Einkommen zu schließen. Die abhängige Variable Einkommen soll durch die beiden unabhängigen Variablen Alter und tägliche Fernsehzeit erklärt werden. Auch hier wird vorausgesetzt, dass die Fehlervariable normalverteilt ist.

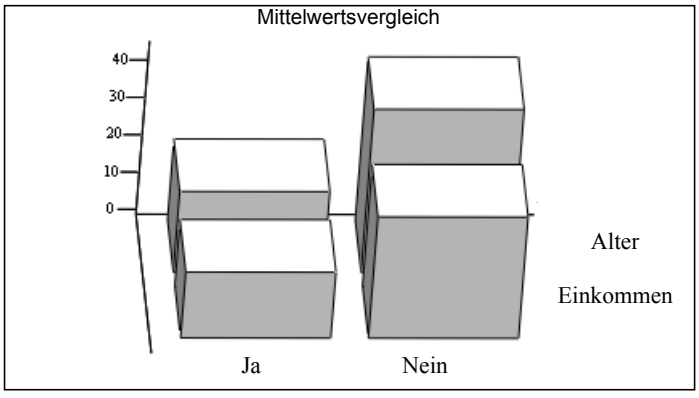
Im Abschnitt 8_3, Faktorenanalyse, werden die 4 Variablen Alter, tägliche Fernsehzeit, tägliche Computerzeit und Einkommen auf 2 Faktoren reduziert. Diese beiden Faktoren, soziodemographischer Faktor und Zeitfaktor, werden so gewählt, dass sie den Informationsgehalt der vier Variablen ausreichend reproduzieren.

8_1 Multivariate Varianzanalyse

Unterscheiden sich die Käufer der Spielkonsole Yoki von den Nichtkäufern im Hinblick auf ihr Alter und monatlichem Einkommen? Eine Befragung von 10 Besuchern einer Werbeveranstaltung brachte folgendes Ergebnis:

(Käufer_Ja = 1 / Käufer_Nein = 2)	Alter	Einkommen
1	17	10
1	25	21
1	18	15
1	22	20
1	30	22
1	18	18
2	23	22
2	35	35
2	61	41
2	56	32

Die Durchschnitte aus den Alters- und Einkommensangaben für die zwei Stichproben zeigt folgende Grafik:



("Kaufabsicht")	"Ja"	"Nein"
"Alter"	21.67	43.75
"Einkommen"	17.67	32.5

Kann man auf Grund dieser Stichprobenergebnisse allgemein behaupten, dass sich Käufer und Nichtkäufer im Hinblick auf ihr Durchschnittsalter und -einkommen unterscheiden?

Ja, die Unterschiede zwischen den Durchschnitten sind groß genug. Es könnte zwar tatsächlich sein, dass das Durchschnittsalter und -einkommen die Käufer und Nichtkäufer unterschiedlich ist. Das Risiko ist dafür mit 5% beschränkt. (Vergleiche den Softwareoutput für dieses Beispiel auf Seite 215).

Wie kommt man zu diesem Ergebnis? Zuerst formuliert man die Hypothesen. Man nimmt als Nullhypothese an, dass das Durchschnittsalter und -einkommen der Käufer und Nichtkäufer in der Grundgesamtheit aller potentiellen Kunden gleich ist. Diese Hypothese wird formal wie folgt angeschrieben:

$$H_0: \begin{pmatrix} \text{Durchschnittsalter} \\ \text{Durchschnittseinkommen} \end{pmatrix}_{\text{Käufer}} = \begin{pmatrix} \text{Durchschnittsalter} \\ \text{Durchschnittseinkommen} \end{pmatrix}_{\text{Nichtkäufer}}$$

Eine Alternative zu dieser Nullhypothese ist

$$H_1: \begin{pmatrix} \text{Durchschnittsalter} \\ \text{Durchschnittseinkommen} \end{pmatrix}_{\text{Käufer}} \neq \begin{pmatrix} \text{Durchschnittsalter} \\ \text{Durchschnittseinkommen} \end{pmatrix}_{\text{Nichtkäufer}}$$

Das Durchschnittsalter und -einkommen der Käufer und Nichtkäufer unterscheidet sich. An Hand der Ergebnisse der Stichprobe entscheidet man sich für die Ablehnung oder Nichtablehnung der Nullhypothese. Dazu berechnet man folgende Testmaßzahl:

$$F_{beob} = \frac{1 - \Lambda}{\Lambda} \cdot \frac{\nu_2}{\nu_1}$$

mit

$$\Lambda = |\mathbf{W}|/|\mathbf{T}|$$

Im Zähler von Λ steht die Determinante aus den Abweichungsquadraten und -produkten innerhalb der Stichproben **W** und im Nenner die Determinante aus den Abweichungsquadraten und -produkten der Gesamtstichprobe **T**. Die Freiheitsgrade ν_1 und ν_2 sind wie folgt definiert:

$$\nu_1 = t$$

$$\nu_2 = n - t - 1$$

t ist die Anzahl der Stichproben und n ist der Umfang der Gesamtstichprobe. Die Berechnung des Zählers ist in folgenden Tabellen für die erste Stichprobe dargestellt:

$x_{i1}^{\langle 1\rangle}$	$x_{i1}^{\langle 1\rangle} - xquer_1^{\langle 1\rangle}$	$\left(x_{i1}^{\langle 1\rangle} - xquer_1^{\langle 1\rangle}\right)^2$
17	–4.667	21.778
25	3.333	11.111
18	–3.667	13.444
22	0.333	0.111
30	8.333	69.444
18	–3.667	13.444
Alter	SUMME	129.333

In der 1. Spalte stehen die Altersangaben der Käufer und in der 2. Spalte die Abweichungen des Alters vom Durchschnitt 21.667. In der 3. Spalte sind die Quadrate von Spalte 2 ausgewiesen. Die Summe dieser letzten Spalte ist 129.333, die Abweichungsquadratsumme des Alters von ihrem Durchschnitt in der Käufer-Stichprobe. Die gleiche Berechnung für das Einkommen zeigt die folgende Tabelle:

$x_{i1}^{\langle 2\rangle}$	$x_{i1}^{\langle 2\rangle} - xquer_1^{\langle 2\rangle}$	$\left(x_{i1}^{\langle 2\rangle} - xquer_1^{\langle 2\rangle}\right)^2$
10	–7.667	58.778
21	3.333	11.111
15	–2.667	7.111
20	2.333	5.444
22	4.333	18.778
18	0.333	0.111
Einkommen	SUMME	101.333

Als dritte Tabelle wird die Berechnung der Abweichungsprodukte zwischen Alter und Einkommen für die Käuferstichprobe gezeigt:

$x_{i1}^{\langle 1\rangle}$	$x_{i1}^{\langle 2\rangle}$	$x_{i1}^{\langle 1\rangle} - xquer_1^{\langle 1\rangle}$	$x_{i1}^{\langle 2\rangle} - xquer_1^{\langle 2\rangle}$	(Spalte_3)·(Spalte_4)
10	17	–7.667	–4.667	35.778
21	25	3.333	3.333	11.111
15	18	–2.667	–3.667	9.778
20	22	2.333	0.333	0.778
22	30	4.333	8.333	36.111
18	18	0.333	–3.667	–1.222
Alter	Einkommen	Käufer_Ja	SUMME	92.334

Für die Käufer-Stichprobe ergibt sich folgende Abweichungsquadrat- und -produktmatrix:

$$\mathbf{W}_1 = \begin{pmatrix} 129.333 & 92.334 \\ 92.334 & 101.333 \end{pmatrix}$$

In der Hauptdiagonale stehen die Abweichungsquadrate, in der Nebendiagonale die Abweichungsprodukte. Für die Nichtkäuferstichprobe erhält man auf die gleiche Art folgende Abweichungsquadrat- und Produktmatrix:

$$\mathbf{W}_2 = \begin{pmatrix} 954.75 & 336.5 \\ 336.5 & 189 \end{pmatrix}$$

Die Summe beider Matrizen ergibt die Matrix der Abweichungsquadrate und -produkte innerhalb der Stichproben

$$\mathbf{W} = \mathbf{W}_1 + \mathbf{W}_2 = \begin{pmatrix} 129.333 & 92.334 \\ 92.334 & 101.333 \end{pmatrix} + \begin{pmatrix} 954.75 & 336.5 \\ 336.5 & 189 \end{pmatrix} = \begin{pmatrix} 1.084 \times 10^3 & 428.834 \\ 428.834 & 290.333 \end{pmatrix}$$

Die Determinante dieser Matrix **W** ist

$$|\mathbf{W}| = 1084.083 \cdot 290.333 - 428.834^2 = 1.308 \times 10^5$$

Die Matrix **T** der Abweichungsquadrate und -produkte insgesamt errechnet man aus der Gesamtstichprobe auf die gleiche Art:

$$\mathbf{T} = \begin{pmatrix} 2254.5 & 1215 \\ 1215 & 818.4 \end{pmatrix}$$

Die Determinante dieser Matrix ist

$$|\mathbf{T}| = 2254.5 \cdot 818.4 - 1215^2 = 3.689 \times 10^5$$

Für Λ ergibt sich nun

$$\Lambda = |\mathbf{W}|/|\mathbf{T}| = \frac{130846.47}{368857.8} = 0.355$$

und für die Testmaßzahl F_{beob}

$$F_{beob} = \frac{1 - \Lambda}{\Lambda} \cdot \frac{\nu_2}{\nu_1} = \frac{1 - 0.355}{0.355} \cdot \frac{7}{2} = 6.359$$

mit

$$\nu_1 = t = 2$$

$$\nu_2 = n - t - 1 = (10 - 2 - 1) = 7$$

Die Testmaßzahl F_{beob} ist F-verteilt mit ν_1 und ν_2 Freiheitsgraden. Die Wahrscheinlichkeit für das Auftreten einer so großen Testmaßzahl F_{beob} wie im vorliegenden Beispiel ist bei Gültigkeit der Nullhypothese gleich

$$F_F(6.359, 2, 7) = 0.973$$

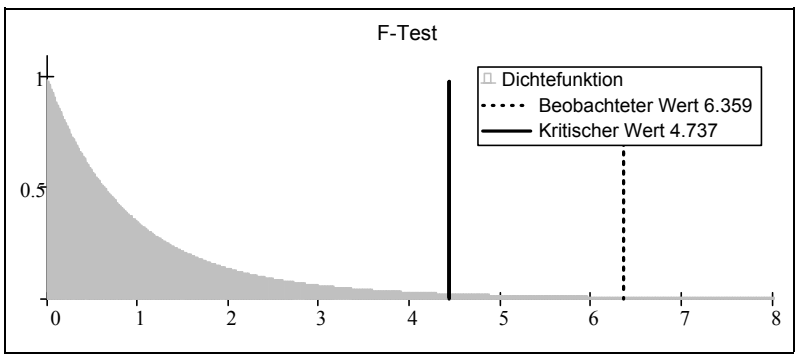
(Diesen Wert liest man aus einer geeigneten Tabelle der Chiquadratverteilung ab oder berechnet ihn mit Hilfe entsprechender Programme.) Die Wahrscheinlichkeit für eine so große oder größere Testmaßzahl ist dann gleich

$$1 - F_F(6.359, 2, 7) = 0.027$$

Da diese Wahrscheinlichkeit kleiner ist, als das Signifikanzniveau von 5% ($\alpha = 0.05$), kann die Nullhypothese abgelehnt werden. Käufer und Nichtkäufer unterscheiden sich signifikant im Hinblick auf ihr Durchschnittsalter oder Durchschnittseinkommen. Das Risiko, dass diese Entscheidung falsch ist, ist höchstens 5%.

Zum selben Ergebnis kommt man, wenn man F_{beob} mit dem kritischen Wert der F-Verteilung vergleicht: Unter Gültigkeit der Nullhypothese sind die 5% seltensten F Werte nach unten beschränkt durch den kritischen Wert F_c gleich 4.737. Da das aus der Stichprobe errechnete F_{beob} von 6.359 größer ist als dieser kritische Wert, kann die Nullhypothese abgelehnt werden.

$$F_{beob} = 6.359 > 4.737 = F_c$$



(F – beobachtet)	(F – kritisch)	(p – Wahrscheinlichkeit)
6.359	4.737	0.027

Wo sind die signifikanten Unterschiede? Unterscheiden sich die Käufer und Nichtkäufer im Hinblick auf ihr Durchschnittsalter oder Durchschnittseinkommen oder in beiden Durchschnitten? Der F-Test kam nur zum Ergebnis, dass Unterschiede bestehen, nicht aber bei welcher Frage. Um dies zu eruieren, muss man Tests durchführen, die die Unterschiede beim Alter und beim Einkommen prüfen.

Der Scheffe´-Test liefert folgendes Ergebnis: Für das Alter sind die Unterschiede zwischen den Durchschnitten der Käufer und Nichtkäufer groß genug, um allgemein zu behaupten, dass sich Käufer und Nichtkäufer im Durchschnittsalter signifikant unterscheiden.

("Alter")	"Kritischer Wert: "	17.328	"5 %"
("Kaufabsicht")	("Kaufabsicht")	"Differenz"	"Signifikanzniveau"
"Ja"	"Nein"	22.083	"signifikant"

Beim Einkommen sind die Unterschiede zwischen Käufer und Nichtkäufer ebenfalls groß genug. Man kann allgemein behaupten, dass sich das Durchschnittseinkommen der Käufer und Nichtkäufer signifikant unterscheidet.

("Einkommen")	"Kritischer Wert: "	8.967	"5 %"
("Kaufabsicht")	("Kaufabsicht")	"Differenz"	"Signifikanzniveau"
"Ja"	"Nein"	14.833	"signifikant"

Zu diesen Ergebnissen kommt man mit Hilfe folgender Formel

$$\mathrm{Diff}_{c,i} = \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \cdot \frac{t-1}{n-t} \cdot W_{i,i} \cdot F_{1-\alpha,\, t-1,\, n-t}}$$

Diese Formel liefert die kritischen Werte für die Mittelwertsdifferenzen. n_1 und n_2 sind die beiden Stichprobenumfänge, n ist der Gesamtstichprobenumfang und t die Zahl der Stichproben. $\mathbf{W}_{ii}$ sind die Abweichungsquadrate der jeweiligen Stichproben innerhalb, also die Diagonalelemente der Matrix **W**. Sind die Mittelwertsdifferenzen (absolut) kleiner als die kritischen Werte, dann kann die Gleichheit der Mittelwerte nicht abgelehnt werden. Nur wenn die beobachteten Mittelwertsdifferenzen größer als die kritischen Werte sind, kann man die Nullhypothese ablehnen. Für das Beispiel sind die kritischen Werte gleich

$$\mathrm{Diff}_{c,1} = \sqrt{\left(\frac{1}{6} + \frac{1}{4}\right) \cdot \frac{2-1}{10-2} \cdot 1084.083 \cdot 5.318} = 17.328$$

$$\mathrm{Diff}_{c,2} = \sqrt{\left(\frac{1}{6} + \frac{1}{4}\right) \cdot \frac{2-1}{10-2} \cdot 290.333 \cdot 5.318} = 8.968$$

wie oben schon ausgewiesen.

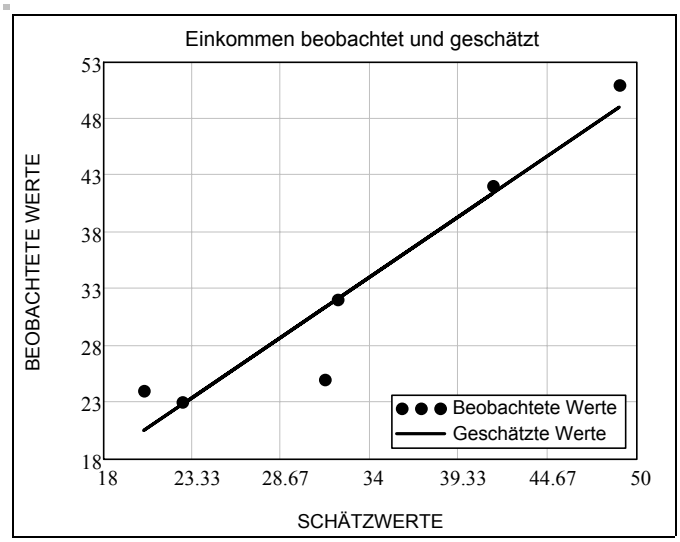
Voraussetzungen für diese multivariate Varianzanalyse sind multivariate Normalverteilung der beobachteten Daten und Homoskedastizität, d.h. Gleichheit der Varianzen und Kovarianzen in der Grundgesamtheit.

8_2 Multivariate Regressionsanalyse

Kann man von der täglichen Fernsehzeit und vom Alter der potentiellen Kunden für die neue Spielkonsole Yoki auf ihr monatliches Einkommen schließen? Eine Befragung von 6 Besuchern einer Werbeveranstaltung für die Spielkonsole brachte folgendes Ergebnis:

2	1	1	4	3	2
36	45	19	62	27	35
32	42	24	51	23	25

In der ersten Zeile steht die tägliche Fernsehzeit (in Stunden) der befragten Person und in der zweite das Alter. Das jeweilige Monatseinkommen (in 100 €) steht in der dritten Zeile. In folgender Grafik sind als Ordinate die beobachteten Einkommen der 6 Besucher als X eingetragen. Auf der Abszisse findet man die entsprechenden Schätzwerte (hier blau eingetragen).



Ist es möglich mit Hilfe einer linearen Regressionsfunktion von der täglichen Fernsehzeit und dem Alter eines potentiellen Kunden auf sein monatliches Einkommen zu schließen? Um das Ergebnis vorwegzunehmen: Ja, die beobachteten Wertepaare liegen eng genug um die Regressionsgerade. Man kann allgemein behaupten, dass die aus den Stichproben errechnete Regressionsfunktion zwischen Fernsehzeit, Alter und Einkommen für alle potentiellen Kunden der Spielkonsole Yoki gilt. Diese lineare Regressionsfunktion der Stichprobe ist allgemein

$$ydach_i = 7.095 + 2.044 \cdot x_{i,1} + 0.808 \cdot x_{i,2}$$

und eingesetzt

$$\text{Schätzwert\_für\_Einkommen}_i = 7.095 - 2.044 \,.\, \text{Fernsehen}_i + 0.808 \,.\, \text{Alter}_i.$$

Mit dieser Funktion kann man für jede Fernsehzeit und jedes Alter der potentiellen Kunden einen Schätzwert für das monatliche Einkommen bestimmen. Für einen 30-Jährigen, der täglich 2 Stunden fernsieht, ergibt sich z. B. durch Einsetzen ein monatliches Einkommen von 2724.7 € (= 27.247 · 100)

$$7.095 - 2.044 \cdot 2 + 0.808 \cdot 30 = 27.247$$

Der Regressionskoeffizient von –2.044 besagt, dass bei einer Erhöhung der Fernsehzeit um eine Stunde das Einkommen um 204.4 € (= 2.044 · 100) abnimmt, wenn man das Alter konstant hält. Der Regressionskoeffizient des Alters von 0.808 drückt aus, dass sich das Einkommen um 80.8 € (= 0.808 · 100) erhöht, wenn das Alter um ein Jahr erhöht wird und die Fernsehzeit konstant bleibt.

Wie kommt man zu diesem Ergebnis? Zuerst berechnet man aus den Stichprobenwerten die lineare Regressionsfunktion und prüft dann, ob die Beobachtungen eng genug um die Regressionsgerade liegen. Anders ausgedrückt: Man prüft, ob die Streuung der beobachteten Fernsehzeiten, Alters- und Einkommensangaben um die Regressionsfunktion klein genug ist, um allgemein zu behaupten, dass zwischen Fernsehzeit, Alter und Einkommen eine lineare Abhängigkeit besteht.

