1 Deskriptive Statistik

1.3 Absolute Häufigkeiten, relative Häufigkeiten und Histogramme

Gruppierbare Daten - wenige Merkmalsausprägungen: Die verschiedene Ausprägungen werden mit a_m $(m = 1, \cdots, M)$ für nominale und ordinale Merkmale bezeichnet und mit x_m $(m = 1, \cdots, M)$ für metrische Merkmale. Die relativen Häufigkeiten $\tilde{p}_m$ können aus den absoluten Häufigkeiten h_m für alle M Merkmalsausprägungen berechnet werden:

$$\tilde{p}_m = \frac{h_m}{\sum_{m=1}^{M} h_m}$$

Bei Kreisdiagrammen werden die einzelnen Ausprägungen Kreissektoren zugeordnet, deren Winkel proportional zu den relativen Häufigkeiten sind. Die Winkel α_k der k-ten Ausprägung werden wie folgt ermittelt:

$$\alpha_k = 360 \cdot \tilde{p}_m$$

Klassifizierbare Daten - viele Merkmalsausprägungen: Das Histogramm besteht aus Rechtecken, die über den Intervallen $[a_{\underline{k}}, a_{\overline{k}}]$ errichtet werden. Die Fläche des Rechtecks entspricht der relativen Häufigkeit $\tilde{p}_k$. Da die Klassen der Länge nach variieren können, müssen die Höhen $\tilde{f}_k$ der Rechtecke wie folgt berechnet werden:

$$\tilde{f}_k = \frac{\tilde{p}_k}{a_{\overline{k}} - a_{\underline{k}}}, \quad k = 1, \ldots, K$$

1.4 Empirische Verteilungsfunktion $\tilde{F}_X(x)$

Rohdaten:

$$I_{x_i}(x) = \begin{cases} 1 & x_i \leq x \\ 0 & \text{sonst} \end{cases} \qquad \tilde{F}_X(x) = \frac{1}{n} \sum_{i=1}^{n} I_{x_i}(x)$$

Gruppierte metrische Daten:

$$I_{x_m}(x) = \begin{cases} 1 & x_m \leq x \\ 0 & \text{sonst} \end{cases} \qquad \tilde{F}_X(x) = \frac{1}{n} \sum_{m=1}^{M} h_m \cdot I_{x_m}(x) = \sum_{m=1}^{M} \tilde{p}_m \cdot I_{x_m}(x)$$

Klassifizierte Daten: Kann von einer Gleichverteilung innerhalb der Klassen ausgegangen werden, kann für beliebige Werte x zwischen der unteren und der oberen Klassengrenze linear interpoliert werden.

$$\tilde{F}_X(x) = \tilde{F}_X(a_{\underline{k}}) + \frac{\tilde{F}_X(a_{\overline{k}}) - \tilde{F}_X(a_{\underline{k}})}{a_{\overline{k}} - a_{\underline{k}}} \cdot (x - a_{\underline{k}})$$

1.5 Deskriptive Lagemaße

1.5.3 Das arithmetische Mittel

Rohdaten: $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$

Gruppierte metrische Daten: $\bar{x} = \frac{1}{n} \sum_{m=1}^{M} h_m \cdot x_m = \sum_{m=1}^{M} \tilde{p}_m \cdot x_m$

Klassifizierte Daten: $\bar{x} = \frac{1}{n} \sum_{k=1}^{K} h_k \cdot c_k = \sum_{k=1}^{K} \tilde{p}_k \cdot c_k$ wobei c_k die Klassenmitte bezeichnet

1.5.4 Das geometrische Mittel

Das geometrische Mittel $\bar{x}_G$ beschreibt den durchschnittlichen Wachstumsfaktor, der über alle Perioden konstant bleibt und den Anfangsbestand B_0 auf den Endbestand B_n anwachsen läßt.

$$B_n = B_0 \cdot (\bar{x}_G)^n \qquad \bar{x}_G = \sqrt[n]{\prod_{i=1}^{n} x_i} = \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n}$$

1.5.5 Der Median

Als *Median* oder *Zentralwert* $x_{0,5}$ bezeichnet man den Wert, für den gilt:

$$\tilde{F}_X(x_{0,5}) = 0{,}5$$

Rohdaten: Zur Berechnung von $x_{0,5}$ werden die Werte x_i einer metrischen Variablen der Größe nach geordnet. Dann ist der Median definiert als:

$$x_{0.5} = \begin{cases} x_{[(n+1)/2]} & n \text{ ungerade} \\ \frac{1}{2}\left(x_{[n/2]} + x_{[(n/2)+1]}\right) & n \text{ gerade} \end{cases}$$

Klassifizierte Daten: Zunächst wird die Klasse k bestimmt, in der der Median liegt.

$$k \text{ ist Medianklasse} \Longleftrightarrow \tilde{F}_X(a_{\underline{k}}) < 0{,}5 \leq \tilde{F}_X(a_{\overline{k}}) \quad .$$

Unter der Annahme der Gleichverteilung der Werte innerhalb dieser Klasse, kann der Median linear interpoliert werden:

$$x_{0,5} = a_{\underline{k}} + \frac{a_{\overline{k}} - a_{\underline{k}}}{\tilde{F}_X(a_{\overline{k}}) - \tilde{F}_X(a_{\underline{k}})} \cdot (0{,}5 - \tilde{F}_X(a_{\underline{k}}))$$

1.5.6 Quantile

Gibt man einen Wert $\alpha \in [0{,}1]$ vor, so läßt sich für ein metrisches Merkmal X der Wert x_α bestimmen, für den gilt:

$$\tilde{F}_X(x_\alpha) = \alpha$$

Rohdaten und gruppierte metrische Daten: Das x_α-Quantil ist für eine geordneten Liste $\{x_{[1]}, \ldots, x_{[n]}\}$ wie folgt definiert:

$$x_\alpha = \begin{cases} x_{[n\cdot\alpha]} & \text{falls } n \cdot \alpha \text{ keine ganze Zahl ist, wird auf die nächste ganze Zahl gerundet} \\ \frac{1}{2}\left(x_{[n\cdot\alpha]} + x_{[n\cdot\alpha+1]}\right) & \text{falls } n \cdot \alpha \text{ ganzzahlig ist} \end{cases}$$

Klassifizierte Daten: Die Quantile werden analog zum Median durch lineare Interpolation bestimmt.

1.6 Streuungsmaße

1.6.1 Die Spannweite

Liegen die Daten als geordnete Liste $\{x_{[1]}, \ldots, x_{[n]}\}$ vor, so ist die Spannweite definiert als:

$$R = x_{\text{max}} - x_{\text{min}} = x_{[n]} - x_{[1]}$$

1.6.2 Mittlere absolute Abweichungen

Mittlere absolute Abweichung vom arithmetischen Mittel

Rohdaten: $MAA = \frac{1}{n}\sum_{i=1}^{n} |x_i - \bar{x}|$

Gruppierte metrische Daten: $MAA = \frac{1}{n}\sum_{m=1}^{M} h_m \cdot |x_m - \bar{x}|$

Klassifizierte Daten: $MAA = \frac{1}{n}\sum_{k=1}^{K} h_k \cdot |c_k - \bar{x}|$

Mittlere absolute Abweichung vom Median

Rohdaten: $MAM = \frac{1}{n}\sum_{i=1}^{n} |x_i - x_{0,5}|$

Gruppierte metrische Daten: $MAM = \frac{1}{n}\sum_{m=1}^{M} h_m \cdot |x_m - x_{0,5}|$

Klassifizierte Daten: $MAM = \frac{1}{n}\sum_{k=1}^{K} h_k \cdot |c_k - x_{0,5}|$

1.6.3 Die Varianz- und Standardabweichung

Rohdaten: $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)$

Gruppierte metrische Daten: $s^2 = \frac{1}{n}\sum_{m=1}^{M} h_m \cdot (x_m - \bar{x})^2 = \frac{1}{n}\left(\sum_{m=1}^{M} h_m \cdot x_m^2 - n\bar{x}^2\right)$

Klassifizierte Daten: $s^2 = \frac{1}{n}\sum_{k=1}^{K} h_k \cdot (c_k - \bar{x})^2 = \frac{1}{n}\left(\sum_{k=1}^{K} h_k \cdot c_k^2 - n\bar{x}^2\right)$

Um auf die ursprüngliche Maßeinheit zu kommen, verwendet man die Standardabweichung: $s = \sqrt{s^2}$

1.6.4 Der Variationskoeffizient

$$v = \frac{s}{\bar{x}}$$

1.6.5 Der Quartilsabstand

$$q = (x_{0,75} - x_{0,25})$$

1.7 Schiefemaße

1.7.1 Fechnersche Lageregel

$$\begin{aligned}
\bar{x} = x_{0,5} = x_{mod} &\quad\Rightarrow\quad \text{symmetrische Verteilung} \\
\bar{x} > x_{0,5} > x_{mod} &\quad\Rightarrow\quad \text{rechtsschiefe Verteilung} \\
\bar{x} < x_{0,5} < x_{mod} &\quad\Rightarrow\quad \text{linksschiefe Verteilung}
\end{aligned}$$

1.7.2 Momentenkoeffizient der Schiefe

$$G = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{s^3}$$

1.8 Konzentrationsmaße

1.8.1 Lorenzkurve

Die n gegebenen Merkmalsausprägungen werden in aufsteigender Reihenfolge sortiert:

$$x_1 \leq x_2 \leq \ldots \leq x_n$$

Kumulierte Anteile der Merkmalsträger u_i:

$$u_i = \frac{i}{n} \quad i = 1, \ldots, n \quad u_0 := 0$$

Kumulierten relative Merkmalssummen v_i:

$$v_i = \frac{\sum_{j=1}^{i} x_j}{\sum_{j=1}^{n} x_j} \quad i = 1, \ldots, n \quad v_0 := 0$$

Die Lorenzkurve ergibt sich dann als Streckenzug durch die Punkte:

$$(0,0), (u_1, v_1), \ldots, (u_n, v_n)$$

1.8.2 Lorenzkurve für klassifizierte Daten

Kennt man die absoluten Häufigkeiten $h_1, \ldots, h_K$ oder die relativen Häufigkeiten $\tilde{p}_1, \ldots, \tilde{p}_K$ der Klassen $[a_{\underline{1}}; a_{\overline{1}}], \ldots, (a_{\underline{K}}; a_{\overline{K}}]$ und die Merkmalssummen, die auf die einzelnen Klassen entfallen (x_k), kann die Lorenzkurve wie folgt berechnet werden.

$$u_i = \sum_{k=1}^{i} \frac{h_k}{n} = \sum_{k=1}^{i} \tilde{p}_k \quad i = 1, \ldots, K \quad u_0 := 0 \qquad\qquad v_i = \frac{\sum_{k=1}^{i} x_k}{\sum_{k=1}^{K} x_k} \quad i = 1, \ldots, K \quad v_0 := 0$$

Ist die Merkmalssumme (x_k), die auf die k-te Klasse entfallen unbekannt, kann die Annahme getroffen werden, dass innerhalb der Klasse keine Konzentration vorliegt und alle Ausprägungen den selben Wert annehmen, nämlich die

Klassenmitte c_k.

$$u_i = \sum_{k=1}^{i} \frac{h_k}{n} = \sum_{k=1}^{i} \tilde{p}_k \quad i = 1, \ldots, K \quad u_0 := 0 \qquad v_i = \frac{\sum_{k=1}^{i} h_k \cdot c_k}{\sum_{k=1}^{K} h_k \cdot c_k} \quad i = 1, \ldots, K \quad v_0 := 0$$

1.8.3 Gini- und Lorenz-Münzner-Koeffizient

Für eine geordnete Liste $x_1 \leq x_2 \leq \ldots \leq x_n$ berechnet sich der Gini-Koeffizient wie folgt:

$$GI = 1 - \frac{1}{n} \sum_{i=1}^{n} (v_{i-1} + v_i)$$

Für klassifizierte Daten gilt:

$$GI = 1 - \frac{1}{n} \sum_{i=1}^{K} (v_{i-1} + v_i) \cdot h_i$$

Normiert man den Gini-Koeffizienten erhält man den Lorenz-Münzner-Koeffizienten.

$$LM = \frac{GI}{GI_{max}} = \frac{n}{n-1} \cdot GI \quad LM \in [0,1]$$

1.9 Preisindizes

1.9.1 Preisindex nach Laspeyres:

$$P_t^L := \frac{\sum_{i=1}^{n} p_{it} \cdot q_{i0}}{\sum_{i=1}^{n} p_{i0} \cdot q_{i0}} \cdot 100$$

1.9.2 Preisindex nach Paasche:

$$P_t^P := \frac{\sum_{i=1}^{n} p_{it} \cdot q_{it}}{\sum_{i=1}^{n} p_{i0} \cdot q_{it}} \cdot 100$$

2 Wahrscheinlichkeitsrechnung

2.3.3 Axiomatische Einführung der Wahrscheinlichkeit

Eigenschaften der Wahrscheinlichkeit:

(a) Für das komplementäre Ereignis von A gilt $P(A^c) = 1 - P(A)$

(b) Gilt für zwei Ereignisse A und B $A \subset B$, so ist $P(A) \leq P(B)$

(c) Für zwei beliebige Ereignisse A und B gilt: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(d) Bilden die Ereignisse $A_1, \ldots, A_n$ eine Zerlegung von Ω, so gilt: $\sum_{i=1}^{n} P(A_i) = 1$

Gesetzte von De Morgan:

$$P(A \cup B)^c = P(A^c \cap B^c) \qquad P(A \cap B)^c = P(A^c \cup B^c)$$

2.4 Bedingte Wahrscheinlichkeit und stochastische Unabhängigkeit

Bedingte Wahrscheinlichkeit:

$$P(A|B) = \begin{cases} \frac{P(A \cap B)}{P(B)} & \text{falls } P(B) > 0 \\ 0 & \text{falls } P(B) = 0 \end{cases}$$

Stochastische Unabhängigkeit von Ereignissen: Die beiden Ereignisse A und B eines Zufallsexperiments sind *stochastisch unabhängig*, wenn gilt:

$$P(A|B) = P(A), \quad \text{falls } P(B) > 0$$

Bedingung für stochastische Unabhängigkeit: A und B sind genau dann stochastisch unabhängig, wenn gilt:

$$P(A \cap B) = P(A) \cdot P(B)$$

Multiplikationsregel:

