

Statistik und
Wahrscheinlichkeitsrechnung

Formelsammlung

Prof. Dr. Roland Schuhr

Institut für Empirische Wirtschaftsforschung
Bereich Statistik
Wirtschaftswissenschaftliche Fakultät der
Universität Leipzig

Stand: 26.05.2012

Inhalt

1	Deskription univariater Datensätze	2
1.1	Eindimensionale Häufigkeitsverteilungen.....	2
1.2	Lagemaße.....	4
1.3	Streuungsmaße.....	5
1.4	Momente und Schiefemaße	7
1.5	Die Bestandteile des Box-Plots.....	7
1.6	Konzentrationsmessung	7
2	Deskription bivariater Datensätze	10
2.1	Bivariate Häufigkeitsverteilungen und Randverteilungen.....	10
2.2	Bedingte Verteilungen	11
2.3	Maßzahlen für bivariate Verteilungen (Korrelationsrechnung).....	11
3	Indexrechnung	14
4	Elementare Wahrscheinlichkeitsrechnung	17
4.1	Zufallsvorgänge und Wahrscheinlichkeitsräume.....	17
4.2	Wahrscheinlichkeitskonzeptionen.....	18
4.3	Bedingte Wahrscheinlichkeit und Unabhängigkeit.....	19
4.4	Kombinatorik.....	20
5	Eindimensionale Zufallsvariablen und deren Wahrscheinlichkeitsverteilungen	21
6	Mehrdimensionale Zufallsvariablen und deren Wahrscheinlichkeitsverteilungen.	24
7	Spezielle Wahrscheinlichkeitsverteilungsmodelle	26
7.1	Diskrete Verteilungsmodelle.....	26
7.2	Stetige Verteilungsmodelle.....	27
7.3	Stichprobenverteilungen	30
8	Gesetze der großen Zahlen, Punktschätzung	31
9	Confidence Intervals and Tests of Hypotheses	33
10	Regressionsanalyse	40
10.1	Lineare Einfachregression	40
10.2	Lineare Mehrfachregression	42
11	Verteilungsfunktionen und Quantile ausgewählter Verteilungen	44

1 Deskription univariater Datensätze

Notation

Merkmalsvariable	X	
Realisationsmöglichkeiten	\tilde{x}_i	$(i = 1, \dots, m)$
Statistische Reihe	x_ν	$(\nu = 1, \dots, n)$
Geordnete statistische Reihe	$x_{(\nu^*)}$, d.h. $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(\nu^*)} \leq \dots \leq x_{(n)}$	$(\nu^* = 1, \dots, n)$
Absolute Häufigkeit	n_i	$(i = 1, \dots, m)$
Absolute Häufigkeitsverteilung	(\tilde{x}_i, n_i)	$(i = 1, \dots, m)$
Relative Häufigkeit	h_i	$(i = 1, \dots, m)$
Relative Häufigkeitsverteilung	(\tilde{x}_i, h_i)	$(i = 1, \dots, m)$
Klassen	K_j	$(j = 1, \dots, m)$
Median	x_{med}	
Arithmetischer Mittelwert	\bar{x}	
Standardabweichung	\tilde{s}_X	
Quartile	$x_{[0.25]}, x_{[0.5]}, x_{[0.75]}$	

1.1 Eindimensionale Häufigkeitsverteilungen

1.1.1 Häufigkeitsverteilungen bei diskreten Variablen (unklassierte Häufigkeitsverteilungen)

Absolute Häufigkeiten

Es sei X eine diskrete Variable mit den Realisationsmöglichkeiten \tilde{x}_i ($i = 1, \dots, m$). Die Anzahl der mit \tilde{x}_i übereinstimmenden Messwerte aus der statistischen Reihe x_ν ($\nu = 1, \dots, n$) heißt die *absolute Häufigkeit von \tilde{x}_i* ; geschrieben

$$n(X = \tilde{x}_i) \quad \text{oder kurz} \quad n_i .$$

Relative Häufigkeiten

$$h(X = \tilde{x}_i) = \frac{n(X = \tilde{x}_i)}{n} \quad \text{oder kurz} \quad h_i = \frac{n_i}{n}$$

Kumulierte Häufigkeiten

Es sei X eine diskrete Variable mit den *geordneten* Realisationsmöglichkeiten $\tilde{x}_1 < \dots < \tilde{x}_i < \dots < \tilde{x}_m$. Man bezeichnet

$$N_i = n_1 + n_2 + \dots + n_i = \sum_{u=1}^i n_u \quad (i = 1, \dots, m)$$

als *kumulierte absolute Häufigkeiten* und

$$H_i = h_1 + h_2 + \dots + h_i = \sum_{u=1}^i h_u \quad (i = 1, \dots, m)$$

als *kumulierte relative Häufigkeiten*.

1.1.2 Klassierte Häufigkeitsverteilungen bei stetigen oder quasistetigen Variablen

Absolute Klassenhäufigkeiten

Die **Anzahl der Messwerte** x_v von X , die in die Klasse K_j fallen, heißt *absolute Häufigkeit der Klasse K_j* ; geschrieben

$$n(X \in K_j) \quad \text{oder kurz} \quad n_j.$$

Die Daten, die in die j -te Klasse K_j fallen, werden im Folgenden auch x_{vj} ($v = 1, \dots, n_j$) geschrieben.

Relative Klassenhäufigkeiten

$$h(X \in K_j) = \frac{n(X \in K_j)}{n} \quad \text{oder kurz} \quad h_j = \frac{n_j}{n}$$

Kumulierte absolute und relative Klassenhäufigkeiten

$$N_j = \sum_{u=1}^j n_u, \quad H_j = \sum_{u=1}^j h_u \quad (j = 1, \dots, m)$$

Halboffene Intervalle und Klassenbreite

Im Falle metrisch skalierte Daten wählt man als Klassen in der Regel *halboffene Intervalle*

$$K_j = [k_{j-1}, k_j) \quad \text{oder} \quad K_j = (k_{j-1}, k_j] \quad (j = 1, \dots, m),$$

wobei $k_0 < k_1 < \dots < k_j < \dots < k_m$ die Anfangs- bzw. Endpunkte der Intervalle sind.

Die Differenz von End- und Anfangspunkt einer Klasse

$$b_j = k_j - k_{j-1}$$

bezeichnet man dann als *Klassenbreite* der j -ten Klasse.

Normierte Klassenhäufigkeiten

$$n_j^* = \frac{n_j}{b_j} \quad \text{bzw.} \quad h_j^* = \frac{h_j}{b_j} \quad (j = 1, \dots, m)$$

Häufigkeitsdichtefunktion

$$d(x) = \begin{cases} h_j^* & \text{falls } x \in K_j \quad (j = 1, \dots, m) \\ 0 & \text{sonst} \end{cases}$$

1.1.3 Empirische Verteilungsfunktion

$$v(x) = h(X \leq x) = \frac{n(X \leq x)}{n} \quad \text{für alle } x \in \mathbb{R}$$

1.2 Lagemaße

Modus unklassierter Daten

x_{mod} heißt *Modus* der Häufigkeitsverteilung, falls

$$n(X = x_{mod}) = \max_i n_i \quad \text{bzw.} \quad h(X = x_{mod}) = \max_i h_i .$$

Modalklasse

K_{mod} heißt *Modalklasse*, falls

$$n^*(X \in K_{mod}) = \max_j n_j^* \quad \text{bzw.} \quad h^*(X \in K_{mod}) = \max_j h_j^* .$$

Modus klassierter Daten

$$x_{mod} = \frac{1}{2}(k_{mod-1} + k_{mod})$$

Quantile (p-Quantil)

Gegeben sei die statistische Reihe x_v ($v=1, \dots, n$) und die zugehörige geordnete statistische Reihe $x_{(v^*)}$ ($v^*=1, \dots, n$). Das p -Quantil oder der $(100 \cdot p)$ -Prozentpunkt des Datensatzes mit $0 \leq p \leq 1$ ist

$$x_{[p]} = \begin{cases} x^{(k)} & \text{falls } n \cdot p \text{ nicht ganzzahlig,} \\ & \text{mit } k = \text{kleinste natürliche Zahl größer } n \cdot p \\ \frac{1}{2}(x^{(k)} + x^{(k+1)}) & \text{falls } n \cdot p \text{ ganzzahlig,} \\ & \text{mit } k = n \cdot p. \end{cases}$$

Median

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

Unteres und oberes Quartil (1. und 3. Quartil)

$$x_{[0.25]} = \begin{cases} x^{(k)} & \text{falls } \frac{n}{4} \text{ nicht ganzzahlig,} \\ & \text{mit } k = \text{kleinste natürliche Zahl größer } \frac{n}{4} \\ \frac{1}{2} \left(x_{\left(\frac{n}{4}\right)} + x_{\left(\frac{n}{4}+1\right)} \right) & \text{falls } \frac{n}{4} \text{ ganzzahlig} \end{cases}$$
$$x_{[0.75]} = \begin{cases} x^{(k)} & \text{falls } \frac{3n}{4} \text{ nicht ganzzahlig,} \\ & \text{mit } k = \text{kleinste natürliche Zahl größer } \frac{3n}{4} \\ \frac{1}{2} \left(x_{\left(\frac{3n}{4}\right)} + x_{\left(\frac{3n}{4}+1\right)} \right) & \text{falls } \frac{3n}{4} \text{ ganzzahlig} \end{cases}$$

Arithmetischer Mittelwert (ungewogenes arithmetisches Mittel, Durchschnitt)

$$\bar{x} = \frac{1}{n} \sum_{v=1}^n x_v = \frac{1}{n} \sum_{i=1}^m n_i \tilde{x}_i = \sum_{i=1}^m h_i \tilde{x}_i$$

Klassenmerkmalssummen und Klassenmittelwerte

$$S_j = \sum_{v=1}^{n_j} x_{vj} \quad , \quad \bar{x}_j = \frac{1}{n_j} \sum_{v=1}^{n_j} x_{vj} = \frac{S_j}{n_j} \quad (j = 1, \dots, m)$$

Gesamtmittelwert

$$\bar{x} = \frac{1}{n} \sum_{j=1}^m S_j = \frac{1}{n} \sum_{j=1}^m n_j \bar{x}_j = \sum_{j=1}^m h_j \bar{x}_j$$

Gewogener arithmetischer Mittelwert (gewogener Durchschnitt)

$$\bar{x} = \sum_{v=1}^n g_v \cdot x_v \quad \text{mit} \quad 0 \leq g_v \leq 1 \quad (v = 1, \dots, n) \quad \text{und} \quad \sum_{v=1}^n g_v = 1$$

Geometrischer Mittelwert

$$\bar{x}_g = \sqrt[n]{x_1 \cdot \dots \cdot x_n} = \sqrt[n]{\prod_{v=1}^n x_v} = \left(\prod_{v=1}^n x_v \right)^{\frac{1}{n}} = \left(\prod_{i=1}^m \tilde{x}_i^{n_i} \right)^{\frac{1}{n}} = \prod_{i=1}^m \tilde{x}_i^{h_i} \quad (0 < x_v < \infty)$$

Harmonischer Mittelwert

$$\bar{x}_h = \frac{n}{\sum_{v=1}^n \frac{1}{x_v}} = \frac{n}{\sum_{i=1}^m n_i \cdot \frac{1}{\tilde{x}_i}} = \frac{1}{\sum_{i=1}^m h_i \cdot \frac{1}{\tilde{x}_i}}$$

1.3 Streuungsmaße

Spannweite

$$W = x_{(n)} - x_{(1)}$$

Quartilsabstand und mittlere Quartilsabstand

$$QA = x_{[0.75]} - x_{[0.25]}, \quad \overline{QA} = \frac{1}{2} (x_{[0.75]} - x_{[0.25]})$$

5-Zahlen-Zusammenfassung

$$\begin{array}{cc} & x_{[0.5]} \\ x_{[0.25]} & x_{[0.75]} \\ x_{(1)} & x_{(n)} \end{array}$$

Mittlere absolute Abweichung vom Median

$$MA_{x_{med}} = \frac{1}{n} \sum_{v=1}^n |x_v - x_{med}| = \frac{1}{n} \sum_{i=1}^m n_i \cdot |\tilde{x}_i - x_{med}| = \sum_{i=1}^m h_i \cdot |\tilde{x}_i - x_{med}|$$

Mittlere absolute Abweichung vom arithmetischen Mittel

$$MA_{\bar{x}} = \frac{1}{n} \sum_{v=1}^n |x_v - \bar{x}| = \frac{1}{n} \sum_{i=1}^m n_i \cdot |\tilde{x}_i - \bar{x}| = \sum_{i=1}^m h_i \cdot |\tilde{x}_i - \bar{x}|$$

Varianz

$$\tilde{s}^2 = \frac{1}{n} \sum_{v=1}^n (x_v - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^m n_i \cdot (\tilde{x}_i - \bar{x})^2 = \sum_{i=1}^m h_i \cdot (\tilde{x}_i - \bar{x})^2$$

Standardabweichung

$$\tilde{s} = \sqrt{\frac{1}{n} \sum_{v=1}^n (x_v - \bar{x})^2} = \sqrt{\frac{1}{n} \sum_{i=1}^m n_i \cdot (\tilde{x}_i - \bar{x})^2} = \sqrt{\sum_{i=1}^m h_i \cdot (\tilde{x}_i - \bar{x})^2}$$

Verschiebungssatz für die Varianz

$$\tilde{s}^2 = \frac{1}{n} \sum_{v=1}^n x_v^2 - \bar{x}^2 = \frac{1}{n} \sum_{i=1}^m n_i \tilde{x}_i^2 - \bar{x}^2 = \sum_{i=1}^m h_i \tilde{x}_i^2 - \bar{x}^2$$

Klassenvarianzen

$$\tilde{s}_j^2 = \frac{1}{n_j} \sum_{v=1}^{n_j} (x_{vj} - \bar{x}_j)^2 = \frac{1}{n_j} \sum_{v=1}^{n_j} x_{vj}^2 - \bar{x}_j^2 \quad (j = 1, \dots, m)$$

Gesamtvarianz klassierter Daten

$$\tilde{s}^2 = \frac{1}{n} \sum_{j=1}^m \sum_{v=1}^{n_j} (x_{vj} - \bar{x})^2 = \underbrace{\sum_{j=1}^m \tilde{s}_j^2 \cdot h_j}_{\text{interne Varianz}} + \underbrace{\sum_{j=1}^m (\bar{x}_j - \bar{x})^2 \cdot h_j}_{\text{externe Varianz}}$$

Variationskoeffizient

$$VK = \frac{\tilde{s}}{\bar{x}}$$

Quartilsdispersionskoeffizient

$$QD = \frac{\overline{QA}}{x_{med}} = \frac{x_{[0.75]} - x_{[0.25]}}{2 \cdot x_{med}}$$

Lineartransformation

$$y_v = a + b \cdot x_v \quad (v = 1, \dots, n)$$

Lageparameter linear transformierter Daten

$$y_{mod} = a + b \cdot x_{mod} \quad , \quad y_{med} = a + b \cdot x_{med} \quad , \quad \bar{y} = a + b \cdot \bar{x}$$

Streuungsmaße linear transformierter Daten

$$\tilde{s}_y^2 = b^2 \tilde{s}_x^2, \quad \tilde{s}_y = |b| \tilde{s}_x, \quad MA_{y_{med}} = |b| MA_{x_{med}}$$

Zentrierung

$$y_v = x_v - \bar{x} \quad (v=1, \dots, n)$$

Zentrierte Daten y_v ($v=1, \dots, n$) besitzen das Mittel $\bar{y} = 0$ und die Standardabweichung $\tilde{s}_y = \tilde{s}_x$.

Standardisierung

$$z_v = \frac{x_v - \bar{x}}{\tilde{s}_x} \quad (v=1, \dots, n)$$

Standardisierte Daten z_v ($v=1, \dots, n$) besitzen das Mittel $\bar{z} = 0$ und die Standardabweichung $\tilde{s}_z = 1$.

1.4 Momente und Schiefemaße

Schiefemomentenkoeffizient

$$g_M = \frac{m_3(\bar{x})}{\tilde{s}^3} = \frac{\frac{1}{n} \sum_{v=1}^n (x_v - \bar{x})^3}{\left(\sqrt{\frac{1}{n} \sum_{v=1}^n (x_v - \bar{x})^2} \right)^3}$$

Schiefequartilskoeffizient von Bowley

$$g_Q = \frac{(x_{[0.75]} - x_{[0.5]}) - (x_{[0.5]} - x_{[0.25]})}{x_{[0.75]} - x_{[0.25]}} = \frac{(x_{[0.75]} - x_{med}) - (x_{med} - x_{[0.25]})}{QA} \quad -1 \leq g_Q \leq 1$$

1.5 Die Bestandteile des Box-Plots

- (i) Eine Skala parallel zur Hauptachse des Plots.
- (ii) Ein Rechteck (Box) vom unteren Quartil $x_{[0.25]}$ bis zum oberen Quartil $x_{[0.75]}$.
- (iii) Senkrechte Striche, die den Median $x_{[0.5]}$, den kleinsten Wert $x_{(u)}$, der größer oder gleich dem unteren Ausreißerzaun $x_{[0.25]} - 1.5 \cdot QA$ ist, und den größten Wert $x_{(o)}$, der kleiner oder gleich dem oberen Ausreißerzaun $x_{[0.75]} + 1.5 \cdot QA$ ist, markieren.
- (iv) Querstriche von der Box zu den beiden äußeren senkrechten Strichen.
- (v) Mit dem Symbol * markierte Messwerte, die außerhalb der Zäune liegen (potentielle Ausreißer).

1.6 Konzentrationsmessung

Zur Vereinfachung der Notation wird im Folgenden unterstellt, dass die Messwerte x_v ($v=1, \dots, n$) einer Variablen X bereits der Größe nach geordnet vorgegeben sind, so dass $0 < x_1 \leq \dots \leq x_v \leq \dots \leq x_n$ gilt.

1.6.1 Maße der relativen Konzentration (Disparitätsmaße)

Berechnung der Lorenzkurve: Statistische Reihe

Der Streckenzug, der Punkte

$$(0,0) , (H_1, Q_1) , (H_2, Q_2) , \dots , (H_n, Q_n) = (1,1)$$

mit

$$Q_v = \frac{\sum_{u=1}^v x_u}{\sum_{u=1}^n x_u} , \quad H_v = \frac{v}{n}$$

verbindet, ist die *Lorenzkurve L* der statistischen Reihe.

Falls Anteilssätze H_v und Q_v in Prozent: $(H_v \cdot 100\% , Q_v \cdot 100\%)$ ($v = 1, \dots, n$).

Berechnung der Lorenzkurve: Unklassierte Häufigkeitsverteilung

Die Variable X habe in einem Datensatz die Skalenwerte \tilde{x}_i ($i = 1, \dots, m$) mit den relativen Häufigkeiten

$$h_i = \frac{n_i}{n} \quad (i = 1, \dots, m)$$

angenommen. Ferner gelte

$$H_i = \sum_{u=1}^i h_u \quad \text{und} \quad Q_i = \sum_{u=1}^i q_u \quad (i = 1, \dots, m)$$

mit

$$q_i = \frac{\tilde{x}_i n_i}{S} = \frac{\tilde{x}_i h_i}{\bar{x}} \quad \text{und} \quad S = \sum_{i=1}^m \tilde{x}_i n_i = n \cdot \bar{x} .$$

Der Streckenzug, der Punkte

$$(0,0) , (H_1, Q_1) , (H_2, Q_2) , \dots , (H_m, Q_m) = (1,1)$$

verbindet, ist die *Lorenzkurve L* der unklassierten Häufigkeitsverteilung.