Für die lineare Regressionsfunktion der Stichprobe

$$\text{ydach}_i = b_0 + b_1 \cdot x_{i,1} + b_2 \cdot x_{i,2}$$

oder in Matrixschreibweise

$$\mathbf{ydach} = \mathbf{X} \,.\, \mathbf{b}$$

werden die Regressionskoeffizienten b_0, b_1 und b_2 mit Hilfe folgender Formel bestimmt:

$$\mathbf{b} = (\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^T \cdot \mathbf{y}$$

b ist der Regressionsvektor

$$\mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix}$$

X die Matrix der unabhängigen Variablen (= Fragen), ergänzt um einen Einsen-Vektor in der ersten Spalte und **y** ist der Vektor der abhängigen Variable (=Frage). Für das Beispiel ist die Matrix **X** gleich

$$\mathbf{X} = \begin{pmatrix} 1 & 2 & 36 \\ 1 & 1 & 45 \\ 1 & 1 & 19 \\ 1 & 4 & 62 \\ 1 & 3 & 27 \\ 1 & 2 & 35 \end{pmatrix}$$

In der ersten Spalte steht der Einsen-Vektor, in der zweiten die Fernsehzeiten und in der dritten Spalte das Alter. Die beobachteten Werte des Einkommens stehen im **y** Vektor

$$\mathbf{y} = \begin{pmatrix} 32 \\ 42 \\ 24 \\ 51 \\ 23 \\ 25 \end{pmatrix}$$

b ist daher

$$\mathbf{b} = \left[\begin{pmatrix} 1 & 2 & 36 \\ 1 & 1 & 45 \\ 1 & 1 & 19 \\ 1 & 4 & 62 \\ 1 & 3 & 27 \\ 1 & 2 & 35 \end{pmatrix}^{T} \cdot \begin{pmatrix} 1 & 2 & 36 \\ 1 & 1 & 45 \\ 1 & 1 & 19 \\ 1 & 4 & 62 \\ 1 & 3 & 27 \\ 1 & 2 & 35 \end{pmatrix} \right]^{-1} \cdot \begin{pmatrix} 1 & 2 & 36 \\ 1 & 1 & 45 \\ 1 & 1 & 19 \\ 1 & 4 & 62 \\ 1 & 3 & 27 \\ 1 & 2 & 35 \end{pmatrix}^{T} \cdot \begin{pmatrix} 32 \\ 42 \\ 24 \\ 51 \\ 23 \\ 25 \end{pmatrix} = \begin{pmatrix} 7.095 \\ -2.044 \\ 0.808 \end{pmatrix}$$

Die lineare Stichprobenfunktion ist daher

$$\text{ydach}_i = 7.095 + 2.044 \cdot x_{i,1} + 0.808 \cdot x_{i,2}$$

ydach ist der Schätzwert für das beobachtete y. Wenn diese Regressionsfunktion allgemein gültig ist, dann müssen die Regressionskoeffizienten **b** in der Grundgesamtheit aller potentiellen Käufer der Spielkonsole von Null verschieden sein. Ein Regressionskoeffizient von Null bedeuten, dass die Steigung der partiellen linearen Funktion Null ist und daher die unabhängige Variable keinen Einfluss auf die abhängige Variable ausübt.

Diese Hypothese wird formal wie folgt angeschrieben:

$$H_0: \boldsymbol{\beta} = \mathbf{0}.$$

Eine Alternative zu dieser Nullhypothese ist die Vermutung, dass zumindest ein Regressionskoeffizient von Null verschieden ist, also zumindest eine Variable einen signifikanten Einfluss auf das Einkommen in der Grundgesamtheit hat. Formal

$$H_1: \boldsymbol{\beta} \neq \mathbf{0}.$$

An Hand der Ergebnisse der Stichprobe entscheidet man sich für die Annahme oder Ablehnung der Null- oder Alternativhypothese. Dazu berechnet man folgende Testmaßzahl:

$$F_{beob} = (\mathbf{b}^T \cdot \mathbf{X}^T \cdot \mathbf{y} - n \cdot yquer^2) / (\mathbf{y}^T \cdot \mathbf{y} - \mathbf{b}^T \cdot \mathbf{X}^T \cdot \mathbf{y}) \cdot (\nu_2/\nu_1)$$

Diese Testmaßzahl folgt einer F-Verteilung mit

$$\nu_1 = p$$

und

$$\nu_2 = n - p - 1$$

Freiheitsgraden. n ist der Umfang der Stichprobe und p die Anzahl unabhängiger Variablen.
Die Zwischenschritte zur Berechnung der Testmaßzahl sind

$$\mathbf{b}^T \cdot \mathbf{X}^T \cdot \mathbf{y} =$$

$$= \begin{pmatrix} 7.095 \\ -2.044 \\ 0.808 \end{pmatrix}^T \cdot \begin{pmatrix} 1 & 2 & 36 \\ 1 & 1 & 45 \\ 1 & 1 & 19 \\ 1 & 4 & 62 \\ 1 & 3 & 27 \\ 1 & 2 & 35 \end{pmatrix}^T \cdot \begin{pmatrix} 32 \\ 42 \\ 24 \\ 51 \\ 23 \\ 25 \end{pmatrix}$$

$$n \cdot yquer^2 = 6 \cdot (32.833)^2 = 6468.035$$

$$\mathbf{y}^T \cdot \mathbf{y} =$$

$$= \begin{pmatrix} 32 \\ 42 \\ 24 \\ 51 \\ 23 \\ 25 \end{pmatrix}^T \cdot \begin{pmatrix} 32 \\ 42 \\ 24 \\ 51 \\ 23 \\ 25 \end{pmatrix} = 7119$$

Die Testmaßzahl kann nun wie folgt berechnen werden:

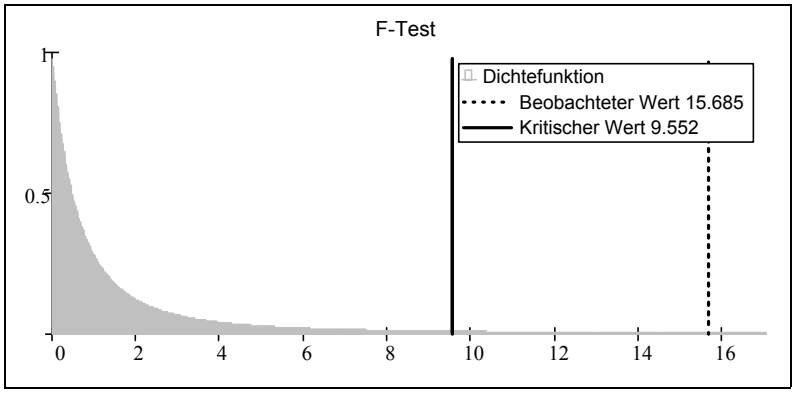
$$F_{beob} = \frac{7061.831 - 6468.167}{7119 - 7061.831} \cdot \frac{6 - 2 - 1}{2} = 15.685$$

Die Wahrscheinlichkeit für das Auftreten einer so großen Testmaßzahl von $F_{beob} = 15.685$ unter Gültigkeit der Nullhypothese ist gleich $F_F(15.685, 2,3) = 0.974$. (Diesen Wert liest man aus einer geeigneten Tabelle der F-Verteilung ab oder berechnet ihn mit Hilfe entsprechender Programme.)

Ist man bereit, in 5 von 100 Fällen eine richtige Nullhypothese abzulehnen, dann ist das Signifikanzniveau α gleich 0.05. Unter Gültigkeit der Nullhypothese sind die 5% seltensten F- Werte nach unten beschränkt durch den kritischen Wert F_c gleich 9.552 . Da das aus der Stichprobe errechnete F_{beob} von 15.685 größer ist als dieser kritische Wert, kann die Nullhypothese abgelehnt werden.

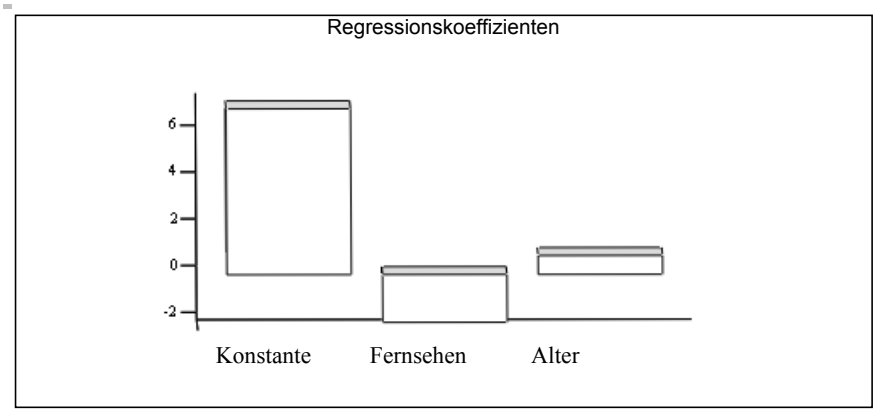
$$F_{beob} = 15.685 > 9.552 = F_c$$

Die Grafik zeigt die F-Verteilung, die kritischen Wert F_c sowie den beobachteten Wert F_{beob}:



(Beobachteter_Wert)	(Kritischer_Wert)	(p – Wahrscheinlichkeit)
15.685	9.552	0.026

Man kann auf Grund der Stichprobe behaupten, dass zwischen der Fernsehzeit und dem Alter als unabhängige Variablen einerseits und dem Einkommen aller potentiellen Kunden als abhängige Variable andererseits allgemein ein linearer Zusammenhang besteht. Tragen beide unabhängigen Variablen signifikant zur Vorhersage der abhängigen Variablen bei? Weichen die Regressionskoeffizienten von Fernsehzeit oder Alter oder beide signifikant von Null ab? Folgende Grafik zeigt die Regressionskoeffizienten einschließlich der Konstanten:



Um zu prüfen, welcher der Regressionskoeffizienten in der Grundgesamtheit von Null verschieden ist, berechnet man die Testmaßzahl

$$t_i = \frac{b_i}{\sqrt{V_{ii}}}$$

$\mathbf{V}_{ii}$ ist das i-te Diagonalelement der Kovarianzmatrix der Regressionskoeffizienten. Diese Matrix wird aus der **X**-Matrix errechnet:

$$\mathbf{V} = s_{res}^2 \cdot (\mathbf{X}^T \cdot \mathbf{X})^{-1}$$

s_{res}^2 ist der Schätzwert für die Fehlervarianz, geschätzt aus der Varianz der Residuen $y_i - ydach_i$.

$$s_{res}^2 = \frac{1}{n - p - 1} \cdot \sum_{i=1}^{n} \left(y_i - ydach_i\right)^2$$

$$= (\mathbf{y}^T \cdot \mathbf{y} - \mathbf{b}^T \cdot \mathbf{X}^T \cdot \mathbf{y}) / (n - p - 1) .$$

Für das Beispiel ist s_{res}^2 gleich

$$s_{res}^2 = \frac{7119 - 7061.831}{6 - 2 - 1} = 19.056$$

und

$$\mathbf{V} = 19.059 \cdot \left[\begin{pmatrix} 1 & 2 & 36 \\ 1 & 1 & 45 \\ 1 & 1 & 19 \\ 1 & 4 & 62 \\ 1 & 3 & 27 \\ 1 & 2 & 35 \end{pmatrix}^T \cdot \begin{pmatrix} 1 & 2 & 36 \\ 1 & 1 & 45 \\ 1 & 1 & 19 \\ 1 & 4 & 62 \\ 1 & 3 & 27 \\ 1 & 2 & 35 \end{pmatrix} \right]^{-1} = \begin{pmatrix} 28.011 & -2.09 & -0.544 \\ -2.09 & 4.12 & -0.183 \\ -0.544 & -0.183 & 0.025 \end{pmatrix}$$

Mit Hilfe der Diagonalelemente von **V** kann man nun prüfen, ob die beiden Regressionskoeffizienten von Fernsehzeit und Alter signifikant von Null abweichen. Die Varianz des Regressionskoeffizienten für die Fernsehzeit ist 4.12 und für das Alter 0.025. t_2 und t_3 sind daher

$$t_2 = \frac{-2.044}{\sqrt{4.12}} = -1.007$$

$$t_3 = \frac{0.808}{\sqrt{0.025}} = 5.110$$

t_i ist studentverteilt mit $n - p - 1$ Freiheitsgraden. Für ein 5%-iges Signifikanzniveau ist der kritische Wert der Studentverteilung für $6 - 2 - 1 = 3$ Freiheitsgrade

$$t_c = 2.353.$$

Da t_3 größer ist als t_c, nicht aber t_2 (absolut genommen), kann man annehmen, dass der Regressionskoeffizient des Alters auch in der Grundgesamtheit von Null verschieden ist, nicht aber der der Fernsehzeit.

Um für einen individuellen Punktschätzwert ein Konfidenzintervall zu bestimmen (wird auch als Prognoseintervall bezeichnet), benötigt man ebenfalls die Streuung der Residuen. Diese Varianz ist gegeben durch

$$s_{res}^2 = \frac{1}{n - p - 1} \cdot \sum_{i=1}^{n} \left(y_i - ydach_i\right)^2 = 19.056$$

Unter der Voraussetzung, dass die Residuen normalverteilt sind, bestimmt man die Konfidenzgrenzen mit Hilfe der Studentverteilung.

$$K_u = \mathbf{x}_0 \cdot \mathbf{b} - t_{1-\alpha/2,\, n-p-1} \cdot s_{res} \cdot (1 + \mathbf{x}_0^T \cdot \mathbf{V} \cdot \mathbf{x}_0)^{1/2}$$

$$K_o = \mathbf{x}_0 \cdot \mathbf{b} + t_{1-\alpha/2,\, n-p-1} \cdot s_{res} \cdot (1 + \mathbf{x}_0^T \cdot \mathbf{V} \cdot \mathbf{x}_0)^{1/2}$$

Für 2 Stunden Fernsehen pro Tag und für ein Alter von 30 Jahren ist $\mathbf{x}_0$ gleich

$$\begin{pmatrix} 1 \\ 2 \\ 30 \end{pmatrix} = \begin{pmatrix} \text{Konstante} \\ \text{Fernsehen} \\ \text{Alter} \end{pmatrix}$$

und $t_{1-a/2,n-p-1}$ für ein Konfidenzniveau von 95% gleich

$$t_{0.95,3} = 2.353$$

und K_u, K_o gleich

$$K_u = 27.247 - 2.353 \cdot \sqrt{19.056} \cdot \sqrt{1 + 4.031} = 4.208$$

$$K_o = 27.247 + 2.353 \cdot \sqrt{19.056} \cdot \sqrt{1 + 4.031} = 50.286$$

mit

$$\mathbf{x}_0^T \cdot \mathbf{V} \cdot \mathbf{x}_0 =$$

$$= \begin{pmatrix} 1 \\ 2 \\ 30 \end{pmatrix}^T \cdot \begin{pmatrix} 28.011 & -2.09 & -0.544 \\ -2.09 & 4.12 & -0.183 \\ -0.544 & -0.183 & 0.025 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 30 \end{pmatrix} = 4.031$$

Mit einem Vertrauen von 95% kann man das monatliche Einkommen eines 30-Jährigen, der täglich 2 Stunden fernsieht, zwischen 420.8 € und 5028.6 € erwarten. Voraussetzung für die Anwendung des F-Tests und die Studentverteilung ist die Normalverteilung der Residuen. Dies kann mit einem Kolmogorov-Smirnov-Test überprüft werden. Den Softwareoutput für dieses Beispiel findet man auf Seite 216.

8_3 Faktorenanalyse

Um das Kaufverhalten potentieller Kunden für die neue Spielkonsole zu eruieren, wurden bei einer Werbeveranstaltung den Besuchern folgende vier Fragen gestellt:

Wie alt sind Sie?

Wie viele Stunden sehen Sie täglich fern?

Wie groß ist Ihr monatliches Einkommen?

Wie viele Stunden verbringen Sie täglich vor dem Computer?

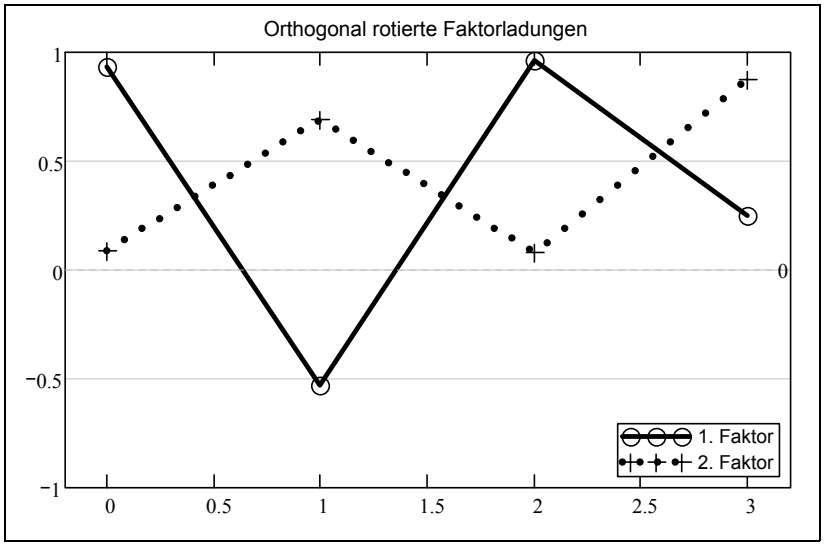
Die Antworten zeigt folgende Tabelle:

36	45	19	62	27	35	45	55	43	35	36	21
2	1	2.5	1	3	0.5	3	5	2	2.5	4.5	7
32	42	24	51	23	25	36	45	50	23	26	16
6	4	3	3	0	2.5	0.5	6	3	0	5	4

Zwischen diesen 4 Variablen bestehen folgende Korrelationen:

(Korrelationen)	Alter	Fernsehen	Einkommen	Computer
Alter	1	–0.316	0.875	0.193
Fernsehen	–0.316	1	–0.393	0.252
Einkommen	0.875	–0.393	1	0.249
Computer	0.193	0.252	0.249	1

Zwischen Alter und Einkommen besteht die stärkste Korrelation (0.875). Zwischen Fernsehen und Einkommen besteht eine negative Korrelation von –0.393. Diese Korrelationen drücken aus, dass man von der ersten Variable auf die zweite schließen kann und umgekehrt, dass also Redundanz vorliegt. Erlaubt diese Redundanz eine Reduktion dieser 4 Variablen auf 2 Faktoren? Kann man zwei hypothetische Variable (= Faktoren) so konstruieren, dass sie den Informationsgehalt der 4 beobachteten Variablen wiedergeben? Ja, dies ist möglich. Folgende Grafik zeigt das Ergebnis der Analyse:



(0)	1	2	3
"Alter"	"Fernsehen"	"Einkommen"	"Computer"

Alter und Einkommen laden den 1. Faktor hoch und Fernsehzeit sowie Computerzeit den 2. Faktor. Hoch laden heißt, die Variablen korrelieren mit dem Faktor mit mindestens 0.7. Bei einem Korrelationskoeffizienten von 0,7 werden rund 50% ($0.7^2 = 0.49$) der gemeinsamen Varianz erklärt. Den 1. Faktor kann man daher als soziodemografischen Faktor bezeichnen, den zweiten als Zeitfaktor. (Vergleiche den Output für dieses Beispiel von Seite 217) Wie kommt man zu diesem Ergebnis? Zuerst prüft man die Korrelationsmatrix der beobachteten Variablen. Wenn alle nichtdiagonalen Elemente Null sind, wenn also eine Einheitsmatrix vorliegt, dann enthält die Korrelationsmatrix keine Redundanz, die eine Reduktion der Variablen auf weniger hypothetische Variablen ermöglicht. Die Nullhypothese ist daher

$$H_0: \mathbf{R} = \mathbf{I}$$

Die Korrelationsmatrix **R** der Grundgesamtheit ist gleich der Einheitsmatrix **I**. Die Alternative besagt, dass zumindest eine Korrelation von Null verschieden ist.

$$H_1: \mathbf{R} \neq \mathbf{I}$$

Nur bei Ablehnung der Nullhypothese ist es sinnvoll, zumindest einen Faktor aus der Korrelationsmatrix zu extrahieren. Die Testmaßzahl zur Prüfung dieser Hypothesen ist

$$\chi 2_{beob} = -\left[n - \frac{(2 \cdot t + 11)}{6}\right] \cdot \ln(\Lambda)$$

mit

$$\Lambda = |\mathbf{R}|.$$

Die Wilk'sche Prüfvariable Λ ist die Determinante der Korrelationsmatrix und t die Zahl der Variablen. Diese Maßzahl ist chiquadratverteilt mit

$$\nu = t \cdot (t - 1) / 2$$

Freiheitsgraden. Für das Beispiel ist Λ gleich

$$\left| \begin{pmatrix} 1 & -0.316 & 0.875 & 0.193 \\ -0.316 & 1 & -0.393 & 0.252 \\ 0.875 & -0.393 & 1 & 0.249 \\ 0.193 & 0.252 & 0.249 & 1 \end{pmatrix} \right| = 0.155$$

und

$$\chi 2_{beob} = -\left[12 - \frac{(2 \cdot 4 + 11)}{6}\right] \cdot \ln(0.155) = 16.468$$

Für ein 5%-iges Signifikanzniveau ist

$$\chi 2_c = 12.592$$

Da $\chi 2_{beob}$ größer ist als der kritische Wert der Chiquadratverteilung, kann man die Nullhypothese ablehnen. Die Korrelationsmatrix ist so redundant, dass sich zumindest die Extraktion eines Faktors lohnt. Um Faktoren aus der Korrelationsmatrix zu extrahieren, berechnet man zuerst die Eigenwerte dieser Matrix. Die Korrelationsmatrix hat folgende 4 Eigenwerte

$$\lambda = (2.134 \quad 1.239 \quad 0.512 \quad 0.115)$$

Nach dem Kaiser-Guttman-Kriterium sind so viele Faktoren gerechtfertigt, wie Eigenwerte größer gleich 1 vorkommen. Zwei Faktoren erklären

$$\frac{\lambda_1 + \lambda_2}{4} = \frac{2.134 + 1.239}{4} = 0.843$$

also 84.3% der Gesamtvarianz. Die Korrelationen der 4 Variablen mit den zwei Faktoren, die so genannten Ladungen, berechnet man mit Hilfe der beiden Eigenwerte und Eigenvektoren. Die Eigenvektoren der beiden Eigenwerte sind

$$\mathbf{E} = \begin{pmatrix} 0.637 & 0.073 \\ -0.367 & 0.615 \\ 0.656 & 0.07 \\ 0.171 & 0.782 \end{pmatrix}$$