$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

Satz von der vollständigen Wahrscheinlichkeit: Bilden die Ereignisse $E_1, \ldots, E_n$ eine Zerlegung von Ω, so gilt für ein beliebiges Ereignis A:

$$P(A) = \sum_{i=1}^{n} P(A|E_i) \cdot P(E_i) = \sum_{i=1}^{n} P(A \cap E_i)$$

Satz von Bayes: Bilden die Ereignisse $E_1, \ldots, E_n$ eine Zerlegung von Ω, so gilt für ein beliebiges Ereignis A mit $P(A) > 0$:

$$P(E_i|A) = \frac{P(A|E_i) \cdot P(E_i)}{\sum\limits_{i=1}^{n} P(A|E_i) \cdot P(E_i)} = \frac{P(A \cap E_i)}{P(A)} \quad \text{für } i = 1, \ldots, n$$

3 Diskrete Verteilungen

3.2 Erwartungswertoperator und Momente

(a) Erwartungswert: $E(X) = \sum\limits_{x_j \in \mathbb{M}} x_j \cdot p(x_j) = \mu_X$

(b) Varianz: $V(X) = E((X - \mu_X)^2) = \sum\limits_{x_j \in \mathbb{M}} (x_j - \mu_X)^2 \cdot p(x_j)$

(c) Standardabweichung: $\sigma(X) = \sqrt{\sigma^2(X)}$

(d) Momentkoeffizient der Schiefe: $S(X) = \dfrac{E(X - \mu_X)^3}{\sigma^3(X)}$

$$S(X) < 0 \quad \text{linksschief}, \quad S(X) = 0 \quad \text{symmetrisch}, \qquad S(X) > 0 \quad \text{rechtsschief}$$

Bedingter Erwartungswert Sei X eine diskrete Zufallsvariable. Dann ist der bedingte Erwartungswert, gegeben dass ein Ereignis A eingetreten ist, definiert als:

$$E(X|A) = \frac{1}{P(A)} \sum_{x_j \in A} x_j \cdot P(X = x_j)$$

Satz vom totalen Erwartungswert Ist X eine diskrete Zufallsvariable und bilden die Ereignisse $E_1, E_2, \ldots, E_n$ eine Zerlegung von Ω gilt:

$$E(X) = E(X|E_1) \cdot P(E_1) + \ldots + E(X|E_n) \cdot P(E_n) = \sum_{i=1}^{n} E(X|E_i) \cdot P(E_i)$$

Rechenregeln für Momente einer lineartransformierten diskreten Zufallsvariablen

(a) Erwartungswert einer Linearkombination von X: $E(a + bX) = a + bE(X)$

(b) Varianz einer Linearkombination von X: $V(a + bX) = b^2 \cdot V(X)$

(c) Verschiebesatz für Varianzen: $V(X) = E(X^2) - \mu^2$

3.3 Quantile von diskreten Zufallsvariablen

Das α-Quantil x_α der Verteilung von X (mit $\alpha \in (0,1)$) ist definiert als: $P(X \leq x_\alpha) \geq \alpha$ und $P(X \geq x_\alpha) \geq 1 - \alpha$

3.4 Kombinatorik

3.4.1 Permutationen

n unterscheidbare Objekte lassen sich auf $1 \cdot 2 \cdot 3 \cdot 4 \cdot \ldots \cdot (n-1) \cdot n = n!$ (n-Fakultät) Arten anordnen. Jede Anordnung der n Objekte oder von n Zahlen wird als Permutation bezeichnet. 0! wird durch 1 festgelegt. Sollen n Objekte auf n Plätzen verteilt werden, wobei k_1 Objekte des Typs 1, k_2 des Typs 2, ..., k_p Objekte des Typs p, mit $n = \sum_{i=1}^{p} k_i$, dann existieren:

$$\frac{n!}{k_1! \cdot k_2! \cdot \ \ldots \ \cdot k_p!} \quad \text{Permutationen.}$$

3.4.2 Kombinationen

n beschreibt die Anzahl der **unterscheidbaren** Objekte, k die Anzahl der zur Verfügung stehenden Plätze (mit $n,k \geq 0$).

Ohne Wiederholung: Kommen die Objekte nur einmal vor, berechnet sich die Anzahl der Kombinationen wie folgt:

$$\binom{n}{k} = \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{k!} = \frac{n!}{k!(n-k)!} \quad , \binom{n}{k} = 0 \quad \text{für} \quad k > n$$

Mit Wiederholung: Können die Objekte hingegen mehrfach bzw. wiederholt ausgewählt werden, so gibt es

$$\binom{n+k-1}{k} = \frac{(n+k-1)!}{k!(n-1)!} \quad \text{Kombinationen.}$$

3.4.3 Variationen

Bei den Variationen wird ebenfalls die Anzahl möglicher Teilmengen berechnet, es wird jedoch zusätzlich die Reihenfolge der Objekte berücksichtigt.

Ohne Wiederholung: Kommt jedes Objekt nur einmal vor, so gibt es

$$\frac{n!}{(n-k)!} \quad \text{Variationen.}$$

Mit Wiederholung: Werden die n Objekte auf k Plätzen verteilt und kann jedes Objekt mehrfach ausgewählt werden, gibt es:

$$n^k \quad \text{Variationen.}$$

3.5 Wichtige diskrete Verteilungen

3.5.1 Die Bernoulliverteilung

Eine bernoulliverteilte Zufallsvariable X mit dem Parameter π ist definiert durch die Wahrscheinlichkeitsfunktion

$$P(X=k) = \begin{cases} \pi & k=1 \\ 1-\pi & k=0 \end{cases}$$

Momente: $E(X) = \pi, \qquad V(X) = \pi(1-\pi), \qquad S(X) = (1-2\pi)/\sqrt{\pi(1-\pi)}$

3.5.2 Die Binomialverteilung

Eine binomialverteilte Zufallsvariable X mit den Parametern n und π hat folgende Wahrscheinlichkeitsfunktion

$$P(X=k) = \binom{n}{k} \pi^k (1-\pi)^{n-k} \quad \text{für} \quad k = 0, \ldots, n$$

Momente: $E(X) = n\pi, \qquad V(X) = n\pi(1-\pi), \qquad S(X) = (1-2\pi)/\sqrt{n\pi(1-\pi)}$

3.5.3 Die hypergeometrische Verteilung

Eine hypergeometrisch verteilte Zufallsvariable X mit den Parametern N,A,n hat folgende Wahrscheinlichkeitsfunktion:

$$P(X=k) = \frac{\binom{A}{k}\binom{N-A}{n-k}}{\binom{N}{n}} \quad \text{für} \quad k = 0, \ldots, n$$

Momente: $E(X) = n\frac{A}{N}, \qquad V(X) = n\frac{A}{N}\left(1 - \frac{A}{N}\right)\left(\frac{N-n}{N-1}\right), \qquad S(X) = \frac{\sqrt{(N-1)}(N-2A)(N-2n)}{(N-2)\sqrt{A \cdot n(N-A)(N-n)}}$

3.5.4 Die Poissonverteilung

Eine poissonverteilte Zufallsvariable X mit dem Parameter $\lambda \in \mathbb{R}_+$ hat folgende Wahrscheinlichkeitsfunktion:

$$P(X=k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!} \quad \text{für} \quad k = 0, \ldots, n$$ **Momente:** $E(X) = \lambda, \qquad V(X) = \lambda, \qquad S(X) = 1/\sqrt{\lambda}$

3.6 Näherungen

Original-verteilung	Näherungs-verteilung	Parameter-transformation	Voraussetzungen
$B_{n,\pi}$	P_λ	$\lambda = n \cdot \pi$	$\pi \leq \frac{1}{10}$
$H_{N,A,n}$	$B_{n,\pi}$	$n = n$ $\pi = \frac{A}{N}$	$n \leq \frac{N}{10}$ $0{,}1 \leq \frac{A}{N} \leq 0{,}9$
$H_{N,A,n}$	P_λ	$\lambda = n \cdot \frac{A}{N}$	$n \leq \frac{N}{10}$ $\frac{A}{N} \leq 0{,}1$ oder $\frac{A}{N} \geq 0{,}9$

4 Stetige Verteilungen

4.2 Erwartungswertoperator und Momente

(a) Erwartungswert: $E(X) = \int_{-\infty}^{\infty} x \cdot f(x)dx = \mu_X$

(b) Varianz: $V(X) = E(X - \mu_X)^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 \cdot f(x)dx$

(c) Standardabweichung: $\sigma(X) = \sqrt{\sigma^2(X)}$

4.6 Wichtige stetige Verteilungen

4.6.1 Gleichverteilung

Dichtefunktion einer auf dem Intervall $[a,b]$ gleichverteilten Zufallssvariable X:

$$f(x) = \begin{cases} \frac{1}{b-a} & x \in [a,b] \\ 0 & \text{sonst} \end{cases}$$

Verteilungsfunktion:

$$F(x) = \begin{cases} 0 & x \in [-\infty, a) \\ \frac{x-a}{b-a} & x \in [a,b] \\ 1 & \text{sonst} \end{cases}$$

Momente: $E(X) = \frac{a+b}{2}$, $V(X) = \frac{(b-a)^2}{12}$, $S(X) = 0$

4.6.2 Exponentialverteilung

Dichtefunktion einer exponentialverteilten Zufallsvariable X mit der Übergangsrate λ:

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda x} & x \geq 0 \\ 0 & \text{sonst} \end{cases}$$

Verteilungsfunktion:

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & \text{sonst} \end{cases}$$

Momente: $E(X) = \frac{1}{\lambda}$, $V(X) = \frac{1}{\lambda^2}$, $S(X) = 2$ **Quantile:** $x_\alpha = \frac{-\ln(1-\alpha)}{\lambda}$

4.6.3 Normalverteilung

Die Dichtefunktion einer normalverteilten Zufallsvariable X:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

besitzt. Für die Parameter $\mu = 0$ und $\sigma^2 = 1$ erhält man die Dichtefunktion der **Standardnormalverteilung**:

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$$

Momente: $E(X) = \mu, \quad V(X) = \sigma^2, \quad S(X) = 0,$ **Quantile:** $x_\alpha = \mu_X + \sigma_X \cdot z_\alpha$

4.7 Näherungen

Original-verteilung	Näherungs-verteilung	Parameter-transformation	Voraussetzungen
$B_{n,\pi}$	$N(\mu,\sigma^2)$	$\mu = n \cdot \pi$ $\sigma^2 = n \cdot \pi(1-\pi)$	$n\pi(1-\pi) \geq 10$
$H_{N,A,n}$	$N(\mu,\sigma^2)$	$\mu = n \cdot \frac{A}{N}$ $\sigma^2 = n \cdot \frac{A}{N} \cdot \left(1 - \frac{A}{N}\right) \cdot \left(\frac{N-n}{N-1}\right)$	$n < \frac{N}{20}$ $0{,}1 \leq \frac{A}{N} \leq 0{,}9$
P_λ	$N(\mu,\sigma^2)$	$\mu = \lambda$ $\sigma^2 = \lambda$	$\lambda \geq 10$

5 Zweidimensionale Zusammenhänge - Exploration

5.2.2 Maßzahl - Empirische Kovarianz und empirische Korrelation

Empirische Kovarianz: $s_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n}\left(\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}\right)$

Empirischer Korrelationskoeffizient: $r_{xy} = \frac{s_{xy}}{\sqrt{s_x^2 \cdot s_y^2}} = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sqrt{\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)\left(\sum_{i=1}^{n} y_i^2 - n\bar{y}^2\right)}}$

5.3.3 Maßzahl - Kontingenzkoeffizienten

Chi-Quadrat-Koeffizient Grundlage für die Definition eines Abhängigkeitsmaßes sind die Differenzen zwischen der empirischen gemeinsamen Verteilung (Gesamtheit der h_{ij}) und der bei Unabhängigkeit zu erwartenden gemeinsamen Verteilung (Gesamtheit der h_{ij}^e):

$$\chi^2 = \sum_{i=1}^{k}\sum_{j=1}^{m} \frac{(h_{ij} - h_{ij}^e)^2}{h_{ij}^e} = n\left(\left(\sum_{i=1}^{k}\sum_{j=1}^{m} \frac{h_{ij}^2}{h_{i\bullet}h_{\bullet j}}\right) - 1\right)$$

Kontingenzkoeffizient: $K = \sqrt{\frac{\chi^2}{n + \chi^2}}$

Korrigierter Kontingenzkoeffizient: $K_{\text{kor}} = \frac{\sqrt{\frac{\chi^2}{n+\chi^2}}}{\sqrt{\frac{\min\{k,m\} - 1}{\min\{k,m\}}}}$

5.4.1 Maßzahl - Rangkorrelationskoeffizient von Spearman

Der Rangkorrelationskoeffizient von Spearman entspricht dem normalen Korrelationskoeffizient von Bravais-Pearson, bei dem statt den metrischen Ausprägungen die Ränge $r(x_i)$ und $r(y_i)$ verwendet werden:

$$r^s_{xy} = \frac{\sum_{i=1}^{n}[r(x_i) - \bar{r}(x)][r(y_i) - \bar{r}(y)]}{\sqrt{\sum_{i=1}^{n}[r(x_i) - \bar{r}(x)]^2 \sum_{i=1}^{n}[r(y_i) - \bar{r}(y)]^2}} = \frac{\sum_{i=1}^{n} r(x_i)r(y_i) - n\bar{r}(x)\bar{r}(y)}{\sqrt{\left(\sum_{i=1}^{n} r(x_i)^2 - n\bar{r}(x)^2\right)\left(\sum_{i=1}^{n} r(y_i)^2 - n\bar{r}(y)^2\right)}}$$

Spezialfall: Keine Bindungen

$$r^s_{xy} = 1 - \frac{6}{n(n^2-1)} \sum_{i=1}^{n}[r^*(x_i) - r^*(y_i)]^2$$

6 Mehrdimensionale Verteilungen

Ist $f_{XY}(x,y)$ die Dichtefunktion des Zufallsvektors (X,Y), so heißt

$$F_{XY}(c,d) = P(X \leq c, Y \leq d) = \sum_{X \leq c} \sum_{Y \leq d} f_{XY}(x,y)$$

die Verteilungsfunktion des Zufallsvektors (X,Y) an der Stelle (c,d), wobei $(c,d) \in \mathbb{R}^2$.

6.1 Grundlagen

Diskrete Randdichte- und Randverteilungsfunktion Ist $f_{XY}(x,y)$ die Dichtefunktion und $F_{XY}(a,b)$ die Verteilungsfunktion des diskreten zweidimensionalen Zufallsvektors (X,Y), dann sind die Randdichtefunktionen gegeben durch:

$$f_X(x_i) = p_X(x_i) = \sum_{j=1}^{m} f_{XY}(x_i,y_j), \qquad f_Y(y_j) = p_Y(y_j) = \sum_{x=1}^{n} f_{XY}(x_i,y_j)$$

und die Randverteilungsfunktionen durch:

$$F_X(a) = P(X \leq a) = P(X \leq a, Y \leq \infty) = \sum_{i=1}^{a} \sum_{j=1}^{m} f_{XY}(x_i,y_j) = F_{XY}(a,\infty)$$

$$F_Y(b) = P(Y \leq b) = P(X \leq \infty, Y \leq b) = \sum_{j=1}^{b} \sum_{i=1}^{n} f_{XY}(x_i,y_j) = F_{XY}(\infty,b)$$

6.2 Erwartungswertoperator und Momente

Gegeben ist eine diskrete zweidimensionale Zufallsvariable (X,Y) mit der Dichtefunktion $f_{XY}(x_i,y_j)$.