Berechnung der Lorenzkurve: Klassierte Häufigkeitsverteilung

Die Werte der Variablen X seien in m Klassen K_j ($j = 1, \dots, m$) eingeteilt. Bekannt seien neben den relativen Klassenhäufigkeiten h_j auch die Klassenmerkmalssummen S_j oder die Klassenmittelwerte \bar{x}_j . Ferner gelte

$$H_j = \sum_{u=1}^j h_u \quad \text{und} \quad Q_j = \sum_{u=1}^j q_u \quad (j = 1, \dots, m)$$

mit

$$q_j = \frac{S_j}{S} = \frac{\bar{x}_j n_j}{S} \quad \text{und} \quad S = \sum_{j=1}^m S_j = \sum_{j=1}^m \bar{x}_j n_j .$$

Der Streckenzug, der Punkte

$$(0,0) , (H_1, Q_1) , (H_2, Q_2) , \dots , (H_m, Q_m) = (1,1)$$

verbindet, ist die *Lorenzkurve L* der klassierten Häufigkeitsverteilung.

Gini-Koeffizient (Gini'sches Konzentrationsmaß)

$$GK = \frac{F}{\frac{1}{2}} = 2F \quad \text{mit} \quad F = \frac{1}{2} - \frac{1}{2} \sum_u h_u (Q_{u-1} + Q_u) \quad 0 \leq GK \leq \frac{n-1}{n}$$

Korrigierter Gini-Koeffizient

$$GK^* = \frac{n}{n-1} GK$$

$$0 \leq GK^* \leq 1$$

Variationskoeffizient

$$VK = \frac{\tilde{s}}{\bar{x}}$$

$$0 \leq VK \leq \sqrt{n-1}$$

1.6.2 Maße der absoluten Konzentration

Konzentrationsverhältnis, Konzentrationskurve

Konzentrationsverhältnis (concentration ratio) der Ordnung κ mit $\kappa \in \{1, \dots, n\}$:

$$CR_\kappa = \frac{\sum_{v=n+1-\kappa}^n x_v}{\sum_{v=1}^n x_v} \equiv \sum_{v=n+1-\kappa}^n q_v \quad \text{mit} \quad q_v = \frac{x_v}{S} \quad 0 \leq CR_\kappa \leq 1$$

Trägt man die Paare (κ, CR_κ) für alle κ als Punkte in ein Koordinatensystem ein, so heißt der verbindende Streckenzug *Konzentrationskurve*.

Zusammenhang zwischen Lorenz- und Konzentrationskurve

$$CR_\kappa = 1 - L\left(\frac{n-\kappa}{n}\right) \quad (\kappa = 1, \dots, n) .$$

Herfindahl-Koeffizient

$$HK = \sum_{v=1}^n q_v^2 \quad \text{mit} \quad q_v = \frac{x_v}{S} \quad \frac{1}{n} \leq HK \leq 1$$

Zusammenhang zwischen Variations- und Herfindahl-Koeffizienten

$$HK = \frac{VK^2 + 1}{n}$$

Rosenbluth-Koeffizient

$$RK = \frac{1}{2A} \quad \frac{1}{n} \leq RK \leq 1$$

$$\text{mit} \quad A = \sum_{v=1}^n v \cdot q_v - \frac{1}{2} \quad \text{und} \quad q_v = \frac{x_v}{S}$$

Zusammenhang zwischen Rosenbluth- und Gini-Koeffizienten

$$RK = \frac{1}{n(1-GK)}$$

2 Deskription bivariater Datensätze

Notation

Merkmalsvariablen	X, Y	
Realisationsmöglichkeiten	\tilde{x}_i	$(i = 1, \dots, m)$
	\tilde{y}_j	$(j = 1, \dots, k)$
Bivariater Datensatz	(x_v, y_v)	$(v = 1, \dots, n)$
Gemeinsame absolute Häufigkeiten	n_{ij}	$(i = 1, \dots, m; j = 1, \dots, k)$
Gemeinsame relative Häufigkeiten	h_{ij}	$(i = 1, \dots, m; j = 1, \dots, k)$
Absolute Randhäufigkeiten	$n_{i\bullet}$	$(i = 1, \dots, m)$
	$n_{\bullet j}$	$(j = 1, \dots, k)$
Bedingte Häufigkeitsverteilungen	$h(\tilde{x}_i / \tilde{y}_j)$	$(i = 1, \dots, m)$
	$h(\tilde{y}_j / \tilde{x}_i)$	$(j = 1, \dots, k)$
Arithmetischen (Rand-) Mittelwerte	\bar{x}, \bar{y}	
Varianzen der Randverteilung	$\tilde{s}_x^2, \tilde{s}_y^2$	
Kovarianz	\tilde{s}_{xy}	
Rangzahlenpaare	$(R(x_v), R(y_v))$	$(v = 1, \dots, n)$

2.1 Bivariate Häufigkeitsverteilungen und Randverteilungen

Gemeinsame absolute Häufigkeiten

Die Anzahl der mit $(\tilde{x}_i, \tilde{y}_j)$ übereinstimmenden Messwertpaare des bivariaten Datensatzes (x_v, y_v) ($v = 1, \dots, n$) heißt die *absolute gemeinsame Häufigkeit* von $(\tilde{x}_i, \tilde{y}_j)$, geschrieben:

$$n(X = \tilde{x}_i, Y = \tilde{y}_j) \quad \text{oder kurz} \quad n_{ij} .$$

Gemeinsame relative Häufigkeiten

$$h(X = \tilde{x}_i, Y = \tilde{y}_j) = \frac{n(X = \tilde{x}_i, Y = \tilde{y}_j)}{n} \quad \text{oder kurz} \quad h_{ij} = \frac{n_{ij}}{n}$$

Kumulierte gemeinsame Häufigkeiten

$$N_{ij} = \sum_{u=1}^i \sum_{v=1}^j n_{uv} \quad , \quad H_{ij} = \sum_{u=1}^i \sum_{v=1}^j h_{uv}$$

Randhäufigkeiten der Merkmalsvariable X

$$n_{i\bullet} = \sum_{j=1}^k n_{ij} = n(X = \tilde{x}_i) \quad , \quad h_{i\bullet} = \sum_{j=1}^k h_{ij} = \frac{n_{i\bullet}}{n} = h(X = \tilde{x}_i) \quad (i = 1, \dots, m)$$

Randhäufigkeiten der Merkmalsvariable Y

$$n_{\bullet j} = \sum_{i=1}^m n_{ij} = n(Y = \tilde{y}_j) \quad , \quad h_{\bullet j} = \sum_{i=1}^m h_{ij} = \frac{n_{\bullet j}}{n} = h(Y = \tilde{y}_j) \quad (j = 1, \dots, k)$$

Bivariate empirische Verteilungsfunktion

$$v(x, y) = h(X \leq x, Y \leq y) = \frac{n(X \leq x, Y \leq y)}{n} \quad \text{für alle } x, y \in \mathbb{R}$$

2.2 Bedingte Verteilungen

Bedingte Häufigkeit von \tilde{x}_i gegeben $Y = \tilde{y}_j$

$$h(\tilde{x}_i / \tilde{y}_j) = \frac{n_{ij}}{n_{\bullet j}} = \frac{h_{ij}}{h_{\bullet j}}$$

Bedingte Häufigkeit von \tilde{y}_j gegeben $X = \tilde{x}_i$

$$h(\tilde{y}_j / \tilde{x}_i) = \frac{n_{ij}}{n_{i\bullet}} = \frac{h_{ij}}{h_{i\bullet}}$$

Bedingter arithmetischer Mittelwert von X gegeben $Y = \tilde{y}_j$

$$\bar{x} | \tilde{y}_j = \sum_{i=1}^m x_i \cdot h(\tilde{x}_i / \tilde{y}_j)$$

Bedingter arithmetischer Mittelwert von Y gegeben $X = \tilde{x}_i$

$$\bar{y} / \tilde{x}_i = \sum_{j=1}^k y_j \cdot h(\tilde{y}_j / \tilde{x}_i)$$

Bedingte Varianz von X gegeben $Y = \tilde{y}_j$

$$\tilde{s}_x^2 / \tilde{y}_j = \sum_{i=1}^m x_i^2 \cdot h(\tilde{x}_i / \tilde{y}_j) - (\bar{x} / \tilde{y}_j)^2$$

Bedingte Varianz von Y gegeben $X = \tilde{x}_i$

$$\tilde{s}_y^2 / \tilde{x}_i = \sum_{j=1}^k y_j^2 \cdot h(\tilde{y}_j / \tilde{x}_i) - (\bar{y} / \tilde{x}_i)^2$$

2.3 Maßzahlen für bivariate Verteilungen (Korrelationsrechnung)

Kovarianz

$$\begin{aligned} \tilde{s}_{xy} &= \frac{1}{n} \sum_{v=1}^n (x_v - \bar{x})(y_v - \bar{y}) = \frac{1}{n} \sum_{i=1}^m \sum_{j=1}^k (\tilde{x}_i - \bar{x}) \cdot (\tilde{y}_j - \bar{y}) \cdot n_{ij} \\ &= \sum_{i=1}^m \sum_{j=1}^k (\tilde{x}_i - \bar{x}) \cdot (\tilde{y}_j - \bar{y}) \cdot h_{ij} \end{aligned}$$

Verschiebungssatz für die Kovarianz

$$\begin{aligned}\tilde{s}_{XY} &= \frac{1}{n} \sum_{v=1}^n x_v \cdot y_v - \bar{x} \cdot \bar{y} = \frac{1}{n} \sum_{i=1}^m \sum_{j=1}^k \tilde{x}_i \cdot \tilde{y}_j \cdot n_{ij} - \bar{x} \cdot \bar{y} \\ &= \sum_{i=1}^m \sum_{j=1}^k \tilde{x}_i \cdot \tilde{y}_j \cdot h_{ij} - \bar{x} \cdot \bar{y} = \overline{xy} - \bar{x} \cdot \bar{y}\end{aligned}$$

Produkt-Moment-Korrelationskoeffizient (Bravais-Pearson-Korrelationskoeffizient)

$$r_{XY} = \frac{\tilde{s}_{XY}}{\tilde{s}_X \cdot \tilde{s}_Y} = \frac{\frac{1}{n} \sum_{v=1}^n (x_v - \bar{x})(y_v - \bar{y})}{\sqrt{\frac{1}{n} \sum_{v=1}^n (x_v - \bar{x})^2} \cdot \sqrt{\frac{1}{n} \sum_{v=1}^n (y_v - \bar{y})^2}} = \frac{\frac{1}{n} \sum_{v=1}^n x_v y_v - \bar{x} \cdot \bar{y}}{\sqrt{\frac{1}{n} \sum_{v=1}^n x_v^2 - \bar{x}^2} \cdot \sqrt{\frac{1}{n} \sum_{v=1}^n y_v^2 - \bar{y}^2}}$$

Eigenschaften des Korrelationskoeffizienten

Der Korrelationskoeffizient r_{XY} ist eine normierte und dimensionslose Maßzahl mit den Eigenschaften:

- (i) $-1 \leq r_{XY} \leq +1$ bzw. $|r_{XY}| \leq 1$ (Normierung);
- (ii) $r_{XY} = \pm 1$ gilt genau dann, wenn es zwei reelle Konstanten a, b mit $b \neq 0$ so gibt, dass

$$y_v = a + bx_v \quad (v = 1, \dots, n)$$

erfüllt ist. Man sagt dann, zwischen X und Y besteht eine *exakte lineare Beziehung*.

Rangkorrelationskoeffizient (nach Spearman)

$$r_{SP} = \frac{\frac{1}{n} \sum_{v=1}^n R(x_v) \cdot R(y_v) - \frac{(n+1)^2}{4}}{\sqrt{\frac{1}{n} \sum_{v=1}^n R(x_v)^2 - \frac{(n+1)^2}{4}} \cdot \sqrt{\frac{1}{n} \sum_{v=1}^n R(y_v)^2 - \frac{(n+1)^2}{4}}}$$

$(R(x_v), R(y_v))$ sind die zu den Messwertpaaren (x_v, y_v) ($v = 1, \dots, n$) gehörigen Rangzahlenpaare.

Näherungswert für Rangkorrelationskoeffizient (exakt falls Daten ohne Bindungen)

$$r_{SP} = 1 - \frac{6 \cdot \sum_{v=1}^n (R(x_v) - R(y_v))^2}{n^3 - n} = 1 - \frac{6 \cdot \sum_{v=1}^n d_v^2}{n^3 - n} \quad \text{mit } d_v = R(x_v) - R(y_v)$$

Eigenschaften des Rangkorrelationskoeffizienten

Der Rangkorrelationskoeffizient r_{SP} ist eine normierte dimensionslose Maßzahl mit den Eigenschaften:

- (i) $-1 \leq r_{SP} \leq +1$ bzw. $|r_{SP}| \leq 1$ (Normierung);
- (ii) $r_{SP} = \pm 1$ gilt genau dann, wenn eine beliebige monoton wachsende bzw. eine beliebige monoton fallende Funktion $f(x)$ existiert, so dass

$$y_v = f(x_v) \quad (v = 1, \dots, n)$$

erfüllt ist. Man sagt dann, zwischen X und Y besteht ein *monotoner Zusammenhang*.

Kontingenzkoeffizient (nach Pearson)

Erwartete gemeinsame absolute Häufigkeit

$$e_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n} = n \cdot h_{i\bullet} \cdot h_{\bullet j}$$

χ^2 – Koeffizient (Chi-Quadrat-Koeffizient)

$$\chi^2 = \sum_{i=1}^m \sum_{j=1}^k \frac{(n_{ij} - e_{ij})^2}{e_{ij}} \quad 0 \leq \chi^2 < \infty$$

Kontingenzkoeffizient (transformierter χ^2 – Koeffizient)

$$K_{XY} = \sqrt{\frac{\chi^2}{n + \chi^2}} \quad 0 \leq K_{XY} \leq K_{XY}^{max} \quad \text{mit} \quad K_{XY}^{max} = \sqrt{\frac{\min\{m, k\} - 1}{\min\{m, k\}}}$$

Korrigierter Kontingenzkoeffizienten

$$K_{XY}^* = \frac{K_{XY}}{K_{XY}^{max}} \quad 0 \leq K_{XY}^* \leq 1$$

Lineartransformationen und Linearkombinationen

Kovarianz linear transformierter Daten

Gegeben seien die Messwertpaare (x_v, y_v) ($v=1, \dots, n$) der metrisch skalierten Variablen X, Y und deren Lineartransformationen (u_v, w_v) ($v=1, \dots, n$) mit

$$u_v = a + b \cdot x_v, \quad w_v = c + d \cdot y_v, \quad a, b, c, d \in \mathbb{R}, \quad b \neq 0, \quad d \neq 0.$$

Zwischen der Kovarianz \tilde{s}_{XY} der Ausgangsdaten und der Kovarianz \tilde{s}_{UW} der transformierten Daten besteht der Zusammenhang:

$$\tilde{s}_{UW} = b \cdot d \cdot \tilde{s}_{XY}.$$

Bravais-Pearson-Korrelationskoeffizient linear transformierter Daten

$$r_{UW} = \frac{b \cdot d}{|b| \cdot |d|} \cdot r_{XY}$$

Mittelwert und Varianz linear kombinierter Daten

Gegeben seien die Messwertpaare (x_v, y_v) ($v=1, \dots, n$) der metrisch skalierten Variablen X, Y mit den univariaten Mittelwerten \bar{x}, \bar{y} und den univariaten Varianzen $\tilde{s}_X^2, \tilde{s}_Y^2$ sowie die Linearkombinationen

$$z_v = a + bx_v + cy_v \quad (v=1, \dots, n) \quad \text{mit} \quad a, b, c \in \mathbb{R}, \quad b \neq 0, \quad c \neq 0.$$

Für das arithmetische Mittel und die Varianz der Werte z_v ($v=1, \dots, n$) gilt:

(i)
$$\bar{z} = a + b\bar{x} + c\bar{y}$$

(ii)
$$\tilde{s}_Z^2 = b^2 \tilde{s}_X^2 + c^2 \tilde{s}_Y^2 + 2bc \tilde{s}_{XY}.$$

3 Indexrechnung

Notation

Zeitreihe	x_t	$(t = 0,1,\dots,T)$
Basiszeit	$t = k$ (meist $t = 0$)	
Berichtszeit	$t \neq k$	
Messzahl, Änderungsfaktor	$m_{k,t}$	$(t = 0,1,\dots,T)$
Änderungsrate	$r_{k,t}$	$(t = 1,2,\dots,T)$
Logarithmische Änderungsrate	$w_{k,t}$	$(t = 1,2,\dots,T)$
Preis des Wirtschaftsguts i im Zeitpunkt t	$p_{i,t}$	
Menge des Wirtschaftsguts i im Zeitpunkt t	$q_{i,t}$	
Index	$I_{0,t}$	$(t = 0,1,\dots,T)$
Preisindex	$P_{0,t}$	$(t = 0,1,\dots,T)$
Mengenindex	$Q_{0,t}$	$(t = 0,1,\dots,T)$
Wertindex	$V_{0,t}$	$(t = 0,1,\dots,T)$

Messzahlen (Änderungs- oder Wachstumsfaktoren)

$$m_{k,t} = \frac{x_t}{x_k} \quad (t = 0,1,\dots,T)$$

Umbasierung von Messzahlen

$$\frac{m_{k,t}}{m_{k,s}} = \frac{x_t}{x_k} \bigg/ \frac{x_s}{x_k} = \frac{x_t}{x_s} = m_{s,t} \quad (t = 0,1,\dots,T)$$

Änderungs- oder Wachstumsraten

$$r_{k,t} = \frac{x_t - x_k}{x_k} = m_{k,t} - 1 \quad (t = 1,2,\dots,T)$$

Logarithmische Änderungs- oder Wachstumsraten

$$w_{k,t} = \ln \frac{x_t}{x_k} = \ln x_t - \ln x_k \quad (t = 1,2,\dots,T)$$

Durchschnittliche Änderungsfaktoren und -raten

$$\bar{m}_g = (m_{0,1} \cdot m_{1,2} \cdot \dots \cdot m_{T-1,T})^{1/T} = \sqrt[T]{m_{0,1} \cdot m_{1,2} \cdot \dots \cdot m_{T-1,T}}, \quad \bar{r} = \bar{m}_g - 1$$