Die Ladungsmatrix **L**, die die Korrelationen der Variablen mit den beiden Faktoren enthält, bestimmt man mit folgender Formel

$$\mathbf{L} = \mathbf{E} \cdot \mathbf{D}^{1/2}$$

$\mathbf{D}^{1/2}$ ist eine Diagonalmatrix mit den positiven Quadratwurzeln aus den Eigenwerten:

$$\mathbf{D}^{1/2} = \begin{pmatrix} \sqrt{2.134} & 0 \\ 0 & \sqrt{1.239} \end{pmatrix}$$

L ist daher gleich

$$\begin{pmatrix} 0.637 & 0.073 \\ -0.367 & 0.615 \\ 0.656 & 0.07 \\ 0.171 & 0.782 \end{pmatrix} \cdot \begin{pmatrix} \sqrt{2.134} & 0 \\ 0 & \sqrt{1.239} \end{pmatrix} = \begin{pmatrix} 0.931 & 0.081 \\ -0.536 & 0.685 \\ 0.958 & 0.078 \\ 0.250 & 0.870 \end{pmatrix}$$

In der ersten Spalte der Ladungsmatrix **L** stehen die Korrelationen (= Ladungen) der 4 Variablen mit dem ersten Faktor, in der zweiten Spalte die mit dem zweiten Faktor. Die Varianz der i-ten Variablen wird in der Faktorenanalyse in zwei Teile zerlegt

$$\mathrm{Var}(X_i) = h_i^2 + u_i^2 = 1$$

h_i^2 ist die so genannte Kommunalität. Sie misst den Anteil der Varianz der i-ten Variablen, der auf die gemeinsamen Faktoren entfällt. u_i^2 ist die Einzelrestvarianz. Aus der Ladungsmatrix **L** schätzt man die Kommunalitäten und Einzelrestvarianzen nach den Formeln

$$h_i^2 = \sum_{i=1}^{k} l_{ij}^2$$

$$u_i^2 = 1 - h_i^2$$

Die Ergebnisse sind in folgender Tabelle zusammengestellt:

(UNROTIERT)	Charge	Charge	Kommunalität	Einzelrestvarianz
Variablen	Faktor_1	Faktor_2	h_i^2	$1 - h_i^2$
Alter	0.931	0.081	0.873	0.127
Fernsehzeit	−0.536	0.685	0.757	0.243
Einkommen	0.958	0.078	0.924	0.076
Computerzeit	0.250	0.870	0.819	0.181

Berücksichtigt man nur Korrelationen größer gleich 0.7 ($0.7^2 = 0.49$), die also 50% der Varianz erklären, dann wird der erste Faktor durch das Alter und Einkommen hoch geladen. Der zweite Faktor wird durch die Computerzeit hoch geladen. Die Korrelation der Fernsehzeit mit den Faktoren liegt unter 0.7.

Um das Ergebnis besser interpretieren zu können, wird die Ladungsmatrix orthogonal rotiert. Man multipliziert dazu die Ladungsmatrix mit einer geeigneten Transformationsmatrix.

$$\mathbf{L}^* = \mathbf{L} \cdot \mathbf{M}$$

Das Ergebnis ist wieder in folgender Tabelle zusammengefasst:

(ROTIERT)	Charge	Charge	Kommunalität	Einzelrestvarianz
Variablen	Faktor_1	Faktor_2	h_i^2	$1 - h_i^2$
Alter	0.934	0.031	0.873	0.127
Fernsehzeit	−0.498	0.713	0.757	0.243
Einkommen	0.961	0.026	0.924	0.076
Computerzeit	0.296	0.855	0.819	0.181

Kommunalität und Einzelrestvarianz haben sich durch die orthogonale Rotation nicht verändert. Auch der erste Faktor wird weiterhin durch die Variablen Alter und Einkommen hoch geladen. Der zweite Faktor wird nun neben der Computerzeit auch durch die Fernsehzeit hoch geladen. Will man Schätzwerte für die nicht direkt beobachtbaren Faktoren bestimmen, dann kann man sie mit Hilfe folgender Formel berechnen:

$$\mathbf{F^*} = \mathbf{Z} \cdot \mathbf{R}^{-1} \cdot \mathbf{L^*}$$

Z ist die Matrix der standardisierten Beobachtungen. Jede Antwort wird vom Mittelwert der Antworten auf die entsprechende Frage abgezogen und durch die Standardabweichung der Antworten dividiert.

$$z_{ij} = \frac{x_{ij} - xquer_j}{s_j}$$

Für das Beispiel erhält man auf diese Art die Matrix der standardisierten Antworten auf diese Fragen:

$$\mathbf{Z} = \begin{pmatrix} -0.185 & -0.464 & -0.066 & 1.452 \\ 0.554 & -1.021 & 0.819 & 0.456 \\ -1.579 & -0.186 & -0.775 & -0.041 \\ 1.948 & -1.021 & 1.616 & -0.041 \\ -0.923 & 0.093 & -0.863 & -1.535 \\ -0.267 & -1.3 & -0.686 & -0.29 \\ 0.554 & 0.093 & 0.288 & -1.286 \\ 1.374 & 1.207 & 1.085 & 1.452 \\ 0.39 & -0.464 & 1.528 & -0.041 \\ -0.267 & -0.186 & -0.863 & -1.535 \\ -0.185 & 0.928 & -0.598 & 0.954 \\ -1.415 & 2.321 & -1.483 & 0.456 \end{pmatrix}$$

Die Matrix der Faktorenschätzwerte F* ist nun

$$\mathbf{F^*} = \mathbf{Z} \begin{pmatrix} 1 & -0.316 & 0.875 & 0.193 \\ -0.316 & 1 & -0.393 & 0.252 \\ 0.875 & -0.393 & 1 & 0.249 \\ 0.193 & 0.252 & 0.249 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 0.934 & 0.031 \\ -0.498 & 0.713 \\ 0.961 & 0.026 \\ 0.296 & 0.855 \end{pmatrix} = \begin{pmatrix} 0.216 & 0.735 \\ 0.909 & -0.207 \\ -1.009 & -0.23 \\ 1.805 & -0.462 \\ -1.053 & -1.086 \\ -0.185 & -0.976 \\ 0.154 & -0.806 \\ 1.052 & 1.792 \\ 0.957 & -0.215 \\ -0.703 & -1.216 \\ -0.408 & 1.157 \\ -1.733 & 1.515 \end{pmatrix}$$

Im Hinblick auf den ersten Faktor (soziodemografischen Faktor) weisen die Besucher Nr. 3, 5, 6, 10, 1 und 12 unterdurchschnittliche Werte auf, beim 2. Faktor (Zeitfaktor) die Besucher 1, 8, 11 und 12 überdurchschnittliche Werte.

9_0 SONSTIGES

Vorschau

Bei metrisch skalierten Antworten muss man bei kleinen Stichproben (n < 30) voraussetzen, dass die Stichprobe aus einer normalverteilten Grundgesamtheit stammt. Im Abschnitt 9_1, Kolmogorov-Smirnov-Test, wird gezeigt, wie man dies prüfen kann.

Bei mehr als einer Frage und metrisch skalierten Antworten gilt dies analog. Hier wird vorausgesetzt, dass die Stichproben aus multivariat normalverteilten Grundgesamtheiten stammen. Eines der möglichen Verfahren, um diese Voraussetzung zu prüfen, wird in Abschnitt 9_2 vorgestellt.

Die zweite häufige Voraussetzung bei metrisch skalierten Antworten und zwei oder mehr Stichproben ist die Homogenität der Varianzen. Die Stichproben müssen aus Grundgesamtheiten mit gleicher Varianz stammen. Testverfahren für diese Voraussetzung findet man im Abschnitt 9_3.

Im Abschnitt 9_4 wird ein Test gezeigt, mit dem man die Varianzen- und Kovarianzenhomogenität bei multivariaten Fragestellungen prüfen kann. Der entsprechende Test heißt Box-M-Test.

Die Wahrscheinlichkeits- bzw. Dichtefunktion und Verteilungsfunktion mehrerer verwendeter Verteilungen sind im letzten Abschnitt 9_5 zusammengestellt.

9_1 Kolmogorov-Smirnov-Test

Wenn der Stichprobenumfang kleiner als 30 ist, dann muss man voraussetzen, dass die Stichprobe aus einer normalverteilten Grundgesamtheit stammt. Diese Voraussetzung kann man mit Hilfe des Kolmogorov-Smirnov-Tests überprüfen. Dazu bestimmt man als Testmaßzahl die maximale Differenz zwischen beobachtetet (standardisierter) Stichprobenverteilung und Standardnormalverteilung:

$$D_{max} = \max\left(\left|F_{beob}(z) - F_{Normal}(z)\right|\right)$$

Die wöchentlichen Zeiten vor dem Computer von 5 Befragten sind

$$(4.4 \quad 12 \quad 9 \quad 9 \quad 5.6)$$

Diese Zeiten werden wie folgt standardisiert: Von jeder Zeit wird der Stichprobendurchschnitt xquer = 8 Stunden abgezogen und durch die Standardabweichung s = 2.71 Stunden dividiert.

$$z = \frac{x_i - xquer}{s}$$

Diese z-Werte findet man in folgenden Tabelle in der 4. Spalte (= "z"-Spalte). In der 5. Spalte sind die entsprechenden Verteilungswerte der Standardnormalverteilung angegeben (= "F(z)"-Spalte). In dieser Spalte findet man als ersten Wert 0.117. D. h. in einer Normalverteilung würden 11,7% der Mietzahlungen höchstens 4.4 Geldeinheiten betragen. Tatsächlich sind es 20% (siehe 3. Spalte "F(beob)"). Den entsprechenden Anteilswert von 0.20 erhält man durch die Division der Häufigkeit 1 (= 2. Spalte "hj") für diese Mietzahlungen durch den Stichprobenumfang von 5. Die Differenz zwischen beobachteten Verteilungswert 0.20 und dem der Normalverteilung von 0.117 ist absolut 0.083 (= 6. Spalte "D").

x_j	h_j	F_{beob}	z_j	F_{z_j}	D_j
4.4	1	0.2	−1.188	0.117	0.083
5.6	1	0.4	−0.792	0.214	0.186
9.0	1	0.6	0.330	0.629	0.029
9.0	1	0.8	0.330	0.629	0.171
12.0	1	1.0	1.320	0.907	0.093
(SUMME)	5	"*"	"*"	D_{max}	0.186

Analog werden die weiteren Differenzen berechnet. Die größte Differenz ist 0.186.

$$D_{max} = \max\left(\left|F_{beob}(z) - F_{Normal}(z)\right|\right) = 0.186$$

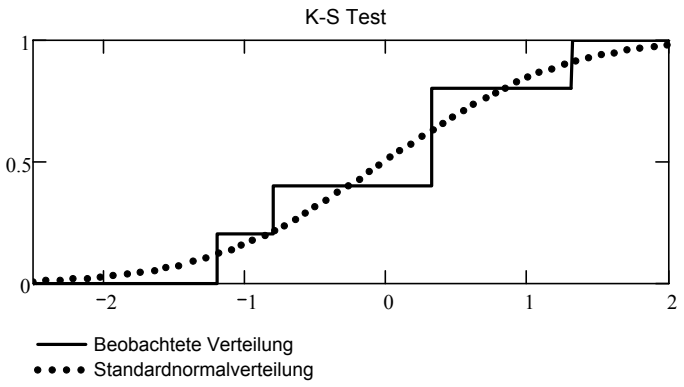
Die Verteilungsfunktion der Testmaßzahl $\sqrt{n}\ D_{max}$ kann durch folgenden Ausdruck gut angenähert werden

$$KS(t) = 1 - 2 \cdot \sum_{j=1}^{10} \left[(-1)^{j-1} \cdot \exp\left(-2 \cdot j^2 \cdot t^2\right)\right]$$

Setzt man die Rechenergebnisse in diese Formel ein, dann erhält man die Wahrscheinlichkeit für das Auftreten einer so großen maximalen Differenz unter Gültigkeit der Nullhypothese:

$$KS\left(\sqrt{5}\cdot 0.186\right) = 0.0448$$

Da diese Wahrscheinlichkeit von 1 abgezogen nicht kleiner gleich dem Signifikanzniveau von 0.05 ist, spricht nichts gegen die Hypothese, dass die beobachtete Verteilung aus einer normalverteilten Grundgesamtheit stammt. Folgende Grafik zeigt die kumulierten Beobachtungen und die entsprechende Standardnormalverteilung.



("KS-beobachtet")	("Wahrscheinlichkeit p")
0.22932	0.95518

9_2 Multivariate Normalverteilung

Für die Durchführung verschiedener Tests ist Voraussetzung, dass die Grundgesamtheit multivariat normalverteilt ist. Dies kann z. B. an Hand der Stichprobenwölbung überprüft werden. Die Stichprobenwölbung wird mit Hilfe der Formel

$$b_{2t} = \frac{1}{n} \cdot \sum_{r=1}^{n} \left[\left(x_r - xquer\right) \cdot S^{-1} \cdot \left(x_r - xquer\right) \right]^2$$

mit

$$\mathbf{S} = (\mathbf{X}^T \cdot \mathbf{X} - \mathbf{xquer} \cdot \mathbf{xquer}^T) / n$$

berechnet. Bei einer Werbeveranstaltung wurde das Alter und Einkommen von 10 Besuchern erhoben. In folgender Tabelle steht in der ersten Spalte das Alter und in der zweiten das Einkommen:

17	10
25	21
18	15
22	20
30	22
18	18
23	22
35	35
61	41
56	32

Wenn man feststellen will, ob das Alter und das Einkommen der 10 Besucher aus einer multivariat normalverteilten Grundgesamtheit stammt, berechnet man zuerst die beiden Mittelwerte:

$$xquer_1 = \frac{1}{10} \cdot (17 + 25 + ... + 56) = 30.5$$

$$xquer_2 = \frac{1}{10} \cdot (10 + 21 + ... + 32) = 23.6$$

Nun kann man die Varianz- Kovarianzmatrix **S** berechnen

$$\mathbf{S} = \frac{1}{10}\cdot\left[\begin{pmatrix} 17 & 10 \\ 25 & 21 \\ 18 & 15 \\ 22 & 20 \\ 30 & 22 \\ 18 & 18 \\ 23 & 22 \\ 35 & 35 \\ 61 & 41 \\ 56 & 32 \end{pmatrix}^{T}\cdot\begin{pmatrix} 17 & 10 \\ 25 & 21 \\ 18 & 15 \\ 22 & 20 \\ 30 & 22 \\ 18 & 18 \\ 23 & 22 \\ 35 & 35 \\ 61 & 41 \\ 56 & 32 \end{pmatrix} - \begin{pmatrix} 30.5 \\ 23.6 \end{pmatrix}\cdot\begin{pmatrix} 30.5 \\ 23.6 \end{pmatrix}^{T}\right] = \begin{pmatrix} 1062.675 & 769.32 \\ 769.32 & 583.104 \end{pmatrix}$$

und in die Formel für die Stichprobenwölbung einsetzen:

$$b_{2t} = \frac{1}{10}\cdot\sum_{r=1}^{10}\left[\left[x_r - \begin{pmatrix} 30.5 \\ 23.6 \end{pmatrix}\right]\cdot\begin{pmatrix} 1062.675 & 769.32 \\ 769.32 & 583.104 \end{pmatrix}^{-1}\cdot\left[x_r - \begin{pmatrix} 30.5 \\ 23.6 \end{pmatrix}\right]\right]^2 = 7.66$$

Da die Maßzahl

$$z_{beob} = \frac{b_{2t} - t\cdot(t+2)}{\sqrt{8\cdot t\cdot\frac{t+2}{n}}}$$

asymptotisch standardnormalverteilt ist, kann man an Hand dieser Maßzahl prüfen, ob für die beiden Fragen nach dem Alter und dem Einkommen die multivariate Normalverteilung abgelehnt werden muss. t ist die Anzahl der Fragen und z_{beob} ist daher

$$z_{beob} = \frac{7.66 - 2\cdot(2+2)}{\sqrt{8\cdot 2\cdot\frac{2+2}{10}}} = -0.134$$

Da für ein 5%-iges Signifikanzniveau die kritischen Werte der Standardnormalverteilung gleich −1.96 und 1.96 sind, kann die Nullhypothese nicht abgelehnt werden.

9_3 Homogenitätstest für Varianzen

Beide Stichprobenumfänge größer 30

Wenn die Stichprobenumfänge größer als 30 sind, dann prüft man die Hypothesen

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_1^2 \neq \sigma_2^2$$

mit Hilfe folgender normalverteilten Testmaßzahl

$$z_{beob} = \frac{1.1513 \cdot \ln\left(\frac{s_1^2}{s_2^2}\right) + \frac{1}{2} \cdot \left(\frac{1}{n_1 - 1} + \frac{1}{n_2 - 1}\right)}{\sqrt{\frac{1}{2} \cdot \left(\frac{1}{n_1 - 1} + \frac{1}{n_2 - 1}\right)}}$$

mit

$$s_1^2 = \frac{1}{n_1 - 1} \cdot \sum_{i=1}^{n_1} \left(x_{i1} - xquer_1\right)^2$$

und

$$s_2^2 = \frac{1}{n_2 - 1} \cdot \sum_{i=1}^{n_2} \left(x_{i2} - xquer_2\right)^2$$

Bei einer Werbeveranstaltung wurden 50 Männer und Frauen befragt, wie viele Stunden sie täglich vor dem Computer verbringen. Aus den Befragungsdaten errechnet man folgende Schätzwerte für die Varianzen:

$$s_1^2 = 6.796$$

$$s_2^2 = 9.715$$

Wenn man diese Ergebnisse in z_{beob} einsetzt, dann ist die Testmaßzahl gleich

$$z_{beob} = \frac{1.1513 \cdot \ln\left(\frac{6.796}{9.715}\right) + \frac{1}{2} \cdot \left(\frac{1}{50 - 1} + \frac{1}{50 - 1}\right)}{\sqrt{\frac{1}{2} \cdot \left(\frac{1}{50 - 1} + \frac{1}{50 - 1}\right)}} = -2.737$$

Für ein Signifikanzniveau von $\alpha = 0.05$ sind die beiden kritischen Werte der Normalverteilung $c_u = -1.96$ und $c_o = 1.96$. Da –2.737 kleiner ist als c_u, muss die Nullhypothese abgelehnt werden. Die beiden Varianzen sind in den Grundgesamtheiten nicht gleich.

Beide Stichprobenumfänge kleiner 30

Kann man voraussetzen, dass die beiden Stichproben aus normalverteilten Grundgesamtheiten stammen, dann verwendet man folgende Testmaßzahl, um zwischen den beiden oben angeführten Hypothesen zu entscheiden:

$$F_{beob} = \frac{s_1^2}{s_2^2} \cdot \frac{v_2}{v_1}$$

mit

$$v_1 = n_1 - 1$$

$$v_2 = n_2 - 1$$

Wurden bei der Werbeveranstaltung nicht je 50 Männer und Frauen befragt, sondern nur 16 Männer und 9 Frauen, dann gilt

$$F_{beob} = \frac{6.796}{9.715} \cdot \frac{8}{15} = 0.373$$

$$v_1 = 16 - 1 = 15$$

$$v_2 = 9 - 1 = 8$$

Für ein 5%-iges Signifikanzniveau und obige Freiheitsgrade ist der kritische Wert der F-Verteilung gleich

$$F_c\left(1 - \alpha, v_1, v_2\right) = F(0.95, 15, 8) = 3.218$$

Da der beobachtete F-Wert von 0.373 kleiner ist als der kritische F-Wert von 3.218, kann die Nullhypothese gleicher Varianzen nicht abgelehnt werden.