(a) **Erwartungswert von X:** $E(X) = \sum_{i=1}^{n} x p_X(x_i) = \mu_X$

(b) **Erwartungswert von Y:** $E(Y) = \sum_{j=1}^{m} y_j p_Y(y_j) = \mu_Y$

(c) **Varianz von X:** $V(X) = E((X-\mu_X)^2) = \sum_{i=1}^{n}(x_i - \mu_X)^2 p_X(x_i) = \sigma^2_X$

(d) **Varianz von Y:** $V(Y) = E((Y-\mu_Y)^2) = \sum_{j=1}^{m}(y_j - \mu_Y)^2 p_Y(y_j) = \sigma^2_Y$

(e) **Standardabweichung von X und Y:** $\sigma_X = \sqrt{\sigma^2_X}$ und $\sigma_Y = \sqrt{\sigma^2_Y}$

(f) **Kovarianz zwischen X und Y:** $COV(X,Y) = \sum_{i=1}^{n}\sum_{j=1}^{m}(x_i - \mu_X)(y_j - \mu_Y) \cdot f_{XY}(x_i,y_j) = \sigma_{XY}$

(g) **Korrelation zwischen X und Y:** $\rho_{XY} = \frac{\sigma_{XY}}{\sqrt{\sigma^2_X\ \sigma^2_Y}} = \frac{\sigma_{XY}}{\sigma_X\ \sigma_Y}$

Rechenregel für den Erwartungswert: Erwartungswert der Summe von n Zufallsvariablen:
$E(X_1 + X_2 + \cdots + X_n) = \mu_1 + \mu_2 + \cdots + \mu_n$

Rechenregeln für Varianzen und Kovarianzen

(a) Verschiebesatz für Kovarianzen: $COV(X,Y) = E[XY] - \mu_X \mu_Y$

(b) Kovarianz von zwei linear transformierten Zufallsvariablen: $COV(aX+b, cY+d) = acCOV(X,Y)$

(c) Varianz der Summe von zwei Zufallsvariablen: $V(X+Y) = V(X) + V(Y) + 2 \cdot COV(X,Y)$

(d) Varianz der Differenz von zwei Zufallsvariablen: $V(X-Y) = V(X) + V(Y) - 2 \cdot COV(X,Y)$

6.4 Multinomialverteilung

Eine k-dimensionale Multinomialverteilung hat die folgende Wahrscheinlichkeitsfunktion:

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k) = \frac{n!}{x_1! x_2! \ldots x_k!} \pi_1^{x_1} \pi_2^{x_2} \ldots \pi_k^{x_k}$$

mit $0 \leq x_i \leq n,\ x_1 + \ldots + x_k = n$, $0 < \pi_i < 1$ und $\pi_1 + \ldots + \pi_k = 1$

Momente $E(X_i) = n\pi_i, \quad V(X_i) = n\pi_i(1-\pi_i) \ \ 1 \leq i \leq k, \quad COV(X_i,X_j) = -n\pi_i\pi_j \ \ i,j = 1,\ldots,k$ und $i \neq j$

6.5 Bivariate Normalverteilung

Die Dichtefunktion einer bivariaten Normalverteilung ist gegeben durch:

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]\right)$$

Bei bivariat normalverteilten Zufallsvariablen folgt aus der Unkorreliertheit die Unabhängigkeit.

7 Punktschätzer und Grenzwertsätze

7.3 Stochastische Ungleichungen

Satz 7.1. ***Ungleichung von Markov***
X ist eine nicht-negative Zufallsvariable mit Erwartungswert $E(X)$. Dann gilt für jedes $a > 0$

$$P(X \geq a) \leq \frac{E(X)}{a}$$

Satz 7.2. ***Ungleichung von Tschebyscheff***
X ist eine Zufallsvariable mit Erwartungswert $E(X)$ und Varianz $V(X)$. Dann gilt für jedes $c > 0$

$$P(|X - E(X)| \geq c) \leq \frac{V(X)}{c^2} \quad .$$

7.4 Grenzwertsätze für Erwartungswerte

Satz 7.3. ***Gesetz der großen Zahlen***
Für beliebig kleines $c > 0$ gilt mit $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

$$P(|\bar{X} - \mu| \leq c) \to 1 \quad \textit{für} \quad n \to \infty \quad \textit{bzw.} \quad P(|\bar{X} - \mu| > c) \to 0 \quad \textit{für} \quad n \to \infty \quad .$$

Man sagt, das Stichprobenmittel konvergiert nach Wahrscheinlichkeit bzw. konvergiert stochastisch gegen μ und schreibt dafür auch kurz:

$$\bar{X} \xrightarrow{p} \mu \quad \textit{bzw.} \quad plim(\bar{X}) = \mu \quad .$$

Satz 7.4. ***Zentraler Grenzwertsatz***
Gegeben sind unabhängig und identisch verteilte Zufallsvariablen X_i mit $E(X_i) = \mu$ und $V(X_i) = \sigma^2$ mit $i = 1, \cdots, n$. Für die standardisierte Summe dieser Zufallsvariablen

$$Z_n = \frac{\left(\sum_{i=1}^{n} X_i\right) - n\mu}{\sqrt{n\sigma^2}}$$

ist $E(Z_n) = 0$ und $V(Z_n) = 1$. Für $n \to \infty$ ist die standardisierte Summe standardnormalverteilt, kurz

$$Z_n \sim N(0{,}1) \quad .$$

Für großes, aber endliches n ist die standardisierte Summe approximativ standardnormalverteilt:

$$Z_n \overset{a}{\sim} N(0{,}1) \quad .$$

7.5 Grenzwertsätze für Wahrscheinlichkeiten

Satz 7.5. ***Bernoulli-Theorem***
Sind $X_i \sim Bernoulli_\pi$ mit $i = 1, \cdots, n$, dann gilt für beliebig kleines $c > 0$ mit $\tilde{P} = \frac{1}{n}\sum_{i=1}^{n} X_i$

$$P(\left|\tilde{P} - \pi\right| \leq c) \to 1 \quad \textit{für} \quad n \to \infty \quad .$$

Satz 7.6. ***Grenzwertsatz von de Moivre***
Gegeben sind unabhängig bernoulliverteilte Zufallsvariablen X_i mit $E(X_i) = \pi$ und $V(X_i) = \pi(1-\pi)$ mit $i = 1, \cdots, n$. Die standardisierte Summe dieser Zufallsvariablen

$$Z_n = \frac{\left(\sum_{i=1}^{n} X_i\right) - n\pi}{\sqrt{n\pi(1-\pi)}}$$

ist für $n \to \infty$ standardnormalverteilt, kurz

$$Z_n \sim N(0{,}1) \quad .$$

Für großes, aber endliches n ist die standardisierte Summe approximativ standardnormalverteilt:

$$Z_n \overset{a}{\sim} N(0{,}1) \quad .$$

Daraus folgt, dass für $\tilde{P} = \frac{1}{n}\sum_{i=1}^{n} X_i$ gilt:

$$\tilde{P} \overset{a}{\sim} N\left(\pi, \frac{1}{n}\pi(1-\pi)\right) \quad .$$

8 Punktschätzung

8.1 Eigenschaften von Schätzfunktionen

Definition 8.1. ***Erwartungstreue und Verzerrung (Bias)***
Eine Schätzstatistik T (Schätzfunktion) für einen Parameter θ heißt erwartungstreu, wenn gilt:

$$E(T) = \theta \quad .$$

Die Verzerrung (bzw. aus dem englischsprachigen auch Bias genannt) ist bestimmt durch

$$Bias(T) = E(T) - \theta \quad .$$

Eine Folge von Schätzern, T_n, heißt asymptotisch erwartungstreu für θ, wenn gilt

$$\lim_{n\to\infty} E(T_n) = \theta \quad .$$

Definition 8.2. ***Konsistenz***
Eine Schätzstatistik T für einen Parameter θ heißt schwach konsistent, wenn für ein beliebig kleines, positives $c \in \mathbb{R}$ gilt:

$$\lim_{n\to\infty} P(|T_n - \theta| < c) = 1 \quad \textit{bzw.} \quad \lim_{n\to\infty} P(|T_n - \theta| \geq c) = 0 \quad .$$

Man sagt die Folge T_n konvergiert nach Wahrscheinlichkeit und schreibt auch kurz:

$$plim(T_n) = \theta \quad \textit{bzw.} \quad T_n \overset{p}{\to} \theta \quad .$$

Schwache Konsistenz gilt für eine erwartungstreue Schätzstatistik, wenn

$$\lim_{n\to\infty} V(T_n) = 0 \quad .$$

Satz 8.1. ***Schätzung von Funktionen eines Parameters (Stetigkeitssatz)***
Sei T_n eine schwach konsistente Folge von Schätzern für θ und g eine stetige Funktion. Dann ist $g(T_n)$ eine schwach konsistente Folge von Schätzern für $g(\theta)$.
Aus $T_n \xrightarrow{p} \theta$ folgt $g(T_n) \xrightarrow{p} g(\theta)$

Definition 8.3. ***Effizienz***
Es seien T_1 und T_2 erwartungstreue Schätzstatistiken für θ. Die Statistik T_1 ist effizienter (wirksamer) als die Statistik T_2, wenn sie eine kleinere Varianz besitzt:

$$V(T_1) < V(T_2) \quad .$$

8.2 Ausgewählte Punktschätzer

8.2.1 Schätzer für die Varianz der Grundgesamtheit

Schätzstatistik, Schätzwert und Momente der Schätzstatistik Gegeben sind n unabhängig und identisch verteilte Stichprobenvariablen $X_i (i = 1, \ldots, n)$. Die korrigierte Stichprobenvarianz

$$\hat{S}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

ist eine erwartungstreue Schätzstatistik für die Varianz der Grundgesamtheit σ^2. Der konkrete Schätzwert ist gegeben durch

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \quad .$$

8.2.2 Schätzer für den Erwartungswert der Grundgesamtheit

Schätzstatistik, Schätzwert und Momente der Schätzstatistik Gegeben sind n unabhängig und identisch verteilte Stichprobenvariablen $X_i (i = 1, \ldots, n)$. Die Schätzstatistik lautet

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad .$$

Der Schätzwert ist das arithmetische Mittel

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad .$$

Die Momente der Schätzstatistik sind: $E(\bar{X}) = \mu$ und $V(\bar{X}) = \sigma^2/n \quad .$

8.2.3 Schätzer für eine Wahrscheinlichkeit der Grundgesamtheit

Schätzstatistik, Schätzwert und Momente der Schätzstatistik Gegeben sind n unabhängig und identisch bernoulliverteilte Stichprobenvariablen $X_i (i = 1, \ldots, n)$, die das Auftreten eines Ereignisses A beschreiben. Die Schätzstatistik lautet

$$\tilde{P} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad .$$

Der Schätzwert ist die relative Häufigkeit des Ereignisse A in der Stichprobe

$$\tilde{p} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{mit} \quad x_i = \begin{cases} 1 & \text{Ereignis } A \text{ ist eingetreten/wurde beobachtet} \\ 0 & \text{sonst} \end{cases} \quad .$$

Die Momente der Schätzstatistik sind gegeben durch: $E(\tilde{P}) = \pi$ und $V(\tilde{P}) = (\pi(1-\pi))/n \quad .$

8.3 Schätzverfahren

8.3.1 Methode der kleinsten Quadrate

Bei der Methode der kleinsten Quadrate (KQ-Methode) wird der Schätzer für den Parameter einer Grundgesamtheit so bestimmt, dass die Summe der quadrierten Abweichungen der einzelnen Beobachtungen vom angepassten Wert möglichst klein ist.

8.3.2 Momentenmethode

Ziel ist die Schätzung der unbekannten Parameter $\theta_1, \dots \theta_k$. Gegeben sei eine Zufallsvariable X mit parametrischer Verteilungsfunktion bzw. parametrischen Verteilungsmodell, d.h. die Verteilungsfunktion hängt von den k Paramteren $\theta_1, \dots \theta_k$ ab. Des Weiteren liegen $X_i (i = 1, \dots n)$ unabhängig und identisch verteilte Stichprobenvariablen aus dieser parametrischen Verteilung vor. Dann ist

- das j-te theoretische Moment von X definiert als:

$$\mu_j = E(X^j), \quad j = 1,2, \dots \quad .$$

- das j-te empirische Moment von X bzw. das j-te Stichprobenmoment definiert als:

$$M_j = \frac{1}{n} \sum_{i=1}^{n} X_i^j \quad j = 1,2, \dots \quad .$$

- 1. Formuliere k theoretische Moment. Diese sind i.A. Funktionen der unbekannten Parameter, d.h. $\mu_j = \mu_j(\theta_1, \dots \theta_k)$.
  2. Setze $\mu_j = \mu_j(\theta_1, \dots \theta_k) = M_j$.
  3. Löse das System von k Gleichungen nach den Parametern auf. Dies liefert die Momentenschätzer von $\theta_1, \dots \theta_k$.

8.3.3 Maximum-Likelihood-Methode

Definition 8.4. ***Likelihood-Funktion, Maximum-Likelihood-Schätzung***
Seien X_i $(i = 1, \dots, n)$ Zufallsvariablen mit gemeinsamer Dichte- bzw. Wahrscheinlichkeitsfunktion $f_{x_1, \cdots, x_n}$, die von k unbekannten Parametern $\theta_1, \dots, \theta_k$ abhängt. Liegen k Realisationen $x_1, \dots, x_n$ vor, so ist die Likelihood-Funktion definiert als:

$$L(\theta_1, \dots, \theta_k) = f_{x_1, \cdots, x_n}(x_1, \dots, x_n;\ \theta_1, \dots, \theta_k) \quad .$$

Die Maximum-Likelihood-Schätzungen für $\theta_1, \dots, \theta_k$ sind diejenigen Werte, für welche L maximal ist, d. h.

$$L(\hat{\theta}_1, \dots, \hat{\theta}_k) = \max_{\theta_1, \dots, \theta_k} L(\theta_1, \dots, \theta_k) = \max_{\theta_1, \dots, \theta_k} \log(L(\theta_1, \dots, \theta_k)).$$

9 Intervallschätzung

9.1 Wichtige stetige Verteilungen in der schließenden Statistik

9.1.1 Chi-Quadrat-Verteilung

Definition

$$\sum_{i=1}^{n} Z_i^2 \sim \chi_n^2 \quad , \quad Z_i \quad \text{unabhängig und identisch standardnormalverteilt}$$

sprich: χ^2-verteilt mit n Freiheitsgraden.