Durchschnittliche logarithmische Änderungsrate

$$\bar{w} = \frac{1}{T} \sum_{t=1}^T w_{t-1,t} = \frac{w_{0,T}}{T}$$

Preisindex nach Laspeyres

$$P_{0,t}^{(La)} = \frac{\sum_{i=1}^n p_{i,t} \cdot q_{i,0}}{\sum_{i=1}^n p_{i,0} \cdot q_{i,0}} = \sum_{i=1}^n \frac{p_{i,t}}{p_{i,0}} \cdot g_{i,0} \quad \text{mit} \quad g_{i,0} = \frac{p_{i,0} \cdot q_{i,0}}{\sum_{i=1}^n p_{i,0} \cdot q_{i,0}} \quad (t = 0, 1, 2, \dots)$$

Preisindex nach Paasche

$$P_{0,t}^{(Pa)} = \frac{\sum_{i=1}^n p_{i,t} \cdot q_{i,t}}{\sum_{i=1}^n p_{i,0} \cdot q_{i,t}} = \frac{1}{\sum_{i=1}^n \frac{p_{i,t}}{p_{i,0}} \cdot g_{i,t}} \quad \text{mit} \quad g_{i,t} = \frac{p_{i,t} \cdot q_{i,t}}{\sum_{i=1}^n p_{i,t} \cdot q_{i,t}} \quad (t = 0, 1, 2, \dots)$$

Fisher-Ideal-Preisindex

$$P_{0,t}^{(Fi)} = \sqrt{P_{0,t}^{(La)} \cdot P_{0,t}^{(Pa)}} \quad (t = 0, 1, 2, \dots)$$

Mengenindex nach Laspeyres

$$Q_{0,t}^{(La)} = \frac{\sum_{i=1}^n q_{i,t} \cdot p_{i,0}}{\sum_{i=1}^n q_{i,0} \cdot p_{i,0}} = \sum_{i=1}^n \frac{q_{i,t}}{q_{i,0}} \cdot g_{i,0} \quad \text{mit} \quad g_{i,0} = \frac{p_{i,0} \cdot q_{i,0}}{\sum_{i=1}^n p_{i,0} \cdot q_{i,0}} \quad (t = 0, 1, 2, \dots)$$

Mengenindex nach Paasche

$$Q_{0,t}^{(Pa)} = \frac{\sum_{i=1}^n q_{i,t} \cdot p_{i,t}}{\sum_{i=1}^n q_{i,0} \cdot p_{i,t}} = \frac{1}{\sum_{i=1}^n \frac{p_{i,t}}{p_{i,0}} \cdot g_{i,t}} \quad \text{mit} \quad g_{i,t} = \frac{p_{i,t} \cdot q_{i,t}}{\sum_{i=1}^n p_{i,t} \cdot q_{i,t}} \quad (t = 0, 1, 2, \dots)$$

Fisher-Ideal-Mengenindex

$$Q_{0,t}^{(Fi)} = \sqrt{Q_{0,t}^{(La)} \cdot Q_{0,t}^{(Pa)}} \quad (t = 0, 1, 2, \dots)$$

Umsatz- oder Wertindex

$$V_{0,t} = \frac{\sum_{i=1}^n p_{i,t} \cdot q_{i,t}}{\sum_{i=1}^n p_{i,0} \cdot q_{i,0}} \quad (t = 0, 1, 2, \dots)$$

Beziehung zwischen Wertindex und Preis- sowie Mengenindizes

$$V_{0,t} = P_{0,t}^{(La)} \cdot Q_{0,t}^{(Pa)} = P_{0,t}^{(Pa)} \cdot Q_{0,t}^{(La)}$$

$$V_{0,t} = P_{0,t}^{(Fi)} \cdot Q_{0,t}^{(Fi)}$$

Gesamtindex und Sektorenindizes

Ein Gesamtindex ist in $j = 1, \dots, m$ Sektoren und jeder Sektor in $i = 1, \dots, n_j$ Untersektoren aufgegliedert. Für die Untersektoren seien Indizes $I_{i,j,0,t}$ vorgegeben, die auch einfache Messzahlen sein können. Aus dem Untersektorenindizes können bei vorgegebenen relativen Gewichten $\gamma_{i,j}^*$ Sektorenindizes $I_{j,0,t}$ berechnet werden:

$$I_{j,0,t} = \sum_{i=1}^{n_j} \gamma_{i,j}^* \cdot I_{i,j,0,t} \quad \text{mit} \quad \sum_{i=1}^{n_j} \gamma_{i,j}^* = 1 \quad (j = 1, \dots, m).$$

Zwischen dem Gesamtindex $I_{0,t}$ und den Sektorenindizes $I_{j,0,t}$ besteht die Beziehung:

$$I_{0,t} = \sum_{j=1}^m \gamma_j \cdot I_{j,0,t} = \sum_{j=1}^m \sum_{i=1}^{n_j} \gamma_{i,j} \cdot I_{i,j,0,t} \quad \text{,}$$

wobei

$$\sum_{j=1}^m \gamma_j = 1 \quad , \quad \gamma_j = \sum_{i=1}^{n_j} \gamma_{i,j} \quad \text{und} \quad \gamma_{i,j} = \gamma_{i,j}^* \cdot \gamma_j$$

ist.

Für gegebene Werte von Gesamtindex und Sektorenindizes sind relative Gewichte γ_j ($j = 1, \dots, m$) eindeutig durch das folgende Gleichungssystem bestimmt:

$$\sum_{j=1}^m \gamma_j \cdot I_{j,0,t} = I_{0,t} \quad (t = 1, \dots, m-1),$$

$$\sum_{j=1}^m \gamma_j = 1 .$$

Umbasierung

Umbasierung des Indexwerts $I_{u,t}$ mit der Basiszeit u auf die neue Basiszeit v im Falle $u > v$:

$$I_{v,t}^* = I_{u,v} \cdot I_{u,t} .$$

Umbasierung des Indexwerts $I_{u,t}$ mit der Basiszeit u auf die neue Basiszeit v im Falle $u < v$:

$$I_{v,t}^{**} = \frac{I_{u,t}}{I_{u,v}} .$$

4 Elementare Wahrscheinlichkeitsrechnung

4.1 Zufallsvorgänge und Wahrscheinlichkeitsräume

Zufallsvorgang, Ergebnis

Ein *Zufallsvorgang* oder *Zufallsexperiment* ist ein Vorgang, dessen Ausgang im Rahmen verschiedener, prinzipiell bekannter Möglichkeiten ungewiss ist und der sich unter Einhaltung bestimmter Vorschriften beliebig oft (zumindest gedanklich) wiederholen lässt. Der nach Ablauf des Vorganges tatsächlich eingetretene Ausgang wird als *Ergebnis* des Zufallsvorganges bezeichnet.

Ergebnisraum

Eine Menge Ω heißt *Ergebnisraum* eines Zufallsvorganges, wenn sie jedes mögliche Ergebnis des Vorganges als Element enthält. Weitere Sprechweisen: *Grundraum*, *Stichprobenraum*.

Ereignis, Ereignisfeld

Ein *Ereignis* A eines Zufallsvorganges ist eine Teilmenge des Ergebnisraumes Ω , d.h. $A \subset \Omega$. Man sagt, ein Ereignis A ist bei der Durchführung eingetreten, wenn das Ergebnis ω des Zufallsvorganges ein Element von A ist, also $\omega \in A$. Die Menge \mathcal{F} aller zulässigen Ereignisse eines Zufallsvorganges heißt *Ereignisfeld*.

Unmögliches und sicheres Ereignis

- (i) $\emptyset \subset \Omega$ heißt das *unmögliche* Ereignis (\emptyset tritt sicher nicht ein);
- (ii) $\Omega \subset \Omega$ heißt das *sichere* Ereignis (Ω tritt sicher ein).

Elementarereignisse

Einelementige Teilmengen $A = \{\omega\}$ von Ω heißen *Elementarereignisse*.

Operieren mit Ereignissen

(i) Gleichheit von Ereignissen

Zwei Ereignisse A und B heißen gleich, in Zeichen $A = B$, wenn gilt:
 $\omega \in A \Leftrightarrow \omega \in B$.

(ii) Teilereignis

Ein Ereignis A heißt Teilereignis des Ereignisses B , in Zeichen $A \subset B$, wenn gilt: $\omega \in A \Rightarrow \omega \in B$. Man sagt, dass Ereignis A zieht das Ereignis B nach sich.

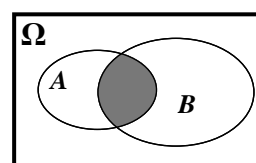
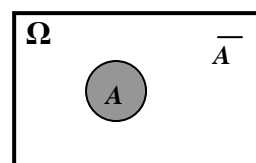
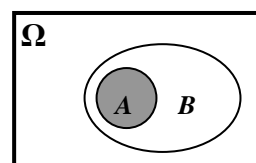
(iii) Komplementärereignis

Ist A ein Ereignis, dann ist auch $\bar{A} = \{\omega \in \Omega \mid \omega \notin A\}$ ein Ereignis und heißt komplementär zu A . Man sagt, \bar{A} tritt genau dann ein, wenn A nicht eintritt.

(iv) Durchschnitt

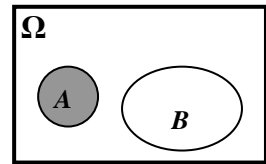
Sind A und B Ereignisse, dann ist auch $A \cap B = \{\omega \in \Omega \mid \omega \in A \wedge \omega \in B\}$ ein Ereignis und heißt Durchschnitt der Ereignisse A und B . Man sagt, die Ereignisse A und B treten zugleich ein.

Venn-Diagramm



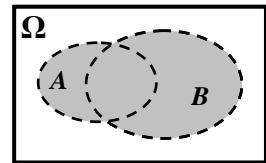
(v) *Disjunkte Ereignisse*

Zwei Ereignisse A und B heißen disjunkt (unverträglich), wenn gilt:
 $A \cap B = \emptyset$. Man sagt, wenn A eintritt, tritt B nicht ein.



(vi) *Vereinigung*

Sind A und B Ereignisse, dann ist auch $A \cup B = \{\omega \in \Omega \mid \omega \in A \vee \omega \in B\}$ ein Ereignis und heißt Vereinigung der Ereignisse A und B . Man sagt, A oder B tritt ein.



W-Maß, W-Raum

Eine Abbildung $P: \mathcal{F} \rightarrow \mathbb{R}$ heißt *Wahrscheinlichkeitsmaß* (*W-Maß*), wenn sie die folgenden Kolmogorow'schen Axiome erfüllt:

Axiom 1 $P(A) \geq 0$ für alle Ereignisse $A \in \mathcal{F}$ (*Nichtnegativitätsaxiom*)

Axiom 2 $P(\Omega) = 1$ (*Normierungsaxiom*)

Axiom 3 Für disjunkte Ereignisse $A, B \in \mathcal{F}$ mit $A \cap B = \emptyset$ gelte:

$$P(A \cup B) = P(A) + P(B) \quad (\textit{Additivität})$$

Die reelle Zahl $P(A)$ heißt *Wahrscheinlichkeit* des Ereignisses A .

Operieren mit Wahrscheinlichkeiten

Es seien (Ω, \mathcal{F}, P) ein W-Raum und $A, B, C \in \mathcal{F}$ Ereignisse. Dann gilt:

(i) $0 \leq P(A) \leq 1$

(ii) $P(\emptyset) = 0$

(iii) $P(\bar{A}) = 1 - P(A)$ (*Gegenwahrscheinlichkeit*)

(iv) $A \subset B \Rightarrow P(A) \leq P(B)$ (*Monotonie*)

(v) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (*Additionssatz für 2 Ereignisse*)

(vi) $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$ (*Additionssatz für 3 Ereignisse*)

4.2 Wahrscheinlichkeitskonzeptionen

Laplace-Wahrscheinlichkeit

Es liege ein Zufallsvorgang mit einem abzählbar-endlichem Ergebnisraum $\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$, $N < \infty$, vor. Das Ereignisfeld sei $\mathcal{F} = P(\Omega)$. Die Elementarereignisse $\{\omega_i\}$ ($i = 1, \dots, N$) seien gleichwahrscheinlich, d.h. sie sind hinsichtlich der Unbestimmtheit ihres Eintretens nicht unterscheidbar.

Dann ist für ein beliebiges Ereignis $A \in \mathcal{F}$ die *Laplace-Wahrscheinlichkeit* für das Eintreten von A festgelegt durch das Verhältnis

$$P(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{N} = \frac{\text{Anzahl der für } A \text{ günstigen Ergebnisse}}{\text{Anzahl aller möglichen Ergebnisse}}.$$

Häufigkeiten von Ereignissen

Bei der n -fach wiederholten Durchführung eines Zufallsvorgangs sei ein Ereignis $n(A)$ -fach eingetreten. Wir bezeichnen $n(A)$ als die *absolute Häufigkeit* und

$$h_n(A) = \frac{n(A)}{n}$$

als die relative Häufigkeit von A .

Statistische Wahrscheinlichkeit

Die *statistische Wahrscheinlichkeit* für das Eintreten eines Ereignisses A ist die Zahl $P(A)$, um die sich die relative Häufigkeit $h_n(A)$ bei wachsender Versuchswiederholungszahl n stabilisiert.

4.3 Bedingte Wahrscheinlichkeit und Unabhängigkeit

Bedingte Wahrscheinlichkeit (von A unter der Bedingung B)

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

Gemeinsames Eintreten zweier Ereignisse (Multiplikationssatz)

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

Gemeinsames Eintreten dreier Ereignisse (Multiplikationssatz)

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

Vollständiges System

Für einen gegebenen W-Raum (Ω, \mathcal{F}, P) seien $A_1, A_2, \dots, A_m \in \mathcal{F}$ paarweise disjunkte Ereignisse, die den Ergebnisraum Ω ganz ausfüllen, d.h. $A_1 \cup A_2 \cup \dots \cup A_m = \Omega$ mit $A_i \cap A_j = \emptyset$ für $i \neq j$. Dann bezeichnet man die Ereignisse als ein *vollständiges System*.

Satz von der totalen Wahrscheinlichkeit

A_1, A_2, \dots, A_m sei ein vollständiges System von Ereignissen und B ein weiteres Ereignis mit $P(A_i) > 0$, $P(B) > 0$. Dann gilt:

$$P(B) = \sum_{i=1}^m P(B|A_i) \cdot P(A_i).$$

Theorem von Bayes

A_1, A_2, \dots, A_m sei ein vollständiges System von Ereignissen und B ein weiteres Ereignis mit $P(A_i) > 0$, $P(B) > 0$. Dann gilt:

$$P(A_k|B) = \frac{P(B|A_k) \cdot P(A_k)}{\sum_{i=1}^m P(B|A_i) \cdot P(A_i)} \quad \text{für } k = 1, \dots, m.$$

Im Spezialfall $m = 2$ gilt mit $A_1 = A$ und $A_2 = \bar{A}$:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A})}.$$

Unabhängigkeit zweier Ereignisse

Die Ereignisse A und B heißen *stochastisch unabhängig* (voneinander), wenn

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

Unabhängigkeit mehrerer Ereignisse

Die Ereignisse A_1, A_2, \dots, A_m heißen (vollständig) *stochastisch unabhängig*, wenn für jede beliebige Auswahl von ganzzahligen Indizes i_1, i_2, \dots, i_k mit

$$1 \leq i_1 < i_2 < \dots < i_k \leq m \quad (k=2,3,\dots,m)$$

die Gleichung

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \dots \cdot P(A_{i_k})$$

erfüllt ist.

4.4 Kombinatorik

Fakultät

$$n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot (n-1) \cdot n$$

Binomialkoeffizient

$$\binom{n}{m} = \frac{n \cdot (n-1) \cdot \dots \cdot (n-m+1)}{m \cdot (m-1) \cdot \dots \cdot 2 \cdot 1} = \frac{n!}{m!(n-m)!}$$

Kombinationen

Eine Auswahl von n Objekten aus N Objekten heißt *Kombination* von n aus N Objekten.

Kombinationen ohne Wiederholung unter Berücksichtigung der Anordnung

Für N gegebene voneinander verschiedene Objekte ist die Anzahl der Kombinationen von n aus N Objekten mit Berücksichtigung der Anordnungen gleich

$$N \cdot (N-1) \cdot (N-2) \cdot \dots \cdot (N-n+1) = \frac{N!}{(N-n)!} .$$

Kombinationen ohne Wiederholung ohne Berücksichtigung der Anordnung

Die Anzahl der Kombinationen von n aus N verschiedenen Objekten ohne Berücksichtigung der Anordnung ist

$$\binom{N}{n} = \frac{N \cdot (N-1) \cdot \dots \cdot (N-n+1)}{n \cdot (n-1) \cdot \dots \cdot 2 \cdot 1} = \frac{N!}{n!(N-n)!} .$$

Kombinationen mit Wiederholung unter Berücksichtigung der Anordnung

Die Anzahl der Kombinationen von n aus N verschiedenen wiederholbaren Objekten mit Berücksichtigung der Anordnung beträgt

$$N^n .$$

Kombinationen mit Wiederholung ohne Berücksichtigung der Anordnung

Die Anzahl der Kombinationen von n aus N verschiedenen wiederholbaren Objekten ohne Berücksichtigung der Anordnung ist

$$\binom{N+n-1}{n} .$$

5 Eindimensionale Zufallsvariablen und deren Wahrscheinlichkeitsverteilungen

Notation

Zufallsvariable	X
Verteilungsfunktion	$F(x)$
Wahrscheinlichkeitsfunktion	$f(x)$
Dichtefunktion	$f(x)$
Mittelwert, Erwartungswert	$\mu, E(X)$
Varianz	σ^2
Standardabweichung	σ

Zufallsvariable

Gegeben sei ein Zufallsexperiment mit dem W-Raum (Ω, \mathcal{F}, P) . Eine Abbildung

$$X : \Omega \rightarrow \mathbb{R}, \omega \mapsto X(\omega)$$

heißt [eindimensionale] Zufallsvariable [über (Ω, \mathcal{F}, P)], wenn für jede reelle Zahl x die Menge

$$\{\omega \in \Omega \mid X(\omega) \leq x\}$$

ein Element von \mathcal{F} (d.h. ein Ereignis) ist. Ist ω das Ergebnis der Durchführung des Zufallsexperiments, dann heißt der zugehörige Wert $X(\omega)$ Realisierung oder Realisation der Zufallsvariablen X .

Verteilungsfunktion

Die Funktion

$$F : \mathbb{R} \rightarrow [0, 1] \quad \text{mit} \quad x \mapsto F(x) = P(X \leq x)$$

heißt Verteilungsfunktion der Zufallsvariablen X . Sie ordnet jeder reellen Zahl $-\infty < x < +\infty$ die Wahrscheinlichkeit $P(X \leq x)$ zu.

Eigenschaften von Verteilungsfunktionen

(i) F ist monoton wachsend, d.h. $x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$;

(ii) $\lim_{x \rightarrow -\infty} F(x) = 0$ und $\lim_{x \rightarrow +\infty} F(x) = 1$;

(iii) F ist rechtsstetig, d.h. $\lim_{x \rightarrow x_0} F(x) = F(x_0)$.