Mehr als 2 Stichproben und n < 30

In diesem Fall kann man folgende Testmaßzahl verwenden, um die Gleichheit der Varianzen in der Grundgesamtheit zu überprüfen:

$$\chi 2_{beob} = \frac{2.3026}{\omega} \cdot \left[(n - k) \cdot \ln\left[\frac{\sum_{i=1}^{k} \left[\left(n_i - 1\right) \cdot s_i^2 \right]}{n - k} \right] - \sum_{i=1}^{k} \left[\left(n_i - 1\right) \cdot \ln\left(s_i^2\right) \right] \right]$$

mit

$$\omega = 1 + \frac{1}{3\cdot(k-1)}\cdot\left(\sum_{i=1}^{k}\frac{1}{n_i - 1} - \frac{1}{n-k}\right)$$

Diese Maßzahl ist chiquadratverteilt mit

$$\nu = k - 1$$

Freiheitsgraden. Bei einer Werbeveranstaltung wurden 6 Käufer der Spielkonsole Yoki, 6 Nichtkäufer und 5 Unentschlossene nach ihrem Alter befragt. Folgende Schätzwerte für die Varianzen wurden berechnet:

$$s_1^2 = 21.555$$

$$s_2^2 = 221.555$$

$$s_3^2 = 146.16$$

Für die Testmaßzahl benötigt man noch ω:

$$\omega = 1 + \frac{1}{3\cdot(3-1)}\cdot\left(\sum_{i=1}^{3}\frac{1}{n_i - 1} - \frac{1}{17-3}\right) = 1.096$$

Die beobachtete Testmaßzahl ist nun

$$\chi 2_{beob} = \frac{2.3026}{1.096}\cdot\left[(17-3)\cdot\ln\left[\frac{\sum_{i=1}^{3}\left[\left(n_i - 1\right)\cdot s_i^2\right]}{17-3}\right] - \sum_{i=1}^{3}\left[\left(n_i - 1\right)\cdot\ln\left(s_i^2\right)\right]\right] = 11.969$$

Für ein Signifikanzniveau von 5% ist der kritische Wert der Chiquadratverteilung für $\nu = 3 - 1 = 2$ Freiheitsgrade gleich 5.991. Dieser Wert ist kleiner als der beobachtete von 11.969. Man kann daher die Nullhypothese ablehnen. Die Varianzen sind in den Grundgesamtheiten nicht gleich.

9_4 Box-M-Test

Für die Durchführung verschiedener Tests ist Voraussetzung, dass die Varianzen und Kovarianzen in den Grundgesamtheiten nicht verschieden sind. Dies kann man an Hand des Box-M-Tests überprüfen. Die entsprechende Testmaßzahl lautet

$$\chi_{beob}^{\,2} = \gamma \cdot \left[\sum_{i=1}^{g} \left(n_i - g \right) \right] \cdot \ln\left(\left| S \right| \right) - \sum_{i=1}^{g} \left[\left(n_i - 1 \right) \cdot \ln\left(\left| S_i \right| \right) \right]$$

mit

$$\gamma = 1 - \frac{2 \cdot p^2 + 3 \cdot p - 1}{6 \cdot (p+1) \cdot (g-1)} \cdot \left[\sum_{i=1}^{g} \frac{1}{\left(n_i - 1 \right)} - \frac{1}{\sum_{i=1}^{g} \left(n_i - 1 \right)} \right]$$

g ist die Anzahl an Stichproben und p die der Variablen. Bei einer Werbeveranstaltung wurde das Alter und Einkommen von Käufern der Spielkonsole Yoki und Nichtkäufern erhoben. In folgender Tabelle steht in der ersten Spalte die Kaufabsicht (1 = Käufer, 2 = Nichtkäufer), in der zweiten Spalte das Alter und in der dritten das Einkommen:

1	17	10
1	25	21
1	18	15
1	22	20
1	30	22
1	18	18
2	23	22
2	35	35
2	61	41
2	56	32

Wenn man feststellen will, ob die Varianz des Alters und des Einkommens sowie die Kovarianz von beiden für Käufer und Nichtkäufer in der Grundgesamtheit gleich sind, dann berechnet man zuerst die beiden Abweichungsquadrat- und -produktmatrizen:

$$\mathbf{W}_1 = \mathbf{X}_1^{\mathrm{T}} \cdot \mathbf{X}_1 - n_1 \cdot \mathbf{xquer}_1 \cdot \mathbf{xquer}_1^{\mathrm{T}} =$$

$$= \begin{pmatrix} 17 & 10 \\ 25 & 21 \\ 18 & 15 \\ 22 & 20 \\ 30 & 22 \\ 18 & 18 \end{pmatrix}^{\mathrm{T}} \cdot \begin{pmatrix} 17 & 10 \\ 25 & 21 \\ 18 & 15 \\ 22 & 20 \\ 30 & 22 \\ 18 & 18 \end{pmatrix} - 6 \cdot \begin{pmatrix} 21.667 \\ 17.667 \end{pmatrix} \cdot \begin{pmatrix} 21.667 \\ 17.667 \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} 129.333 & 92.333 \\ 92.333 & 101.333 \end{pmatrix}$$

$$\mathbf{W}_2 = \begin{pmatrix} 954.75 & 336.5 \\ 336.5 & 189 \end{pmatrix}$$

Die beiden Varianz- Kovarianzmatrizen sind

$$\mathbf{S}_1 = \mathbf{W}_1 / (n_1 - 1) = \begin{pmatrix} 25.867 & 18.467 \\ 18.467 & 20.267 \end{pmatrix}$$

$$\mathbf{S}_2 = \mathbf{W}_2 / (n_2 - 1) = \begin{pmatrix} 318.25 & 112.167 \\ 112.167 & 63 \end{pmatrix}$$

Die Varianz- Kovarianzmatrix insgesamt ist

$$\mathbf{S} = (\mathbf{W}_1 + \mathbf{W}_2) / ((n_1 - 1) + (n_2 - 1)) = \begin{pmatrix} 135.51 & 53.604 \\ 53.604 & 36.292 \end{pmatrix}$$

χ_{beob}^2 ist nun

$$\chi_{beob}^2 = 0.295 \cdot \left[\sum_{i=1}^{2} (n_i - 2) \right] \cdot \ln\left[\left| \begin{pmatrix} 135.51 & 53.604 \\ 53.604 & 36.292 \end{pmatrix} \right| \right] - \sum_{i=1}^{2} \left[(n_i - 1) \cdot \ln(|S_i|) \right] = 5.764$$

mit

$$\gamma = 1 - \frac{2 \cdot 2^2 + 3 \cdot 2 - 1}{6 \cdot (2 + 1) \cdot (2 - 1)} \cdot \left[\sum_{i=1}^{2} \frac{1}{(n_i - 1)} - \frac{1}{\sum_{i=1}^{2} (n_i - 1)} \right] = 0.295$$

Diese Testmaßzahl ist chiquadratverteilt mit

$$\nu = \frac{p \cdot (p + 1) \cdot (k - 1)}{2} = \frac{2 \cdot (2 + 1) \cdot (2 - 1)}{2} = 3$$

Freiheitsgraden. Der entsprechende kritische Wert für ein 5% Signifikanzniveau ist 7.815. Da die beobachtete Testmaßzahl mit 5.764 kleiner ist als dieser kritische Wert, kann die Nullhypothese nicht abgelehnt werden. Die Stichprobenergebnisse sprechen nicht für ungleiche Varianzen und Kovarianzen in der Grundgesamtheit.

9_5 Verteilungen

Standardnormalverteilung:

Die Standard-Normalverteilung hat die Dichtefunktion

$$f(z) = \frac{1}{\sqrt{2\cdot\pi}}\cdot e^{\frac{-z^2}{2}}$$

mit der Euler'schen Zahl e = 2.718… und der Kreiszahl π = 3.142… Die Verteilungsfunktion ist

$$F(m) = \int_0^m f(z)\,dz$$

Binomialverteilung:

Die Wahrscheinlichkeitsfunktion der Binomialverteilung ist

$$f(p,x,n) = \frac{n!}{x!\cdot(n-x)!}\cdot p^x\cdot(1-p)^{n-x} = \binom{n}{x}\cdot p^x\cdot(1-p)^{n-x}$$

und die Verteilungsfunktion ist

$$F(p,m,n) = \sum_{k=0}^{m}\left[\frac{n!}{k!\,(n-k)!}\cdot p^k\cdot(1-p)^{n-k}\right]$$

für m = 0, 1, ..., n.

Hypergeometrische Verteilung:

Zieht man aus einer dichotomen Grundgesamtheit vom Umfang n mit n_1-Einheiten der 1. Ausprägung und n_2-Einheiten der 2. Ausprägung eine Zufallsstichprobe ohne Zurücklegen des Umfanges n_1, dann ist die Zufallsvariable X, nämlich die Anzahl der Einheiten, die sowohl die Merkmalsausprägung 1 in der ersten und zweiten Verteilung besitzen, hypergeometrisch verteilt mit der Wahrscheinlichkeitsfunktion

$$W(X=x) = \frac{\binom{n_{.1}}{x}\cdot\binom{n-n_{.1}}{n_{.1}-x}}{\binom{n}{n_{.1}}} = \frac{n_{1.}!\cdot n_{.1}!\cdot n_{2.}!\cdot n_{.2}!}{n!\cdot h_{11}!\cdot h_{12}!\cdot h_{21}!\cdot h_{22}!}$$

und der Verteilungsfunktion

$$F(x) = \sum_{k=0}^{x}\frac{n_{1.}!\cdot n_{.1}!\cdot n_{2.}!\cdot n_{.2}!}{n!\cdot k!\cdot(n_{1.}-k)!\cdot(n_{.1}-k)!\cdot(n_{.2}-n_{1.}+k)!}$$

Student oder t-Verteilung:

Die Dichtefunktion der t-Verteilung ist

$$f(x,\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\cdot\Gamma\left(\frac{1}{2}\right)\cdot\sqrt{\nu}}\cdot\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

mit der Eulersche Gammafunktion $\Gamma(n) = (n-1)!$.

Die Verteilungsfunktion ist

$$F(x,\nu) = \int_0^x f(t,\nu)\,dt$$

Chiquadratverteilung:

Die Dichtefunktion der $\chi 2$-Verteilung ist gegeben durch

$$f(x,\nu) = \frac{e^{\frac{-x}{2}}}{2\cdot\Gamma\left(\frac{\nu}{2}\right)}\cdot\left(\frac{x}{2}\right)^{\frac{\nu}{2}-1}$$

mit der Euler'schen $e = 2.718$ und der Eulersche Gammafunktion $\Gamma(n) = (n-1)!$ Die Verteilungsfunktion ist

$$F(x,\nu) = \int_0^x f(t,\nu)\,dt$$

F-Verteilung:

Die F-Verteilung hat die Dichtefunktion

$$f(x,\nu_1,\nu_2) = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\cdot\Gamma\left(\frac{\nu_2}{2}\right)}\cdot\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}}\cdot\frac{x^{\frac{\nu_1}{2}-1}}{\left(1+\frac{\nu_1}{\nu_2}\cdot x\right)^{\frac{\nu_1+\nu_2}{2}}}$$

Mit $\Gamma(n) = (n-1)!$ und den Freiheitsgraden ν_1 und ν_2. Die Verteilungsfunktion ist

$$F(m,\nu_1,\nu_2) = \int_0^m f(x,\nu_1,\nu_2)\,dx$$

SOFTWAREPROGRAMME
FÜR DIE DARGESTELLTEN STATISTISCHEN VERFAHREN

Die Beschreibung der einzelnen statistischen Auswertungsverfahren an Hand einfacher Beispiele zeigt, dass selbst für einfache Verfahren der Rechenaufwand groß ist. Mit Hilfe der von den Autoren entwickelten Software wird dem Anwender nicht nur die Berechnung abgenommen, sondern auch die Interpretation der Berechnungsergebnisse. Ob ein Unterschied zwischen 2 oder mehr Stichprobenergebnissen schon verallgemeinert werden kann oder ob ein Zusammenhang nicht nur in den Stichproben vorhanden ist, sondern auch für die Grundgesamtheiten gilt wird nicht nur an Hand von Testmaßzahlen berechnet, sondern auch wörtlich interpretiert.

Für die behandelten statistischen Verfahren werden im Folgenden die Softwareei- und -ausgaben für die einzelnen Beispiele gezeigt. Um diese zu erhalten, ist es nur notwendig, die Rohdaten oder auch Häufigkeitstabellen, Kreuztabellen oder Korrelationsmatrizen einzugeben, sowie die gewünschte Auswertungsart.

Zuerst sind die Fragen und Antworten der auszuwertenden Erhebung unter der Bezeichnung

I) DATENEINGABE:

einzugeben. Die Eingabe selbst ist durch einen türkis unterlegten Balken mit dem Titel "FRAGE UND ANTWORT" gekennzeichnet:

▶ FRAGE UND ANTWORTEN

Durch doppeltes Anklicken öffnet sich der Bereich. Er ist nun oben und unten durch einen türkis unterlegten Balken mit der Bezeichnung "FRAGE UND ANTWORT" begrenzt und erlaubt die Eingaben der auszuwertenden Frage und der entsprechenden Antworten. Im folgenden Beispiel wird der Teil der Eingabe für eine Frage gezeigt:

FRAGE UND ANTWORTEN

Frage:

Zuerst gibt man die auszuwertende Frage in eine neue Tabelle "FRAGEN" in der ersten Zeile und ersten Spalte ein, dann folgen in der gleichen Zeile die vorgesehenen Antworten jeweils in einer eigenen Spalte und schließlich als letzte Spalte dieser Zeile wird ein Stichwort für die Frage eingegeben. (Für eine neue Tabelle klicken Sie auf "Einfügen", "Daten" und "Tabelle")

FRAGEN :=

	0	1	2	3
0	"Konsole kaufen?"	"Nein"	"Ja"	"Kaufabsicht"

▲ FRAGE UND ANTWORTEN

Der obere Begrenzungsbalken eines Bereiches hat am Beginn vor der Beschriftung "FRAGE UND ANTWORT" ein kleines schwarzes Dreieck dessen Spitze nach unten zeigt und der untere Balken hat ein Dreieck, dessen Spitze nach oben zeigt. Durch doppeltes Anklicken kann jeder Bereich auch wieder geschlossen werden. Im geschlossenen Zustand zeigt das schwarze Dreieck mit seiner Spitze nach rechts. Die zweite Eingabe betrifft die Auswertung:

II) AUSWERTUNG:

Hier wird z. B. eingegeben, ob ein einseitiges oder zweiseitiges Konfidenzintervalle oder ein einseitig oder zweiseitig Test zu berechnen ist. Die Eingabe ist wieder durch einen türkis unterlegten Balken gekennzeichnet wie z. B

▶ WELCHE ANTWORT, KONFIDENZINTERVALL?

Durch Doppelklicken kann dieser Bereich wieder geöffnet werden und die notwendigen Eingaben erfolgen. Voreinstellungen, die normalerweise ungeöffnet übernommen werden können, sind in einem zweiten Balken untergebracht. Im Folgenden ist ein Beispiel eines solchen Voreinstellungs-Bereichs gezeigt:

VOREINSTELLUNGEN

Voreinstellungen:

"Konfidenzniveau"	0.95
"Maximale Zeilenlänge"	80
"Fehlende Werte"	0

▲ VOREINSTELLUNGEN

Nach diesen Eingabebereichen folgen unter

III) AUSWERTUNGSERGEBNIS:

die gewünschten Auswertungsergebnisse. Diese sind im Folgenden für jedes berechnete Beispiel einzeln dargestellt. Im vierten Teil findet man noch jeweils Detailergebnisse zu den einzelnen Verfahren.

IV) DETAILERGEBNISSE:

Diese Detailergebnisse sind unter einem rosa unterlegten Balken verborgen und können bei Bedarf durch einen Doppelklick geöffnet werden.

▶ DETAILS

1_1 ANTEILSWERT, SCHÄTZVERFAHREN:

Ziel: Häufigkeitstabelle, Grafik, Anteilswerte, Punktschätzwert und Konfidenzintervall

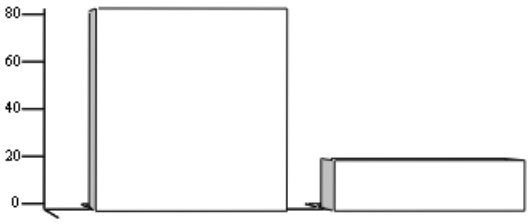
I) DATENEINGABE:

II) AUSWERTUNG:

III) AUSWERTUNGSERGEBNIS:

AUSWERTUNG
der Antworten auf die
Frage = "Werden Sie die Spielkonsole Yoki kaufen?"

HÄUFIGKEITSVERTEILUNG (in %)



$$L^T = \begin{pmatrix} \text{"Antworten:"} & \text{"Nein"} & \text{"Ja"} \\ \text{"in \%: "} & 80 & 20 \end{pmatrix}$$

Die_Antwort = "Ja" auf die Frage = "Werden Sie die Spielkonsole Yoki kaufen?" wurde von = 20 Prozent der Stichprobenbefragten gewählt.

$$T = \begin{pmatrix} \text{"Kaufabsicht"} & \text{"Häufigkeit"} & \text{"in Prozent"} \\ \text{"Nein"} & 8 & 80 \\ \text{"Ja"} & 2 & 20 \\ \text{"Summe"} & 10 & 100 \end{pmatrix}$$

Man kann mit_ = 95 % Vertrauen annehmen, dass in der Grundgesamtheit der Prozentsatz für die_Antwort = "Ja" zwischen den Grenzen 3 % und 56 % liegt."

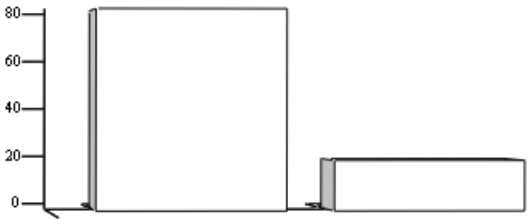
1_2 ANTEILSWERT, TEST, 1 STICHPROBE

Ziel: Häufigkeitstabelle, Grafik, Anteilswert, Unterschied zum Anteilswert der Grundgesamtheit

UNTERSCHIED

zwischen dem Anteilswert_von = 0.2 der
Antworten = "Ja" auf die
Frage = "Werden Sie die Spielkonsole Yoki kaufen?"
und = "dem Anteilswert der Grundgesamtheit von 0,4."

HÄUFIGKEITSVERTEILUNG (in %)



$$Le = \begin{pmatrix} \text{"Nr.:"} & 1 & 2 \\ \text{"Kaufabsicht"} & \text{"Nein"} & \text{"Ja"} \end{pmatrix}$$

Auf die Frage = "Werden Sie die Spielkonsole Yoki kaufen?" haben = 20 % die Antwort_ = "Ja" gewählt.

$$T = \begin{pmatrix} \text{"Kaufabsicht"} & \text{"Häufigkeit"} & \text{"in Prozent"} \\ \text{"Nein"} & 8 & 80 \\ \text{"Ja"} & 2 & 20 \\ \text{"Summe"} & 10 & 100 \end{pmatrix}$$

Kann man auf Grund dieses Stichprobenergebnisses behaupten, dass die Stichprobe aus einer Grundgesamtheit stammt mit einem Prozentsatz von = "genau 40 % für die Antwort "Ja"? "

Man kann die Behauptung = "nicht ablehnen, " dass die Stichprobe aus einer Grundgesamtheit stammt mit einem Prozentsatz von_ = "genau 40 % für die Antwort Ja. " auf die Frage = "Werden Sie die Spielkonsole Yoki kaufen?"

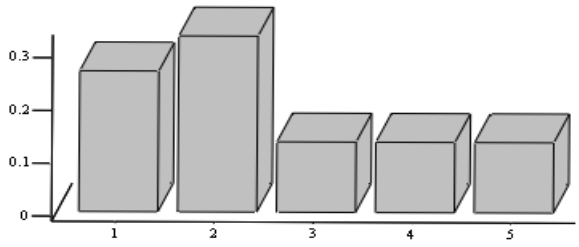
Das Risiko, dass diese Entscheidung falsch
ist = "hängt von der kronkreten Alternativhypothese ab (siehe Gütefunktion)."
(Die Wahrscheinlichkeit p = 0.121)

1_3 ZENTRALWERT, SCHÄTZVERFAHREN

Ziel: Häufigkeitstabelle, Grafik, Anteilswerte, Punktschätzwert und Konfidenzintervall

AUSWERTUNG
der Antworten auf die
Frage = "Wie beurteilen Sie das Design des Produktes?"

HÄUFIGKEITSVERTEILUNG (in %)



$$L^T = \begin{pmatrix} \text{"Nr.:"} & 1 & 2 & 3 & 4 & 5 \\ \text{"Antworten:"} & \text{"sehr gut"} & \text{"gut"} & \text{"mittelmäßig"} & \text{"schlecht"} & \text{"sehr schlecht"} \\ \text{"in \%: "} & 27 & 33 & 13 & 13 & 13 \end{pmatrix}$$

Die Hälfte der Befragten der Stichprobe haben auf die
Frage = "Wie beurteilen Sie das Design des Produktes?" die Antwortmöglichkeit höchstens = "gut"
angegeben und die andere Hälfte der Befragten die Antwortmöglichkeit mindestens = "gut" .

$$T = \begin{pmatrix} \text{"Designbeurteilung"} & \text{"Häufigkeit"} & \text{"Kum. Häuf."} & \text{"in Prozent"} & \text{"Kum. Proz."} \\ \text{"sehr gut"} & 4.00 & 4.00 & 26.67 & 26.67 \\ \text{"gut"} & 5.00 & 9.00 & 33.33 & 60.00 \\ \text{"mittelmäßig"} & 2.00 & 11.00 & 13.33 & 73.33 \\ \text{"schlecht"} & 2.00 & 13.00 & 13.33 & 86.67 \\ \text{"sehr schlecht"} & 2.00 & 15.00 & 13.33 & 100.00 \\ \text{"Summe"} & 15.00 & \text{"*"} & 100.00 & \text{"*"} \end{pmatrix}$$

In der Grundgesamtheit aller Personen, die die Stichprobe repräsentiert, kann man nicht erwarten, dass der Zentralwert auch genau = "gut" ist. Man kann jedoch mit_ = 95 % Vertrauen erwarten, dass der Zentralwert in der Grundgesamtheit_ = "zwischen "sehr gut" und "gut" liegt."