Momente

$$E(X) = n, \qquad V(X) = 2n, \qquad S(X) = \frac{2\sqrt{2}}{\sqrt{n}} \quad .$$

9.1.2 t-Verteilung

Definition Gegeben sind die unabhängigen Zufallsvariablen $Z \sim N(0,1)$ und $Y \sim \chi_n^2$. Dann ist

$$\frac{Z}{\sqrt{Y/n}} \sim t_n \quad ,$$

sprich: t-verteilt mit n Freiheitsgraden. Für $n \to \infty$ konvergiert die t-Verteilung gegen die Standardnormalverteilung, ab $n \geq 30$ ist die t-Verteilung in guter Näherung standardnormalverteilt.

Momente

$$E(X) = 0, \qquad V(X) = \frac{n}{n-2} \quad \text{für} \quad n \geq 3, \qquad S(X) = 0 \quad .$$

Satz 9.2. ***Quantile der t-Verteilung***
Gegeben ist die t_n-verteilte Zufallsvariable X. Dann gilt für ein beliebiges α-Quantil $t_{n;\alpha}$ dieser Zufallsvariablen

$$t_{n;\alpha} = -t_{n;1-\alpha} \quad .$$

9.1.3 Verteilungen der Schätzer für Varianz, Erwartungswert und Wahrscheinlichkeit

Verteilung der Schätzstatistik für die Varianz der Grundgesamtheit Die Verteilung des Schätzers $\hat{S}^2$ selbst ist unbekannt. Sind die X_i $(i = 1, \ldots, n)$ unabhängig und identisch normalverteilt, dann gilt

$$\frac{(n-1)\hat{S}^2}{\sigma^2} \sim \chi^2_{n-1} \quad .$$

Verteilung der Schätzstatistik für den Erwartungswert der Grundgesamtheit Bekannte Varianz σ^2:

$$\sqrt{n}\frac{\bar{X} - \mu}{\sigma} \sim N(0,1) \quad .$$

Varianz σ^2 unbekannt. σ^2 wird über $\hat{S}^2$ geschätzt und X_i $(i = 1, \ldots, n)$ normalverteilt:

$$\sqrt{n}\frac{\bar{X} - \mu}{\hat{S}} \sim t_{n-1} \quad .$$

Varianz unbekannt. σ^2 wird über $\hat{S}^2$ geschätzt und X_i $(i = 1, \ldots, n)$ unabhängig und identisch beliebig verteilt

$$\sqrt{n}\frac{\bar{X} - \mu}{\hat{S}} \overset{a}{\sim} N(0,1) \quad \text{für} \quad n \geq 30 \quad .$$

Verteilung der Schätzstatistik für die Wahrscheinlichkeit der Grundgesamtheit Die unbekannte Varianz der Schätzstatistik wir unter Verwendung von $\tilde{P}$ geschätzt. Dann gilt

$$\frac{\tilde{P} - \pi}{\sqrt{\frac{\tilde{P}(1-\tilde{P})}{n}}} \overset{a}{\sim} N(0,1) \quad \text{für} \quad n \geq 30 \quad .$$

9.2 Konfidenzintervall für den Erwartungswert

Fall 1: Konfidenzintervall für μ bei $X_i \overset{u.i.v.}{\sim} N(\mu,\sigma^2)$, σ^2 unbekannt. Die Varianz wird durch $\hat{S}^2$ geschätzt.

- Das $(1-\alpha)$-Konfidenzintervall für μ ist gegeben durch:

  $$KI_{\mu;1-\alpha} = [G_{\underline{\mu}}; G_{\overline{\mu}}] \quad .$$

  mit $G_{\underline{\mu}} = \bar{X} - t_{n-1;1-\frac{\alpha}{2}} \cdot \frac{\hat{S}}{\sqrt{n}}$ und $G_{\overline{\mu}} = \bar{X} + t_{n-1;1-\frac{\alpha}{2}} \cdot \frac{\hat{S}}{\sqrt{n}}$.

- Die geschätzten Intervallgrenzen sind gegeben durch

  $$g_{\underline{\mu}} = \bar{x} - t_{n-1;1-\frac{\alpha}{2}} \cdot \frac{\hat{\sigma}}{\sqrt{n}} \quad \text{und} \quad g_{\overline{\mu}} = \bar{x} + t_{n-1;1-\frac{\alpha}{2}} \cdot \frac{\hat{\sigma}}{\sqrt{n}} \quad .$$

Fall 2: Konfidenzintervall für μ bei $X_i \overset{u.i.v.}{\sim}$ beliebig, $n \geq 30$, σ^2 unbekannt. Die Varianz wird durch $\hat{S}^2$ geschätzt.

- Das $(1-\alpha)$-Konfidenzintervall für μ ist gegeben durch:

  $$KI_{\mu;1-\alpha} = [G_{\underline{\mu}}; G_{\overline{\mu}}] \quad .$$

  mit $G_{\underline{\mu}} = \bar{X} - z_{1-\frac{\alpha}{2}} \cdot \frac{\hat{S}}{\sqrt{n}}$ und $G_{\overline{\mu}} = \bar{X} + z_{1-\frac{\alpha}{2}} \cdot \frac{\hat{S}}{\sqrt{n}}$.

- Die geschätzten Intervallgrenzen sind gegeben durch

  $$g_{\underline{\mu}} = \bar{x} - z_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\sigma}}{\sqrt{n}} \quad \text{und} \quad g_{\overline{\mu}} = \bar{x} + z_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\sigma}}{\sqrt{n}} \quad .$$

9.3 Konfidenzintervall für eine Wahrscheinlichkeit

Konfidenzintervall für π bei $X_i \overset{u.i.v.}{\sim}$ Bernoulli, $n \geq 30$

- Das $(1-\alpha)$-Konfidenzintervall für π ist gegeben durch:

$$KI_{\pi;1-\alpha} = [G_{\underline{\pi}}; G_{\overline{\pi}}] \quad .$$

mit $G_{\underline{\pi}} = \tilde{P} - \sqrt{\frac{1}{n}\tilde{P}(1-\tilde{P})} \cdot z_{1-\frac{\alpha}{2}}$ und $G_{\overline{\pi}} = \tilde{P} + \sqrt{\frac{1}{n}\tilde{P}(1-\tilde{P})} \cdot z_{1-\frac{\alpha}{2}}$.

- Die geschätzten Intervallgrenzen sind gegeben durch

$$g_{\underline{\pi}} = \tilde{p} - \sqrt{\frac{1}{n}\tilde{p}(1-\tilde{p})} \cdot z_{1-\frac{\alpha}{2}} \quad \text{und} \quad g_{\overline{\pi}} = \tilde{p} + \sqrt{\frac{1}{n}\tilde{p}(1-\tilde{p})} \cdot z_{1-\frac{\alpha}{2}} \quad .$$

9.4 Konfidenzintervall für die Varianz

Konfidenzintervall für σ^2 bei $X_i \overset{u.i.v.}{\sim} N(\mu,\sigma^2)$

- Das $(1-\alpha)$-Konfidenzintervall für σ^2 ist gegeben durch:

$$KI_{\sigma^2;1-\alpha} = [G_{\underline{\sigma}^2}; G_{\overline{\sigma}^2}] \quad .$$

mit $G_{\underline{\sigma}^2} = \dfrac{(n-1)\hat{S}^2}{\chi^2_{n-1;1-\frac{\alpha}{2}}}$ und $G_{\overline{\sigma}^2} = \dfrac{(n-1)\hat{S}^2}{\chi^2_{n-1;\frac{\alpha}{2}}}$.

- Die geschätzten Intervallgrenzen sind gegeben durch

$$g_{\underline{\sigma}^2} = \frac{(n-1)\hat{\sigma}^2}{\chi^2_{n-1;1-\frac{\alpha}{2}}} \quad \text{und} \quad g_{\overline{\sigma}^2} = \frac{(n-1)\hat{\sigma}^2}{\chi^2_{n-1;\frac{\alpha}{2}}} \quad .$$

10 Grundlagen von Signifikanztests

10.1 Grundidee und Prinzipien eines zweiseitigen statistischen Tests

Die Durchführung eines solchen statistischen Tests erfolgt immer gemäß der folgenden 6 Schritte, welche die **Grundstruktur eines statistischen Tests** ausmachen.

1. Schritt: Identifikation des statistischen Problems. Zunächst wird das statistische Problem identifiziert. Dies bedeutet, dass wir den Zufallsvorgang (Grundgesamtheit und Zufallsvariablen) identifizieren, der zur Situation unter Unsicherheit führt, und das reale Problem mit einem oder mehreren Parametern der entsprechenden Grundgesamtheit verknüpfen.

2. Schritt: Verteilungsmodell. Es wird ein passendes statistisches Modell (Verteilungsmodell) formuliert.

3. Schritt: Formulierung des statistischen Testproblems. Das reale Problem wird mit Hilfe der Parameter der Grundgesamtheit als statistisches Testproblem formuliert. Dies erfolgt in Form von *Hypothesen*. Eine der Hypothesen enthält die Vermutung über den Parameter/die Parameter der Grundgesamtheit, dies ist die sogenannte *Nullhypothese*, H_0, die andere ist das genaue Gegenteil davon und wird *Alternativhypothese*, H_1, genannt.

4. Schritt: Festlegung des Signifikanzniveaus. Das *Test- bzw. Signifikanzniveau* α wird festgelegt. Dieses gibt an, wie wahrscheinlich eine irrtümliche Ablehnung der Nullhypothese höchstens sein darf.

5. Schritt: Festlegung der Teststatistik. Es wird eine geeignete *Teststatistik*, auch *Prüfgröße* genannt, gebildet. Diese ist eine Zufallsvariable, deren Verteilung unter der Nullhypothese bekannt sein muss. Diese Verteilung wird auch *Testverteilung* bzw. *Prüfverteilung* oder auch *Nullverteilung* genannt. Auf Basis dieser Testverteilung werden *kritische Werte*, auch *kritische Grenzen* genannt, bestimmt, welche die Annahme- und Ablehnungsbereiche dieser Teststatistik festlegen. Der *Annahmebereich* ist diejenige Menge aller Werte der Teststatistik, für die H_0 beibehalten wird. Die kritischen Grenzen gehören dabei auch immer zum Annahmebereich. Der *Ablehnungbereich* ist die Menge aller Werte der Teststatistik, für die H_0 abgelehnt und H_1 angenommen wird. Er enthält also alle möglichen Realisationen des Punktschätzers, die so weit von der ursprünglichen Vermutung entfernt liegen, dass bei einer Realisation des Punktschätzer in diesem Bereich die Nullhypothese abgelehnt wird, da mit hoher Wahrscheinlichkeit eine systematische und nicht nur zufällige Abweichung von der ursprünglichen Vermutung vorliegt.

6. Schritt: Realisation der Teststatistik und Testentscheidung. Für die Stichprobe $X_i(i = 1,\vdots,n)$ wird der realisierte Wert der Teststatistik berechnet. Basierend auf diesem Wert wird mittels der Annahme- und Ablehnungsbereiche, die sich aus Schritt 4 und Schritt 5 ergeben, die Testentscheidung getroffen. Es gilt hierbei folgende *Entscheidungsregel*: Fällt der auf Basis der vorliegenden Stichprobe berechnete Punktschätzer in den Ablehnungsbereich, wird die Nullhypothese *abgelehnt* (auch *verworfen* oder *nicht beibehalten*). Fällt dieser in den Annahmebereich, wird die Nullhypothese *beibehalten* (auch *nicht verworfen* oder *nicht abgelehnt*).

11 Ausgewählte Signifikanztests

11.1 Test für den Erwartungswert einer Grundgesamtheit

11.1.1 Bekannte Varianz

Annahmen:

- Die Stichprobenvariablen $X_i(i = 1, \cdots, n)$ sind unabhängig und identisch verteilt.
- Die Varianz der Stichprobenvariablen $V(X_i) = \sigma^2$ ist bekannt.
- Für kleine Stichproben ($n < 30$) muss zusätzlich unterstellt werden, dass die Stichprobenvariablen X_i normalverteilt sind. Für große Stichproben ($n \geq 30$) kann aufgrund des zentralen Grenzwertsatzes auf die zusätzliche Annahme der Normalverteilung verzichtet werden.

Teststatistik:

$$Z^* = \sqrt{n} \cdot \frac{\bar{X} - \mu_0}{\sigma} \overset{H_0}{\sim} N(0, 1) \quad .$$

Schätzwert der Teststatistik Z^*:

$$z^* = \sqrt{n} \cdot \frac{\bar{x} - \mu_0}{\sigma} \quad .$$

11.1.2 Unbekannte Varianz und normalverteilte Stichprobe

- Die Stichprobenvariablen $X_i(i = 1, \cdots, n)$ sind unabhängig und identisch normalverteilt.

Teststatistik:

$$T^* = \sqrt{n} \cdot \frac{\bar{X} - \mu_0}{\hat{S}} \overset{H_0}{\sim} t_{n-1} \quad .$$

Schätzwert der Teststatistik T^*:

$$t^* = \sqrt{n} \cdot \frac{\bar{x} - \mu_0}{\hat{\sigma}} \quad .$$

11.1.3 Unbekannte Varianz und große Stichprobe

Annahmen:

- Die Stichprobenvariablen $X_i(i = 1, \cdots, n)$ sind unabhängig und identisch verteilt.
- Es liegt eine große Stichprobe mit $n \geq 30$ Beobachtungen vor.

Teststatistik:

$$Z^* = \sqrt{n} \cdot \frac{\bar{X} - \mu_0}{\hat{S}} \overset{H_0}{\sim} N(0, 1) \quad .$$

Schätzwert der Teststatistik Z^*:

$$z^* = \sqrt{n} \cdot \frac{\bar{x} - \mu_0}{\hat{\sigma}} \quad .$$

11.2 Test für die Wahrscheinlichkeit einer Grundgesamtheit

Annahmen:

- Die Stichprobenvariablen $X_i(i = 1, \cdots, n)$ sind unabhängig und identisch bernoulliverteilt.
- Es liegt eine große Stichprobe mit $n \geq 30$ Beobachtungen vor.