Operieren mit Verteilungsfunktionen

F sei die Verteilungsfunktion der Zufallsvariablen X und a, b seien reelle Zahlen mit $a < b$. Es gilt:

(i) $P(X > a) = 1 - F(a)$

(ii) $P(a < X \leq b) = F(b) - F(a)$

(iii) $P(X = a) =$ Höhe des Sprunges von F im Punkt a

(iv) $P(X < a) = F(a) - P(X = a)$

(v) $P(X \geq a) = 1 - F(a) + P(X = a)$

(vi) $P(a \leq X \leq b) = F(b) - F(a) + P(X = a)$

(vii) $P(a < X < b) = F(b) - F(a) - P(X = b)$

(viii) $P(a \leq X < b) = F(b) - F(a) + P(X = a) - P(X = b)$.

Diskrete Zufallsvariable

Eine Zufallsvariable X heißt *diskret*, wenn ihr Bildbereich endlich oder abzählbar unendlich viele Werte enthält. Wir bezeichnen diese Realisationsmöglichkeiten einfach mit x_1, x_2, x_3, \dots . Die Werte nennt man auch *Sprungstellen* und ihre Wahrscheinlichkeiten *Sprunghöhen*.

Wahrscheinlichkeitsfunktion

Es sei X eine diskrete Zufallsvariable. Die Funktion $f: \mathbb{R} \rightarrow [0, 1]$ mit

$$f(x) = P(X = x) = \begin{cases} p_i & \text{falls } x = x_i \text{ (} i=1,2,3,\dots \text{)} \\ 0 & \text{sonst (alle übrigen reellen } x \text{)} \end{cases}$$

heißt die *Wahrscheinlichkeitsfunktion* von X .

Stetige Zufallsvariable, Dichtefunktion

Eine Zufallsvariable X heißt *stetig*, wenn eine nicht-negative Funktion $f(x)$ (d.h. $f(x) \geq 0$ für alle $x \in \mathbb{R}$) existiert, die für jede reelle Zahl x die Gleichung

$$F(x) = \int_{-\infty}^x f(t) dt$$

erfüllt, wobei $F(x) = P(X \leq x)$ die Verteilungsfunktion von X ist. Die Funktion $f(x)$ heißt *Dichtefunktion* oder kurz *Dichte* von X .

Eigenschaften von Dichtefunktionen

(i) $\int_{-\infty}^{\infty} f(x) dx = 1;$

(ii) $P(a < X \leq b) = \int_a^b f(x) dx$ für beliebige $a, b \in \mathbb{R}$ mit $a \leq b$;

(iii) $F'(x) = f(x)$ falls $F(x)$ differenzierbar im Punkt x .

Mittelwert, Erwartungswert

$$E(X) = \mu = \begin{cases} \sum_i x_i \cdot f(x_i) & \text{falls } X \text{ diskret mit den Sprungstellen } x_i \text{ (} i = 1, 2, \dots \text{)} \\ \int_{-\infty}^{\infty} x \cdot f(x) dx & \text{falls } X \text{ stetig} \end{cases}$$

Varianz

$$\text{Var}(X) = E[(X - \mu)^2] = \sigma^2 = \begin{cases} \sum_i (x_i - \mu)^2 \cdot f(x_i) & \text{falls } X \text{ diskret mit den Sprungstellen } x_i \text{ (} i = 1, 2, \dots \text{)} \\ \int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x) dx & \text{falls } X \text{ stetig} \end{cases}$$

Standardabweichung

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

Verschiebungssatz der Varianz

$$\text{Var}(X) = E(X^2) - E(X)^2 = E(X^2) - \mu^2$$

Symmetrische Verteilung

Eine Zufallsvariable X mit der Wahrscheinlichkeits- bzw. Dichtefunktion $f(x)$ heißt *symmetrisch verteilt bezüglich einer Zahl c* , wenn für alle $x \in \mathbb{R}$ gilt: $f(c-x) = f(c+x)$.

Symmetriepunkt einer symmetrischen Verteilung

Besitzt eine Zufallsvariable den Mittelwert μ und eine bezüglich der Zahl c symmetrische Verteilung, dann ist $c = \mu$.

Mittelwert und Varianz lineartransformierter Zufallsvariablen

Es sei X eine Zufallsvariable mit dem Mittelwert $E(X) = \mu_X$ und der Varianz $\text{Var}(X) = \sigma_X^2$, dann gilt für die lineartransformierte Variable $Y = a_0 + a_1X$ mit $a_1 \neq 0$:

$$E(Y) = a_0 + a_1E(X) \quad \text{bzw.} \quad \mu_Y = a_0 + a_1\mu_X,$$

$$\text{Var}(Y) = a_1^2\text{Var}(X) \quad \text{bzw.} \quad \sigma_Y^2 = a_1^2\sigma_X^2.$$

Standardisierte Zufallsvariable

$$Z = \frac{X - E(X)}{\sqrt{\text{Var}(X)}} \quad \text{bzw.} \quad Z = \frac{X - \mu}{\sigma}$$

Quantile, Median, Quartile

Jede reelle Zahl $x_{[p]}$, die für eine vorgegebene Zahl p mit $0 < p < 1$ die Ungleichungen

$$P(X \leq x_{[p]}) \geq p \quad \text{und} \quad P(X \geq x_{[p]}) \geq 1 - p$$

beide erfüllt, heißt *p-Quantil* oder *(p·100)%-Punkt* (der Wahrscheinlichkeitsverteilung) der Zufallsvariablen X . Im Fall einer stetigen Zufallsvariablen vereinfacht sich die Bedingung zu

$$P(X \leq x_{[p]}) = F(x_{[p]}) = p.$$

$x_{[p]}$ muss nicht eindeutig bestimmt sein.

Das 0.5-Quantil (50%-Punkt) heißt *Median* und wird als Lagemaß einer Verteilung genutzt.

Die 0.25- und 0.75-Quantile (25%- und 75%-Punkte) heißen *unteres* bzw. *oberes Quartil*.

Ungleichung von Tschebyschev

Es sei X eine Zufallsvariable mit dem Mittelwert $E(X) = \mu$ und der Varianz $\text{Var}(X) = \sigma^2$. Dann gilt für jedes reelle $\varepsilon > 0$:

$$P(|X - \mu| < \varepsilon) \geq 1 - \frac{\sigma^2}{\varepsilon^2} \quad \text{oder äquivalent} \quad P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}.$$

6 Mehrdimensionale Zufallsvariablen und deren Wahrscheinlichkeitsverteilungen

Zufallsvektor

Mit X_1, \dots, X_n seien $n \geq 2$ Zufallsvariablen über denselben W-Raum (Ω, \mathcal{F}, P) bezeichnet. Das n -Tupel

$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

heißt n -Zufallsvektor.

Der Zufallsvektor \mathbf{X} ordnet einem Ergebnis $\omega \in \Omega$ des zugrundeliegenden Zufallsvorgangs ein n -Zahlentupel

$$\mathbf{X}(\omega) = \begin{pmatrix} X_1(\omega) \\ X_2(\omega) \\ \vdots \\ X_n(\omega) \end{pmatrix} \in \mathbb{R}^n$$

zu, das als *Realisation* von \mathbf{X} bezeichnet wird.

Gemeinsame Verteilungsfunktion

Die Funktion $F: \mathbb{R}^n \rightarrow [0, 1]$ der n reellwertigen Veränderlichen x_1, \dots, x_n mit

$$F_{x_1, \dots, x_n}(x_1, \dots, x_n) = P((X_1 \leq x_1) \cap \dots \cap (X_n \leq x_n)) \equiv P(X_1 \leq x_1, \dots, X_n \leq x_n)$$

heißt *Verteilungsfunktion des Zufallsvektor \mathbf{X}* oder *gemeinsame Verteilungsfunktion* der Zufallsvariablen X_1, \dots, X_n .

Diskreter Zufallsvektor

Ein 2-Zufallsvektor $\mathbf{X} = (X, Y)'$ heißt *diskret*, wenn er endlich oder höchstens abzählbar unendlich viele Wertepaare $(x_i, y_j)' \in \mathbb{R}^2$ mit $i = 1, 2, \dots$ und $j = 1, 2, \dots$ annehmen kann.

Wahrscheinlichkeitsfunktion

Die für alle $(x, y)' \in \mathbb{R}^2$ definierte Funktion $f_{XY}(x, y)$ mit

$$f_{x,y}(x, y) = P(X = x, Y = y) = \begin{cases} p_{ij} & \text{falls } x = x_i, y = y_j \quad (i = 1, 2, \dots; j = 1, 2, \dots) \\ 0 & \text{sonst} \end{cases}$$

heißt *Wahrscheinlichkeitsfunktion des Zufallsvektors* oder *gemeinsame Wahrscheinlichkeitsfunktion* der Zufallsvariablen X, Y .

Stetiger Zufallsvektor, Dichtefunktion

Ein 2-Zufallsvektor $\mathbf{X} = (X, Y)'$ heißt *stetig*, wenn sich die zugehörige Verteilungsfunktion durch ein Zweifach-Integral der Form

$$F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(t_1, t_2) dt_2 dt_1$$

darstellen lässt. Hierbei ist $f_{XY}(x, y)$ eine für alle $(x, y)' \in \mathbb{R}^2$ definierte nicht-negative Funktion und heißt *Dichtefunktion des Zufallsvektors* oder *gemeinsame Dichte* der Zufallsvariablen X, Y .

Randverteilungsfunktion

$$F_X(x) = F_{X,Y}(x, \infty) \quad \text{bzw.} \quad F_Y(y) = F_{X,Y}(\infty, y)$$

Unabhängigkeit von Zufallsvariablen

Die Komponenten eines n -Zufallsvektors $\mathbf{X} = (X_1, \dots, X_n)'$ heißen *stochastisch unabhängig*, wenn für alle reellen Zahlen x_1, \dots, x_n gilt:

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdot \dots \cdot F_{X_n}(x_n).$$

Ist dies nicht der Fall, so heißen die Zufallsvariablen *stochastisch abhängig*.

Unabhängigkeit von Zufallsvariablen

Die Komponenten eines n -Zufallsvektors $\mathbf{X} = (X_1, \dots, X_n)'$ sind genau dann stochastisch unabhängig, wenn für ihre gemeinsame Wahrscheinlichkeitsfunktion bzw. Dichte für alle reellen x_1, \dots, x_n gilt:

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdot \dots \cdot f_{X_n}(x_n).$$

Kovarianz

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = \sigma_{XY}$$

Korrelationskoeffizient (normierte Kovarianz)

$$\text{Corr}(X, Y) = \frac{\sigma_{XY}}{\sigma_X \cdot \sigma_Y} = \rho_{XY}$$

Verschiebungssatz für die Kovarianz

$$\sigma_{XY} = E(X \cdot Y) - E(X) \cdot E(Y) = E(X \cdot Y) - \mu_X \cdot \mu_Y.$$

Zusammenhang zwischen stochastischer Unabhängigkeit und Korrelation

Wenn die Zufallsvariablen X und Y stochastisch unabhängig sind, dann sind sie auch unkorreliert, d.h. $\sigma_{XY} = \rho_{XY} = 0$. Die Umkehrung der Aussage gilt nicht allgemein.

Mittelwert und Varianz der Linearkombination zweier Zufallsvariablen

Es seien $\mathbf{X} = (X, Y)'$ ein 2-Zufallsvektor, $a_1, a_2 \neq 0$ reelle Konstanten und $Z = a_1X + a_2Y$ eine Linearkombination von X und Y . Dann ist

$$\mu_Z = a_1\mu_X + a_2\mu_Y \quad \text{und} \quad \sigma_Z^2 = a_1^2\sigma_X^2 + a_2^2\sigma_Y^2 + 2a_1a_2\sigma_{XY}.$$

Mittelwert und Varianz der Linearkombination mehrerer Zufallsvariablen

Es seien X_1, \dots, X_n Zufallsvariablen und a_1, \dots, a_n reelle Konstanten. Dann gilt

$$E\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i E(X_i),$$
$$\text{Var}\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 \text{Var}(X_i) + \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n a_i a_j \text{Cov}(X_i, X_j).$$

7 Spezielle Wahrscheinlichkeitsverteilungsmodelle

7.1 Diskrete Verteilungsmodelle

Gleichverteilung oder uniforme Verteilung [$X \sim U(m)$]

Wahrscheinlichkeitsfunktion	$f(x) = \begin{cases} \frac{1}{m} & \text{falls } x = x_i \ (i=1, \dots, m) \\ 0 & \text{sonst.} \end{cases}$
Parameter	$m \in \mathbb{N}$
Mittelwert, Varianz	$E(X) = \frac{1}{m} \sum_{i=1}^m x_i, \quad \text{Var}(X) = \frac{1}{m} \sum_{i=1}^m \left(x_i - \frac{1}{m} \sum_{j=1}^m x_j \right)^2$

Binomialverteilung [$X \sim B(n, p)$]

Wahrscheinlichkeitsfunktion	$f(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{falls } x = 0, 1, \dots, n \\ 0 & \text{sonst} \end{cases}$
Parameter	n, p mit $n \in \mathbb{N}$, $p \in \mathbb{R}$ und $0 < p < 1$
Mittelwert, Varianz	$E(X) = n \cdot p, \quad \text{Var}(X) = n \cdot p \cdot (1-p)$
Reproduktivität	$X \sim B(n_1, p)$ und $Y \sim B(n_2, p)$ seien unabhängig, dann gilt $X + Y \sim B(n_1 + n_2, p)$.

Poisson-Verteilung [$X \sim Po(\lambda)$]

Wahrscheinlichkeitsfunktion	$f(x) = \begin{cases} \frac{\lambda^x}{x!} \cdot e^{-\lambda} & \text{falls } x = 0, 1, 2, \dots \\ 0 & \text{sonst} \end{cases}$
Parameter	$\lambda \in \mathbb{R}, \lambda > 0$
Mittelwert, Varianz	$E(X) = \text{Var}(X) = \lambda$
Reproduktivität	$X \sim Po(\lambda_1)$ und $Y \sim Po(\lambda_2)$ seien unabhängig, dann ist $X + Y \sim Po(\lambda_1 + \lambda_2)$.

Geometrische Verteilung [$X \sim G(p)$]

Wahrscheinlichkeitsfunktion	$f(x) = \begin{cases} p \cdot (1-p)^x & \text{falls } x = 0, 1, 2, \dots \\ 0 & \text{sonst} \end{cases}$
Verteilungsfunktion	$F(x) = \begin{cases} 0 & \text{falls } x < 0 \\ 1 - (1-p)^{\text{int}(x)+1} & \text{falls } x \geq 0 \end{cases}$ ($\text{int}(x)$ ist der ganzzahlige Anteil von x)
Parameter	$p \in \mathbb{R}$ und $0 < p < 1$
Mittelwert, Varianz	$E(X) = \frac{1-p}{p}, \quad \text{Var}(X) = \frac{1-p}{p^2}$

Hypergeometrische Verteilung [$X \sim H(n, N, M)$]

Wahrscheinlichkeitsfunktion	$f(x) = \begin{cases} \frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}} & \text{falls } x = a, a+1, \dots, b-1, b \\ 0 & \text{sonst} \end{cases}$
Parameter	$n, N, M \in \mathbb{N}$ mit $n \leq N$, $M \leq N$, $a = \max\{0, [n - (N - M)]\}$, $b = \min\{n, M\}$
Mittelwert, Varianz	$E(X) = n \cdot \frac{M}{N} = n \cdot p$ $\text{Var}(X) = n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right) \cdot \frac{N-n}{N-1} = n \cdot p \cdot (1-p) \cdot \frac{N-n}{N-1} \quad \text{mit } p = M/N$

7.2 Stetige Verteilungsmodelle

Rechteckverteilung [$X \sim R(\alpha, \beta)$]

Dichtefunktion	$f(x) = \begin{cases} \frac{1}{\beta-\alpha} & \text{falls } \alpha \leq x \leq \beta \\ 0 & \text{sonst} \end{cases}$
Verteilungsfunktion	$F(x) = \begin{cases} 0 & \text{falls } x < \alpha \\ \frac{x-\alpha}{\beta-\alpha} & \text{falls } \alpha \leq x \leq \beta \\ 1 & \text{falls } x > \beta \end{cases}$
Parameter	$\alpha, \beta \in \mathbb{R}$ mit $\alpha < \beta$
Mittelwert, Varianz	$E(X) = \frac{\alpha + \beta}{2}, \quad \text{Var}(X) = \frac{(\beta - \alpha)^2}{12}$

Exponentialverteilung [$X \sim Ex(\lambda)$]

Dichtefunktion	$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{falls } x \geq 0 \\ 0 & \text{sonst.} \end{cases}$
Verteilungsfunktion	$F(x) = \begin{cases} 0 & \text{falls } x < 0 \\ 1 - e^{-\lambda x} & \text{falls } x \geq 0 \end{cases}$
Parameter	$\lambda \in \mathbb{R}, \lambda > 0$
Mittelwert, Varianz	$E(X) = \frac{1}{\lambda}, \quad \text{Var}(X) = \frac{1}{\lambda^2}$
Markov-Eigenschaft	$P(X \leq x + s X > x) = P(X \leq s)$ für alle reellen $x, s > 0$

Normalverteilung (Gauss-Verteilung) [$X \sim N(\mu, \sigma^2)$]

Dichtefunktion	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (-\infty < x < \infty)$
Verteilungsfunktion	$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$

Parameter	$\mu, \sigma^2 \in \mathbb{R}, \sigma^2 > 0$
Mittelwert, Varianz	$E(X) = \mu, \quad \text{Var}(X) = \sigma^2$
Reproduktivität	Es seien $X \sim N(\mu_X, \sigma_X^2)$ und $Y \sim N(\mu_Y, \sigma_Y^2)$ unabhängige Zufallsvariablen, dann gilt für alle reellen $a, b \neq 0$: $aX + bY \sim N(a\mu_X + b\mu_Y, a^2\sigma_X^2 + b^2\sigma_Y^2)$.

Standardnormalverteilung [$Z \sim N(0,1)$]

Dichtefunktion	$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad (-\infty < z < \infty)$
Verteilungsfunktion	$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{t^2}{2}} dt$
Parameter	$\mu = 0, \sigma^2 = 1,$
Mittelwert, Varianz	$E(Z) = 0, \quad \text{Var}(Z) = 1$

Verteilungsfunktionen/Quantile der Normal- und der Standardnormalverteilung

Zwischen dem p -Quantil $x_{[p]}$ der $N(\mu, \sigma^2)$ -Verteilung und dem p -Quantil $z_{[p]}$ der Standardnormalverteilung besteht die Beziehung

$$z_{[p]} = \frac{x_{[p]} - \mu}{\sigma} \quad \text{oder} \quad x_{[p]} = \mu + \sigma \cdot z_{[p]}.$$

Ferner gilt für alle reellen Zahlen x :

$$F(x) = P(X \leq x) = \Phi\left(\underbrace{\frac{x - \mu}{\sigma}}_{=z}\right) = \Phi(z).$$

Außerdem gilt

$$P(a < X \leq b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right).$$

Symmetrie der Standardnormalverteilung

Die Standardnormalverteilung ist symmetrisch bezüglich dem Punkt $\mu = 0$. Hieraus resultiert

$$\Phi(z) = 1 - \Phi(-z) \quad \text{und} \quad z_{[p]} = -z_{[1-p]}.$$

Zentrales Schwankungsintervall

Sei $X \sim N(\mu, \sigma^2)$. Das Zahlenintervall

$$\left[\mu - z_{[1-\frac{\alpha}{2}]} \cdot \sigma, \mu + z_{[1-\frac{\alpha}{2}]} \cdot \sigma \right]$$

mit

$$P\left(\mu - z_{[1-\frac{\alpha}{2}]} \cdot \sigma \leq X \leq \mu + z_{[1-\frac{\alpha}{2}]} \cdot \sigma\right) = 1 - \alpha$$

heißt zentrales Schwankungsintervall oder $z_{[1-\frac{\alpha}{2}]}$ -faches σ -Intervall.