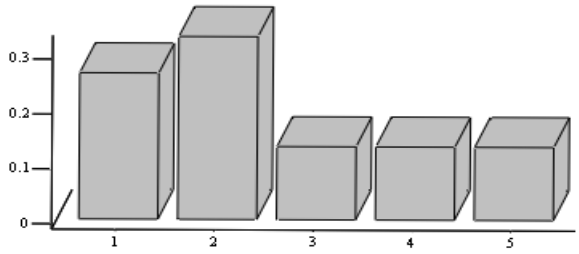
Auf die Frage = "Wie beurteilen Sie das Design des Produktes?" wählten die meisten Befragten der Stichprobe die Antwortmöglichkeit = "gut" . Von allen Befragten der Stichprobe sind dies = 33 Prozent. In der Grundgesamtheit, die die Stichprobe repräsentiert, kann man mit_ = 95 % Vertrauen erwarten, dass der Prozentsatz für diese_Antwort = "zwischen den Grenzen 12 und 62 % liegt."

1_4 ZENTRALWERT, TEST, 1 STICHPROBE

Ziel: Häufigkeitstabelle, Grafik, Anteilswert, Unterschied zu Zentralwert der Grundgesamtheit

UNTERSCHIED
zwischen dem
Stichproben- Zentralwert = "gut"
der Antworten auf die
Frage = "Wie beurteilen Sie das Design des Produktes?"
und dem Zentralwert der Grundgesamtheit
Zentralwert_ = "mittelmäßig"

HÄUFIGKEITSVERTEILUNG (in %)



$$L^T = \begin{pmatrix} \text{"Nr.:"} & 1 & 2 & 3 & 4 & 5 \\ \text{"Antworten:"} & \text{"sehr gut"} & \text{"gut"} & \text{"mittelmäßig"} & \text{"schlecht"} & \text{"sehr schlecht"} \\ \text{"in \%: "} & 27 & 33 & 13 & 13 & 13 \end{pmatrix}$$

Der Zentralwert der Antworten auf die
Frage = "Wie beurteilen Sie das Design des Produktes?" ist = "gut" .

$$T = \begin{pmatrix} \text{"Designbeurteilung"} & \text{"Häufigkeit"} & \text{"Kum. Häuf."} & \text{"in Prozent"} & \text{"Kum. Proz."} \\ \text{"sehr gut"} & 4.00 & 4.00 & 26.67 & 26.67 \\ \text{"gut"} & 5.00 & 9.00 & 33.33 & 60.00 \\ \text{"mittelmäßig"} & 2.00 & 11.00 & 13.33 & 73.33 \\ \text{"schlecht"} & 2.00 & 13.00 & 13.33 & 86.67 \\ \text{"sehr schlecht"} & 2.00 & 15.00 & 13.33 & 100.00 \\ \text{"Summe"} & 15.00 & \text{"*"} & 100.00 & \text{"*"} \end{pmatrix}$$

Kann man auf Grund dieses Stichprobenergebnisses behaupten, dass die Stichprobe aus einer Grundgesamtheit stammt, mit dem _Zentralwert = "mittelmäßig?"
Man kann die

$$\text{Behauptung} = \begin{pmatrix} \text{"nicht ablehnen, dass die Stichprobe aus einer Grundgesamtheit stammt, "} \\ \text{"mit dem Zentralwert "mittelmäßig"} \\ \text{" auf die Frage: Wie beurteilen Sie das Design des Produktes?"} \end{pmatrix}$$

Das Risiko, dass diese Entscheidung falsch ist_ = "kann nicht angegeben werden (Beta-Fehler). "

1_5 DURCHSCHNITT, SCHÄTZVERFAHREN

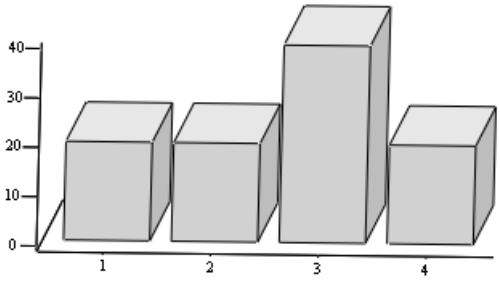
Ziel: Häufigkeitstabelle, Grafik, Durchschnitt, Punktschätzwert und Konfidenzintervall

AUSWERTUNG

der Antworten der

Frage = "Wie viele Stunden verbringen Sie pro Woche vor dem Computer?"

HÄUFIGKEITSVERTEILUNG (in %)



$$Le = \begin{pmatrix} \text{"Nr."} & 1 & 2 & 3 & 4 & \text{"Intervall"} \\ \text{"Intervallmittelpunkte"} & 2 & 6 & 8 & 15 & \text{"Zeit"} \end{pmatrix}$$

Auf die Frage = "Wie viele Stunden verbringen Sie pro Woche vor dem Computer?" antworteten die Befragten der Stichprobe im Schnitt mit = 8 . Im Mittel weichen die Antworten der Befragten um = 3.03 von diesem Durchschnitt = 8 ab.

$$T = \begin{pmatrix} \text{"Unter-"} & \text{"Ober-"} & \text{"Mittel-"} & \text{"Häufig-"} & \text{"Prozente"} & \text{"Kumulierte"} \\ \text{"grenze"} & \text{"grenze"} & \text{"punkt"} & \text{"keiten"} & \text{"%"} & \text{"%"} \\ 0 & 5 & 2 & 1 & 20 & 20 \\ 5 & 7 & 6 & 1 & 20 & 40 \\ 7 & 10 & 8 & 2 & 40 & 80 \\ 10 & 20 & 15 & 1 & 20 & 100 \\ \text{"Zeit"} & \text{"Zeit"} & \text{"Summe"} & 5 & 100 & \text{"*"} \end{pmatrix}$$

In der Grundgesamtheit aller Personen, die die Stichprobe repräsentiert, kann man nicht erwarten, dass der Durchschnitt ebenfalls = 8 ist. Man kann jedoch mit_ = 95 % Vertrauen annehmen, dass der Durchschnitt in der Grundgesamtheit = " zwischen den Grenzen 4,24 und 11,76 liegt."

1_6 DURCHSCHNITTSWERT, t-TEST, 1 STICHPROBE

Ziel: Häufigkeitstabelle, Grafik, Durchschnitt, Unterschied zu Durchschnitt der Grundgesamtheit

UNTERSCHIED
zwischen dem
Stichprobendurchschnitt = 8
der
Frage = "Wie viele Stunden verbringen Sie pro Woche vor dem Computer?"
und = "dem Durchschnitt der Grundgesamtheit von höchstens 3."

Der Durchschnitt aus den Antworten auf die
Frage = "Wie viele Stunden verbringen Sie pro Woche vor dem Computer?" ist = 8 in der Stichprobe mit einer Standardabweichung_von= 3.03 und einem Umfang von n = 5 . .

$$T = \begin{pmatrix} \text{"Unter-"} & \text{"Ober-"} & \text{"Mittel-"} & \text{"Häufig-"} & \text{"Prozente"} & \text{"Kumulierte"} \\ \text{"grenze"} & \text{"grenze"} & \text{"punkt"} & \text{"keiten"} & \text{"%"} & \text{"%"} \\ 0 & 5 & 2 & 1 & 20 & 20 \\ 5 & 7 & 6 & 1 & 20 & 40 \\ 7 & 10 & 8 & 2 & 40 & 80 \\ 10 & 20 & 15 & 1 & 20 & 100 \\ \text{"Zeit"} & \text{"Zeit"} & \text{"Summe"} & 5 & 100 & \text{"*"} \end{pmatrix}$$

Kann man auf Grund dieses Stichprobenergebnisses behaupten, dass die Stichprobe aus einer Grundgesamtheit stammt, mit = "dem Durchschnitt von höchstens 3?"

Ergebnis:

$$Es = \begin{pmatrix} \text{"Man kann nicht annehmen, dass die Stichprobe aus einer Grundgesamtheit stammt, "} \\ \text{"mit dem Durchschnitt von höchstens 3"} \end{pmatrix}$$

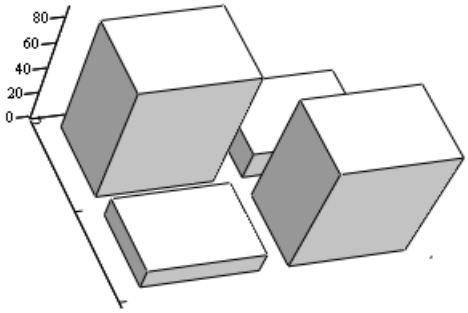
Das Risiko, dass diese Entscheidung falsch ist_ = "ist höchstens 5 %." (Die Wahrscheinlichkeit p = 0.01)

2_1 ANTEILSWERTDIFFERENZ, 2 UNABHÄNGIGE STICHPROBEN, z-TEST

Ziel: Kreuztabelle, Grafik, Anteilswerte, Unterschied zwischen den Anteilswerten

UNTERSCHIED
zwischen den Antworten der
Frage1 = "Geschlecht?"
und der
Frage2 = "Werden Sie die Spielkonsole kaufen?"

2-DIMENSIONALE VERTEILUNG (in %)



$$S = \begin{pmatrix} \text{"Geshclecht"} & \text{"Geshclecht"} \\ \text{"Männlich"} & \text{"Weiblich"} \end{pmatrix} \qquad L^T = \begin{pmatrix} \text{"Ja"} & 1 \\ \text{"Nein"} & 2 \end{pmatrix}$$

Zwischen den Stichproben = "Geshclecht - Männlich" und "Geshclecht - Weiblich" , ausgezählt nach der _Antwort = "Ja" der Frage2 = "Werden Sie die Spielkonsole kaufen?" , bestehen offensichtlich Unterschiede in den jeweiligen Prozentsätzen für diese Antwort. Für die Stichprobe_1 = "Geshclecht - Männlich" ist der Anteil der Befragten mit dem Merkmal = "Ja 85,7 %" und in der Stichprobe_2_ = "Geshclecht - Weiblich 20 %"

$$KT2 = \begin{pmatrix} \text{"Geshclecht"} & \text{"Männlich"} & \text{"Weiblich"} \\ \text{"Kaufabsicht"} & \text{"in %"} & \text{"in %"} \\ \text{"Ja"} & 86 & 20 \\ \text{"Nein"} & 14 & 80 \\ \text{"SUMME RELATIV"} & 100 & 100 \end{pmatrix}$$

Auf Grund der Stichprobenergebnisse kann man = "behaupten, "
das in den beiden Grundgesamtheiten = "Geshclecht - Männlich" und "Geshclecht - Weiblich"
signifikante Anteilswertunterschiede im Hinblick auf die Antwort_ = "Ja" bestehen. Der Anteil der Personen mit dieser Antwort kann in der Grundgesamtheit_1 = "Geshclecht - Männlich" größer als der" Anteil in der Grundgesamtheit_2 = "Geshclecht - Weiblich" angenommen werden.

Das Risiko, dass diese Entscheidung falsch ist_ = "ist höchstens 5 %." (Die Wahrscheinlichkeit p = 0.011)

2_2 ANTEILSWERTDIFFERENZEN, 2 UND MEHR UNABHÄNGIGE STICHPROBEN, CHIQUADRAT-TEST

Ziel: Kreuztabelle, Grafik, Anteilswerte, Unterschied zwischen den Anteilswerten

UNTERSCHIEDE
zwischen den Antworten auf die
Frage1 = "Familienstand?"
und die
Frage2 = "Werden Sie die Spielkonsole Yoki kaufen?"

Zwischen den Stichproben der_Frage1 = ("Familienstand - Verheiratet" "Familienstand - Nichtverheiratet") , ausgezählt nach den_Antworten = ("Ja" "Nein" "Weiß nicht") der Frage2 = "Werden Sie die Spielkonsole Yoki kaufen?" , bestehen offensichtlich Unterschiede in den jeweiligen Prozentsätzen für die einzelnen Antworten.

KT2 =

"Familienstand"	"Verheiratet"	"Nichtverheiratet"
"Kaufabsichr"	"in %"	"in %"
"Ja"	71	33
"Nein"	5	33
"Weiß nicht"	24	33
"SUMME"	100	100

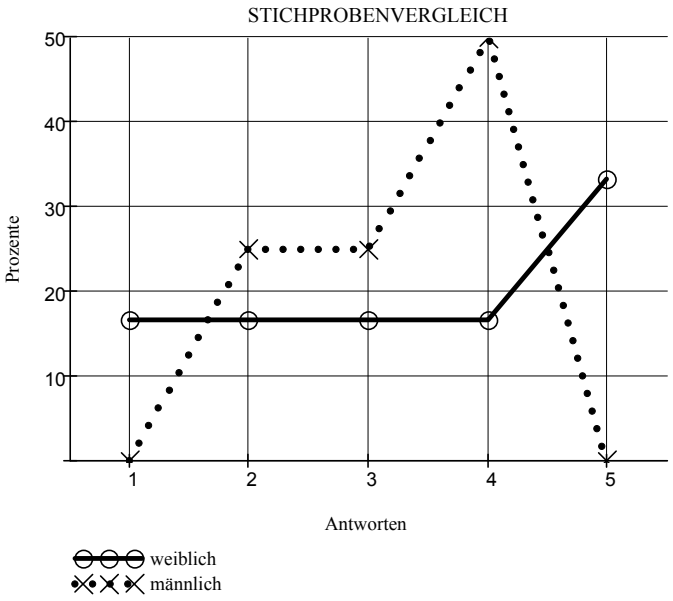
Kann man auf Grund dieser Stichprobenergebnisse behaupten, dass im Hinblick auf die Stichproben = ("Familienstand - Verheiratet" "Familienstand - Nichtverheiratet") signifikante Anteilswertunterschiede in den Ausprägungen = ("Ja" "Nein" "Weiß nicht") bestehen?
Testergebnis = "Ja, es bestehen signifikante Anteilswertunterschiede. "

2_3 VERTEILUNGSDIFFERENZ, 2 UNABHÄNGIGE STICHPROBEN, U-TEST, ORDINAL

Ziel: Kreuztabelle, Grafik, Anteilswerte, Unterschied zwischen den beiden Verteilungen

UNTERSCHIEDE
zwischen den Antworten auf die
Frage1 = "Geschlecht?"
und die
Frage2 = "Wie beruteilen Sie das Design?"

$$FL = \begin{bmatrix} \text{"Geschlecht"} \\ \begin{pmatrix} \text{"rot = weiblich"} \\ \text{"blau = männlich"} \end{pmatrix} \end{bmatrix}$$



$$Le = \begin{pmatrix} \text{"design"} & 1 & 2 & 3 & 4 & 5 \\ \text{"Geschlecht"} & \text{"sehr gut"} & \text{"gut"} & \text{"mittelmäßig"} & \text{"schlecht"} & \text{"sehr schlecht"} \end{pmatrix}$$

Zwischen den Stichproben = "Geschlecht - "weiblich" und Geschlecht - "männlich" " , ausgezählt nach den Ausprägungen = ("sehr gut" "gut" "mittelmäßig" "schlecht" "sehr schlecht") der Frage2 = "Wie beruteilen Sie das Design?" , bestehen offensichtlich Unterschiede in den jeweiligen Prozentsätzen für die einzelnen Antworten in den Stichproben.

KT1 =

"Geschlecht"	"weiblich"	"männlich"
"design"	"absolut"	"absolut"
"sehr gut"	1	0
"gut"	1	1
"mittelmäßig"	1	1
"schlecht"	1	2
"sehr schlecht"	2	0
"SUMME"	6	4

Auf Grund der Stichprobenergebnissen kann man = "nicht annehmen," dass diese Verteilungsunterschiede in den Prozentsätzen der jeweiligen Antworten auch in der Grundgesamtheit aller Personen existieren, aus der die Stichproben stammen. Die Unterschiede sind insgesamt = "nicht groß genug," um sie als signifikant zu bezeichnen.

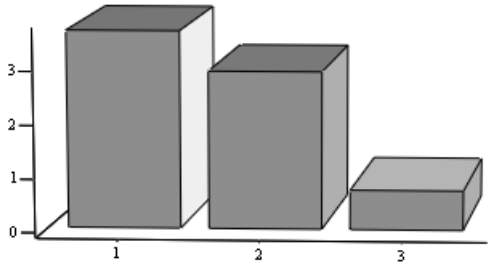
Das Risiko, dass diese Entscheidung falsch
ist_ = "hängt von der kronkreten Alternativhypothese ab (siehe Gütefunktion)." (Die Wahrscheinlichkeit
p = 0.586)

2_4 DURCHSCHNITTSDIFFERENZ, 2 UNABHÄNGIGE STICHPROBEN, TEST

Ziel: Kreuztabellen, Grafik, Durchschnitte, Unterschiede zwischen Durchschnitten

UNTERSCHIED
zwischen den 2 Stichproben der
Frage1 = "Geschlecht - Männlich" und "Geschlecht - Weiblich"
und den Antworten der
Frage2 = "Wie lange sitzen Sie pro Tag vor dem Computer?"

MITTELWERTSVERGLEICHE



$$Le = \begin{pmatrix} "+" & 1 & 2 & 3 \\ "Geschlecht" & "Männlich" & "Weiblich" & "Differenz" \\ "Computerzeit" & 3.688 & 2.944 & 0.743 \end{pmatrix}$$

Zwischen den Stichproben = "Geschlecht - Männlich" und "Geschlecht - Weiblich" , ausgezählt nach den Antworten der Frage2 = "Wie lange sitzen Sie pro Tag vor dem Computer?" , bestehen offensichtlich Unterschiede in den Durchschnitten. Für die Stichprobe_1 = "Geschlecht - Männlich" ist der Durchschnitt für die Frage2 = "Wie lange sitzen Sie pro Tag vor dem Computer?" gleich = 3.688 und für die Stichprobe_2 = "Geschlecht - Weiblich" ist er = 2.944

$$KT = \begin{pmatrix} "Computerzeit" & "Computerzeit" & "Geschlecht" & "Geschlecht" \\ "Untergrenze" & "Obergrenze" & "Männlich" & "Weiblich" \\ 0 & 2 & 4 & 4 \\ 2 & 4 & 6 & 1 \\ 4 & 6 & 2 & 2 \\ 6 & 8 & 2 & 1 \\ 8 & 9 & 2 & 1 \\ "absolut" & "Summe" & 16 & 9 \end{pmatrix}$$

Kann man behaupten, dass diese Unterschiede in den Durchschnitten auch in der Grundgesamtheit aller Personen existieren, aus der die beiden Stichproben stammen? Die Antwort ist = "Nein, man kann keine signifikanten Unterschiede nachweisen." .

Das Risiko, dass diese Entscheidung falsch
ist_ = "hängt von der kronkreten Alternativhypothese ab (siehe Gütefunktion)."
(Die Wahrscheinlichkeit p = 0.274)

3_1 ANTEILSWERTDIFFERENZ, 2 ABHÄNGIGE STICHPROBEN, McNEMAR TEST

Ziel: Kreuztabellen, Grafik, Anteilswerte, Unterschiede zwischen Anteilswerten

UNTERSCHIED
zwischen den Antworten der
Frage1 = "Einstellung zu Konsolenspielen?"
und der
Frage2 = "Wann war die Befragung?"

Zwischen den Stichproben = "Vor Werbeveranstaltung" und "Nach Werbeveranstaltung" , ausgezählt nach der Antwort_ = "positiv" , bestehen offensichtlich Unterschiede in den jeweiligen Prozentsätzen für diese Antwort. Für die Stichprobe_1 = "Vor Werbeveranstaltung" ist der Anteil der Befragten mit dem Merkmal_ = "positiv 62,5 %" und in der Stichprobe_2_ = "Nach Werbeveranstaltung - positiv 50 %"

$$KT1 = \begin{pmatrix} \text{"Einstellung"} & \text{"Vor Werbeveranstaltung"} & \text{"Nach Werbeveranstaltung"} \\ \text{"Befragung"} & \text{"absolut"} & \text{"absolut"} \\ \text{"positiv"} & 5 & 1 \\ \text{"negativ"} & 3 & 1 \\ \text{"SUMME ABSOLUT"} & 8 & 2 \end{pmatrix}$$

Auf Grund dieser Stichprobenergebnisse kann man = "nicht behaupten," dass diese Unterschiede in den Prozentsätzen der jeweiligen Antworten auch in der Grundgesamtheit aller Personen existieren, aus der die Stichprobe stammt. Der Unterschied ist insgesamt = "nicht groß genug," um ihn als signifikant zu bezeichnen (McNemar Test).