Teststatistik:

$$Z^* = \frac{\tilde{P} - \pi_0}{\sqrt{\frac{1}{n} \cdot \pi_0(1 - \pi_0)}} \overset{H_0}{\sim} N(0, 1) \quad .$$

Schätzwert der Teststatistik Z^* ergibt:

$$z^* = \frac{\tilde{p} - \pi_0}{\sqrt{\frac{1}{n} \cdot \pi_0(1 - \pi_0)}} \quad .$$

11.3 Test für den Median einer Grundgesamtheit: Vorzeichen-Test

Annahmen:

- Die Stichprobenvariablen $X_i(i = 1, \cdots, n)$ sind unabhängig und identisch verteilt mit stetiger Verteilungsfunktion.
- Es liegt eine große Stichprobe mit $n \geq 30$ Beobachtungen vor.

Teststatistik und Nullverteilung: Sei $Y_i = I_{(\delta_0,\infty)}(X_i)$, wobei $I_{(\delta_0,\infty)}(\cdot)$ die Indikatorfunktion bezeichnet. Sie nimmt den Wert 1 an, wenn $X > \delta_0$, und den Wert 0 an, wenn $X \leq \delta_0$.
Die Schätzstatistik

$$\tilde{P}_Y = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

ist ein erwartungstreuer und konsistenter Schätzer für $\pi_Y = P(Y_i = 1)$.

Teststatistik:

$$Z^* = \frac{\tilde{P}_Y - 0.5}{\sqrt{\frac{1}{n}\cdot 0.25}} \overset{H_0}{\sim} N(0,1) \quad .$$

Schätzwert der Teststatistik Z^*:

$$z^* = \frac{\tilde{p}_Y - 0.5}{\sqrt{\frac{1}{n}\cdot 0.25}} \quad .$$

11.4 Test für die Varianz einer Grundgesamtheit

Annahme: Die Stichprobenvariablen $X_i (i = 1, \cdots, n)$ sind unabhängig und identisch normalverteilt.

Teststatistik:

$$\frac{(n-1)\hat{S}^2}{\sigma_0^2} \overset{H_0}{\sim} \chi^2_{n-1} \quad .$$

Schätzwert der Teststatistik:

$$\chi^{2*} = \frac{(n-1)\hat{\sigma}^2}{\sigma_0^2} \quad .$$

11.5 Test für Erwartungswertunterschiede bei unabhängigen Grundgesamtheiten

11.5.1 Bekannte Varianzen

Annahmen:

- Die Varianz der unabhängigen und identisch verteilten Stichprobenvariablen $X_i (i = 1, \cdots, n)$ ist bekannt und gegeben durch $V(X_i) = \sigma_X^2$.
- Die Varianz der unabhängigen und identisch verteilten Stichprobenvariablen $Y_j (j = 1, \cdots, m)$ ist bekannt und gegeben durch $V(Y_j) = \sigma_Y^2$.
- Die Stichprobenvariablen X_i und Y_j sind unabhängig voneinander.
- Für kleine Stichproben ($n,m < 30$) muss zusätzlich unterstellt werden, dass die Stichprobenvariablen X_i und Y_j normalverteilt sind. Für große Stichproben ($n,m \geq 30$) kann aufgrund des zentralen Grenzwertsatzes auf die zusätzliche Annahme der Normalverteilung verzichtet werden.

Teststatistik:

$$D^* = \frac{\bar{X} - \bar{Y} - \delta_0}{\sqrt{\frac{1}{n}\cdot\sigma_X^2 + \frac{1}{m}\cdot\sigma_Y^2}} \overset{H_0}{\sim} N(0,1) \quad .$$

Schätzwert der Teststatistik D^*:

$$d^* = \frac{\bar{x} - \bar{y} - \delta_0}{\sqrt{\frac{1}{n}\cdot\sigma_X^2 + \frac{1}{m}\cdot\sigma_Y^2}} \quad .$$

11.5.2 Unbekannte Varianzen und normalverteilte Stichproben

Annahmen:

- Die Stichprobenvariablen $X_i (i = 1, \ldots, n)$ sind unabhängig und identisch normalverteilt.
- Die Stichprobenvariablen $Y_j (j = 1, \ldots, m)$ sind unabhängig und identisch normalverteilt.
- Die Stichprobenvariablen X_i und Y_j sind unabhängig voneinander.

Teststatistik D^*:

$$D^* = \frac{\bar{X} - \bar{Y} - \delta_0}{\sqrt{\frac{1}{n}\cdot\hat{S}_X^2 + \frac{1}{m}\cdot\hat{S}_Y^2}} \overset{H_0}{\sim} t_{df} \qquad \text{mit} \qquad df = \left\lfloor \frac{\left(\hat{S}_X^2/n + \hat{S}_Y^2/m\right)^2}{\frac{\left(\hat{S}_X^2/n\right)^2}{n-1} + \frac{\left(\hat{S}_Y^2/m\right)^2}{m-1}} \right\rfloor$$

Dabei bedeutet die Schreibweise $\lfloor c \rfloor$, dass die Zahl c auf die nächste ganze Zahl abgerundet wird. Als Schätzwert der Teststatistik D^* und als geschätzte Freiheitsgrade der t-Verteilung ergeben sich

$$d^* = \frac{\bar{x} - \bar{y} - \delta_0}{\sqrt{\frac{1}{n}\cdot\hat{\sigma}_X^2 + \frac{1}{m}\cdot\hat{\sigma}_Y^2}} \qquad \text{und} \qquad \widehat{df} = \left\lfloor \frac{\left(\hat{\sigma}_X^2/n + \hat{\sigma}_Y^2/m\right)^2}{\frac{\left(\hat{\sigma}_X^2/n\right)^2}{n-1} + \frac{\left(\hat{\sigma}_Y^2/m\right)^2}{m-1}} \right\rfloor$$

Bemerkung: Wird zusätzlich unterstellt, dass die Varianzen in den Grundgesamtheiten identisch sind, kann die Varianz der Differenz zwischen $\bar{X}$ und $\bar{Y}$ genauer geschätzt werden durch

$$V(\bar{X}-\bar{Y}) = \left(\frac{1}{n}+\frac{1}{m}\right)\frac{(n-1)\hat{S}_X^2+(m-1)\hat{S}_Y^2}{n+m-2} \quad .$$

11.5.3 Unbekannte Varianzen und große Stichproben

Annahmen:

- Die Stichprobenvariablen $X_i(i=1,\ldots,n)$ sind unabhängig und identisch verteilt.
- Die Stichprobenvariablen $Y_j(j=1,\ldots,m)$ sind unabhängig und identisch verteilt.
- Die Stichprobenvariablen X_i und Y_j sind unabhängig voneinander.
- In beiden Stichproben liegen mindestens 30 Beobachtungen vor.

Teststatistik:

$$D^* = \frac{\bar{X}-\bar{Y}-\delta_0}{\sqrt{\frac{1}{n}\cdot\hat{S}_X^2+\frac{1}{m}\cdot\hat{S}_Y^2}} \overset{H_0}{\sim} N(0,1) \quad .$$

Schätzwert der Teststatistik D^*:

$$d^* = \frac{\bar{x}-\bar{y}-\delta_0}{\sqrt{\frac{1}{n}\cdot\hat{\sigma}_X^2+\frac{1}{m}\cdot\hat{\sigma}_Y^2}} \quad .$$

Bemerkung: Wird zusätzlich unterstellt, dass die Varianzen in den Grundgesamtheiten identisch sind, kann die Varianz der Differenz zwischen $\bar{X}$ und $\bar{Y}$ wieder genauer geschätzt werden durch

$$V(\bar{X}-\bar{Y}) = \left(\frac{1}{n}+\frac{1}{m}\right)\frac{(n-1)\hat{S}_X^2+(m-1)\hat{S}_Y^2}{n+m-2} \quad .$$

11.6 Test für Wahrscheinlichkeitsunterschiede bei unabhängigen Grundgesamtheiten

Annahmen:

- Die Stichprobenvariablen $X_i(i=1,\cdots,n)$ sind unabhängig und identisch bernoulliverteilt.
- Die Stichprobenvariablen $Y_j(j=1,\cdots,m)$ sind unabhängig und identisch bernoulliverteilt.
- Die Stichprobenvariablen X_i und Y_j sind unabhängig voneinander.
- In beiden Stichproben liegen mindestens 30 Beobachtungen vor, d. h. $n \geq 30$ und $m \geq 30$.

Teststatistik D^* :

$$D^* = \frac{\tilde{P}_X-\tilde{P}_Y-\delta_0}{\sqrt{\frac{\tilde{P}_X(1-\tilde{P}_X)}{n}+\frac{\tilde{P}_Y(1-\tilde{P}_Y)}{m}}} \overset{H_0}{\sim} N(0,1) \quad .$$

Schätzwert der Teststatistik D^*:

$$d^* = \frac{\tilde{p}_X-\tilde{p}_Y-\delta_0}{\sqrt{\frac{\tilde{p}_X(1-\tilde{p}_X)}{n}+\frac{\tilde{p}_Y(1-\tilde{p}_Y)}{m}}} \quad .$$

11.8 Test auf Unabhängigkeit von zwei metrisch skalierten Merkmalen

Annahme: Die Stichprobenvariablen $(X_i,Y_i)(i=1,\ldots,n)$ sind bivariat normalverteilt.

Hypothesenpaar: $H_0 : \rho_{XY} = 0$ gegen $H_1 : \rho_{XY} \neq 0$.

Teststatistik:

$$T^* = \frac{R_{xy}}{\sqrt{\frac{1-R_{xy}^2}{n-2}}} \overset{H_0}{\sim} t_{n-2} \quad .$$

Schätzwert der Teststatistik T^*:

$$t^* = \frac{r_{xy}}{\sqrt{\frac{1-r_{xy}^2}{n-2}}} \quad .$$

11.9 Test für den Korrelationskoeffizienten einer Grundgesamtheit

Annahmen:

- Die bivariaten Stichprobenvariablen $(X_i,Y_i)(i=1,\ldots,n)$ sind unabhängig und identisch bivariat normalverteilt.
- Die Stichprobengröße ist $n \geq 30$.

Teststatistik:

$$Z^* = \frac{\sqrt{n-3}}{2} \cdot \left[\ln\left(\frac{1+R_{xy}}{1-R_{xy}}\right) - \ln\left(\frac{1+\rho_0}{1-\rho_0}\right) \right] \overset{H_0}{\sim} N(0,1) \quad .$$

Schätzwert der Teststatistik Z^*:

$$z^* = \frac{\sqrt{n-3}}{2} \cdot \left[\ln\left(\frac{1+r_{xy}}{1-r_{xy}}\right) - \ln\left(\frac{1+\rho_0}{1-\rho_0}\right) \right] \quad .$$

11.10 Test auf Unabhängigkeit von zwei nominal skalierten Merkmalen

Annahmen:

- Gegeben ist ein nominal skaliertes Merkmal X mit k verschiedenen Kategorien und ein nominal skaliertes Merkmal Y mit m verschiedenen Kategorien.
- Die gemeinsame Häufigkeitsverteilung der Merkmale X und Y liegt in Form einer Kontingenztabelle vor.
- Eines der beiden Merkmale kann dabei auch aus einer Skalenniveautransformation hervorgegangen sein.

Hypothesenpaar: H_0 : *Merkmale X und Y sind unabhängig* gegen H_1 : *Merkmale X und Y sind abhängig* .

was auch dargestellt werden kann als:

$H_0 : P(X=i,Y=j) = P(X=i) \cdot P(Y=j)$ für alle $i = 1,...,k$ und $j = 1,...,m$ gegen
$H_1 : P(X=i,Y=j) \neq P(X=i) \cdot P(Y=j)$ für mindestens ein Paar (i,j) mit $i = 1,...,k$ und $j = 1,...,m$

Teststatistik χ^{2*}:

$$\chi^{2*} = \sum_{i=1}^{k} \sum_{j=1}^{m} \frac{(h_{ij} - h_{ij}^e)^2}{h_{ij}^e} = n \left(\left(\sum_{i=1}^{k} \sum_{j=1}^{m} \frac{h_{ij}^2}{h_{i\bullet} h_{\bullet j}} \right) - 1 \right)$$

$$\chi^{2*} \overset{H_0}{\sim} \chi^2_{(k-1)(m-1)} \quad .$$

Entscheidungsregel: H_0 wird bei einem Signifikanzniveau α verworfen, wenn gilt: $\chi^{2*} > \chi^2_{(k-1)(m-1);1-\alpha}$.

11.11 Anpassungstest für diskrete Verteilungen

Annahme: In einer Stichprobe liegen die Realisationen einer diskreten Zufallsvariablen X vor.

Hypothesenpaar: H_0 : *X folgt einer bestimmten diskreten Verteilung* gegen H_1 : *Verteilung von X unbekannt* .

Teststatistik: Der Schätzwert der Teststatistik ist gegeben durch

$$\chi^{2*} = \sum_{i=1}^{r} \frac{(h_i - h_i^e)^2}{h_i^e} \quad .$$

Ist die Nullhypothese zutreffend und es liegen tatsächlich Realisationen der vermuteten diskreten Verteilung vor, dann gilt für die Verteilung des als Teststatistik aufgefassten χ^{2*}-Koeffizienten

$$\chi^{2*} \overset{H_0}{\sim} \chi^2_{r-1-p} \quad .$$

Der Parameter p ist hier die Anzahl der Parameter, die ggf. für die Spezifikation der in der Nullhypothese angenommenen Verteilung aus der Stichprobe geschätzt werden müssen. Die Nullverteilung der Teststatistik ist hinreichend gut, wenn $h_{ij}^e \geq 1$, für alle (i,j) mit $i = 1, \ldots k$ und $j = 1, \ldots m$ und $h_{ij}^e \geq 5$, für und mindestens 80% aller (i,j) mit $i = 1, \ldots, k$ und $j = 1, \ldots, m$.

Entscheidungsregel: H_0 wird bei einem Signifikanzniveau α verworfen, wenn gilt: $\chi^{2*} \geq \chi^2_{1-\alpha;r-1-p}$.

12 Lineares Einfachregressionsmodell

12.2 Parameterschätzung mit der KQ-Methode

$$QS(\beta_0, \beta_1) = \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2 \quad \longrightarrow \min_{\beta_0, \beta_1} \quad .$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} = \frac{\frac{1}{n}\left(\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}\right)}{\frac{1}{n}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)} = \frac{s_{xy}}{s_x^2}$$

12.6 Anpassungsgüte der Regression

Definition 12.1. ***Bestimmtheitsmaß R^2***
Das Bestimmtheitsmaß ist definiert als

$$R^2 = \frac{ESS}{SST} = \frac{SST - RSS}{SST} = 1 - \frac{RSS}{SST} \quad mit \quad SST = \sum_{i=1}^{n}(y_i - \bar{y})^2, ESS = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 \quad und \quad RSS = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Satz 12.1. ***Berechung von R^2 im Einfachregressionsmodell***
Im Einfachregressionmodell entspricht das Bestimmtheitsmaß der quadrierten empirischen Korrelation zwischen der abhängigen Variable und der Kovariable:

$$R^2 = \left(R_{xy}\right)^2 \quad . \tag{12.1}$$

12.7 Die KQ-Annahmen

KQ-Annahmen:

(A1) $E(U_i|X_i) = 0 \quad .$

(A2) $(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$ sind identisch und unabhängig verteilt.