Grenzwertsatz von Lindeberg-Lévy

X_1, \dots, X_n seien unabhängige und identisch verteilte Zufallsvariablen, die alle den Mittelwert μ und die Varianz σ^2 besitzen. Dann konvergiert die Verteilungsfunktion $F_{Z_n}(z) = P(Z_n \leq z)$ der Zufallsvariablen

$$Z_n = \frac{\sum_{i=1}^n X_i - n \cdot \mu}{\sqrt{n} \cdot \sigma}$$

mit wachsender Summandenzahl n gegen die Verteilungsfunktion $\Phi(z)$ der $N(0,1)$ -Verteilung:

$$\lim_{n \rightarrow \infty} F_{Z_n}(z) = \Phi(z) \quad (-\infty < z < \infty) .$$

Grenzwertsatz von DeMoivre-Laplace

Es sei X binomialverteilt mit den Parametern n und p sowie dem Erwartungswert $\mu = np$ und der Varianz $\sigma^2 = np(1-p)$. Dann konvergiert die Verteilungsfunktion der Zufallsvariable

$$Z_n = \frac{X - np}{\sqrt{np(1-p)}}$$

für $n \rightarrow \infty$ gegen die Verteilungsfunktion $\Phi(z)$ der $N(0,1)$ -Verteilung.

Approximation der Binomial- durch die Normalverteilung mit Stetigkeitskorrektur

Für $X \sim B(n, p)$ und ganzzahlige x gilt:

$$F(x) = P(X \leq x) \approx \Phi\left(\frac{x + 0.5 - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}}\right)$$

$$b(x/n, p) = P(X = x) \approx \Phi\left(\frac{x + 0.5 - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}}\right) - \Phi\left(\frac{x - 0.5 - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}}\right) .$$

Als Faustregel sollte $n > \frac{9}{p(1-p)}$ gelten.

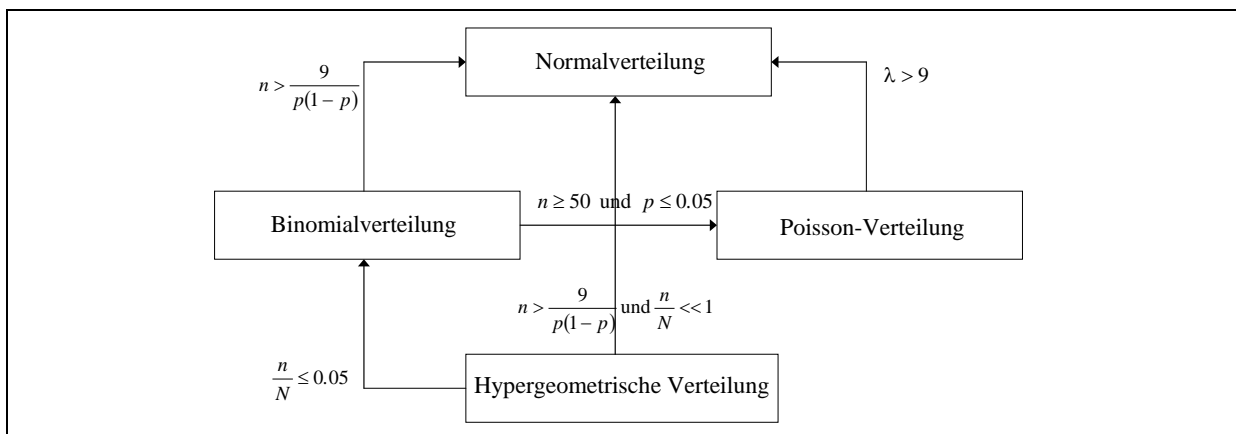
Approximation der Poisson- durch die Normalverteilung mit Stetigkeitskorrektur

Für $X \sim Po(\lambda)$ gilt bei großem $\lambda (= np)$ und ganzzahlige x

$$F(x) = P(X \leq x) \approx \Phi\left(\frac{x + 0.5 - \lambda}{\sqrt{\lambda}}\right) .$$

Man beachte, dass $E(X) = Var(X) = \lambda$ ist. Als Faustregel sollte $\lambda > 9$ gelten.

Approximationsbeziehungen zwischen verschiedenen Verteilungen



7.3 Stichprobenverteilungen

Chi-Quadrat-Verteilung

Z_1, \dots, Z_n seien n stochastisch unabhängige, jeweils $N(0,1)$ -verteilte Zufallsvariablen. Die Verteilung der Quadratsumme

$$\chi^2 = \sum_{i=1}^n Z_i^2$$

heißt *Chi-Quadrat-Verteilung mit n Freiheitsgraden* oder kurz $\chi^2(n)$ -Verteilung. Die sogenannte Anzahl n der Freiheitsgrade ist der Parameter der Verteilung.

t-Verteilung (STUDENT-Verteilung)

Z sei $N(0,1)$ -verteilt, χ^2 besitze eine $\chi^2(n)$ -Verteilung, und beide Zufallsvariablen seien unabhängig voneinander. Die Verteilung der Zufallsvariable

$$T = \frac{Z}{\sqrt{\chi^2/n}}$$

heißt *t-Verteilung mit n Freiheitsgraden* oder kurz $t(n)$ -Verteilung. Die Anzahl n der Freiheitsgrade ist der Parameter der Verteilung.

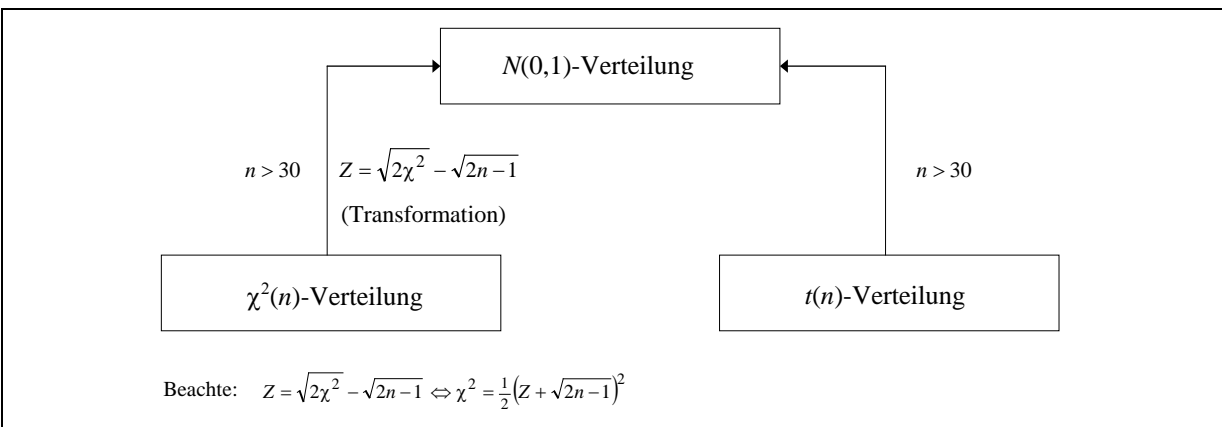
F-Verteilung (Fisher-Verteilung)

Die Zufallsvariable χ_1^2 sei $\chi^2(n)$ -verteilt, die Zufallsvariable χ_2^2 sei $\chi^2(m)$ -verteilt, und beide Zufallsvariablen seien unabhängig voneinander. Die Verteilung der Zufallsvariable

$$F = \frac{\chi_1^2/n}{\chi_2^2/m}$$

heißt *F-Verteilung mit n und m Freiheitsgraden* oder kurz $F(n,m)$ -Verteilung. Hierbei sind n und m positive ganze Zahlen. Sie sind die Parameter der Verteilung.

Approximationsbeziehungen zwischen verschiedenen Verteilungen



8 Gesetze der großen Zahlen, Punktschätzung

Einfache Zufallsstichprobe

Sei X eine Zufallsvariable mit der Verteilungsfunktion $F(x)=P(X\leq x)$. Ein Vektor $\mathbf{X}=(X_1,\dots,X_n)'$ von n Stichprobenvariablen mit den Eigenschaften

- (i) die Variablen X_1,\dots,X_n sind alle identisch wie X verteilt, d.h. sie besitzen alle die gleiche Verteilungsfunktion F ;
- (ii) die Variablen X_1,\dots,X_n sind stochastisch unabhängig;

heißt *einfache Zufallsstichprobe* (mathematische Stichprobe) vom Umfang n aus der Grundgesamtheit X mit der Verteilungsfunktion F . Die Eigenschaften der Stichprobenvariablen bezeichnet man auch als *i.i.d. (independent and identically distributed)*. Eine beobachtete Realisierung $\mathbf{x}=(x_1,\dots,x_n)'$ von \mathbf{X} nennen wir *konkrete Stichprobe* oder einfach beobachtete Stichprobe.

Stichprobenfunktion, Statistik

Eine Zufallsvariable $T=T(X_1,\dots,X_n)$, die als Funktion von Stichprobenvariablen X_1,\dots,X_n definiert ist, heißt *Stichprobenfunktion* oder *Statistik*. Die Wahrscheinlichkeitsverteilung von T nennen wir *Stichprobenverteilung* von T . Liegt eine konkrete Stichprobenrealisation x_1,\dots,x_n vor, dann nimmt die Stichprobenfunktion den konkreten Wert $t=T(x_1,\dots,x_n)$ als Realisation an.

(Schwaches) Gesetz der großen Zahlen

X_1,\dots,X_n sei eine einfache Zufallsstichprobe vom Umfang n aus einer Grundgesamtheit X mit dem Mittel $E(X)=\mu$ und der Varianz $Var(X)=\sigma^2$. Ferner sei $\bar{X}_n=\frac{1}{n}\sum_{v=1}^n X_v$ der arithmetische Mittelwert der Stichprobenvariablen. Dann gilt für beliebig kleine $\varepsilon>0$:

$$\lim_{n\rightarrow\infty} P(|\bar{X}_n - \mu| < \varepsilon) = 1.$$

BERNOULLIS Gesetz der großen Zahlen

X_1,\dots,X_n sei eine einfache Zufallsstichprobe vom Umfang n aus einer dichotomen Grundgesamtheit X . Das interessierende Ereignis A trete mit der Wahrscheinlichkeit $P(A)=p$ ein. Ferner sei $H_n(A)$ die relative Stichprobenhäufigkeit von A . Dann gilt für beliebig kleine $\varepsilon>0$:

$$\lim_{n\rightarrow\infty} P(|H_n(A) - p| < \varepsilon) = 1.$$

Hauptsatz der mathematischen Statistik (in abgeschwächter Form)

$V_n(x)$ sei die empirische Verteilungsfunktion einer einfachen Zufallsstichprobe vom Umfang n aus der Grundgesamtheit X mit der Verteilungsfunktion $F(x)$. Für beliebig kleine reelle $\varepsilon>0$ und jede reelle Zahl x gilt:

$$\lim_{n\rightarrow\infty} P(|V_n(x) - F(x)| < \varepsilon) = 1.$$

Schätzfunktion, Schätzwert

Gegeben sei eine Realisierung x_1,\dots,x_n der einfachen Zufallsstichprobe X_1,\dots,X_n vom Umfang n aus der Grundgesamtheit X mit der unbekanntem Kenngröße θ . Eine Stichprobenfunktion $T_{n,\theta}=T(X_1,\dots,X_n)$, deren Realisierung $t_{n,\theta}=T(x_1,\dots,x_n)$ als Näherung der Kenngröße θ dient, heißt *Schätzfunktion* oder kurz *Schätzer* für θ . Die Realisierung $t_{n,\theta}$ des Schätzers heißt *Schätzwert* für θ . Den Schätzwert schreibt man häufig auch $\hat{\theta}$.

Erwartungstreue, Verzerrung, Bias

Eine Schätzfunktion $T_{n,\theta}$ heißt *erwartungstreu* (oder *unverzerrt*) Schätzung für θ , falls $E(T_{n,\theta}) = \theta$ gilt. Andernfalls heißt $T_{n,\theta}$ *verzerrt*, und die Differenz $E(T_{n,\theta}) - \theta = B(T_{n,\theta})$ heißt der *Bias* oder der *systematische Fehler* von $T_{n,\theta}$. Gilt $\lim_{n \rightarrow \infty} B(T_{n,\theta}) = 0$, so nennt man $T_{n,\theta}$ *asymptotisch erwartungstreu*.

Relative und absolute Effizienz

Sind $T_{n,\theta}$ und $T_{n,\theta}^*$ zwei erwartungstreue Schätzfunktionen für die Kennzahl θ , so heißt $T_{n,\theta}^*$ *effizienter* als $T_{n,\theta}$, wenn für alle möglichen Werte von θ

$$\text{Var}(T_{n,\theta}^*) \leq \text{Var}(T_{n,\theta})$$

erfüllt ist und für mindestens einen Wert das ' $<$ ' Zeichen gilt. $T_{n,\theta}^*$ heißt *absolut effizient* oder *effizientester* Schätzer für θ , falls es keinen anderen erwartungstreuen Schätzer $T_{n,\theta}$ für θ gibt, der effizienter als $T_{n,\theta}^*$ ist.

Konsistenz, stochastische Konvergenz

Eine Schätzfunktion $T_{n,\theta} = T(X_1, \dots, X_n)$ für die Kennzahl θ heißt (schwach) *konsistent*, wenn für jedes reelles $\varepsilon > 0$ gilt:

$$\lim_{n \rightarrow \infty} P(|T_{n,\theta} - \theta| < \varepsilon) = 1.$$

Man sagt dann auch, $T_{n,\theta}$ konvergiert stochastisch (konvergiert in Wahrscheinlichkeit) gegen θ und schreibt symbolisch $\text{plim} T_{n,\theta} = \theta$. Hierbei steht *plim* für „probability limit“.

Hinreichende Bedingung für Konsistenz

Gilt $E(T_{n,\theta}) = \theta$ oder zumindest $\lim_{n \rightarrow \infty} E(T_{n,\theta}) = \theta$ und $\lim_{n \rightarrow \infty} \text{Var}(T_{n,\theta}) = 0$, dann ist die Schätzfunktion $T_{n,\theta}$ (schwach) konsistent.

Likelihoodfunktion, Maximum– Likelihood–Schätzung

Sei x_1, \dots, x_n eine Realisation der einfachen Zufallsstichprobe X_1, \dots, X_n aus der Grundgesamtheit X mit der Wahrscheinlichkeits- bzw. Dichtefunktion $f_X(x/\theta)$. Die Funktion

$$L(\theta/x_1, \dots, x_n) = \prod_{v=1}^n f_X(x_v/\theta)$$

für alle zulässigen Werte von θ heißt die *Likelihoodfunktion der konkreten Stichprobe* x_1, \dots, x_n . Gilt in der Stelle $\theta = \hat{\theta}$

$$L(\hat{\theta}/x_1, \dots, x_n) \geq L(\theta/x_1, \dots, x_n) \quad \text{für alle möglichen Werte von } \theta,$$

so ist $\hat{\theta} \equiv t_{n,\theta} = T(x_1, \dots, x_n)$ der *Maximum-Likelihood-(ML)-Schätzwert* für θ . Die zugehörige Stichprobenfunktion $T_{n,\theta} = T(X_1, \dots, X_n)$ heißt der *Maximum-Likelihood-(ML)-Schätzer* für θ .

Zur praktischen Bestimmung von ML-Schätzern bedient man sich häufig der Differentialrechnung. Hierbei kann man sich zunutze machen, dass $L(\theta/x_1, \dots, x_n)$ genau dort ein Maximum besitzt, wo auch die logarithmierte Likelihoodfunktion oder *Loglikelihoodfunktion*

$$\ln L(\theta/x_1, \dots, x_n) = \ln \prod_{v=1}^n f_X(x_v/\theta) = \sum_{v=1}^n \ln f_X(x_v/\theta)$$

ein Maximum besitzt. Eine Summe ist i.d.R. leichter zu differenzieren als ein Produkt.

9 Confidence Intervals and Tests of Hypotheses

Symbols

$N(\mu, \sigma^2)$	normal distribution with mean μ and variance σ^2
$N(0,1)$	standard normal distribution
$B(1, p)$	Bernoulli distribution
$z_{[p]}$	p -quantile of the standard normal distribution (notice: $z_{[p]} = -z_{[1-p]}$)
$t_{[p;n]}$	p -quantile of the t -distribution with n degrees of freedom (notice: $t_{[p;n]} = -t_{[1-p;n]}$)
$\chi^2_{[p;n]}$	p -quantile of the χ^2 -distribution with n degrees of freedom
$f_{[p;n,m]}$	p -quantile of the F -distribution with n and m degrees of freedom
α	significance level
x_{med}	median of a distribution
$\bar{X} = \frac{1}{n} \sum_{v=1}^n X_v$	sample mean
$S^2 = \frac{1}{n-1} \sum_{v=1}^n (X_v - \bar{X})^2$ $= \frac{1}{n-1} \left(\sum_{v=1}^n X_v^2 - n\bar{X}^2 \right)$	sample variance

Interval Estimation

Confidence interval for the mean of a normal distribution

Sample: n independent sample variables X_1, \dots, X_n from a $N(\mu, \sigma^2)$ -population; μ, σ^2 unknown

100(1- α)% confidence interval:

$$\left[\bar{X} - t_{[1-\alpha/2;n-1]} \cdot \frac{S}{\sqrt{n}}; \bar{X} + t_{[1-\alpha/2;n-1]} \cdot \frac{S}{\sqrt{n}} \right]$$

Remark: For $n > 30$ $t_{[p;n-1]}$ can be approximated by the standard normal p -quantile $z_{[p]}$.

Confidence interval for the variance of a normal distribution

Sample: n independent sample variables X_1, \dots, X_n from a $N(\mu, \sigma^2)$ -population; μ, σ^2 unknown

100(1- α)% confidence interval:

$$\left[\frac{(n-1)S^2}{\chi^2_{[1-\alpha/2;n-1]}}; \frac{(n-1)S^2}{\chi^2_{[\alpha/2;n-1]}} \right]$$

Large sample confidence interval for the mean of a distribution (arbitrary distribution)

Sample: n independent sample variables X_1, \dots, X_n from a population with unknown mean μ and variance σ^2 ;
 n large

Approximate $100(1-\alpha)\%$ confidence interval:

$$\left[\bar{X} - z_{[1-\alpha/2]} \cdot \frac{S}{\sqrt{n}}; \bar{X} + z_{[1-\alpha/2]} \cdot \frac{S}{\sqrt{n}} \right]$$

Required sample size for estimating the mean of a distribution (arbitrary distribution)

Sample: Independent sample variables from a population with unknown mean μ and known variance σ^2 ;
 L^* is an upper bound for the length of the confidence interval

Lower bound for the sample size:

$$n \geq \left(\frac{2 \cdot z_{[1-\alpha/2]} \cdot \sigma}{L^*} \right)^2$$

Remark: The resulting bound should be large to justify the formula by the central limit theorem.