Die Wahrscheinlichkeit, dass diese Schlussfolgerung falsch ist_ = "hängt von der kronkreten Alternativhypothese ab (siehe Gütefunktion)." (Die Wahrscheinlichkeit p = 0.655)

3_2 ANTEILSWERTDIFFERENZEN, 2 ABHÄNGIGE STICHPROBEN, WILCOXON-TEST, ORDINAL

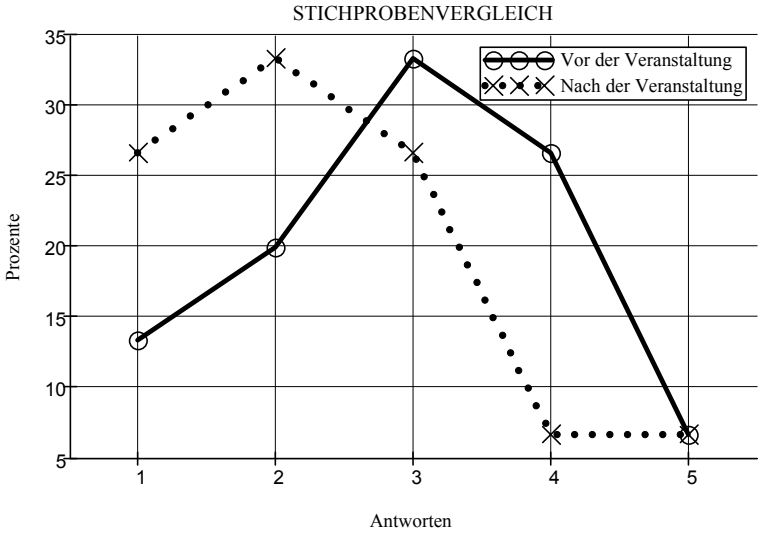
Ziel: Kreuztabelle, Grafik, Anteilswerte, Unterschied zwischen den beiden Verteilungen

UNTERSCHIEDE

zwischen den Antworten auf die Fragen

$$FF = \begin{pmatrix} \text{"Beurteilung des Designs vor der Veranmstaltung?"} \\ \text{"Beurteilung des Designs nach der Veranmstaltung?"} \end{pmatrix}$$

$$FL = \begin{pmatrix} \text{"rot = Vor der Veranstaltung"} \\ \text{"blau = Nach der Veranstaltung"} \end{pmatrix}$$



$$LLL = \begin{pmatrix} \text{"sehr gut"} & \text{"gut"} & \text{"befriedigend"} & \text{"genügend"} & \text{"nicht genügend"} \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}$$

Zwischen den Antworten der Stichproben_ = "Vor der Veranstaltung und Nach der Veranstaltung" , bestehen offensichtlich Unterschiede in den beiden Stichproben.

$$KT1 = \begin{pmatrix} "*" & "Vor\ der\ Veranstaltung" & "Nach\ der\ Veranstaltung" \\ "*" & "absolut" & "absolut" \\ "sehr\ gut" & 2 & 4 \\ "gut" & 3 & 5 \\ "befriedigend" & 5 & 4 \\ "genügend" & 4 & 1 \\ "nicht\ genügend" & 1 & 1 \\ "SUMME" & 15 & 15 \end{pmatrix}$$

Auf Grund der Stichprobenergebnissen kann man = "nicht annehmen," dass diese Verteilungsunterschiede in den Prozentsätzen der jeweiligen Antworten auch in der Grundgesamtheit aller Personen existieren, aus der die Stichproben_ = "Vor der Veranstaltung und Nach der Veranstaltung" stammen. Die Unterschiede sind insgesamt = "nicht groß genug," um sie als signifikant zu bezeichnen.

Das Risiko, dass diese Entscheidung falsch
ist_ = "hängt von der kronkreten Alternativhypothese ab (siehe Gütefunktion)." (p = 0.018)

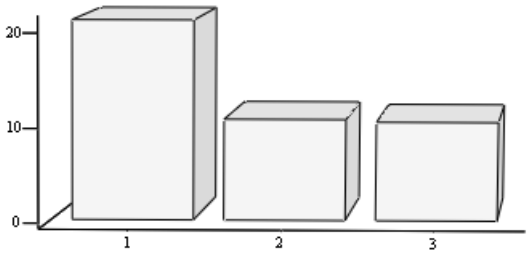
3_3 DURCHSCHNITTSDIFFERENZ, 2 ABHÄNGIGE STICHPROBEN, t-TEST

Ziel: Kreuztabellen, Grafik, Durchschnitte, Unterschiede zwischen Durchschnitten

UNTERSCHIED
zwischen den Antworten auf die verbundenen Fragen

$$F = \begin{pmatrix} \text{"Spielzeit für das Spiel Fantasy beim 1. Versuch"} \\ \text{"Spielzeit für das Spiel Fantasy beim 3 Versuch"} \end{pmatrix}$$

MITTELWERTSVERGLEICHE



$$Le = \begin{pmatrix} 1 & 2 & 3 \\ \text{"1. Versuch"} & \text{"2. Versuch"} & \text{"Differenz"} \\ 21.2 & 10.733 & 10.467 \end{pmatrix}$$

Zwischen den Antworten = "1. Versuch" und "2. Versuch " , bestehen offensichtlich Unterschiede in den jeweiligen Mittelwerten.

$$TT = \begin{pmatrix} \text{"Unter- "} & \text{"Obergrenze"} & \text{"Mitte"} & \text{"1. Versuch"} & \text{"2. Versuch"} \\ 4 & 12 & 8 & 3 & 8 \\ 12 & 20 & 16 & 3 & 6 \\ 20 & 36 & 28 & 9 & 1 \\ \text{"ABSOLUT"} & \text{"*"} & \text{"Summe"} & 15 & 15 \end{pmatrix}$$

Auf Grund dieser Stichprobenergebnissen kann man = "annehmen, " dass diese Unterschiede in den Durchschnitten der jeweiligen Antworten auch in der Grundgesamtheit aller Personen existieren, aus der die Stichproben stammen. Die Unterschiede sind insgesamt = "groß genug," um sie als signifikant zu bezeichnen (t-Test).

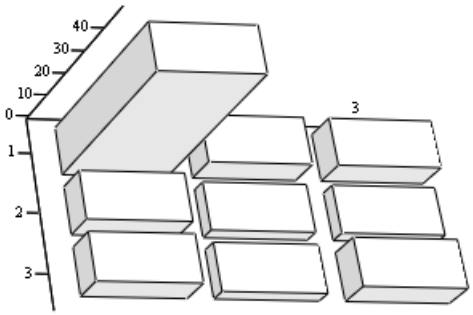
Das Risiko, dass diese Entscheidung falsch ist_ = "ist höchstens 5 %." (Die Wahrscheinlichkeit p = 0)

4_1 KONTINGENZKOEFFIZIENTEN, ZUSAMMENHANG, CHIQUADRAT-TEST

Ziel: Kreuztabelle, Grafik, Anteilswerte, Zusammenhang zwischen den Anteilswerten

ZUSAMMENHANG
zwischen den Antworten der
Frage1 = "Wie beurteilen Sie das Design?"
und der
Frage2 = "Wie beurteilen Sie die Veranstaltung"

KREUZTABELLE



$$L = \begin{bmatrix} \text{"Design"} \\ \begin{pmatrix} \text{"gut"} & \text{"mittel"} & \text{"schlecht"} \\ 1 & 2 & 3 \end{pmatrix} \end{bmatrix}$$

Zwischen der Frage1 = "Wie beurteilen Sie das Design?" und der Frage2 = "Wie beurteilen Sie die Veranstaltung" bestehen offensichtlich ein Zusammenhang in den Stichprobenergebnissen.

$$KT1 = \begin{pmatrix} \text{"Design"} & \text{"gut"} & \text{"mittel"} & \text{"schlecht"} \\ \text{"Veranstaltung"} & \text{"absolut"} & \text{"absolut"} & \text{"absolut"} \\ \text{"gut"} & 12 & 2 & 2 \\ \text{"mittel"} & 2 & 1 & 1 \\ \text{"schlecht"} & 2 & 1 & 2 \\ \text{"SUMME"} & 16 & 4 & 5 \end{pmatrix}$$

Auf Grund dieser Stichprobenergebnissen kann man = "nicht annehmen," dass diese Zusammenhänge in den Prozentsätzen der jeweiligen Antworten auch in der Grundgesamtheit aller Personen existieren, aus der die Stichprobe stammt. Der Zusammenhang in der Stichprobe ist insgesamt = "nicht groß genug," um ihn als signifikant zu bezeichnen.

Die Wahrscheinlichkeit, dass diese Schlussfolgerung falsch ist = "kann nicht angegeben werden (Beta-Fehler).," (Die Wahrscheinlichkeit $p = 0.6$)

4_2 RANGKORRELATIONSKOEFFIZIENT, ZUSAMMENHANG, z-TEST, ORDINAL

Ziel: Grafik, Rangkorrelationskoeffizient, Zusammenhang zwischen den Stichproben

ZUSAMMENHANG

zwischen den Antworten der

Frage1 = "Wie beurteilen Sie das Design?"

und der

Frage2 = "Wie beurteilen Sie die Veranstaltung?"

Zwischen der Frage1 = "Wie beurteilen Sie das Design?" und der Frage2 = "Wie beurteilen Sie die Veranstaltung?" besteht offensichtlich ein Zusammenhang in den Stichprobenergebnissen und zwar:

Zz = ("Je mehr "Design" desto mehr "Veranstaltung" und umgekehrt"
"Positiver Zusammenhang: r = 0,667")

KT1 =

"Design"	"sehr gut"	"gut"	"befriedigend"	"genügend"	"nicht genügend"
"Veranstaltung"	"absolut"	"absolut"	"absolut"	"absolut"	"absolut"
"sehr gut"	1	0	0	0	0
"gut"	0	1	0	0	0
"mittelmäßig"	0	0	0	1	0
"schlecht"	0	0	0	1	0
"sehr schlecht"	0	0	1	0	0
"SUMME"	1	1	1	2	0

Gilt dieser Zusammenhang auch für die beiden Grundgesamtheiten? Auf Grund dieser Stichprobenergebnisse kann man = "nicht annehmen," dass diese Zusammenhänge in den jeweiligen Antworten auch in den Grundgesamtheiten aller Personen existieren, aus der die Stichproben stammen. Der Zusammenhang in den
Stichproben ist insgesamt = "nicht groß genug," um ihn als signifikant zu bezeichnen.

Die Wahrscheinlichkeit, dass diese Schlussfolgerung falsch
ist_ = "hängt von der kronkreten Alternativhypothese ab (siehe Gütefunktion)." (Die Wahrscheinlichkeit p = 0.109)

4_3 MASSKORRELATIONSKOEFFIZIENT, ZUSAMMENHANG, z-TEST

Ziel: Grafik, Korrelationskoeffizient, Zusammenhang zwischen den Stichproben

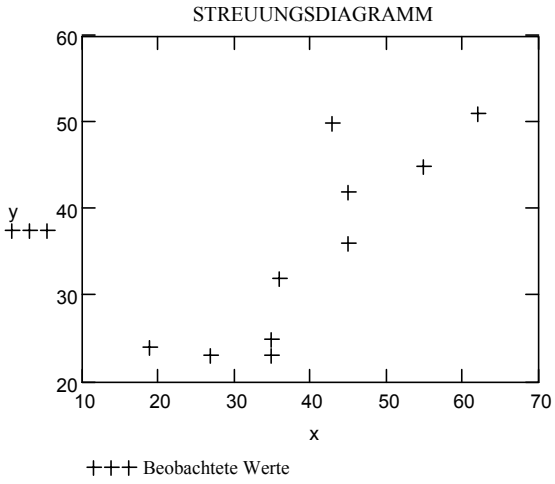
ZUSAMMENHANG

zwischen den Antworten der

Frage1 = "Wie alt sind Sie?"

und der

Frage2 = ""Wie hoch ist Ihr monatliches Einkommer



Lx = "Alter" Ly = "Einkommen"

Zwischen der Frage1 = "Wie alt sind Sie?" und der Frage2 = ""Wie hoch ist Ihr monatliches Einkommer besteht offensichtlich ein Zusammenhang in den Stichprobenergebnissen und zwar.

$$Zz = \begin{pmatrix} \text{"Je mehr "Alter" desto mehr "Einkommen" und umgekehr} \\ \text{"Positiver Zusammenhang: r = 0,851"} \end{pmatrix}$$

Gilt dieser Zusammenhang auch für die beiden Grundgesamtheiten? Auf Grund der Stichprobenergebnissen kann man = "annehmen, "dass dieser lineare Zusammenhang in den jeweiligen Antworten auch in den Grundgesamtheiten aller Personen existieren, aus der die Stichproben stammen. Der Zusammenhang in den
Stichproben ist insgesamt = "groß genug," um ihn als signifikant zu bezeichnen.

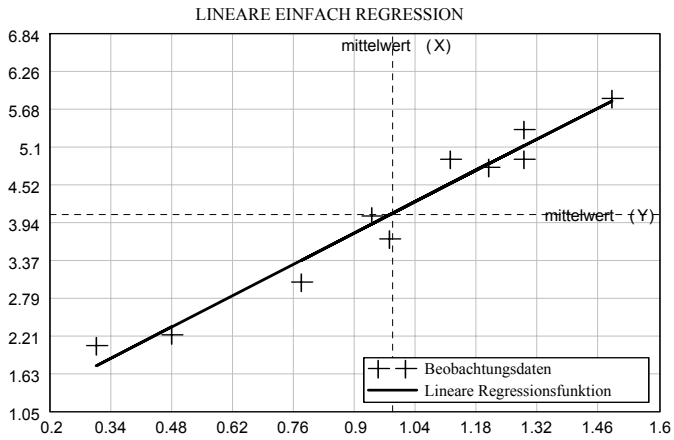
Die Wahrscheinlichkeit, dass diese Schlussfolgerung falsch ist_ = "ist höchstens 5 %." (Die Wahrscheinlichkeit p = 0.0018)

4_4 REGRESSIONSKOEFFIZIENTEN, ZUSAMMENHANG, z-TEST

Ziel: Grafik, Test der Regressionskoeffizienten, Prognose

LINEARE ABHÄNGIGKEIT

zwischen den Antworten der unabhängigen
Frage1 = "Wie alt sind Sie?"
und den Antworten der abhängigen
Frage2 = ""Wie hoch ist Ihr monatliches Einkommen?"



Zwischen der abhängigen Variable = "Einkommen" und der unabhängigen Variablen = "Alter" besteht offensichtlich in der Stichprobe ein Zusammenhang. Die aus der Stichprobe ermittelte Regressionsfunktion = " [Einkommen = 4,585 + (0,759) * Alter] " . Sie besagt, dass man mit Hilfe der unabhängigen Variablen = "Alter" und geeigneter Gewichtung einen Schätzwert für die abhängige Variable = "Einkommen" berechnen kann. Wenn sich der Wert der unabhängigen Variablen = "Alter" um eine Einheit erhöht, dann verändert sich der Wert der abhängigen Variable = "Einkommen" um = 0.166 Einheiten.

Auf Grund dieser Stichprobenergebnissen kann man = "annehmen, " dass dieser lineare Zusammenhang zwischen der abhängigen Variablen und der unabhängigen Variablen auch in der Grundgesamtheit aller Personen existiert, aus der die Stichprobe stammt. Der Anteil der erklärten Varianz der abhängigen Variablen ist insgesamt = "groß genug," um ihn als signifikant zu bezeichnen. Die Wahrscheinlichkeit, dass diese Schlussfolgerung falsch ist = ("ist höchstens" 5 "%") (Die Wahrscheinlichkeit sig = 0.0018)

Kann man auf Grund dieses Stichprobenergebnisses behaupten, dass die Stichprobe aus einer Grundgesamtheit stammt, mit = "dem Regressionskoffizienten beta_1 3?"
Ergebnis:

Es = ("Man kann nicht annehmen, dass die Stichprobe aus einer Grundgesamtheit stammt, "
"mit dem Regressionskoeffizienten beta_1 = 3")

Das Risiko, dass diese Entscheidung falsch ist_ = "ist höchstens 5 %." (Die Wahrscheinlichkeit p = 0 **).**

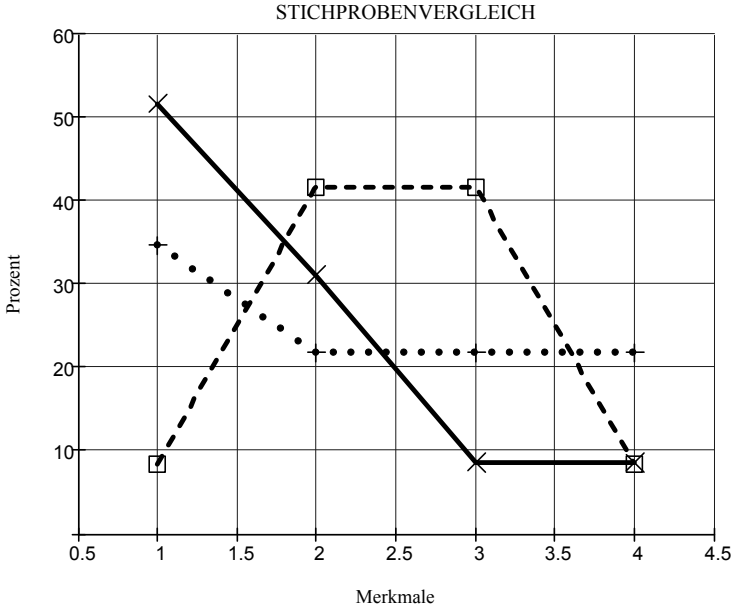
5_1 ANTEILSWERTDIFFERENZEN, 2 UND MEHR UNABHÄNGIGE STICHPROBEN, CHIQUADRAT-TEST

Ziel: Kreuztabelle, Grafik, Anteilswerte, Unterschied zwischen den Anteilswerten

UNTERSCHIEDE

zwischen den Antworten auf die
Frage1 = "Werden Sie die Spielkonsole Yoki kaufen?"
und die
Frage2 = "Wie beurteilen Sie die Handlichkeit der Spielkonsole?"

$$FL = \begin{bmatrix} \text{"Kaufabsicht"} \\ \begin{bmatrix} (\text{"rot"} \quad \text{"="} \quad \text{"Ja"}) \\ (\text{"blau"} \quad \text{"="} \quad \text{"Nein"}) \\ (\text{"grün"} \quad \text{"="} \quad \text{"Weiß nicht"}) \end{bmatrix} \end{bmatrix}$$



$$Le = \begin{pmatrix} \text{"Handlichkeit"} & 1 & 2 & 3 & 4 \\ \text{"Kaufabsicht"} & \text{"Sehr gut"} & \text{"Gut"} & \text{"Mittelmäßig"} & \text{"Schlecht"} \end{pmatrix}$$

Zwischen den Stichproben der_Frage1 = ("Kaufabsicht - Ja" "Kaufabsicht - Nein" "Kaufabsicht - Weiß nicht") , ausgezählt nach den_Antworten = ("Sehr gut" "Gut" "Mittelmäßig" "Schlecht") der Frage2 = "Wie beurteilen Sie die Handlichkeit der Spielkonsole?" , bestehen offensichtlich Unterschiede in den jeweiligen Prozentsätzen für die einzelnen Antworten.

$$KT1 = \begin{pmatrix} \text{"Kaufabsicht"} & \text{"Ja"} & \text{"Nein"} & \text{"Weiß nicht"} \\ \text{"Handlichkeit"} & \text{"absolut"} & \text{"absolut"} & \text{"absolut"} \\ \text{"Sehr gut"} & 30 & 8 & 1 \\ \text{"Gut"} & 18 & 5 & 5 \\ \text{"Mittelmäßig"} & 5 & 5 & 5 \\ \text{"Schlecht"} & 5 & 5 & 1 \\ \text{"SUMME"} & 58 & 23 & 12 \end{pmatrix}$$

Kann man auf Grund dieser Stichprobenergebnisse behaupten, dass im Hinblick auf die Stichproben = ("Kaufabsicht - Ja" "Kaufabsicht - Nein" "Kaufabsicht - Weiß nicht") signifikante Anteilswertunterschiede in den Ausprägungen = ("Sehr gut" "Gut" "Mittelmäßig" "Schlecht") bestehen? Testergebnis = "Ja, es bestehen signifikante Anteilswertunterschiede. " (Die Wahrscheinlichkeit p = 0.015)

Wenn die Unterschiede allgemein groß genug sind, um sie als signifikant zu bezeichnen, dann will man wissen, zwischen welchen Stichprobenpaaren die Anteilswertunterschiede signifikant sind. Bei folgenden Antworten sind die Abweichungen vom Durchschnitt signifikant:(Signifikanzniveau von = "5 %")