(A3) $0 < E(X_i^4) < \infty, \quad 0 < E(Y_i^4) < \infty \quad .$

12.8 Die Eigenschaften des KQ-Schätzers

Eigenschaften der KQ-Schätzer: Basierend auf den KQ-Annahmen (A1) - (A3) kann gezeigt werden, dass für die KQ-Schätzer von β_0 und β_1 des einfachen linearen Regressionsmodells gilt:

a) $E(\hat{\beta}_0) = \beta_0$ und $E(\hat{\beta}_1) = \beta_1$, d. h. $\hat{\beta}_0$ und $\hat{\beta}_1$ sind erwartungstreue Schätzer,

b) $\hat{\beta}_0$ und $\hat{\beta}_1$ sind in großen Stichproben gemeinsam normalverteilt.

c) In großen Stichproben sind $\hat{\beta}_0 \overset{a}{\sim} N(\beta_0, \sigma^2_{\hat{\beta}_0})$ und $\hat{\beta}_1 \overset{a}{\sim} N(\beta_1, \sigma^2_{\hat{\beta}_1})$,
wobei $\sigma^2_{\hat{\beta}_0} = \frac{V(H_i U_i)}{n[E(H_i^2)]^2}$ mit $H_i = 1 - \frac{X_i \mu_x}{E(X_i^2)}$ und $\sigma^2_{\hat{\beta}_1} = \frac{V[(X_i - \mu_x)U_i]}{nV(X_i)^2}$

d) $\hat{\beta}_0$ und $\hat{\beta}_1$ sind konsistent.

12.9 Hypothesentests und Konfidenzintervalle im einfachen Regressionsmodell

Konsistente Schätzer der Varianz der KQ-Schätzer: Im einfachen linearen Regressionsmodell unter den Annahmen (A1)-(A3) sind

$$\hat{\sigma}^2_{\hat{\beta}_0} = \frac{1}{n}\frac{\frac{1}{n-2}\sum_{i=1}^{n}\hat{H}_i^2\hat{U}_i^2}{\left(\frac{1}{n}\sum_{i=1}^{n}\hat{H}_i^2\right)^2} \quad \text{mit} \quad \hat{H}_i = 1 - \left(\frac{\bar{X}}{\frac{1}{n}\sum_{i=1}^{n}X_i^2}\right)X_i \quad \text{und} \quad \hat{\sigma}^2_{\hat{\beta}_1} = \frac{1}{n}\frac{\frac{1}{n-2}\sum_{i=1}^{n}(X_i-\bar{X})^2\hat{U}_i^2}{\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2\right)^2}$$

konsistente Schätzer für $\sigma^2_{\hat{\beta}_0}$ und $\sigma^2_{\hat{\beta}_1}$.

(Approximative) Tests bezüglich der Regressionskoeffizienten:

Annahmen: Es liegt ein einfaches lineares Regressionsmodell vor und die Annahmen (A1)-(A3) seien erfüllt. Des Weiteren liegt eine große Stichprobe vor.

Teststatistik und Nullverteilung:

$$Z^*_{\hat{\beta}_0} = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2_{\hat{\beta}_0}}} \overset{H_0}{\sim} N(0,1) \quad \text{bzw.} \quad Z^*_{\hat{\beta}_1} = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\hat{\sigma}^2_{\hat{\beta}_1}}} \overset{H_0}{\sim} N(0,1) \quad .$$

Schätzwerte der Teststatistiken:

$$z^*_{\hat{\beta}_0} = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2_{\hat{\beta}_0}}} \quad \text{bzw.} \quad z^*_{\hat{\beta}_1} = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\hat{\sigma}^2_{\hat{\beta}_1}}} \quad .$$

(Approximative) $(1-\alpha)$ **Konfidenzintervalle für die Regressionskoeffizienten**

$$\left[\hat{\beta}_0 - z_{1-\frac{\alpha}{2}}\sqrt{\hat{\sigma}^2_{\hat{\beta}_0}},\ \hat{\beta}_0 + z_{1-\frac{\alpha}{2}}\sqrt{\hat{\sigma}^2_{\hat{\beta}_0}}\right] \quad \text{bzw.} \quad \left[\hat{\beta}_1 - z_{1-\frac{\alpha}{2}}\sqrt{\hat{\sigma}^2_{\hat{\beta}_1}},\ \hat{\beta}_1 + z_{1-\frac{\alpha}{2}}\sqrt{\hat{\sigma}^2_{\hat{\beta}_1}}\right] \quad .$$

12.10 Verteilung des KQ-Schätzers unter erweiterten Annahmen

Definition 12.2. ***Heteroskedastizität und Homoskedastizität***
Im linearen Regressionsmodell wird der Fehlerterm als „homoskedastisch" bezeichnet, wenn die bedingte Varianz von U_i gegeben $X_i = x$ konstant ist, also wenn $V(U_i|X_i = x) = \sigma^2_U$ für $i = 1, \ldots, n$ und jeden Wert x. Andernfalls wird der Fehlerterm als „heteroskedastisch" bezeichnet.

Eigenschaften des KQ-Schätzers bei homoskedastischem Fehlerterm: Gegeben sei ein einfaches lineares Regressionsmodell mit den Annahmen (A1) - (A3). Zusätzlich liegen homoskedastische Fehler und eine große Stichprobe vor. Dann sind die KQ-Schätzer weiterhin erwartungstreu, konsistent und (approximativ) gemeinsam normalverteilt. Die Randverteilungen sind:

$$\hat{\beta}_0 \overset{a}{\sim} N(\beta_0, \sigma^2_{\hat{\beta}_0}) \quad \text{und} \quad \hat{\beta}_1 \overset{a}{\sim} N(\beta_1, \sigma^2_{\hat{\beta}_1})$$

wobei $\sigma^2_{\hat{\beta}_0} = \dfrac{E(X_i^2)\sigma^2_U}{n\sigma^2_X}$ und $\sigma^2_{\hat{\beta}_1} = \dfrac{\sigma^2_U}{n\sigma^2_X}$ mit $\sigma^2_U = V(U_i)$ und $\sigma^2_X = V(X_i)$.

Konsistente Varianzschätzer unter homoskedastischen Fehlertermen: Konsistente Schätzer für $\sigma^2_{\hat{\beta}_0}$ und $\sigma^2_{\hat{\beta}_1}$ sind gegeben durch:

$$\tilde{\sigma}^2_{\hat{\beta}_0} = \frac{\frac{1}{n}\sum_{i=1}^{n}X_i^2\hat{\sigma}^2_U}{\sum_{i=1}^{n}(X_i-\bar{X})^2} \quad \text{bzw.} \quad \tilde{\sigma}^2_{\hat{\beta}_1} = \frac{\hat{\sigma}^2_U}{\sum_{i=1}^{n}(X_i-\bar{X})^2} \quad \text{wobei} \quad \hat{\sigma}^2_U = \frac{1}{n-2}\sum_{i=1}^{n}\hat{U}_i^2 \quad .$$

(Approximative) Tests bezüglich der Regressionskoeffizienten unter homoskedastischen Fehlern

Annahmen: Es liegt ein einfaches lineares Regressionsmodell vor und die Annahmen (A1)-(A3) seien erfüllt. Des Weiteren liegt eine große Stichprobe vor und die Fehlerterme sind homoskedastisch.

Teststatistik und Nullverteilung:

$$Z^*_{\beta_0} = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\tilde{\sigma}^2_{\hat{\beta}_0}}} \overset{H_0}{\sim} N(0,1) \quad \text{bzw.} \quad Z^*_{\beta_1} = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\tilde{\sigma}^2_{\hat{\beta}_1}}} \overset{H_0}{\sim} N(0,1) \quad .$$

Schätzwerte der Teststatistiken:

$$z^*_{\beta_0} = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\tilde{\sigma}^2_{\hat{\beta}_0}}} \quad \text{bzw.} \quad z^*_{\beta_1} = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\tilde{\sigma}^2_{\hat{\beta}_1}}} \quad .$$

12.11 Verteilung des KQ-Schätzers unter den „klassischen Annahmen“

Klassische KQ-Annahmen Gegeben sei das einfache lineare Regressionsmodell:

$$Y_i = \beta_0 + \beta_1 x_i + U_i \quad i = 1, \ldots, n.$$

mit den folgenden Annahmen

(B1) $E(U_i) = 0$ für $i = 1, \cdots, n$

(B2) Die Kovariable x_i ist deterministisch und weist eine positive Streuung auf ($s_x > 0$).

(B3) $V(U_i) = \sigma^2_U$ für $i = 1, \cdots, n$
(homoskedastische Fehlerterme für den Fall deterministischer Regressoren)

(B4) $COV(U_i, U_j) = 0 \quad \forall i \neq j$
(kein Zusammenhang zwischen den Fehlern unterschiedlicher Beobachtungen)

(B5) $U_i \sim N(0, \sigma^2)$ für $i = 1, \cdots, n$.

Tests bzgl. der Regressionskoeffizienten unter den klassischen Annahmen:

Annahmen: Es liegt ein einfaches lineares Regressionsmodell

$$Y_i = \beta_0 + \beta_1 x_i + U_i \quad i = 1, \ldots, n$$

vor und die Annahmen (B1) - (B5) seien erfüllt.

Teststatistik und Nullverteilung: Die KQ-Schätzer $\hat{\beta}_0$ und $\hat{\beta}_1$ sind erwartungstreue und konsistente Schätzer der Regressionskoeffizienten β_0 und β_1. Des Weiteren sind die Schätzstatistiken $\tilde{\sigma}^2_{\hat{\beta}_0}$ und $\tilde{\sigma}^2_{\hat{\beta}_1}$ erwartungstreue und konsistente Schätzer von $\sigma^2_{\hat{\beta}_0}$ und $\sigma^2_{\hat{\beta}_1}$.

$$T^*_{\beta_0} = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\tilde{\sigma}^2_{\hat{\beta}_0}}} \overset{H_0}{\sim} t_{n-2} \quad \text{bzw.} \quad T^*_{\beta_1} = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\tilde{\sigma}^2_{\hat{\beta}_1}}} \overset{H_0}{\sim} t_{n-2} \quad .$$

Schätzwerte der Teststatistiken:

$$t^*_{\beta_0} = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\tilde{\sigma}^2_{\hat{\beta}_0}}} \quad \text{bzw.} \quad t^*_{\beta_1} = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\tilde{\sigma}^2_{\hat{\beta}_1}}} \quad .$$

Konfidenzintervalle Die $(1-\alpha)$ Konfidenzintervalle der Regressionskoeffizienten β_0 und β_1 sind gegeben durch

$$\left[\hat{\beta}_0 - t_{n-2;1-\frac{\alpha}{2}} \sqrt{\tilde{\sigma}^2_{\hat{\beta}_0}}, \quad \hat{\beta}_0 + t_{n-2;1-\frac{\alpha}{2}} \sqrt{\tilde{\sigma}^2_{\hat{\beta}_0}}\right] \quad \text{bzw.} \quad \left[\hat{\beta}_1 - t_{n-2;1-\frac{\alpha}{2}} \sqrt{\tilde{\sigma}^2_{\hat{\beta}_1}}, \quad \hat{\beta}_1 + t_{n-2;1-\frac{\alpha}{2}} \sqrt{\tilde{\sigma}^2_{\hat{\beta}_1}}\right] \quad .$$

13 Lineares Regressionsmodell

13.1 Das lineare Regressionsmodell in Matrixnotation

$$\boldsymbol{x}_i = \begin{pmatrix} 1 \\ x_{i1} \\ \vdots \\ x_{ik} \end{pmatrix} \text{ für } i = 1, \ldots, n \qquad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} \qquad \hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{pmatrix} \qquad \boldsymbol{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \qquad \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

$$\boldsymbol{U} = \begin{pmatrix} U_1 \\ U_2 \\ \vdots \\ U_n \end{pmatrix} \qquad \boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \qquad \hat{\boldsymbol{U}} = \begin{pmatrix} \hat{U}_1 \\ \hat{U}_2 \\ \vdots \\ \hat{U}_n \end{pmatrix} .$$

$$Y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + U_i, \qquad i = 1, \ldots, n.$$

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_1' \\ \boldsymbol{x}_2' \\ \vdots \\ \boldsymbol{x}_n' \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix} .$$

Multiples lineares Regressionsmodell mit deterministischen Regressoren:

$$\boldsymbol{Y} = \boldsymbol{x}\boldsymbol{\beta} + \boldsymbol{U} \quad .$$

Annahmen:

(C1) $E(\boldsymbol{U}) = \boldsymbol{0}$

(C2) Designmatrix $\boldsymbol{x}$ hat vollen Spaltenrang und ist deterministisch.

(C3) $COV(\boldsymbol{U}) = \sigma^2 \boldsymbol{I}$

(C4) $\boldsymbol{U} \sim \boldsymbol{N}(\boldsymbol{0}, \sigma_U^2 \boldsymbol{I})$

13.2 Parameterschätzung mit der KQ-Methode

Zielfunktion:

$$QS(\boldsymbol{\beta}) = \sum_{i=1}^{n} (y_i - \boldsymbol{x}_i'\boldsymbol{\beta})^2 \quad .$$

Der $(k+1) \times 1$-Vektor der KQ-Schätzer für $\beta_0, \ldots, \beta_k$:

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{x}'\boldsymbol{x})^{-1}\boldsymbol{x}'\boldsymbol{y}$$

Streuung der Modellfehler Angepassten Werte:

$$\hat{\boldsymbol{y}} = \boldsymbol{x}\hat{\boldsymbol{\beta}}$$

Residuen:

$$\hat{\boldsymbol{u}} = \boldsymbol{y} - \hat{\boldsymbol{y}} \quad .$$

13.3 Parameterinterpretation

Der marginale Effekt der j-ten Kovariable ($j = 1, \ldots k$) ist gegeben durch

$$\frac{\partial f(\boldsymbol{x_i})}{\partial x_{ij}} = \frac{\partial(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik})}{\partial x_{ij}} = \beta_j \quad .$$

13.6 Statistische Eigenschaften der geschätzten Parameter

Fehlerstreuungsschätzer

$$\hat{\sigma}_U^2 = \frac{1}{n-(k+1)} \cdot \sum_{i=1}^{n} \hat{U}_i^2 = \frac{1}{n-(k+1)} \cdot \hat{\boldsymbol{U}}'\hat{\boldsymbol{U}} = \frac{1}{n-(k+1)} \cdot RSS \quad .$$

Satz 13.2. ***Statistische Eigenschaften des Fehlerstreuungsschätzers***
Der Streuungschätzer $\hat{S}_U^2$ hat die folgenden Eigenschaften:

(a) $E(\hat{\sigma}_U^2) = \sigma_U^2$

(b) $\frac{n-(k+1)}{\sigma^2}\hat{\sigma}_U^2 \sim \chi^2_{n-(k+1)}$

Satz 13.4. ***Statistische Eigenschaften des KQ-Schätzers***
Unter den Annahmen (B1) - (B4) gilt für den KQ-Schätzer $\hat{\boldsymbol{\beta}}$ *:*

(a) $E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$,

(b) $COV(\hat{\boldsymbol{\beta}}) = \sigma_U^2(\boldsymbol{x}'\boldsymbol{x})^{-1}$.