Large sample confidence interval for the probability of a Bernoulli distribution

Sample: n independent sample variables X_1, \dots, X_n from a Bernoulli population with unknown probability p ;
 n large

Approximate $100(1-\alpha)\%$ confidence interval:

$$\left[\bar{X} - z_{[1-\alpha/2]} \cdot \sqrt{\frac{\bar{X} \cdot (1 - \bar{X})}{n}}; \bar{X} + z_{[1-\alpha/2]} \cdot \sqrt{\frac{\bar{X} \cdot (1 - \bar{X})}{n}} \right]$$

Required sample size for estimating the probability of a Bernoulli distribution

Sample: Independent sample variables from a Bernoulli population with unknown probability p ; L^* is an upper bound for the length of the confidence interval

Lower bound for the sample size:

$$n \geq \left(\frac{z_{[1-\alpha/2]}}{L^*} \right)^2$$

Remark: The resulting bound should be large to justify the formula by the central limit theorem.

Confidence interval for the difference of means of two normal distributions (equal variances)

Samples: Two independent samples of n independent, $N(\mu_x, \sigma_x^2)$ -distributed sample variables X_1, \dots, X_n and
 m independent, $N(\mu_y, \sigma_y^2)$ -distributed sample variables Y_1, \dots, Y_m ; $\mu_x, \mu_y, \sigma_x^2, \sigma_y^2$ unknown, but
 $\sigma_x^2 = \sigma_y^2$

$100(1-\alpha)\%$ confidence interval:

$$\left[(\bar{X} - \bar{Y}) - t_{[1-\alpha/2; n+m-2]} \cdot S_p \cdot \sqrt{\frac{1}{n} + \frac{1}{m}}; (\bar{X} - \bar{Y}) + t_{[1-\alpha/2; n+m-2]} \cdot S_p \cdot \sqrt{\frac{1}{n} + \frac{1}{m}} \right], \text{ where } S_p^2 = \frac{(n-1)S_x^2 + (m-1)S_y^2}{n+m-2}$$

Confidence interval for the difference of means of two normal distributions (unequal variances)

Samples: Two independent samples of n independent, $N(\mu_x, \sigma_x^2)$ -distributed sample variables X_1, \dots, X_n and m independent, $N(\mu_y, \sigma_y^2)$ -distributed sample variables; $\mu_x, \mu_y, \sigma_x^2, \sigma_y^2$ unknown, but $\sigma_x^2 \neq \sigma_y^2$

100(1- α)% confidence interval:

$$\left[(\bar{X} - \bar{Y}) - t_{[1-\alpha/2; df]} \cdot \sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}; (\bar{X} - \bar{Y}) + t_{[1-\alpha/2; df]} \cdot \sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}} \right], \text{ where } df = \frac{(1+Q)^2}{\frac{Q^2}{n-1} + \frac{1}{m-1}} \text{ and } Q = \frac{S_x^2 \cdot m}{S_y^2 \cdot n}$$

Confidence interval for the ratio of variances of two normal distributions

Samples: Two independent samples of n independent, $N(\mu_x, \sigma_x^2)$ -distributed sample variables X_1, \dots, X_n and m independent, $N(\mu_y, \sigma_y^2)$ -distributed sample variables Y_1, \dots, Y_m ; $\mu_x, \mu_y, \sigma_x^2, \sigma_y^2$ unknown

100(1- α)% confidence interval:

$$\left[\frac{S_x^2}{S_y^2} \cdot \frac{1}{f_{[\alpha/2; n-1, m-1]}}; \frac{S_x^2}{S_y^2} \cdot \frac{1}{f_{[1-\alpha/2; n-1, m-1]}} \right]$$

Large sample confidence interval for the difference of means of two distributions (arbitrary distribution)

Samples: Two independent samples of n independent sample variables X_1, \dots, X_n and m independent sample variables Y_1, \dots, Y_m from populations with unknown means μ_x, μ_y and variances σ_x^2, σ_y^2 ; n and m large

Approximate 100(1- α)% confidence interval:

$$\left[(\bar{X} - \bar{Y}) - z_{[1-\alpha/2]} \cdot \sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}; (\bar{X} - \bar{Y}) + z_{[1-\alpha/2]} \cdot \sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}} \right]$$

Tests of Hypotheses

t-test (normal distribution)

Sample: n independent sample variables X_1, \dots, X_n from a $N(\mu, \sigma^2)$ -population; μ, σ^2 unknown

Hypotheses	test-statistic; distribution of the statistic under the null hypothesis	critical values
H_0 H_1		
$\mu = \mu_0$ $\mu \neq \mu_0$	$\sqrt{n} \frac{\bar{X} - \mu_0}{S}$; t -distribution with $n-1$ degrees of freedom	$t_{[\alpha/2; n-1]}$, $t_{[1-\alpha/2; n-1]}$
$\mu \leq \mu_0$ $\mu > \mu_0$	(for $n > 30$ the t -distribution can be approximated by $N(0,1)$)	$t_{[1-\alpha; n-1]}$
$\mu \geq \mu_0$ $\mu < \mu_0$		$t_{[\alpha; n-1]}$

Chi-square-test (normal distribution)

Sample: n independent sample variables X_1, \dots, X_n from a $N(\mu, \sigma^2)$ -population; μ, σ^2 unknown

hypotheses H_0 H_1	test-statistic; distribution of the statistic under the null hypothesis	critical values
$\sigma^2 = \sigma_0^2$ $\sigma^2 \neq \sigma_0^2$	$\frac{(n-1)S^2}{\sigma_0^2}$; χ^2 -distribution with $n-1$ degrees of freedom	$\chi_{[\alpha/2; n-1]}^2, \chi_{[1-\alpha/2; n-1]}^2$
$\sigma^2 \leq \sigma_0^2$ $\sigma^2 > \sigma_0^2$		$\chi_{[1-\alpha; n-1]}^2$
$\sigma^2 \geq \sigma_0^2$ $\sigma^2 < \sigma_0^2$		$\chi_{[\alpha; n-1]}^2$

Large samples Gauß-test for the mean of an arbitrary distribution

Sample: n independent sample variables X_1, \dots, X_n from a population with unknown mean μ and variance σ^2 ; n large

hypotheses H_0 H_1	test-statistic; distribution of the statistic under the null hypothesis	critical values
$\mu = \mu_0$ $\mu \neq \mu_0$	$\sqrt{n} \frac{\bar{X} - \mu_0}{S}$; asymptotic $N(0,1)$ -distribution	$z_{[\alpha/2]}, z_{[1-\alpha/2]}$
$\mu \leq \mu_0$ $\mu > \mu_0$		$z_{[1-\alpha]}$
$\mu \geq \mu_0$ $\mu < \mu_0$		$z_{[\alpha]}$

Large samples Gauß-test for the probability of a Bernoulli distribution

Sample: n independent sample variables X_1, \dots, X_n from a Bernoulli population with unknown probability p ; n large

hypotheses H_0 H_1	test-statistic; distribution of the statistic under the null hypothesis	critical values
$p = p_0$ $p \neq p_0$	$\sqrt{n} \frac{\bar{X} - p_0}{\sqrt{p_0(1-p_0)}}$; asymptotic $N(0,1)$ -distribution	$z_{[\alpha/2]}, z_{[1-\alpha/2]}$
$p \leq p_0$ $p > p_0$		$z_{[1-\alpha]}$
$p \geq p_0$ $p < p_0$		$z_{[\alpha]}$

Two sample t-test (normal distributions)

Samples: Two independent samples of independent, $N(\mu_x, \sigma_x^2)$ -distributed sample variables X_1, \dots, X_n and independent, $N(\mu_y, \sigma_y^2)$ -distributed sample variables Y_1, \dots, Y_m ; $\mu_x, \mu_y, \sigma_x^2, \sigma_y^2$ unknown, but $\sigma_x^2 = \sigma_y^2$

hypotheses H_0 H_1	test-statistic; distribution of the statistic under the null hypothesis	critical values
$\mu_x = \mu_y$ $\mu_x \neq \mu_y$	$\frac{\bar{X} - \bar{Y}}{\sqrt{S_p^2 \left(\frac{1}{n} + \frac{1}{m} \right)}}$ where $S_p^2 = \frac{(n-1)S_x^2 + (m-1)S_y^2}{n+m-2}$; t -distribution with $n+m-2$ degrees of freedom	$t_{[\alpha/2; n+m-2]}, t_{[1-\alpha/2; n+m-2]}$
$\mu_x \leq \mu_y$ $\mu_x > \mu_y$		$t_{[1-\alpha; n+m-2]}$
$\mu_x \geq \mu_y$ $\mu_x < \mu_y$		$t_{[\alpha; n+m-2]}$

Two sample Welch-test (normal distributions)

Samples: Two independent samples of independent, $N(\mu_x, \sigma_x^2)$ -distributed sample variables X_1, \dots, X_n and independent, $N(\mu_y, \sigma_y^2)$ -distributed sample variables Y_1, \dots, Y_m ; $\mu_x, \mu_y, \sigma_x^2, \sigma_y^2$ unknown, but $\sigma_x^2 \neq \sigma_y^2$

hypotheses H_0 H_1	test-statistic; distribution of the statistic under the null hypothesis	critical values
$\mu_x = \mu_y$ $\mu_x \neq \mu_y$	$\frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}}$ $\left(df = \frac{(1+Q)^2}{\frac{Q^2}{n-1} + \frac{1}{m-1}}, \quad Q = \frac{S_x^2 \cdot m}{S_y^2 \cdot n} \right);$ t-distribution with df degrees of freedom	$t_{[\alpha/2; df]}$, $t_{[1-\alpha/2; df]}$
$\mu_x \leq \mu_y$ $\mu_x > \mu_y$		$t_{[1-\alpha; df]}$
$\mu_x \geq \mu_y$ $\mu_x < \mu_y$		$t_{[\alpha; df]}$

Two sample F-test (normal distributions)

Samples: Two independent samples of independent, $N(\mu_x, \sigma_x^2)$ -distributed sample variables X_1, \dots, X_n and independent, $N(\mu_y, \sigma_y^2)$ -distributed sample variables Y_1, \dots, Y_m ; $\mu_x, \mu_y, \sigma_x^2, \sigma_y^2$ known

hypotheses H_0 H_1	test-statistic; distribution of the statistic under the null hypothesis	critical values
$\sigma_x^2 = \sigma_y^2$ $\sigma_x^2 \neq \sigma_y^2$	$\frac{S_x^2}{S_y^2};$ F-distribution with $n-1$ and $m-1$ degrees of freedom	$f_{[\alpha/2; n-1, m-1]}$, $f_{[1-\alpha/2; n-1, m-1]}$
$\sigma_x^2 \leq \sigma_y^2$ $\sigma_x^2 > \sigma_y^2$		$f_{[1-\alpha; n-1, m-1]}$
$\sigma_x^2 \geq \sigma_y^2$ $\sigma_x^2 < \sigma_y^2$		$f_{[\alpha; n-1, m-1]}$

Two sample Gauß-test for large samples (arbitrary distributions)

Samples: Two independent samples of n independent sample variables X_1, \dots, X_n and m independent sample variables Y_1, \dots, Y_m from populations with unknown moments $\mu_x, \mu_y, \sigma_x^2, \sigma_y^2$; n and m large

hypotheses H_0 H_1	test-statistic; distribution of the statistic under the null hypothesis	large sample critical values
$\mu_x = \mu_y$ $\mu_x \neq \mu_y$	$\frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}}$; asymptotic $N(0,1)$ -distribution	$z_{[\alpha/2]}$, $z_{[1-\alpha/2]}$
$\mu_x \leq \mu_y$ $\mu_x > \mu_y$		$z_{[1-\alpha]}$
$\mu_x \geq \mu_y$ $\mu_x < \mu_y$		$z_{[\alpha]}$

Two sample Gauß-test for large samples (Bernoulli distributions)

Samples: Two independent samples of n independent sample variables X_1, \dots, X_n and m independent sample variables Y_1, \dots, Y_m from Bernoulli populations with unknown probabilities p_x, p_y ; n and m large

hypotheses H_0 H_1	test-statistic; distribution of the statistic under the null hypothesis	large sample critical values
$p_x = p_y$ $p_x \neq p_y$	$\frac{\bar{X} - \bar{Y}}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n} + \frac{1}{m}\right)}}$ where $\hat{p} = \frac{n \cdot \bar{X} + m \cdot \bar{Y}}{n + m}$; asymptotic $N(0,1)$ -distribution	$z_{[\alpha/2]}$, $z_{[1-\alpha/2]}$
$p_x \leq p_y$ $p_x > p_y$		$z_{[1-\alpha]}$
$p_x \geq p_y$ $p_x < p_y$		$z_{[\alpha]}$

Wilcoxon signed ranks test (distribution-free)

Sample: n independent sample variables X_1, \dots, X_n from a *continuously* and *symmetrically* distributed population with median x_{med}

Symbols: $D_v = X_v - x_{med_0}$, $R(|D_v|) = \text{rank of } |D_v|$, $Z_v = \begin{cases} 1 & X_v > x_{med_0} \\ 0 & X_v < x_{med_0} \end{cases}$ for $v=1, \dots, n$

hypotheses		test-statistics; distribution of the statistics under the null hypothesis	large sample critical values
H_0	H_1		
$x_{med} = x_{med_0}$	$x_{med} \neq x_{med_0}$	$T = \sum_{v=1}^n R(D_v) \cdot Z_v$; non-standard distribution, critical values are tabled for small sample sizes $n \leq 30$ in literature $\frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$; asymptotic $N(0,1)$ -distribution	$z_{[\alpha/2]}$, $z_{[1-\alpha/2]}$
$x_{med} \leq x_{med_0}$	$x_{med} > x_{med_0}$		$z_{[1-\alpha]}$
$x_{med} \geq x_{med_0}$	$x_{med} < x_{med_0}$		$z_{[\alpha]}$

Wilcoxon sum of ranks test (distribution-free)

Samples: Two independent samples of n independent sample variables X_1, \dots, X_n and m independent sample variables Y_1, \dots, Y_m from *continuously* distributed populations with medians x_{med}, y_{med}

Symbol: $R(X_v) = \text{rank of } X_v$ ($v=1, \dots, n$) in a pooled sample $X_1, \dots, X_n, Y_1, Y_2, \dots, Y_m$

hypotheses		test-statistics; distribution of the statistics under the null hypothesis	large sample critical values
H_0	H_1		
$x_{med} = y_{med}$	$x_{med} \neq y_{med}$	$T = \sum_{v=1}^n R(X_v)$; non-standard distribution, critical values are tabled for small sample sizes $n \leq 30$ in literature $\frac{T - \frac{n(n+m+1)}{2}}{\sqrt{\frac{n \cdot m \cdot (n+m+1)}{12}}}$; asymptotic $N(0,1)$ -distribution	$z_{[\alpha/2]}$, $z_{[1-\alpha/2]}$
$x_{med} \leq y_{med}$	$x_{med} > y_{med}$		$z_{[1-\alpha]}$
$x_{med} \geq y_{med}$	$x_{med} < y_{med}$		$z_{[\alpha]}$

Signs test (distribution-free)

Sample: n independent sample variables X_1, \dots, X_n from a *continuously* distributed population with median x_{med}

Symbol: $D_v = X_v - x_{med_0}$ for $v = 1, \dots, n$

hypotheses		test-statistics; distribution of the statistics under the null hypothesis	large sample critical values
H_0	H_1		
$x_{med} = x_{med_0}$	$x_{med} \neq x_{med_0}$	$T = \text{number of negative differences } D_v$; binomial distribution with parameters n and $p=0.5$ $\frac{T - n/2}{\sqrt{n/4}}$; asymptotic $N(0,1)$ -distribution	$z_{[\alpha/2]}$, $z_{[1-\alpha/2]}$
$x_{med} \leq x_{med_0}$	$x_{med} > x_{med_0}$		$z_{[1-\alpha]}$
$x_{med} \geq x_{med_0}$	$x_{med} < x_{med_0}$		$z_{[\alpha]}$

Kolmogorov-Smirnov Goodness-of-Fit test

Sample: n independent sample variables X_1, \dots, X_n from a population with unknown *continuous* cumulative distribution function $F(x)$

Symbol: $\hat{F}(x) = \frac{1}{n}$ (number of X_i less than or equal to x) for all $-\infty < x < \infty$ (sample cumulative distribution function)

hypotheses H_0 H_1	test-statistic; distribution of the statistic under the null hypothesis	critical value
$F(x) = F_0(x)$ $F(x) \neq F_0(x)$ $F_0(x)$ is completely specified under the null	$\sup_{-\infty < x < \infty} \hat{F}(x) - F_0(x) $; non-standard distribution , critical values $c_{[1-\alpha]}$ are tabled in literature	$c_{[1-\alpha]}$

Chi-Square Goodness-of-Fit test

Sample: n independent sample variables X_1, \dots, X_n from a population with unknown cumulative distribution function $F(x)$

Symbols: n_1, \dots, n_k = sample frequencies of k classes or categories; $n \cdot p_1, \dots, n \cdot p_k$ = expected sample frequencies under the null

hypotheses H_0 H_1	test-statistic; distribution of the statistic under the null hypothesis	large sample critical value
$F(x) = F_0(x)$ $F(x) \neq F_0(x)$ $F_0(x)$ is completely specified under the null	$\sum_{j=1}^k \frac{(n_j - n \cdot p_j)^2}{n \cdot p_j}$; asymptotic χ^2 -distribution with $k-1$ degrees of freedom	$\chi^2_{[1-\alpha; k-1]}$
H_0 H_1		
$F(x) = F_0(x)$ $F(x) \neq F_0(x)$ r parameters of $F_0(x)$ estimated by maximum-likelihood	$\sum_{j=1}^k \frac{(n_j - n \cdot p_j)^2}{n \cdot p_j}$; asymptotic χ^2 -distribution with $k-r-1$ degrees of freedom	$\chi^2_{[1-\alpha; k-r-1]}$
(for finite sample sizes n the χ^2 -approximations are reliable provided $n \cdot p_j > 5$ for all $j=1, \dots, k$)		

Chi-Square test of independence

Sample: Sample of n independent pairs of sample variables $(X_1, Y_1), \dots, (X_n, Y_n)$ from the discrete population X with m different categories and the discrete population Y with k different categories

Symbols: n_{11}, \dots, n_{mk} = sample frequencies of $m \cdot k$ pairs of categories; $n_{i\bullet}, n_{\bullet j}$ = marginal sample frequencies of single categories; e_{11}, \dots, e_{mk} = expected sample frequencies under the null

hypotheses H_0 H_1	test-statistic; distribution of the statistic under the null hypothesis	large sample critical value
X und Y are independent X und Y are not independent	$\sum_{i=1}^m \sum_{j=1}^k \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$ where $e_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}$ asymptotic χ^2 -distribution with $(m-1)(k-1)$ degrees of freedom	$\chi^2_{[1-\alpha; (m-1)(k-1)]}$