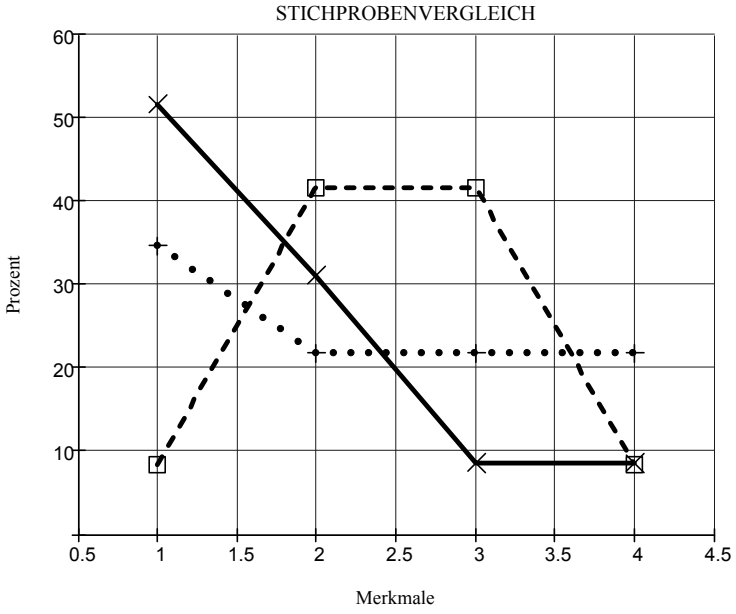
$$ss = \begin{bmatrix} \text{"Befragte mit dem Merkmal "} & \text{"Ja"} \\ \text{" geben überdurchschnittlich oft die Antworten"} & (\text{"Sehr gut"}) \\ \text{"Befragte mit dem Merkmal "} & \text{"Ja"} \\ \text{" geben unterdurchschnittlich oft die Antworten"} & (\text{"Mittelmäßig"}) \\ \text{"Befragte mit dem Merkmal "} & \text{"Nein"} \\ \text{" geben überdurchschnittlich oft die Antworten"} & (\text{"Schlecht"}) \\ \text{"Befragte mit dem Merkmal "} & \text{"Weiß nicht"} \\ \text{" geben überdurchschnittlich oft die Antworten"} & \begin{pmatrix} \text{"Gut"} \\ \text{"Mittelmäßig"} \end{pmatrix} \\ \text{"Befragte mit dem Merkmal "} & \text{"Weiß nicht"} \\ \text{" geben unterdurchschnittlich oft die Antworten"} & (\text{"Sehr gut"}) \end{bmatrix}$$

5_2 VERTEILUNGSDIFFERENZEN, 3 UND MEHR UNABHÄNGIGE STICHPROBEN, H-TEST

Ziel: Kreuztabellen, Grafik, Anteilswerte, Unterschied zwischen den Verteilungen

UNTERSCHIEDE
zwischen den Antworten auf die
Frage1 = "Wie beurteilen Sie die Handlichkeit der Spielkonsole?"
und die
Frage2 = "Werden Sie die Spielkonsole Yoki kaufen?"

$$FL = \begin{bmatrix} \text{"Kaufabsicht"} \\ \begin{bmatrix} (\text{"rot"} \quad \text{"="} \quad \text{"Ja"}) \\ (\text{"blau"} \quad \text{"="} \quad \text{"Nein"}) \\ (\text{"grün"} \quad \text{"="} \quad \text{"Weiß nicht"}) \end{bmatrix} \end{bmatrix}$$



$$Le = \begin{pmatrix} \text{"Handlichkeit"} & 1 & 2 & 3 & 4 \\ \text{"Kaufabsicht"} & \text{"Sehr gut"} & \text{"Gut"} & \text{"Mittelmäßig"} & \text{"Schlecht"} \end{pmatrix}$$

Zwischen den Stichproben der_Frage1 = ("Kaufabsicht - Ja" "Kaufabsicht - Nein" "Kaufabsicht - Weiß nicht") , ausgezählt nach den_Antworten = ("Sehr gut" "Gut" "Mittelmäßig" "Schlecht") der Frage2 = "Werden Sie die Spielkonsole Yoki kaufen?" , bestehen offensichtlich Unterschiede in den jeweiligen Prozentsätzen für die einzelnen Antworten.

$$KT1 = \begin{pmatrix} \text{"Kaufabsicht"} & \text{"Ja"} & \text{"Nein"} & \text{"Weiß nicht"} \\ \text{"Handlichkeit"} & \text{"absolut"} & \text{"absolut"} & \text{"absolut"} \\ \text{"Sehr gut"} & 30 & 8 & 1 \\ \text{"Gut"} & 18 & 5 & 5 \\ \text{"Mittelmäßig"} & 5 & 5 & 5 \\ \text{"Schlecht"} & 5 & 5 & 1 \\ \text{"SUMME"} & 58 & 23 & 12 \end{pmatrix}$$

Kann man auf Grund dieser Stichprobenergebnisse behaupten, dass im Hinblick auf die Stichproben = ("Kaufabsicht - Ja" "Kaufabsicht - Nein" "Kaufabsicht - Weiß nicht") signifikante Anteilswertunterschiede in den Ausprägungen = ("Sehr gut" "Gut" "Mittelmäßig" "Schlecht") bestehen? Testergebnis = "Ja, es bestehen signifikante Verteilungsunterschiede. " (H-Test, Wahrscheinlichkeit p = 0.009)

Wenn die Unterschiede allgemein groß genug sind, um sie als signifikant zu bezeichnen, dann will man wissen, zwischen welchen Stichprobenpaaren die Anteilswertunterschiede signifikant sind. Bei folgenden Antworten sind die Abweichungen vom Durchschnitt signifikant:(Signifikanzniveau von = "5 %" , U-Test)

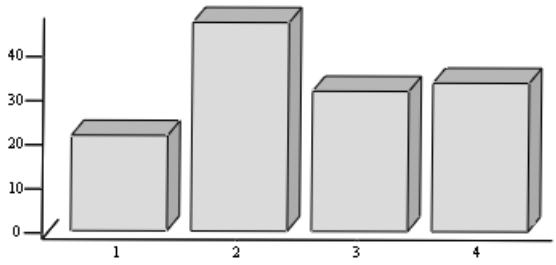
$$ss = \begin{bmatrix} (\text{"Kaufabsicht-Ja"} & \text{"Kaufabsicht-Nein"} & 0.023) \\ (\text{"Kaufabsicht-Ja"} & \text{"Kaufabsicht-Weiß nicht"} & 0.002) \end{bmatrix}$$

5_3 DURCHSCHNITTSDIFFERENZEN, 3 UND MEHR UNABHÄNGIGE STICHPROBEN, F-TEST

Ziel: Kreuztabellen, Grafik, Anteilswerte, Unterschied zwischen den Durchschnitten

UNTERSCHIED
zwischen den Antworten der
Frage1 = "Werden Sie die Spielkonsole kaufen?"
und der
Frage2 = "Wie alt sind Sie?"

MITTELWERTSVERGLEICHE



$$\text{MITTELWERTE} = \begin{pmatrix} \text{"Nr.:"} & 1 & 2 & 3 & \text{"*"} \\ \text{"Kaufabsicht"} & \text{"Ja"} & \text{"Nein"} & \text{"Weiß nicht"} & \text{"Insgesamt"} \\ \text{"Alter"} & 21.667 & 47.333 & 31.8 & 33.706 \end{pmatrix}$$

Zwischen den Stichproben^{T} = ("Kaufabsicht - Ja" "Kaufabsicht - Nein" "Kaufabsicht - Weiß nicht") , ausgezählt nach den Antworten der Frage2 = "Wie alt sind Sie?" , bestehen offensichtlich Unterschiede in den jeweiligen Durchschnitten der Stichproben.

Auf Grund der Stichprobenergebnissen kann man = "annehmen, " dass diese Unterschiede in den Durchschnitten der jeweiligen Antworten auch in der Grundgesamtheit aller Personen existieren, aus der die Stichproben stammen. Die Unterschiede sind insgesamt = "groß genug," um sie als signifikant zu bezeichnen. (Die Wahrscheinlichkeit $p = 0.011$)

Wenn die Unterschiede allgemein groß genug sind, um sie als signifikant zu bezeichnen, dann will man wissen, zwischen welchen Stichprobenpaaren die Durchschnittsunterschiede signifikant sind. Bei folgenden Stichprobenpaaren sind die Unterschiede der Durchschnitte signifikant:

ss = [("Kaufabsicht-Ja" "Kaufabsicht-Nein" −25.667)]

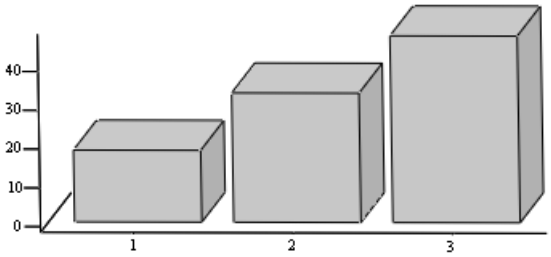
Signifikanzniveau von = "5 %" (Die Wahrscheinlichkeit $p = 0.011$)

6_1 ANTEILSWERTDIFFERENZEN, 3 UND MEHR ABHÄNGIGE STICHPROBEN, COCHRAN-TEST, NOMINAL

Ziel: Kreuztabelle, Grafik, Anteilswerte, Unterschied zwischen den Anteilswerten

UNTERSCHIED
zwischen den Antworten der
FRAGE = "Erreichten Sie die Sollzeit beim Computerspiel Fantasy?"
und den
Antworten = ("1. Versuch" "2. Versuch" "3. Versuch")

STICHPROBENVERGLEICH



$$Legende1^{T} = \begin{pmatrix} \text{"1. Versuch"} & \text{"2. Versuch"} & \text{"3. Versuch"} \\ 19 & 33 & 48 \end{pmatrix}$$

Zwischen den Antworten = ("1. Versuch" "2. Versuch" "3. Versuch") , bestehen in den Stichproben offensichtlich Unterschiede in den jeweiligen Prozentsätzen.

$$T = \begin{pmatrix} \text{"Sollzeit"} & \text{"Häufigkeit"} & \text{"in \% "} \\ \text{"1. Versuch"} & 5 & 19 \\ \text{"2. Versuch"} & 9 & 33 \\ \text{"3. Versuch"} & 13 & 48 \\ \text{"Summe"} & 27 & 100 \end{pmatrix}$$

Auf Grund dieser Stichprobenergebnissen kann man = "annehmen, " dass diese Unterschiede in den Prozentsätzen der jeweiligen Antworten auch in der Grundgesamtheit aller Personen existieren, aus der die Stichproben stammen. Die Unterschiede sind insgesamt = "groß genug," um sie als signifikant zu bezeichnen (Cochran Test).

Wenn die Unterschiede allgemein groß genug sind, um sie als signifikant zu bezeichnen, dann will man wissen, zwischen welchen Stichprobenpaaren die Anteilswertunterschiede signifikant sind. Bei folgenden Antworten sind die Abweichungen signifikant: (Signifikanzniveau von = "5 %")

$$ss = \begin{bmatrix} (\text{"1. Versuch und 2. Versuch"} & \text{"signifikant"}) \\ (\text{"1. Versuch und 3. Versuch"} & \text{"signifikant"}) \\ (\text{"2. Versuch und 3. Versuch"} & \text{"signifikant"}) \end{bmatrix}$$

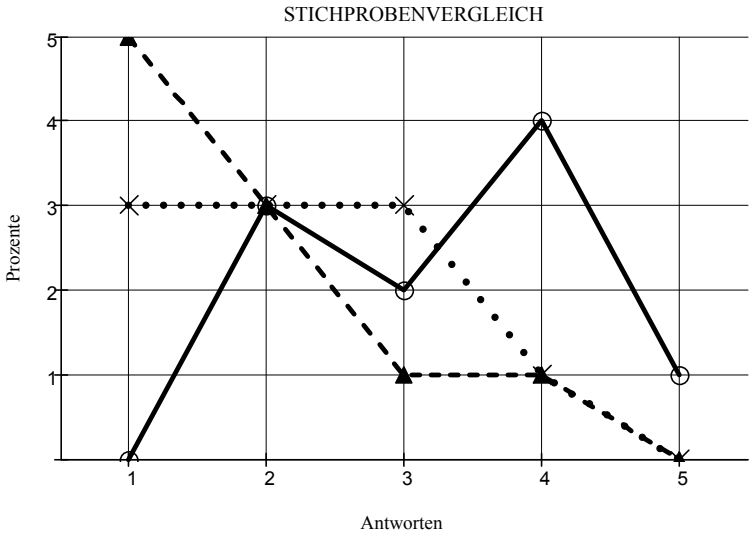
6_2 VERTEILUNGSDIFFERENZEN, 3 UND MEHR UNABHÄNGIGE STICHPROBEN, FRIEDMAN-TEST, ORDINAL

Ziel: Kreuztabellen, Grafik, Anteilswerte, Unterschied zwischen den Verteilungen

UNTERSCHIED
zwischen den Antworten auf die verbundenen Fragen

$$F = \begin{pmatrix} \text{"Wie beurteilen Sie das Produkt nach der 1. Werbesendung?"} \\ \text{"Wie beurteilen Sie das Produkt nach der 2. Werbesendung?"} \\ \text{"Wie beurteilen Sie das Produkt nach der 3. Werbesendung?"} \end{pmatrix}$$

$$FL = \begin{pmatrix} \text{"rot = 1. Werbesendung"} \\ \text{"blau = 2. Werbesendung"} \\ \text{"grün = 3. Werbesendung"} \end{pmatrix}$$



$$LLL = \begin{pmatrix} \text{"sehr gut"} & \text{"gut"} & \text{"neutral"} & \text{"schlecht"} & \text{"sehr schlecht"} \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}$$

Zwischen den $\text{Stichproben}^{T} = (\text{"1. Werbesendung"} \quad \text{"2. Werbesendung"} \quad \text{"3. Werbesendung"})$, bestehen offensichtlich Unterschiede in den Verteilungen und Zentralwerten der Stichproben.

$$KT1 = \begin{pmatrix} "*" & "1.\ Werbesendung" & "2.\ Werbesendung" & "3.\ Werbesendung" \\ "*" & "absolut" & "absolut" & "absolut" \\ "sehr\ gut" & 0 & 3 & 5 \\ "gut" & 3 & 3 & 3 \\ "neutral" & 2 & 3 & 1 \\ "schlecht" & 4 & 1 & 1 \\ "sehr\ schlecht" & 1 & 0 & 0 \\ "SUMME\ " & 10 & 10 & 10 \end{pmatrix}$$

Auf Grund der Stichprobenergebnissen kann man = "annehmen, " dass diese Unterschiede auch in der Grundgesamtheit aller Personen existieren, aus der die Stichproben stammen. Die Unterschiede sind insgesamt = "groß genug," um sie als signifikant zu bezeichnen (Friedman Test).

Wenn die Unterschiede allgemein groß genug sind, um sie als signifikant zu bezeichnen, dann will man wissen, zwischen welchen Stichprobenpaaren die Anteilswertunterschiede signifikant sind. Bei folgenden Antworten sind die Abweichungen signifikant: (Signifikanzniveau von = "5 %")

$$ss = \begin{bmatrix} ("1.\ Werbesendung\ und\ 2.\ Werbesendung" & "signifikant") \\ ("1.\ Werbesendung\ und\ 3.\ Werbesendung" & "signifikant") \end{bmatrix}$$

6_3 DURCHSCHNITTSDIFFERENZEN, 3 UND MEHR ABHÄNGIGE STICHPROBEN, F-TEST

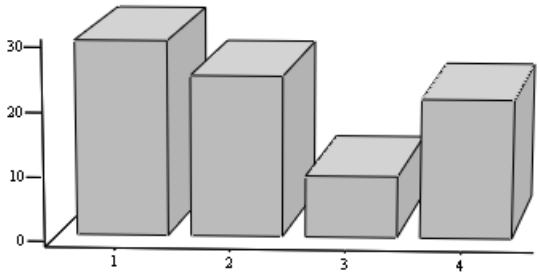
Ziel: Kreuztabellen, Grafik, Mittelwerte, Unterschied zwischen den Mittelwerten

UNTERSCHIED

zwischen den Antworten auf die Fragen

$$F = \begin{pmatrix} \text{"Wie lang benötigten Sie für das Computerspiel Fantasy beim ersten Versuch?"} \\ \text{"Wie lang benötigten Sie für das Computerspiel Fantasy beim zweiten Versuch?"} \\ \text{"Wie lang benötigten Sie für das Computerspiel Fantasy beim dritten Versuch?"} \end{pmatrix}$$

MITTELWERTSVERGLEICHE



$$\text{Mittelwerte}^T = \begin{pmatrix} 1 & 2 & 3 & 4 \\ \text{"1. Versuch"} & \text{"2. Versuch"} & \text{"3. Versuch"} & \text{"Insgesamt"} \\ 30.3 & 25 & 9.6 & 21.633 \end{pmatrix}$$

Zwischen den $\text{Stichproben}^T = (\text{"1. Versuch"} \quad \text{"2. Versuch"} \quad \text{"3. Versuch"})$, bestehen offensichtlich Unterschiede in den jeweiligen Durchschnitten der Stichproben.

Kann man allgemein behaupten, dass im Hinblick auf die Fragen signifikante Mittelwertsunterschiede bestehen? Auf Grund der Stichprobenergebnissen kann $\text{man} = \text{"annehmen, "}$ dass diese Unterschiede in den Durchschnitten der jeweiligen Antworten auch in der Grundgesamtheit aller Personen existieren, aus der die Stichproben stammen. Die Unterschiede sind $\text{insgesamt} = \text{"groß genug,"}$ um sie als signifikant zu bezeichnen.

Wenn die Unterschiede allgemein groß genug sind, um sie als signifikant zu bezeichnen, dann will man wissen, zwischen welchen Stichprobenpaaren die Durchschnittsunterschiede signifikant sind. Bei folgenden Stichprobenpaaren sind die Unterschiede der Durchschnitte signifikant:

$\text{ss} = [(\text{"Mittelwert 1 - 1. Versuch"} \quad \text{"Mittelwert 1 - 3. Versuch"} \quad 20.7)]$

Signifikanzniveau $\text{von} = \text{"5 %"}$

7_1 LOGLINEARE MODELLE

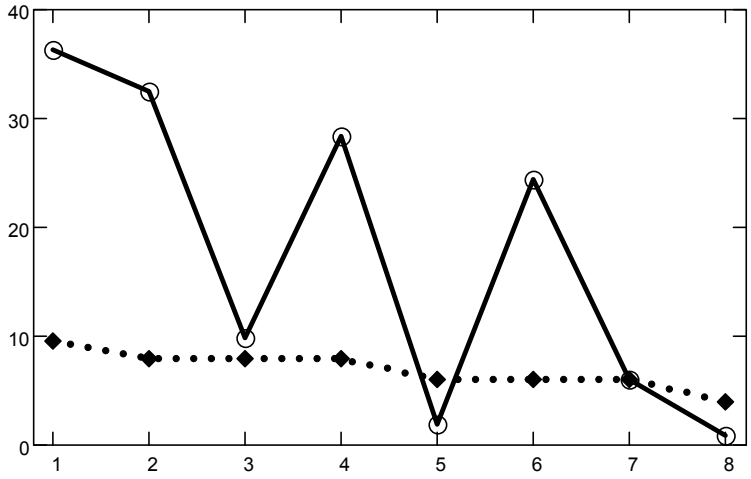
Ziel: Kreuztabelle, Grafik, Anteilswerte, Abhängigkeiten zwischen 3 nominalen Variablen

Loglineares Modell

zwischen

$$\text{den\_Fragen} = \begin{pmatrix} \text{"Geschlecht?"} \\ \text{"Werden Sie die Spielkonsole Yoki kaufen?"} \\ \text{"Familienstand?"} \end{pmatrix}$$

Gewähltes Modell $nr = 5$



$$L = \begin{pmatrix} 1 & \text{"[x1][x2][x3]"} \\ 2 & \text{"[x1][x2,x3]"} \\ 3 & \text{"[x2][x1,x3]"} \\ 4 & \text{"[x3][x1,x2]"} \\ 5 & \text{"[x1,x2][x1,x3]"} \\ 6 & \text{"[x1,x2][x2,x3]"} \\ 7 & \text{"[x1,x3][x2,x3]"} \\ 8 & \text{"[x1,x2][x1,x3][x2,x3]"} \end{pmatrix}$$

$$\text{Bestes\_Modell} = \begin{bmatrix} 5 \\ \text{"[x1,x2][x1,x3]"} \\ (\text{"Geschlecht - Kaufabsicht"} \quad \text{"Geschlecht - Familienstand"}) \end{bmatrix}$$

Wie können die Abhängigkeiten zwischen den drei
Variablen = "Geschlecht-[x1]", "Kaufabsicht-[x2]" und "Familienstand-[x3]"
am besten erklärt werden?