(c) $\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \sigma_U^2(\boldsymbol{x}'\boldsymbol{x})^{-1})$.

Die Tupel der Matrix $(\boldsymbol{x}'\boldsymbol{x})^{-1}$ werden im folgenden durch $c_{ij}(i,j = 0,\ldots k)$ gekennzeichnet, wobei i der Zeilenindex und j der Spaltenindex ist:

$$(\boldsymbol{x}'\boldsymbol{x})^{-1} = \begin{pmatrix} c_{00} & c_{01} & \cdots & c_{0k} \\ c_{10} & c_{11} & \cdots & c_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ c_{k0} & c_{k1} & \cdots & c_{kk} \end{pmatrix} \quad .$$

Die geschätzte Kovarianzmatrix der Regressionskoeffizienten:

$$\widehat{COV(\hat{\boldsymbol{\beta}})} = \begin{pmatrix} \hat{\sigma}_U^2 c_{00} & \hat{\sigma}_U^2 c_{01} & \cdots & \hat{\sigma}_U^2 c_{0k} \\ \hat{\sigma}_U^2 c_{10} & \hat{\sigma}_U^2 c_{11} & \cdots & \hat{\sigma}_U^2 c_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\sigma}_U^2 c_{k0} & \hat{\sigma}_U^2 c_{k1} & \cdots & \hat{\sigma}_U^2 c_{kk} \end{pmatrix} = \hat{\sigma}_U^2(\boldsymbol{x}'\boldsymbol{x})^{-1} \quad .$$

13.7 Konfidenzintervalle für einzelne Regressionskoeffizienten

$(1-\alpha)$ Konfidenzintervall:

$$KI_{\beta_j;1-\alpha} = [G_{\underline{\beta}_j}; G_{\overline{\beta}_j}] \quad .$$

Geschätzte $(1-\alpha)$ Intervallgrenzen:

$$g_{\underline{\beta}_j} = \hat{\beta}_j - \sqrt{\hat{\sigma}_U^2 c_{jj}} \cdot t_{n-(k+1);1-\frac{\alpha}{2}} \quad \text{und} \quad g_{\overline{\beta}_j} = \hat{\beta}_j + \sqrt{\hat{\sigma}_U^2 c_{jj}} \cdot t_{n-(k+1);1-\frac{\alpha}{2}} \quad .$$

13.8 Tests für einzelne Regressionskoeffizienten

Annahmen:

- Es liegt ein lineares multiples Regressionsmodell vor und die Annahmen $(C1)$ bis $(C4)$ sind erfüllt.

Teststatistik:

$$T^* = \frac{\hat{\beta}_j - \beta_{j,0}}{\sqrt{\hat{\sigma}_U^2 c_{jj}}} \overset{H_0}{\sim} t_{n-(k+1)} \quad .$$

Schätzwert der Teststatistik T^*:

$$t^* = \frac{\hat{\beta}_j - \beta_{j,0}}{\sqrt{\hat{\sigma}_U^2 c_{jj}}} \quad .$$

B Tabellen

B.1 Quantile der Standardnormalverteilung

α	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%
+0,0%	$-\infty$	-1,2816	-0,8416	-0,5244	-0,2533	0,0000	0,2533	0,5244	0,8416	1,2816
+0,5%	-2,5758	-1,2536	-0,8239	-0,5101	-0,2404	0,0125	0,2663	0,5388	0,8596	1,3106
+1,0%	-2,3263	-1,2265	-0,8064	-0,4959	-0,2275	0,0251	0,2793	0,5534	0,8779	1,3408
+1,5%	-2,1701	-1,2004	-0,7892	-0,4817	-0,2147	0,0376	0,2924	0,5681	0,8965	1,3722
+2,0%	-2,0537	-1,1750	-0,7722	-0,4677	-0,2019	0,0502	0,3055	0,5828	0,9154	1,4051
+2,5%	-1,9600	-1,1503	-0,7554	-0,4538	-0,1891	0,0627	0,3186	0,5978	0,9346	1,4395
+3,0%	-1,8808	-1,1264	-0,7388	-0,4399	-0,1764	0,0753	0,3319	0,6128	0,9542	1,4758
+3,5%	-1,8119	-1,1031	-0,7225	-0,4261	-0,1637	0,0878	0,3451	0,6280	0,9741	1,5141
+4,0%	-1,7507	-1,0803	-0,7063	-0,4125	-0,1510	0,1004	0,3585	0,6433	0,9945	1,5548
+4,5%	-1,6954	-1,0581	-0,6903	-0,3989	-0,1383	0,1130	0,3719	0,6588	1,0152	1,5982
+5,0%	-1,6449	-1,0364	-0,6745	-0,3853	-0,1257	0,1257	0,3853	0,6745	1,0364	1,6449
+5,5%	-1,5982	-1,0152	-0,6588	-0,3719	-0,1130	0,1383	0,3989	0,6903	1,0581	1,6954
+6,0%	-1,5548	-0,9945	-0,6433	-0,3585	-0,1004	0,1510	0,4125	0,7063	1,0803	1,7507
+6,5%	-1,5141	-0,9741	-0,6280	-0,3451	-0,0878	0,1637	0,4261	0,7225	1,1031	1,8119
+7,0%	-1,4758	-0,9542	-0,6128	-0,3319	-0,0753	0,1764	0,4399	0,7388	1,1264	1,8808
+7,5%	-1,4395	-0,9346	-0,5978	-0,3186	-0,0627	0,1891	0,4538	0,7554	1,1503	1,9600
+8,0%	-1,4051	-0,9154	-0,5828	-0,3055	-0,0502	0,2019	0,4677	0,7722	1,1750	2,0537
+8,5%	-1,3722	-0,8965	-0,5681	-0,2924	-0,0376	0,2147	0,4817	0,7892	1,2004	2,1701
+9,0%	-1,3408	-0,8779	-0,5534	-0,2793	-0,0251	0,2275	0,4959	0,8064	1,2265	2,3263
+9,5%	-1,3106	-0,8596	-0,5388	-0,2663	-0,0125	0,2404	0,5101	0,8239	1,2536	2,5758

Feiner abgestufte Quantile am Verteilungsrand

α	0,05%	0,1%	0,25%	99,75%	99,90%	99,95%
z_α	-3,2905	-3,0902	-2,8070	2,8070	3,0902	3,2905

Umkehrungsregel: $z_\alpha = -z_{1-\alpha}$

B.2 Verteilungsfunktion der Standardnormalverteilung - negative Argumente

z	-0,09	-0,08	-0,07	-0,06	-0,05	-0,04	-0,03	-0,02	-0,01	0,00
-3,0	0,00100	0,00104	0,00107	0,00111	0,00114	0,00118	0,00122	0,00126	0,00131	0,00135
-2,9	0,00139	0,00144	0,00149	0,00154	0,00159	0,00164	0,00169	0,00175	0,00181	0,00187
-2,8	0,00193	0,00199	0,00205	0,00212	0,00219	0,00226	0,00233	0,00240	0,00248	0,00256
-2,7	0,00264	0,00272	0,00280	0,00289	0,00298	0,00307	0,00317	0,00326	0,00336	0,00347
-2,6	0,00357	0,00368	0,00379	0,00391	0,00402	0,00415	0,00427	0,00440	0,00453	0,00466
-2,5	0,00480	0,00494	0,00508	0,00523	0,00539	0,00554	0,00570	0,00587	0,00604	0,00621
-2,4	0,00639	0,00657	0,00676	0,00695	0,00714	0,00734	0,00755	0,00776	0,00798	0,00820
-2,3	0,00842	0,00866	0,00889	0,00914	0,00939	0,00964	0,00990	0,01017	0,01044	0,01072
-2,2	0,01101	0,01130	0,01160	0,01191	0,01222	0,01255	0,01287	0,01321	0,01355	0,01390
-2,1	0,01426	0,01463	0,01500	0,01539	0,01578	0,01618	0,01659	0,01700	0,01743	0,01786
-2,0	0,01831	0,01876	0,01923	0,01970	0,02018	0,02068	0,02118	0,02169	0,02222	0,02275
-1,9	0,02330	0,02385	0,02442	0,02500	0,02559	0,02619	0,02680	0,02743	0,02807	0,02872
-1,8	0,02938	0,03005	0,03074	0,03144	0,03216	0,03288	0,03362	0,03438	0,03515	0,03593
-1,7	0,03673	0,03754	0,03836	0,03920	0,04006	0,04093	0,04182	0,04272	0,04363	0,04457
-1,6	0,04551	0,04648	0,04746	0,04846	0,04947	0,05050	0,05155	0,05262	0,05370	0,05480
-1,5	0,05592	0,05705	0,05821	0,05938	0,06057	0,06178	0,06301	0,06426	0,06552	0,06681
-1,4	0,06811	0,06944	0,07078	0,07215	0,07353	0,07493	0,07636	0,07780	0,07927	0,08076
-1,3	0,08226	0,08379	0,08534	0,08691	0,08851	0,09012	0,09176	0,09342	0,09510	0,09680
-1,2	0,09853	0,10027	0,10204	0,10383	0,10565	0,10749	0,10935	0,11123	0,11314	0,11507
-1,1	0,11702	0,11900	0,12100	0,12302	0,12507	0,12714	0,12924	0,13136	0,13350	0,13567
-1,0	0,13786	0,14007	0,14231	0,14457	0,14686	0,14917	0,15151	0,15386	0,15625	0,15866
-0,9	0,16109	0,16354	0,16602	0,16853	0,17106	0,17361	0,17619	0,17879	0,18141	0,18406
-0,8	0,18673	0,18943	0,19215	0,19489	0,19766	0,20045	0,20327	0,20611	0,20897	0,21186
-0,7	0,21476	0,21770	0,22065	0,22363	0,22663	0,22965	0,23270	0,23576	0,23885	0,24196
-0,6	0,24510	0,24825	0,25143	0,25463	0,25785	0,26109	0,26435	0,26763	0,27093	0,27425
-0,5	0,27760	0,28096	0,28434	0,28774	0,29116	0,29460	0,29806	0,30153	0,30503	0,30854
-0,4	0,31207	0,31561	0,31918	0,32276	0,32636	0,32997	0,33360	0,33724	0,34090	0,34458
-0,3	0,34827	0,35197	0,35569	0,35942	0,36317	0,36693	0,37070	0,37448	0,37828	0,38209
-0,2	0,38591	0,38974	0,39358	0,39743	0,40129	0,40517	0,40905	0,41294	0,41683	0,42074
-0,1	0,42465	0,42858	0,43251	0,43644	0,44038	0,44433	0,44828	0,45224	0,45620	0,46017
0,0	0,46414	0,46812	0,47210	0,47608	0,48006	0,48405	0,48803	0,49202	0,49601	0,50000

Erweiterung der Tafel: $\Phi(-z) = 1 - \Phi(z)$

B.3 Verteilungsfunktion der Standardnormalverteilung - positive Argumente

z	+ 0,00	+ 0,01	+ 0,02	+ 0,03	+ 0,04	+ 0,05	+ 0,06	+ 0,07	+ 0,08	+ 0,09
0,0	0,50000	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,52790	0,53188	0,53586
0,1	0,53983	0,54380	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0,2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0,3	0,61791	0,62172	0,62552	0,62930	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0,4	0,65542	0,65910	0,66276	0,66640	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0,5	0,69146	0,69497	0,69847	0,70194	0,70540	0,70884	0,71226	0,71566	0,71904	0,72240
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,75490
0,7	0,75804	0,76115	0,76424	0,76730	0,77035	0,77337	0,77637	0,77935	0,78230	0,78524
0,8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0,9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1,0	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,1	0,86433	0,86650	0,86864	0,87076	0,87286	0,87493	0,87698	0,87900	0,88100	0,88298
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1,3	0,90320	0,90490	0,90658	0,90824	0,90988	0,91149	0,91309	0,91466	0,91621	0,91774
1,4	0,91924	0,92073	0,92220	0,92364	0,92507	0,92647	0,92785	0,92922	0,93056	0,93189
1,5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	0,94520	0,94630	0,94738	0,94845	0,94950	0,95053	0,95154	0,95254	0,95352	0,95449
1,7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,96080	0,96164	0,96246	0,96327
1,8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,9	0,97128	0,97193	0,97257	0,97320	0,97381	0,97441	0,97500	0,97558	0,97615	0,97670
2,0	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,98030	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,98300	0,98341	0,98382	0,98422	0,98461	0,98500	0,98537	0,98574
2,2	0,98610	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,98840	0,98870	0,98899
2,3	0,98928	0,98956	0,98983	0,99010	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,4	0,99180	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,5	0,99379	0,99396	0,99413	0,99430	0,99446	0,99461	0,99477	0,99492	0,99506	0,99520
2,6	0,99534	0,99547	0,99560	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2,7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,99720	0,99728	0,99736
2,8	0,99744	0,99752	0,99760	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,9	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861
3,0	0,99865	0,99869	0,99874	0,99878	0,99882	0,99886	0,99889	0,99893	0,99896	0,99900

Erweiterung der Tafel: $\Phi(-z) = 1 - \Phi(z)$

B.4 Verteilungsfunktion der Poissonverteilung

Die Verteilungsfunktion der Poissonverteilung ist gegeben mit:

$$F_{P_\lambda}(x) = P(X \leq x) = \sum_{k=0}^{x} P(X = k) = \sum_{k=0}^{x} e^{-\lambda} \frac{\lambda^k}{k!}$$