10 Regressionsanalyse

10.1 Lineare Einfachregression

Voraussetzungen des (klassischen) linearen Regressionsmodells

A-1	Verfügbar sind n Messwertpaare zweier metrischer Variablen X und Y . Die Wertepaare fassen wir als Realisierungen von Stichprobenvariablenpaaren (x_v, y_v) ($v=1, \dots, n$) auf. Die zum Regressor X gehörigen Variablen x_1, \dots, x_n sind deterministisch; ihre Werte sind fest vorgegeben und nicht alle gleich, so dass $\sum_{v=1}^n (x_v - \bar{x})^2 > 0$ erfüllt ist. Die zum Regressanden Y gehörigen Variablen y_1, \dots, y_n sind Zufallsvariablen. y_v steht für den potentiellen Messwert des Regressanden Y bei gegebenem Wert x_v .
A-2	Zwischen den Zufallsvariablen y_v und den festen Werten x_v besteht der lineare Zusammenhang $y_v = \beta_0 + \beta_1 x_v + \varepsilon_v \quad (v=1, \dots, n).$
A-3	Die zufälligen Störvariablen $\varepsilon_1, \dots, \varepsilon_n$ sind stochastisch unabhängig und identisch verteilt mit $E(\varepsilon_v) = 0 \quad (v=1, \dots, n),$ $\text{Var}(\varepsilon_v) = \sigma^2 = \text{const.} \quad (v=1, \dots, n).$ $\text{Cov}(\varepsilon_v, \varepsilon_k) = 0. \quad \text{für alle } v \neq k.$
A-4*	$\varepsilon_v \sim N(0, \sigma^2) \quad (v=1, \dots, n).$

KQ- und ML-Schätzfunktionen für die Parameter des linearen Regressionsmodells

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

$$\hat{\beta}_1 = \frac{\sum_{v=1}^n (x_v - \bar{x})(y_v - \bar{y})}{\sum_{v=1}^n (x_v - \bar{x})^2} = \frac{\sum_{v=1}^n x_v y_v - n \bar{x} \bar{y}}{\sum_{v=1}^n x_v^2 - n \bar{x}^2}$$

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{v=1}^n \hat{\varepsilon}_v^2 \quad \text{mit} \quad \hat{\varepsilon}_v = y_v - \hat{\beta}_0 - \hat{\beta}_1 x_v \quad (v=1, \dots, n)$$

Streuungszerlegung

$$\underbrace{\sum_{v=1}^n (y_v - \bar{y})^2}_{TSS} = \underbrace{\sum_{v=1}^n (y_v - \hat{y}_v)^2}_{RSS} + \underbrace{\sum_{v=1}^n (\hat{y}_v - \bar{y})^2}_{ESS} = \underbrace{\sum_{v=1}^n \hat{\varepsilon}_v^2}_{RSS} + \underbrace{\sum_{v=1}^n (\hat{y}_v - \bar{y})^2}_{ESS}$$

Bestimmtheitsmaß

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} = r_{xy}^2 \quad 0 \leq R^2 \leq 1$$

Standardfehler (Standardabweichungen) der geschätzten Regressionskoeffizienten

$$\hat{\sigma}_{\hat{\beta}_0} = \sqrt{\frac{\hat{\sigma}^2 \cdot \sum_{v=1}^n x_v^2}{n \cdot \sum_{v=1}^n (x_v - \bar{x})^2}} = \hat{\sigma} \cdot \sqrt{\frac{\frac{1}{n} \sum_{v=1}^n x_v^2}{\sum_{v=1}^n x_v^2 - n\bar{x}^2}}$$

$$\hat{\sigma}_{\hat{\beta}_1} = \sqrt{\frac{\hat{\sigma}^2}{\sum_{v=1}^n (x_v - \bar{x})^2}} = \hat{\sigma} \cdot \sqrt{\frac{1}{\sum_{v=1}^n x_v^2 - n\bar{x}^2}}$$

Konfidenzintervalle zum Konfidenzniveau $1 - \alpha$ für die Regressionskoeffizienten

$$\left[\hat{\beta}_0 - t_{[1-\alpha/2; n-2]} \cdot \hat{\sigma}_{\hat{\beta}_0}, \hat{\beta}_0 + t_{[1-\alpha/2; n-2]} \cdot \hat{\sigma}_{\hat{\beta}_0} \right]$$

$$\left[\hat{\beta}_1 - t_{[1-\alpha/2; n-2]} \cdot \hat{\sigma}_{\hat{\beta}_1}, \hat{\beta}_1 + t_{[1-\alpha/2; n-2]} \cdot \hat{\sigma}_{\hat{\beta}_1} \right]$$

Hypothesentests für die Regressionskoeffizienten

Hypothesen		Teststatistik und deren Verteilung unter Gültigkeit von H_0	kritische Werte
H_0	H_1		
$\beta_0 = b_0$	$\beta_0 \neq b_0$	$T_1 = \frac{\hat{\beta}_0 - b_0}{\hat{\sigma}_{\hat{\beta}_0}} \sim t(n-2)$	$\pm t_{[1-\alpha/2; n-2]}$
$\beta_0 \leq b_0$	$\beta_0 > b_0$		$t_{[1-\alpha; n-2]}$
$\beta_0 \geq b_0$	$\beta_0 < b_0$		$t_{[\alpha; n-2]}$
$\beta_1 = b_1$	$\beta_1 \neq b_1$	$T_2 = \frac{\hat{\beta}_1 - b_1}{\hat{\sigma}_{\hat{\beta}_1}} \sim t(n-2)$	$\pm t_{[1-\alpha/2; n-2]}$
$\beta_1 \leq b_1$	$\beta_1 > b_1$		$t_{[1-\alpha; n-2]}$
$\beta_1 \geq b_1$	$\beta_1 < b_1$		$t_{[\alpha; n-2]}$

Prognosefunktion (Punktprognose) für y_0

$$\hat{y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0$$

Konfidenzintervall (Prognoseintervall) für y_0 zum Konfidenzniveau $1 - \alpha$

$$\left[\hat{y}_0 - t_{[1-\alpha/2; n-2]} \cdot \hat{\sigma}_{\hat{y}_0 - y_0}, \hat{y}_0 + t_{[1-\alpha/2; n-2]} \cdot \hat{\sigma}_{\hat{y}_0 - y_0} \right]$$

mit

$$\hat{\sigma}_{\hat{y}_0 - y_0} = \sqrt{\hat{\sigma}^2 \cdot \left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{v=1}^n (x_v - \bar{x})^2} \right)} \quad \text{und} \quad \sum_{v=1}^n (x_v - \bar{x})^2 = \sum_{v=1}^n x_v^2 - n\bar{x}^2$$

10.2 Lineare Mehrfachregression

Voraussetzungen des (klassischen) linearen Regressionsmodells

A-1	$y = X\beta + \varepsilon$
	mit $y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, $X = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix}$, $\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}$, $\varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$.
A-2	X ist eine nicht-stochastische (feste) Matrix. Es gilt $rg(X) = p + 1$; d.h. X hat vollen Spaltenrang und ist daher spaltenregulär.
A-3	Die Störungen ε sind zufällige Größen mit $E(\varepsilon) = \mathbf{0}$ und $Cov(\varepsilon) = \sigma^2 I$.
A-4*	$\varepsilon \sim N_n(\mathbf{0}, \sigma^2 I)$.

KQ- und ML-Schätzfunktionen für die Parameter des linearen Regressionsmodells

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

$$\hat{\sigma}^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n-p-1} \quad \text{mit} \quad \hat{\varepsilon} = y - X\hat{\beta} = (I - H)y \quad \text{und} \quad H = X(X'X)^{-1} X'$$

Streuungszerlegung

$$\underbrace{\sum_{v=1}^n (y_v - \bar{y})^2}_{TSS} = \underbrace{\sum_{v=1}^n (y_v - \hat{y}_v)^2}_{RSS} + \underbrace{\sum_{v=1}^n (\hat{y}_v - \bar{y})^2}_{ESS} = \underbrace{\sum_{v=1}^n \hat{\varepsilon}_v^2}_{RSS} + \underbrace{\sum_{v=1}^n (\hat{y}_v - \bar{y})^2}_{ESS}$$

Bestimmtheitsmaße

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} \quad 0 \leq R^2 \leq 1$$

$$\bar{R}^2 = 1 - \frac{\frac{RSS}{n-p-1}}{\frac{TSS}{n-1}} = 1 - \frac{n-1}{n-p-1} \cdot (1 - R^2)$$

Kovarianzmatrix und Standardfehler der geschätzten Regressionskoeffizienten

$$\hat{\Sigma}_{\hat{\beta}} = \hat{\sigma}^2 (X'X)^{-1}$$

$$\hat{\sigma}_{\hat{\beta}_i} = \hat{\sigma} \cdot \sqrt{a_{jj}} \quad \text{mit} \quad j = i + 1 \quad \text{und} \quad i = 0, 1, \dots, p$$

a_{jj} kennzeichnet das j -te Diagonalelement der Matrix $(X'X)^{-1}$.

Konfidenzintervalle zum Konfidenzniveau $1 - \alpha$ für den i -ten Regressionskoeffizienten

$$\left[\hat{\beta}_i - t_{[1-\alpha/2; n-p-1]} \cdot \hat{\sigma}_{\hat{\beta}_i}, \hat{\beta}_i + t_{[1-\alpha/2; n-p-1]} \cdot \hat{\sigma}_{\hat{\beta}_i} \right]$$

Hypothesentests für den i -ten Regressionskoeffizienten

Hypothesen		Teststatistik und deren Verteilung unter Gültigkeit von H_0	kritische Werte
H_0	H_1		
$\beta_i = b$	$\beta_i \neq b$	$T_1 = \frac{\hat{\beta}_i - b}{\hat{\sigma}_{\hat{\beta}_i}} \sim t(n - p - 1)$	$\pm t_{[1-\alpha/2; n-p-1]}$
$\beta_i \leq b$	$\beta_i > b$		$t_{[1-\alpha; n-p-1]}$
$\beta_i \geq b$	$\beta_i < b$		$t_{[\alpha; n-p-1]}$

Prognosefunktion (Punktprognose) für y_0

$$\hat{y}_0 = \mathbf{x}'_0 \hat{\boldsymbol{\beta}} \quad \text{mit} \quad \mathbf{x}'_0 = (1, x_{01}, \dots, x_{0p})$$

Konfidenzintervall (Prognoseintervall) für y_0 zum Konfidenzniveau $1 - \alpha$

$$\left[\hat{y}_0 - t_{[1-\alpha/2; n-p-1]} \cdot \sqrt{\hat{\sigma}^2 + \hat{\sigma}^2 \mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0}, \hat{y}_0 + t_{[1-\alpha/2; n-p-1]} \cdot \sqrt{\hat{\sigma}^2 + \hat{\sigma}^2 \mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0} \right]$$

Binomialverteilung

Verteilungsfunktionen $F(x|n, p)$ ausgewählter Binomialverteilungen

$$[F(x|n, p) = 1 - F(n - x - 1|n, 1 - p) \text{ für } p > 0.5]$$

n	x	p										
		0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	
2	0	0.9025	0.8100	0.7225	0.6400	0.5625	0.4900	0.4225	0.3600	0.3025	0.2500	
	1	0.9975	0.9900	0.9775	0.9600	0.9375	0.9100	0.8775	0.8400	0.7975	0.7500	
	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	
3	0	0.8574	0.7290	0.6141	0.5120	0.4219	0.3430	0.2746	0.2160	0.1664	0.1250	
	1	0.9927	0.9720	0.9393	0.8960	0.8438	0.7840	0.7183	0.6480	0.5748	0.5000	
	2	0.9999	0.9990	0.9966	0.9920	0.9844	0.9730	0.9571	0.9360	0.9089	0.8750	
3	3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	
	4	0	0.8145	0.6561	0.5220	0.4096	0.3164	0.2401	0.1785	0.1296	0.0915	0.0625
	1	0.9860	0.9477	0.8905	0.8192	0.7383	0.6517	0.5630	0.4752	0.3910	0.3125	
4	2	0.9995	0.9963	0.9880	0.9728	0.9492	0.9163	0.8735	0.8208	0.7585	0.6875	
	3	1.0000	0.9999	0.9995	0.9984	0.9961	0.9919	0.9850	0.9744	0.9590	0.9375	
	4		1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	
5	0	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.1160	0.0778	0.0503	0.0313	
	1	0.9774	0.9185	0.8352	0.7373	0.6328	0.5282	0.4284	0.3370	0.2562	0.1875	
	2	0.9988	0.9914	0.9734	0.9421	0.8965	0.8369	0.7648	0.6826	0.5931	0.5000	
5	3	1.0000	0.9995	0.9978	0.9933	0.9844	0.9692	0.9460	0.9130	0.8688	0.8125	
	4		1.0000	0.9999	0.9997	0.9990	0.9976	0.9947	0.9898	0.9815	0.9687	
	5			1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	
6	0	0.7351	0.5314	0.3771	0.2621	0.1780	0.1176	0.0754	0.0467	0.0277	0.0156	
	1	0.9672	0.8857	0.7765	0.6554	0.5339	0.4202	0.3191	0.2333	0.1636	0.1094	
	2	0.9978	0.9842	0.9527	0.9011	0.8306	0.7443	0.6471	0.5443	0.4415	0.3438	
6	3	0.9999	0.9987	0.9941	0.9830	0.9624	0.9295	0.8826	0.8208	0.7447	0.6562	
	4	1.0000	0.9999	0.9996	0.9984	0.9954	0.9891	0.9777	0.9590	0.9308	0.8906	
	5		1.0000	1.0000	0.9999	0.9998	0.9993	0.9982	0.9959	0.9917	0.9844	
6	6				1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	
	7	0	0.6983	0.4783	0.3206	0.2097	0.1335	0.0824	0.0490	0.0280	0.0152	0.0078
	1	0.9556	0.8503	0.7166	0.5767	0.4449	0.3294	0.2338	0.1586	0.1024	0.0625	
7	2	0.9962	0.9743	0.9262	0.8520	0.7564	0.6471	0.5323	0.4199	0.3164	0.2266	
	3	0.9998	0.9973	0.9879	0.9667	0.9294	0.8740	0.8002	0.7102	0.6083	0.5000	
	4	1.0000	0.9998	0.9988	0.9953	0.9871	0.9712	0.9444	0.9037	0.8471	0.7734	
7	5		1.0000	0.9999	0.9996	0.9987	0.9962	0.9910	0.9812	0.9643	0.9375	
	6			1.0000	1.0000	0.9999	0.9998	0.9994	0.9984	0.9963	0.9922	
	7					1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	
8	0	0.6634	0.4305	0.2725	0.1678	0.1001	0.0577	0.0319	0.0168	0.0084	0.0039	
	1	0.9428	0.8131	0.6572	0.5033	0.3671	0.2553	0.1691	0.1064	0.0632	0.0352	
	2	0.9942	0.9619	0.8948	0.7969	0.6785	0.5518	0.4278	0.3154	0.2201	0.1445	
8	3	0.9996	0.9950	0.9786	0.9437	0.8862	0.8059	0.7064	0.5941	0.4770	0.3633	
	4	1.0000	0.9996	0.9971	0.9896	0.9727	0.9420	0.8939	0.8263	0.7396	0.6367	
	5		1.0000	0.9998	0.9988	0.9958	0.9887	0.9747	0.9502	0.9115	0.8555	
8	6			1.0000	0.9999	0.9996	0.9987	0.9964	0.9915	0.9819	0.9648	
	7				1.0000	1.0000	0.9999	0.9998	0.9993	0.9983	0.9961	
	8						1.0000	1.0000	1.0000	1.0000	1.0000	
9	0	0.6302	0.3874	0.2316	0.1342	0.0751	0.0404	0.0207	0.0101	0.0046	0.0020	
	1	0.9288	0.7748	0.5995	0.4362	0.3003	0.1960	0.1211	0.0705	0.0385	0.0195	
	2	0.9916	0.9470	0.8591	0.7382	0.6007	0.4628	0.3373	0.2318	0.1495	0.0898	
9	3	0.9994	0.9917	0.9661	0.9144	0.8343	0.7297	0.6089	0.4826	0.3614	0.2539	
	4	1.0000	0.9991	0.9944	0.9804	0.9511	0.9012	0.8283	0.7334	0.6214	0.5000	
	5		0.9999	0.9994	0.9969	0.9900	0.9747	0.9464	0.9006	0.8342	0.7461	
9	6		1.0000	1.0000	0.9997	0.9987	0.9957	0.9888	0.9750	0.9502	0.9102	
	7				1.0000	0.9999	0.9996	0.9986	0.9962	0.9909	0.9805	
	8					1.0000	1.0000	0.9999	0.9997	0.9992	0.9980	
10	0	0.5987	0.3487	0.1969	0.1074	0.0563	0.0283	0.0135	0.0060	0.0025	0.0010	
	1	0.9139	0.7361	0.5443	0.3758	0.2440	0.1493	0.0860	0.0464	0.0233	0.0107	
	2	0.9885	0.9298	0.8202	0.6778	0.5256	0.3828	0.2616	0.1673	0.0996	0.0547	
10	3	0.9990	0.9872	0.9500	0.8791	0.7759	0.6496	0.5138	0.3823	0.2660	0.1719	
	4	0.9999	0.9984	0.9901	0.9672	0.9219	0.8497	0.7515	0.6331	0.5044	0.3770	
	5	1.0000	0.9999	0.9986	0.9936	0.9803	0.9527	0.9051	0.8338	0.7384	0.6230	
10	6		1.0000	0.9999	0.9991	0.9965	0.9894	0.9740	0.9452	0.8980	0.8281	
	7			1.0000	0.9999	0.9996	0.9984	0.9952	0.9877	0.9726	0.9453	
	8				1.0000	1.0000	0.9999	0.9995	0.9983	0.9955	0.9893	
10	9						1.0000	1.0000	0.9999	0.9997	0.9990	
	10								1.0000	1.0000	1.0000	