Das gewählte Modell = "[x1,x2][x1,x3]" (nr = 5) zur Erklärung der Abhängigkeiten zwischen diesen 3 Variablen hat nachstehende Eigenschaften: Folgende Variablen sind jeweils voneinander abhängig = "Geschlecht - Kaufabsicht und Geschlecht - Familienstand" .

Die Abhängigkeit der Variablen zeigt sich in den Wechselwirkungen zwischen den Antworten der jeweiligen Fragen. Folgende Antworten treten signifikant überdurchschnittlich oft gemeinsam auf
(Signifikanzniveau von = "5 %"):

$$\text{ob} = \begin{bmatrix} \begin{bmatrix} \text{"Geschlecht - Kaufabsicht"} \\ \begin{pmatrix} \text{"Männlich"} \\ \text{"Ja"} \end{pmatrix} \\ \begin{pmatrix} \text{"Weiblich"} \\ \text{"Nein"} \end{pmatrix} \end{bmatrix} & \begin{bmatrix} \text{"Geschlecht - Familienstand"} \\ \begin{pmatrix} \text{"Männlich"} \\ \text{"Ledig"} \end{pmatrix} \\ \begin{pmatrix} \text{"Weiblich"} \\ \text{"Verheiratet"} \end{pmatrix} \end{bmatrix} \end{bmatrix}$$

Folgende Antworten treten signifikant unterdurchschnittlich oft gemeinsam auf
(Signifikanzniveau von = "5 %"):

$$\text{un} = \begin{bmatrix} \begin{bmatrix} \text{"Geschlecht - Kaufabsicht"} \\ \begin{pmatrix} \text{"Männlich"} \\ \text{"Nein"} \end{pmatrix} \\ \begin{pmatrix} \text{"Weiblich"} \\ \text{"Ja"} \end{pmatrix} \end{bmatrix} & \begin{bmatrix} \text{"Geschlecht - Familienstand"} \\ \begin{pmatrix} \text{"Männlich"} \\ \text{"Verheiratet"} \end{pmatrix} \\ \begin{pmatrix} \text{"Weiblich"} \\ \text{"Ledig"} \end{pmatrix} \end{bmatrix} \end{bmatrix}$$

Für das verwendete Modell

$$\text{gilt} = \begin{pmatrix} \text{"Die Annahme, daß das gewählte Modell die Abhängigkeitsstruktur zwischen den"} \\ \text{"Fragen ausreichend beschreibt, kann nicht abgelehnt werden. Die Unterschiede "} \\ \text{"zwischen beobachteten und geschätzten Daten sind nicht signifikant. "} \end{pmatrix}$$

7_2 LOGISTISCHE REGRESSION

Ziel: Regression zwischen einer abhängigen und mehrere unabhängigen nominalen Variablen

Logistische Regression

zwischen

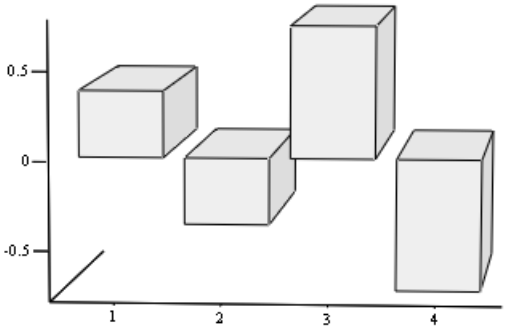
der_Antwort = "Ja"

auf_die_Frage = "Werden Sie die Spielkonsole Yoki kaufen?"

und

$$\text{den\_Fragen} = \begin{pmatrix} \text{"Geschlecht?"} \\ \text{"Familienstand?"} \end{pmatrix}$$

Über- und unterdurchschnittliche Antworten



$$\text{Le} = \begin{pmatrix} 1 & \text{"Geschlecht - Männlich"} & 0.3744 \\ 2 & \text{"Geschlecht - Weiblich"} & -0.3744 \\ 3 & \text{"Familienstand - Ledig"} & 0.7422 \\ 4 & \text{"Familienstand - Verheiratet"} & -0.7422 \end{pmatrix}$$

Wie gut können die unabhängigen Variablen = ("Geschlecht" "Familienstand") den Anteilswert der _Antwort = "Ja" auf die Frage = "Werden Sie die Spielkonsole Yoki kaufen?" erklären? Für das gewählte Modell gilt = "Auf Grund der Stichprobenergebnisse kann man annehmen, " dass zwischen beobachteten und geschätzten Anteilen signifikante Unterschiede bestehen. Das Modell ist für die Anteilsschätzung = "nicht geeignet! "

Auf die Schätzung des Anteilswertes des Merkmals = "Ja" haben folgende unabhängige Variablen einen signifikanten Einfluß = ("Geschlecht" "Familienstand") .

Die Anteilsschätzung der _Antwort = "Ja" auf die Frage = "Werden Sie die Spielkonsole Yoki kaufen?" wird durch die folgenden Merkmale der unabhängigen Variablen signifikant über- bzw. unterdurchschnittlich stark beeinflusst:
(Signifikanzniveau von = "5 %"):

$$\text{Über} = \begin{bmatrix} \text{"Befragte mit den Merkmalen"} \\ \begin{pmatrix} \text{"Geschlecht - Männlich"} \\ \text{"Familienstand - Ledig"} \end{pmatrix} \\ \text{"geben auf die Frage"} \\ \text{"Werden Sie die Spielkonsole Yoki kaufen?"} \\ \text{"signifikant überdurchschnittlich oft die Antwort"} \\ \text{"Ja"} \end{bmatrix}$$

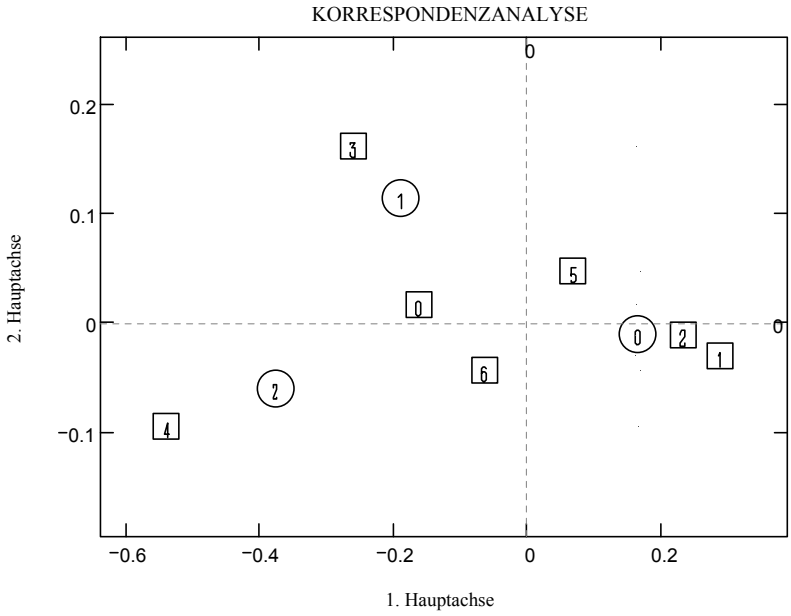
$$\text{Unter} = \begin{bmatrix} \text{"Im Gegensatz dazu geben Befragte mit den Antworten"} \\ \begin{pmatrix} \text{"Geschlecht - Weiblich"} \\ \text{"Familienstand - Verheiratet"} \end{pmatrix} \\ \text{"auf diese Frage"} \\ \text{"Werden Sie die Spielkonsole Yoki kaufen?"} \\ \text{"signifikant unterdurchschnittlich oft die Antwort"} \\ \text{"Ja"} \end{bmatrix}$$

7_3 KORRESPONDENZANALYSE

Ziel: Reduktion nominaler Merkmale auf 2 Dimensionen

KORRESPONDENZANALYSE
zwischen den Antworten der
Frage1 = "Werden Sie die Spielkonsole kaufen?"
und den Antworten auf die Fragen

$$F = \begin{pmatrix} \text{"Geschlecht?"} \\ \text{"Wie alt sind Sie?"} \\ \text{"Familienstand?"} \end{pmatrix}$$



Berücksichtigt man nur jene Merkmalsausprägungen, die mindestens $100 \cdot rr = 70$ Prozent der Variation der einzelnen Faktoren (Hauptachsen) erklären, dann weisen folgende Merkmalsausprägungen Gemeinsamkeiten auf d. h. sie liegen im gleichen Quadranten L1 bis L4:

$$\begin{pmatrix} L2 & L1 \\ L3 & L4 \end{pmatrix}$$

(
Z = Zeilenmerkmal, S = Spaltenmerkmal, die Zahl am Ende verweist auf den Punkt in der Grafik)

L1 = "keine"

$$L2 = \begin{pmatrix} \text{"Z - Geschlecht - Weiblich"} \\ \text{"Z - Alter - 25 bis 50 Jahre"} \\ \text{"S - Kaufabsicht - Nein"} \end{pmatrix}$$

$$L3 = \begin{pmatrix} \text{"Z - Alter - 50 Jahre und mehr"} \\ \text{"S - Kaufabsicht - Weiss nicht"} \end{pmatrix}$$

$$L4 = \begin{pmatrix} \text{"Z - Geschlecht - Männlich"} \\ \text{"Z - Alter - bis 25 Jahre"} \\ \text{"S - Kaufabsicht - Ja"} \end{pmatrix}$$

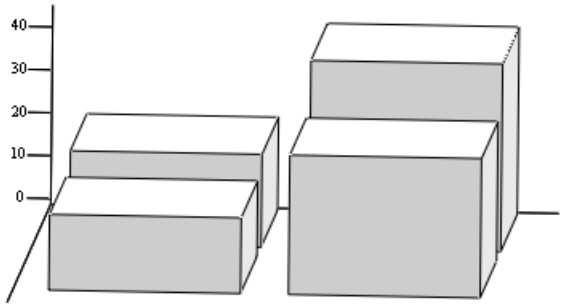
8_1 MULTIVARIATE VARIANZANALYSE (MANOVA):

Ziel: Grafik, Mittelwerte, Unterschied zwischen den Mittelwerten

UNTERSCHIED
zwischen den Antworten auf die
Frage1 = "Werden Sie die Spielkonsole Yoki kaufen?"
und den Fragen

$$F = \begin{pmatrix} \text{"Wie alt sind Sie?"} \\ \text{"Wie hoch ist Ihr monatliches Einkommen (in 100 €)?"} \end{pmatrix}$$

MITTELWERTSVERGLEICHE



$$\text{Mittelwerte} = \begin{pmatrix} \text{"Kaufabsicht"} & \text{"Ja"} & \text{"Nein"} \\ \text{"Alter"} & 21.67 & 43.75 \\ \text{"Einkommen"} & 17.67 & 32.5 \end{pmatrix}$$

Zwischen den Stichproben = ("Kaufabsicht - Ja" "Kaufabsicht - Nein") , ausgezählt nach den Ausprägungen der_Fragen = ("Alter" "Einkommen") , bestehen offensichtlich Unterschiede in den jeweiligen Durchschnitten.

Auf Grund der Stichprobenergebnissen kann man = "annehmen, " dass diese Unterschiede in den Durchschnitten auch in der Grundgesamtheit aller Personen existieren, aus der die Stichproben = ("Kaufabsicht - Ja" "Kaufabsicht - Nein") stammen. Die Unterschiede sind insgesamt = "groß genug," um sie als signifikant zu bezeichnen. Die Wahrscheinlichkeit, dass diese Schlussfolgerung falsch ist = ("ist höchstens" 5 "%")

8_2 MULTIVARIATE REGRESSIONSANALYSE

Ziel: Linearer Zusammenhang zwischen einer abhängigen und mehreren unabhängigen Variablen

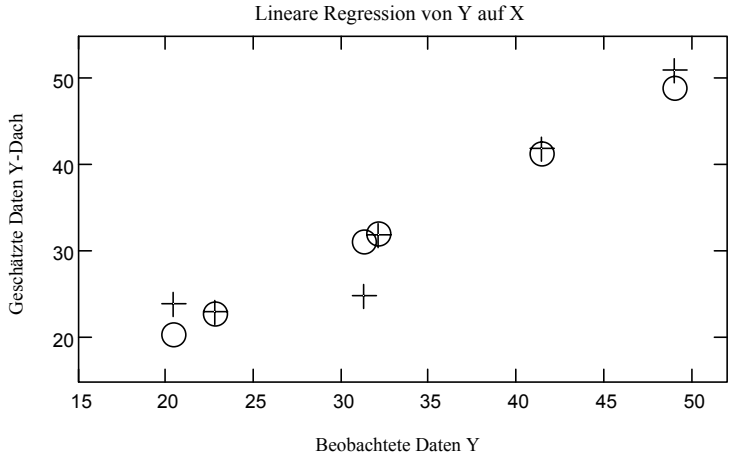
MULTIVARIATE REGRESSIONSANALYSE

zwischen den Antworten der unabhängigen Fragen

$$FF = \begin{pmatrix} \text{"Wieviel Zeit verbringen Sie pro Tag mit Fernsehen (in Stunden)?"} \\ \text{"Wie alt sind Sie?"} \end{pmatrix}$$

und den Antworten der abhängigen

Frage_y = "Wie hoch ist Ihr monatliches Einkommen?"



Zwischen der abhängigen Variable = "Einkommen" und den unabhängigen Variablen = ("Fernsehen" "Alter") besteht offensichtlich in der Stichprobe ein Zusammenhang. Die aus der Stichprobe ermittelte Regressionsfunktion = "Einkommen = 7,095 + (-2,044) * Fernsehen + (0,808) * Alter"
Sie besagt, dass man mit Hilfe der unabhängigen Variablen = ("Fernsehen" "Alter") und geeigneter Gewichtung einen Schätzwert für die abhängige Variable = "Einkommen" berechnen kann. Die Gewichte sind die Regressionskoeffizienten:

$$RK^T = \begin{pmatrix} 7.095 & -2.044 & 0.808 \\ \text{"Konstante"} & \text{"Fernsehen"} & \text{"Alter"} \end{pmatrix}$$

Wenn sich z. B. das = "Fernsehen" um eine Einheit erhöht, dann verändert sich der Wert der abhängigen Variable = "Einkommen" um = –2.044 Einheiten. Die unabhängigen Variablen = ("Fernsehen" "Alter") erklären = 91.271 Prozent der Varianz der abhängigen Variable = "Einkommen" (Bestimmtheitsmaß = 0.913).
Auf Grund dieser Stichprobenergebnissen kann man = "annehmen, " dass dieser lineare Zusammenhang zwischen der abhängigen Variablen und der unabhängigen Variablen auch in der Grundgesamtheit aller Personen existiert, aus der die Stichprobe stammt. Der Anteil der erklärten Varianz der abhängigen Variablen ist insgesamt = "groß genug," um ihn als signifikant zu bezeichnen.
Die Wahrscheinlichkeit, dass diese Schlussfolgerung falsch ist = ("ist höchstens" 5 "%")

8_3 FAKTORENANALYSE

Ziel: Reduktion vieler Variablen auf wenige Faktoren

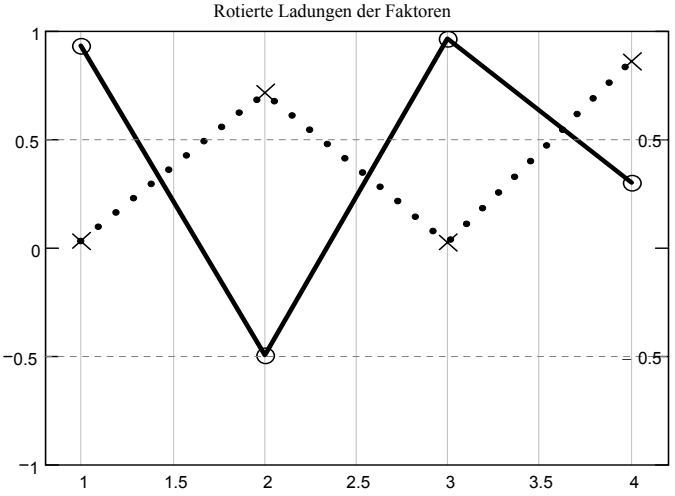
FAKTORENANALYSE

für die Fragen

$$F = \begin{pmatrix} \text{"Wie alt sind Sie?"} \\ \text{"Wieviel Zeit verbringen Sie pro Tag mit Fernsehen (in Stunden)?"} \\ \text{"Wie hoch ist Ihr monatliches Einkommen?"} \\ \text{"Wieviel Zeit verbringen Sie pro Tag am Computer 8in Stun den)?"} \end{pmatrix}$$

Anzahl der Faktoren $= 2$

Rotierte Ladungen der Faktoren



$$LV = \begin{pmatrix} 1 & 2 & 3 & 4 \\ \text{"Alter"} & \text{"Fernsehen"} & \text{"Einkommen"} & \text{"Computer"} \end{pmatrix}$$

Die $= 2$ Faktoren werden durch folgende Variablen hoch geladen, wobei Ladungen größer gleich $= 0.7$ berücksichtigt werden:

$$H = \begin{bmatrix} (\text{"Faktor"} \quad 1) & \begin{bmatrix} (\text{"Alter"} \quad 0.934) \\ (\text{"Einkommen"} \quad 0.961) \end{bmatrix} \\ (\text{"Faktor"} \quad 2) & \begin{bmatrix} (\text{"Fernsehen"} \quad 0.713) \\ (\text{"Computer"} \quad 0.855) \end{bmatrix} \end{bmatrix}$$

Auf Grund des Sphärentests von Bartlett ist es = "gerechtfertigt," zumindest einen Faktor zu extrahieren.
Ein Modell mit $= 2$ Faktoren zur Erklärung der Beobachtungsdaten = "kann nicht abgeleht werden."

LITERATURVERZEICHNIS

Bamberg. G., Baur, F. (2002). Statistik. 12. Auflage. Odenbourg.

Bohley, P. (2000). Statistik. 7. Auflage. Oldenbourg.

Bortz, J. (2004). Statistik. 6. Auflage. Springer.

Bosch, K. (1999). Grundzüge der Statistik. Einführung mit Übungen. 2. Auflage. Oldenbourg.

Buttler, G., Fickel, N. (2002). Einführung in die Statistik. 2. Auflage. Rowohlt Tb.

Bühner, M. (2004). Einführung in die Test- und Fragebogenkonstruktion . Pearson Studium.

Dehling, H., Haupt, B. (2004). Einführung in die Wahrscheinlichkeitstheorie und Statistik. 2. Auflage. Springer.

Fahrmeir, L., u.a. (2004). Statistik, Der Weg zur Datenanalyse. 5. Auflage. Springer.

Hackl, P., Katzenbeisser, W. (2000). Statistik für Sozial- und Wirtschaftswissenschaften. 11. Auflage. Oldenbourg.

Hippmann, H.-D. (2003). Statistik für Wirtschafts- und Sozialwissenschaftler. 3. Auflage. Schäffer-Poeschel.

Hartung, J., u.a. (2005). Statistik. Lehr- und Handbuch der angewandten Statistik . 14. Auflage. Oldenbourg.

Janssen, J., Laatz, W. (2005). Statistische Datenanalyse mit SPSS für Windows. 4. Auflage. Springer.

Kohn, W. (2004). Statistik. Datenanalysis und Wahrscheinlichkeitsrechnung . 11. Auflage. Springer.

Krämer, W. (2005). Statistik verstehen. Eine Gebrauchsanweisung. Piper.

Kühnel, S.-M., Krebs, D. (2006). Statistik für die Sozialwissenschsaften. Grundlagen, Methoden, Anwendungen. 3. Auflage. Rowohlt Tb.

Lehn, J., Wegmann, H. (2006). Einführung in die Statistik. 5. Auflage. Teubner.

Marinell, G., (1995). Multivariate Verfahren. 5. Auflage. Oldenbourg.

Marinell, G., (1987). Statistik. 2. Auflage. Oldenbourg.

Marinell, G., (1986). Statistische Auswertung. 3. Auflage. Oldenbourg.

Pospeschill, M. (2006). Statistische Methoden. Spektrum Akademischer Verlag.

Quatember, A. (2005). Statistik ohne Angst vor Formeln. Ein Lehrbuch für Wirtschafts- und Sozialwissenschaftler . Pearson Studium.

Rinne, H. (2003). Taschenbuch der Statistik. Für Wirtschafts- und Sozialwissenschaften. 3. Auflage. Deutsch (Harri).

Sachs, L., Hedderich, J. (2006). Angewandte Statistik. Anwendung statistischer Methoden. 12. Auflage. Springer.

Siegel, S. (2001). Nichtparametrische statistische Methoden . 5. Auflage. Klotz, Eschborn.

Storm, R. (2001). Wahrscheinlichkeitsrechnung, mathematische Statistik und statistische Qualitätskontrolle. Mit 120 Beispielen. 11. Auflage. Fachbuchverlag Leipzig.

Tutz, G. (2000). Die Analyse kategorialer Daten. Oldenbourg.

STICHWORTVERZEICHNIS