$x \setminus \lambda$	0,05	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0	
0	0,95123	0,90484	0,81873	0,74082	0,67032	0,60653	0,54881	0,49659	0,44933	0,40657	0,36788	0
1	0,99879	0,99532	0,98248	0,96306	0,93845	0,90980	0,87810	0,84420	0,80879	0,77248	0,73576	1
2	0,99998	0,99985	0,99885	0,99640	0,99207	0,98561	0,97688	0,96586	0,95258	0,93714	0,91970	2
3	1,00000	1,00000	0,99994	0,99973	0,99922	0,99825	0,99664	0,99425	0,99092	0,98654	0,98101	3
4			1,00000	0,99998	0,99994	0,99983	0,99961	0,99921	0,99859	0,99766	0,99634	4
5				1,00000	1,00000	0,99999	0,99996	0,99991	0,99982	0,99966	0,99941	5
6						1,00000	1,00000	0,99999	0,99998	0,99996	0,99992	6
7								1,00000	1,00000	1,00000	0,99999	7
8											1,00000	8

$x \setminus \lambda$	1,5	2,0	3,0	4,0	5,0	6,0	7,0	8,0	9,0	10	15	
0	0,22313	0,13534	0,04979	0,01832	0,00674	0,00248	0,00091	0,00034	0,00012	0,00005	0,00000	0
1	0,55783	0,40601	0,19915	0,09158	0,04043	0,01735	0,00730	0,00302	0,00123	0,00050	0,00000	1
2	0,80885	0,67668	0,42319	0,23810	0,12465	0,06197	0,02964	0,01375	0,00623	0,00277	0,00004	2
3	0,93436	0,85712	0,64723	0,43347	0,26503	0,15120	0,08177	0,04238	0,02123	0,01034	0,00021	3
4	0,98142	0,94735	0,81526	0,62884	0,44049	0,28506	0,17299	0,09963	0,05496	0,02925	0,00086	4
5	0,99554	0,98344	0,91608	0,78513	0,61596	0,44568	0,30071	0,19124	0,11569	0,06709	0,00279	5
6	0,99907	0,99547	0,96649	0,88933	0,76218	0,60630	0,44971	0,31337	0,20678	0,13014	0,00763	6
7	0,99983	0,99890	0,98810	0,94887	0,86663	0,74398	0,59871	0,45296	0,32390	0,22022	0,01800	7
8	0,99997	0,99976	0,99620	0,97864	0,93191	0,84724	0,72909	0,59255	0,45565	0,33282	0,03745	8
9	1,00000	0,99995	0,99890	0,99187	0,96817	0,91608	0,83050	0,71662	0,58741	0,45793	0,06985	9
10		0,99999	0,99971	0,99716	0,98630	0,95738	0,90148	0,81589	0,70599	0,58304	0,11846	10
11		1,00000	0,99993	0,99908	0,99455	0,97991	0,94665	0,88808	0,80301	0,69678	0,18475	11
12			0,99998	0,99973	0,99798	0,99117	0,97300	0,93620	0,87577	0,79156	0,26761	12
13			1,00000	0,99992	0,99930	0,99637	0,98719	0,96582	0,92615	0,86446	0,36322	13
14				0,99998	0,99977	0,99860	0,99428	0,98274	0,95853	0,91654	0,46565	14
15				1,00000	0,99993	0,99949	0,99759	0,99177	0,97796	0,95126	0,56809	15
16					0,99998	0,99983	0,99904	0,99628	0,98889	0,97296	0,66412	16
17					0,99999	0,99994	0,99964	0,99841	0,99468	0,98572	0,74886	17
18					1,00000	0,99998	0,99987	0,99935	0,99757	0,99281	0,81947	18
19						0,99999	0,99996	0,99975	0,99894	0,99655	0,87522	19
20						1,00000	0,99999	0,99991	0,99956	0,99841	0,91703	20
21							1,00000	0,99997	0,99983	0,99930	0,94689	21
22								0,99999	0,99993	0,99970	0,96726	22
23								1,00000	0,99998	0,99988	0,98054	23
24									0,99999	0,99995	0,98884	24
25									1,00000	0,99998	0,99382	25
26										0,99999	0,99669	26
27										1,00000	0,99828	27
28											0,99914	28
29											0,99958	29
30											0,99980	30
31											0,99991	31
32											0,99996	32
33											0,99998	33
34											0,99999	34
35											1,00000	35

Ablesebeispiel: $F_{P_{1,5}}(4) = 0{,}98142$

B.5 Quantile der t-Verteilung

df \ α	0,5%	1%	2,5%	5%	10%	90%	95%	97,5%	99%	99,5%
1	-63,6567	-31,8205	-12,7062	-6,3138	-3,0777	3,0777	6,3138	12,7062	31,8205	63,6567
2	-9,9248	-6,9646	-4,3027	-2,9200	-1,8856	1,8856	2,9200	4,3027	6,9646	9,9248
3	-5,8409	-4,5407	-3,1824	-2,3534	-1,6377	1,6377	2,3534	3,1824	4,5407	5,8409
4	-4,6041	-3,7469	-2,7764	-2,1318	-1,5332	1,5332	2,1318	2,7764	3,7469	4,6041
5	-4,0321	-3,3649	-2,5706	-2,0150	-1,4759	1,4759	2,0150	2,5706	3,3649	4,0321
6	-3,7074	-3,1427	-2,4469	-1,9432	-1,4398	1,4398	1,9432	2,4469	3,1427	3,7074
7	-3,4995	-2,9980	-2,3646	-1,8946	-1,4149	1,4149	1,8946	2,3646	2,9980	3,4995
8	-3,3554	-2,8965	-2,3060	-1,8595	-1,3968	1,3968	1,8595	2,3060	2,8965	3,3554
9	-3,2498	-2,8214	-2,2622	-1,8331	-1,3830	1,3830	1,8331	2,2622	2,8214	3,2498
10	-3,1693	-2,7638	-2,2281	-1,8125	-1,3722	1,3722	1,8125	2,2281	2,7638	3,1693
11	-3,1058	-2,7181	-2,2010	-1,7959	-1,3634	1,3634	1,7959	2,2010	2,7181	3,1058
12	-3,0545	-2,6810	-2,1788	-1,7823	-1,3562	1,3562	1,7823	2,1788	2,6810	3,0545
13	-3,0123	-2,6503	-2,1604	-1,7709	-1,3502	1,3502	1,7709	2,1604	2,6503	3,0123
14	-2,9768	-2,6245	-2,1448	-1,7613	-1,3450	1,3450	1,7613	2,1448	2,6245	2,9768
15	-2,9467	-2,6025	-2,1314	-1,7531	-1,3406	1,3406	1,7531	2,1314	2,6025	2,9467
16	-2,9208	-2,5835	-2,1199	-1,7459	-1,3368	1,3368	1,7459	2,1199	2,5835	2,9208
17	-2,8982	-2,5669	-2,1098	-1,7396	-1,3334	1,3334	1,7396	2,1098	2,5669	2,8982
18	-2,8784	-2,5524	-2,1009	-1,7341	-1,3304	1,3304	1,7341	2,1009	2,5524	2,8784
19	-2,8609	-2,5395	-2,0930	-1,7291	-1,3277	1,3277	1,7291	2,0930	2,5395	2,8609
20	-2,8453	-2,5280	-2,0860	-1,7247	-1,3253	1,3253	1,7247	2,0860	2,5280	2,8453
21	-2,8314	-2,5176	-2,0796	-1,7207	-1,3232	1,3232	1,7207	2,0796	2,5176	2,8314
22	-2,8188	-2,5083	-2,0739	-1,7171	-1,3212	1,3212	1,7171	2,0739	2,5083	2,8188
23	-2,8073	-2,4999	-2,0687	-1,7139	-1,3195	1,3195	1,7139	2,0687	2,4999	2,8073
24	-2,7969	-2,4922	-2,0639	-1,7109	-1,3178	1,3178	1,7109	2,0639	2,4922	2,7969
25	-2,7874	-2,4851	-2,0595	-1,7081	-1,3163	1,3163	1,7081	2,0595	2,4851	2,7874
26	-2,7787	-2,4786	-2,0555	-1,7056	-1,3150	1,3150	1,7056	2,0555	2,4786	2,7787
27	-2,7707	-2,4727	-2,0518	-1,7033	-1,3137	1,3137	1,7033	2,0518	2,4727	2,7707
28	-2,7633	-2,4671	-2,0484	-1,7011	-1,3125	1,3125	1,7011	2,0484	2,4671	2,7633
29	-2,7564	-2,4620	-2,0452	-1,6991	-1,3114	1,3114	1,6991	2,0452	2,4620	2,7564
30	-2,7500	-2,4573	-2,0423	-1,6973	-1,3104	1,3104	1,6973	2,0423	2,4573	2,7500
40	-2,7045	-2,4233	-2,0211	-1,6839	-1,3031	1,3031	1,6839	2,0211	2,4233	2,7045
50	-2,6778	-2,4033	-2,0086	-1,6759	-1,2987	1,2987	1,6759	2,0086	2,4033	2,6778
60	-2,6603	-2,3901	-2,0003	-1,6706	-1,2958	1,2958	1,6706	2,0003	2,3901	2,6603
70	-2,6479	-2,3808	-1,9944	-1,6669	-1,2938	1,2938	1,6669	1,9944	2,3808	2,6479
80	-2,6387	-2,3739	-1,9901	-1,6641	-1,2922	1,2922	1,6641	1,9901	2,3739	2,6387
90	-2,6316	-2,3685	-1,9867	-1,6620	-1,2910	1,2910	1,6620	1,9867	2,3685	2,6316
100	-2,6259	-2,3642	-1,9840	-1,6602	-1,2901	1,2901	1,6602	1,9840	2,3642	2,6259

Umkehrungsregel: $t_{df;\alpha} = -t_{df;1-\alpha}$

Ablesebeispiel: $t_{20;0,95} = 1{,}7247$

B.6 Quantile der χ^2-Verteilung

df \ α	0,5%	1%	2,5%	5%	10%	90%	95%	97,5%	99%	99,5%
1	0,0000	0,0002	0,0010	0,0039	0,0158	2,7055	3,8415	5,0239	6,6349	7,8794
2	0,0100	0,0201	0,0506	0,1026	0,2107	4,6052	5,9915	7,3778	9,2103	10,5966
3	0,0717	0,1148	0,2158	0,3518	0,5844	6,2514	7,8147	9,3484	11,3449	12,8382
4	0,2070	0,2971	0,4844	0,7107	1,0636	7,7794	9,4877	11,1433	13,2767	14,8603
5	0,4117	0,5543	0,8312	1,1455	1,6103	9,2364	11,0705	12,8325	15,0863	16,7496
6	0,6757	0,8721	1,2373	1,6354	2,2041	10,6446	12,5916	14,4494	16,8119	18,5476
7	0,9893	1,2390	1,6899	2,1673	2,8331	12,0170	14,0671	16,0128	18,4753	20,2777
8	1,3444	1,6465	2,1797	2,7326	3,4895	13,3616	15,5073	17,5345	20,0902	21,9550
9	1,7349	2,0879	2,7004	3,3251	4,1682	14,6837	16,9190	19,0228	21,6660	23,5894
10	2,1559	2,5582	3,2470	3,9403	4,8652	15,9872	18,3070	20,4832	23,2093	25,1882
11	2,6032	3,0535	3,8157	4,5748	5,5778	17,2750	19,6751	21,9200	24,7250	26,7568
12	3,0738	3,5706	4,4038	5,2260	6,3038	18,5493	21,0261	23,3367	26,2170	28,2995
13	3,5650	4,1069	5,0088	5,8919	7,0415	19,8119	22,3620	24,7356	27,6882	29,8195
14	4,0747	4,6604	5,6287	6,5706	7,7895	21,0641	23,6848	26,1189	29,1412	31,3193
15	4,6009	5,2293	6,2621	7,2609	8,5468	22,3071	24,9958	27,4884	30,5779	32,8013
16	5,1422	5,8122	6,9077	7,9616	9,3122	23,5418	26,2962	28,8454	31,9999	34,2672
17	5,6972	6,4078	7,5642	8,6718	10,0852	24,7690	27,5871	30,1910	33,4087	35,7185
18	6,2648	7,0149	8,2307	9,3905	10,8649	25,9894	28,8693	31,5264	34,8053	37,1565
19	6,8440	7,6327	8,9065	10,1170	11,6509	27,2036	30,1435	32,8523	36,1909	38,5823
20	7,4338	8,2604	9,5908	10,8508	12,4426	28,4120	31,4104	34,1696	37,5662	39,9968
21	8,0337	8,8972	10,2829	11,5913	13,2396	29,6151	32,6706	35,4789	38,9322	41,4011
22	8,6427	9,5425	10,9823	12,3380	14,0415	30,8133	33,9244	36,7807	40,2894	42,7957
23	9,2604	10,1957	11,6886	13,0905	14,8480	32,0069	35,1725	38,0756	41,6384	44,1813
24	9,8862	10,8564	12,4012	13,8484	15,6587	33,1962	36,4150	39,3641	42,9798	45,5585
25	10,5197	11,5240	13,1197	14,6114	16,4734	34,3816	37,6525	40,6465	44,3141	46,9279
26	11,1602	12,1981	13,8439	15,3792	17,2919	35,5632	38,8851	41,9232	45,6417	48,2899
27	11,8076	12,8785	14,5734	16,1514	18,1139	36,7412	40,1133	43,1945	46,9629	49,6449
28	12,4613	13,5647	15,3079	16,9279	18,9392	37,9159	41,3371	44,4608	48,2782	50,9934
29	13,1211	14,2565	16,0471	17,7084	19,7677	39,0875	42,5570	45,7223	49,5879	52,3356
30	13,7867	14,9535	16,7908	18,4927	20,5992	40,2560	43,7730	46,9792	50,8922	53,6720
40	20,7065	22,1643	24,4330	26,5093	29,0505	51,8051	55,7585	59,3417	63,6907	66,7660
50	27,9907	29,7067	32,3574	34,7643	37,6886	63,1671	67,5048	71,4202	76,1539	79,4900
60	35,5345	37,4849	40,4817	43,1880	46,4589	74,3970	79,0819	83,2977	88,3794	91,9517
70	43,2752	45,4417	48,7576	51,7393	55,3289	85,5270	90,5312	95,0232	100,4252	104,2149
80	51,1719	53,5401	57,1532	60,3915	64,2778	96,5782	101,8795	106,6286	112,3288	116,3211
90	59,1963	61,7541	65,6466	69,1260	73,2911	107,5650	113,1453	118,1359	124,1163	128,2989
100	67,3276	70,0649	74,2219	77,9295	82,3581	118,4980	124,3421	129,5612	135,8067	140,1695

Ablesebeispiel: $\chi^2_{21;0,9} = 29,6151$