Fortsetzung 1 Binomialverteilung

n	x	p									
		0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
11	0	0.5688	0.3138	0.1673	0.0859	0.0422	0.0198	0.0088	0.0036	0.0014	0.0005
	1	0.8981	0.6974	0.4922	0.3221	0.1971	0.1130	0.0606	0.0302	0.0139	0.0059
	2	0.9848	0.9104	0.7788	0.6174	0.4552	0.3127	0.2001	0.1189	0.0652	0.0327
	3	0.9984	0.9815	0.9306	0.8389	0.7133	0.5696	0.4256	0.2963	0.1911	0.1133
	4	0.9999	0.9972	0.9841	0.9496	0.8854	0.7897	0.6683	0.5328	0.3971	0.2744
	5	1.0000	0.9997	0.9973	0.9883	0.9657	0.9218	0.8513	0.7535	0.6331	0.5000
	6		1.0000	0.9997	0.9980	0.9924	0.9784	0.9499	0.9006	0.8262	0.7256
	7			1.0000	0.9998	0.9988	0.9957	0.9878	0.9707	0.9390	0.8867
	8				1.0000	0.9999	0.9994	0.9980	0.9941	0.9852	0.9673
	9					1.0000	1.0000	0.9998	0.9993	0.9978	0.9941
	10							1.0000	1.0000	0.9998	0.9995
11									1.0000	1.0000	
12	0	0.5404	0.2824	0.1422	0.0687	0.0317	0.0138	0.0057	0.0022	0.0008	0.0002
	1	0.8816	0.6590	0.4435	0.2749	0.1584	0.0850	0.0424	0.0196	0.0083	0.0032
	2	0.9804	0.8891	0.7358	0.5583	0.3907	0.2528	0.1513	0.0834	0.0421	0.0193
	3	0.9978	0.9744	0.9078	0.7946	0.6488	0.4925	0.3467	0.2253	0.1345	0.0730
	4	0.9998	0.9957	0.9761	0.9274	0.8424	0.7237	0.5833	0.4382	0.3044	0.1938
	5	1.0000	0.9995	0.9954	0.9806	0.9456	0.8822	0.7873	0.6652	0.5269	0.3872
	6		0.9999	0.9993	0.9961	0.9857	0.9614	0.9154	0.8418	0.7393	0.6128
	7		1.0000	0.9999	0.9994	0.9972	0.9905	0.9745	0.9427	0.8883	0.8062
	8			1.0000	0.9999	0.9996	0.9983	0.9944	0.9847	0.9644	0.9270
	9				1.0000	1.0000	0.9998	0.9992	0.9972	0.9921	0.9807
	10						1.0000	0.9999	0.9997	0.9989	0.9968
	11							1.0000	1.0000	0.9999	0.9998
12									1.0000	1.0000	
13	0	0.5133	0.2542	0.1209	0.0550	0.0238	0.0097	0.0037	0.0013	0.0004	0.0001
	1	0.8646	0.6213	0.3983	0.2336	0.1267	0.0637	0.0296	0.0126	0.0049	0.0017
	2	0.9755	0.8661	0.6920	0.5017	0.3326	0.2025	0.1132	0.0579	0.0269	0.0112
	3	0.9969	0.9658	0.8820	0.7473	0.5843	0.4206	0.2783	0.1686	0.0929	0.0461
	4	0.9997	0.9935	0.9658	0.9009	0.7940	0.6543	0.5005	0.3530	0.2279	0.1334
	5	1.0000	0.9991	0.9925	0.9700	0.9198	0.8346	0.7159	0.5744	0.4268	0.2905
	6		0.9999	0.9987	0.9930	0.9757	0.9376	0.8705	0.7712	0.6437	0.5000
	7		1.0000	0.9998	0.9988	0.9944	0.9818	0.9538	0.9023	0.8212	0.7095
	8			1.0000	0.9998	0.9990	0.9960	0.9874	0.9679	0.9302	0.8666
	9				1.0000	0.9999	0.9993	0.9975	0.9922	0.9797	0.9539
	10					1.0000	0.9999	0.9997	0.9987	0.9959	0.9888
	11						1.0000	1.0000	0.9999	0.9995	0.9983
	12								1.0000	1.0000	0.9999
13										1.0000	
14	0	0.4877	0.2288	0.1028	0.0440	0.0178	0.0068	0.0024	0.0008	0.0002	0.0001
	1	0.8470	0.5846	0.3567	0.1979	0.1010	0.0475	0.0205	0.0081	0.0029	0.0009
	2	0.9699	0.8416	0.6479	0.4481	0.2811	0.1608	0.0839	0.0398	0.0170	0.0065
	3	0.9958	0.9559	0.8535	0.6982	0.5213	0.3552	0.2205	0.1243	0.0632	0.0287
	4	0.9996	0.9908	0.9533	0.8702	0.7415	0.5842	0.4227	0.2793	0.1672	0.0898
	5	1.0000	0.9985	0.9885	0.9561	0.8883	0.7805	0.6405	0.4859	0.3373	0.2120
	6		0.9998	0.9978	0.9884	0.9617	0.9067	0.8164	0.6925	0.5461	0.3953
	7		1.0000	0.9997	0.9976	0.9897	0.9685	0.9247	0.8499	0.7414	0.6047
	8			1.0000	0.9996	0.9978	0.9917	0.9757	0.9417	0.8811	0.7880
	9				1.0000	0.9997	0.9983	0.9940	0.9825	0.9574	0.9102
	10					1.0000	0.9998	0.9989	0.9961	0.9886	0.9713
	11						1.0000	0.9999	0.9994	0.9978	0.9935
	12							1.0000	0.9999	0.9997	0.9991
	13								1.0000	1.0000	0.9999
14										1.0000	
15	0	0.4633	0.2059	0.0873	0.0352	0.0134	0.0047	0.0016	0.0005	0.0001	0.0000
	1	0.8290	0.5490	0.3186	0.1671	0.0802	0.0353	0.0142	0.0052	0.0017	0.0005
	2	0.9638	0.8159	0.6042	0.3980	0.2361	0.1268	0.0617	0.0271	0.0106	0.0037
	3	0.9945	0.9444	0.8227	0.6482	0.4613	0.2969	0.1727	0.0905	0.0424	0.0176
	4	0.9994	0.9873	0.9383	0.8358	0.6865	0.5155	0.3519	0.2173	0.1204	0.0592
	5	0.9999	0.9978	0.9832	0.9389	0.8516	0.7216	0.5643	0.4032	0.2608	0.1509
	6	1.0000	0.9997	0.9964	0.9819	0.9434	0.8689	0.7548	0.6098	0.4522	0.3036
	7		1.0000	0.9994	0.9958	0.9827	0.9500	0.8868	0.7869	0.6535	0.5000
	8			0.9999	0.9992	0.9958	0.9848	0.9578	0.9050	0.8182	0.6964
	9			1.0000	0.9999	0.9992	0.9963	0.9876	0.9662	0.9231	0.8491
	10				1.0000	0.9999	0.9993	0.9972	0.9907	0.9745	0.9408
	11					1.0000	0.9999	0.9995	0.9981	0.9937	0.9824
	12						1.0000	0.9999	0.9997	0.9989	0.9963
	13							1.0000	1.0000	0.9999	0.9995
	14									1.0000	1.0000
15										1.0000	

Fortsetzung 2 Binomialverteilung

<i>n</i>	<i>x</i>	<i>p</i>									
		0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
20	0	0.3585	0.1216	0.0388	0.0115	0.0032	0.0008	0.0002	0.0000	0.0000	
	1	0.7358	0.3917	0.1756	0.0692	0.0243	0.0076	0.0021	0.0005	0.0001	0.0000
	2	0.9245	0.6769	0.4049	0.2061	0.0913	0.0355	0.0121	0.0036	0.0009	0.0002
	3	0.9841	0.8670	0.6477	0.4114	0.2252	0.1071	0.0444	0.0160	0.0049	0.0013
	4	0.9974	0.9568	0.8298	0.6296	0.4148	0.2375	0.1182	0.0510	0.0189	0.0059
	5	0.9997	0.9887	0.9327	0.8042	0.6172	0.4164	0.2454	0.1256	0.0553	0.0207
	6	1.0000	0.9976	0.9781	0.9133	0.7858	0.6080	0.4166	0.2500	0.1299	0.0577
	7		0.9996	0.9941	0.9679	0.8982	0.7723	0.6010	0.4159	0.2520	0.1316
	8		0.9999	0.9987	0.9900	0.9591	0.8867	0.7624	0.5956	0.4143	0.2517
	9		1.0000	0.9998	0.9974	0.9861	0.9520	0.8782	0.7553	0.5914	0.4119
	10			1.0000	0.9994	0.9961	0.9829	0.9468	0.8725	0.7507	0.5881
	11				0.9999	0.9991	0.9949	0.9804	0.9435	0.8692	0.7483
	12				1.0000	0.9998	0.9987	0.9940	0.9790	0.9420	0.8684
	13					1.0000	0.9997	0.9985	0.9935	0.9786	0.9423
	14						1.0000	0.9997	0.9984	0.9936	0.9793
	15							1.0000	0.9997	0.9985	0.9941
	16								1.0000	0.9997	0.9987
	17									1.0000	0.9998
	18										1.0000
	19										
20											
25	0	0.2774	0.0718	0.0172	0.0038	0.0008	0.0001	0.0000			
	1	0.6424	0.2712	0.0931	0.0274	0.0070	0.0016	0.0003	0.0000	0.0000	
	2	0.8729	0.5371	0.2537	0.0982	0.0321	0.0090	0.0021	0.0004	0.0001	
	3	0.9659	0.7636	0.4711	0.2340	0.0962	0.0332	0.0097	0.0024	0.0005	0.0000
	4	0.9928	0.9020	0.6821	0.4207	0.2137	0.0905	0.0321	0.0095	0.0023	0.0005
	5	0.9988	0.9666	0.8385	0.6167	0.3783	0.1935	0.0826	0.0294	0.0086	0.0020
	6	0.9998	0.9905	0.9305	0.7800	0.5611	0.3407	0.1734	0.0736	0.0258	0.0073
	7	1.0000	0.9977	0.9745	0.8909	0.7265	0.5118	0.3061	0.1536	0.0639	0.0216
	8		0.9995	0.9920	0.9532	0.8506	0.6769	0.4668	0.2735	0.1340	0.0539
	9		0.9999	0.9979	0.9827	0.9287	0.8106	0.6303	0.4246	0.2424	0.1148
	10		1.0000	0.9995	0.9944	0.9703	0.9022	0.7712	0.5858	0.3843	0.2122
	11			0.9999	0.9985	0.9893	0.9558	0.8746	0.7323	0.5426	0.3450
	12			1.0000	0.9996	0.9966	0.9825	0.9396	0.8462	0.6937	0.5000
	13				0.9999	0.9991	0.9940	0.9745	0.9222	0.8173	0.6550
	14				1.0000	0.9998	0.9982	0.9907	0.9656	0.9040	0.7878
	15					1.0000	0.9995	0.9971	0.9868	0.9560	0.8852
	16						0.9999	0.9992	0.9957	0.9826	0.9461
	17						1.0000	0.9998	0.9988	0.9942	0.9784
	18							1.0000	0.9997	0.9984	0.9927
	19								0.9999	0.9996	0.9980
	20								1.0000	0.9999	0.9995
	21									1.0000	0.9999
	22										1.0000

Poisson-Verteilung

Verteilungsfunktionen ausgewählter Poisson-Verteilungen

x	λ									
	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
0	0.6065	0.3679	0.2231	0.1353	0.0821	0.0498	0.0302	0.0183	0.0111	0.0067
1	0.9098	0.7358	0.5578	0.4060	0.2873	0.1991	0.1359	0.0916	0.0611	0.0404
2	0.9856	0.9197	0.8088	0.6767	0.5438	0.4232	0.3208	0.2381	0.1736	0.1247
3	0.9982	0.9810	0.9344	0.8571	0.7576	0.6472	0.5366	0.4335	0.3423	0.2650
4	0.9998	0.9963	0.9814	0.9473	0.8912	0.8153	0.7254	0.6288	0.5321	0.4405
5	1.0000	0.9994	0.9955	0.9834	0.9580	0.9161	0.8576	0.7851	0.7029	0.6160
6		0.9999	0.9991	0.9955	0.9858	0.9665	0.9347	0.8893	0.8311	0.7622
7		1.0000	0.9998	0.9989	0.9958	0.9881	0.9733	0.9489	0.9134	0.8666
8			1.0000	0.9998	0.9989	0.9962	0.9901	0.9786	0.9597	0.9319
9				1.0000	0.9997	0.9989	0.9967	0.9919	0.9829	0.9682
10					0.9999	0.9997	0.9990	0.9972	0.9933	0.9863
11					1.0000	0.9999	0.9997	0.9991	0.9976	0.9945
12						1.0000	0.9999	0.9997	0.9992	0.9980
13							1.0000	0.9999	0.9997	0.9993
14								1.0000	0.9999	0.9998
15									1.0000	0.9999
16										1.0000

x	λ									
	6.0	7.0	8.0	9.0	10.0	11.0	12.0	13.0	14.0	15.0
0	0.0025	0.0009	0.0003	0.0001						
1	0.0174	0.0073	0.0030	0.0012	0.0005	0.0002				
2	0.0620	0.0296	0.0138	0.0062	0.0028	0.0012	0.0005	0.0002	0.0001	
3	0.1512	0.0818	0.0424	0.0212	0.0103	0.0049	0.0023	0.0011	0.0005	0.0002
4	0.2851	0.1730	0.0996	0.0550	0.0293	0.0151	0.0076	0.0037	0.0018	0.0009
5	0.4457	0.3007	0.1912	0.1157	0.0671	0.0375	0.0203	0.0107	0.0055	0.0028
6	0.6063	0.4497	0.3134	0.2068	0.1301	0.0786	0.0458	0.0259	0.0142	0.0076
7	0.7440	0.5987	0.4530	0.3239	0.2202	0.1432	0.0895	0.0540	0.0316	0.0180
8	0.8472	0.7291	0.5925	0.4557	0.3328	0.2320	0.1550	0.0998	0.0621	0.0375
9	0.9161	0.8305	0.7166	0.5874	0.4579	0.3405	0.2424	0.1658	0.1094	0.0699
10	0.9574	0.9015	0.8159	0.7060	0.5830	0.4599	0.3472	0.2517	0.1757	0.1185
11	0.9799	0.9467	0.8881	0.8030	0.6968	0.5793	0.4616	0.3532	0.2600	0.1848
12	0.9912	0.9730	0.9362	0.8758	0.7916	0.6887	0.5760	0.4631	0.3585	0.2676
13	0.9964	0.9872	0.9658	0.9261	0.8645	0.7813	0.6815	0.5730	0.4644	0.3632
14	0.9986	0.9943	0.9827	0.9585	0.9165	0.8540	0.7720	0.6751	0.5704	0.4657
15	0.9995	0.9976	0.9918	0.9780	0.9513	0.9074	0.8444	0.7636	0.6694	0.5681
16	0.9998	0.9990	0.9963	0.9889	0.9730	0.9441	0.8987	0.8355	0.7559	0.6641
17	0.9999	0.9996	0.9984	0.9947	0.9857	0.9678	0.9370	0.8905	0.8272	0.7489
18	1.0000	0.9999	0.9993	0.9976	0.9928	0.9823	0.9626	0.9302	0.8826	0.8195
19		1.0000	0.9997	0.9989	0.9965	0.9907	0.9787	0.9573	0.9235	0.8752
20			0.9999	0.9996	0.9984	0.9953	0.9884	0.9750	0.9521	0.9170
21			1.0000	0.9998	0.9993	0.9977	0.9939	0.9859	0.9712	0.9469
22				0.9999	0.9997	0.9990	0.9970	0.9924	0.9833	0.9673
23				1.0000	0.9999	0.9995	0.9985	0.9960	0.9907	0.9805
24					1.0000	0.9998	0.9993	0.9980	0.9950	0.9888
25						0.9999	0.9997	0.9990	0.9974	0.9938
26						1.0000	0.9999	0.9995	0.9987	0.9967
27							0.9999	0.9998	0.9994	0.9983
28							1.0000	0.9999	0.9997	0.9991
29								1.0000	0.9999	0.9996
30									0.9999	0.9998
31									1.0000	0.9999
32										1.0000

Normalverteilung

Verteilungsfunktion $\Phi(z)$ der Standardnormalverteilung

$$[\Phi(-z) = 1 - \Phi(z)]$$

Beispiel: $\Phi(2.36) = 0.99087$; Man liest diesen Wert im Schnittpunkt von Zeile 2.3 mit Spalte 0.06 ab.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.50000	.50399	.50798	.51197	.51596	.51994	.52393	.52791	.53189	.53586
0.1	.53983	.54380	.54776	.55172	.55568	.55962	.56356	.56750	.57143	.57535
0.2	.57926	.58317	.58707	.59096	.59484	.59871	.60257	.60642	.61027	.61410
0.3	.61792	.62172	.62552	.62931	.63308	.63684	.64058	.64431	.64803	.65174
0.4	.65543	.65910	.66276	.66641	.67004	.67365	.67725	.68083	.68439	.68794
0.5	.69147	.69498	.69847	.70195	.70541	.70885	.71227	.71567	.71905	.72241
0.6	.72575	.72907	.73238	.73566	.73892	.74216	.74538	.74858	.75175	.75491
0.7	.75804	.76115	.76424	.76731	.77036	.77338	.77638	.77936	.78231	.78524
0.8	.78815	.79103	.79390	.79674	.79955	.80234	.80511	.80785	.81058	.81327
0.9	.81594	.81859	.82122	.82382	.82640	.82895	.83148	.83398	.83646	.83892
1.0	.84135	.84376	.84614	.84850	.85084	.85315	.85543	.85770	.85993	.86215
1.1	.86434	.86651	.86865	.87077	.87286	.87493	.87698	.87900	.88100	.88298
1.2	.88494	.88687	.88877	.89066	.89252	.89436	.89617	.89796	.89973	.90148
1.3	.90320	.90491	.90659	.90825	.90988	.91150	.91309	.91466	.91621	.91774
1.4	.91925	.92074	.92220	.92365	.92507	.92648	.92786	.92922	.93057	.93189
1.5	.93320	.93448	.93575	.93700	.93822	.93943	.94063	.94180	.94295	.94409
1.6	.94521	.94631	.94739	.94845	.94950	.95053	.95155	.95255	.95353	.95449
1.7	.95544	.95637	.95729	.95819	.95908	.95995	.96080	.96164	.96247	.96328
1.8	.96407	.96486	.96563	.96638	.96712	.96785	.96856	.96926	.96995	.97063
1.9	.97129	.97194	.97258	.97320	.97382	.97442	.97501	.97559	.97615	.97671
2.0	.97725	.97779	.97831	.97883	.97933	.97982	.98031	.98078	.98124	.98170
2.1	.98214	.98258	.98300	.98342	.98383	.98423	.98462	.98500	.98538	.98574
2.2	.98610	.98645	.98680	.98713	.98746	.98778	.98809	.98840	.98870	.98899
2.3	.98928	.98956	.98983	.99010	.99036	.99062	.99087	.99111	.99135	.99158
2.4	.99181	.99203	.99224	.99246	.99266	.99286	.99306	.99325	.99344	.99362
2.5	.99380	.99397	.99414	.99430	.99446	.99462	.99477	.99492	.99506	.99521
2.6	.99534	.99548	.99561	.99574	.99586	.99598	.99610	.99621	.99632	.99643
2.7	.99654	.99664	.99674	.99684	.99693	.99703	.99711	.99720	.99729	.99737
2.8	.99745	.99753	.99760	.99768	.99775	.99782	.99789	.99795	.99802	.99808
2.9	.99814	.99820	.99825	.99831	.99836	.99842	.99847	.99852	.99856	.99861
3.0	.99866	.99870	.99874	.99878	.99882	.99886	.99890	.99893	.99897	.99900
3.1	.99904	.99907	.99910	.99913	.99916	.99919	.99922	.99924	.99927	.99929
3.2	.99932	.99934	.99936	.99939	.99941	.99943	.99945	.99947	.99949	.99950
3.3	.99952	.99954	.99955	.99957	.99959	.99960	.99962	.99963	.99964	.99966
3.4	.99967	.99968	.99969	.99970	.99971	.99972	.99973	.99974	.99975	.99976
3.5	.99977	.99978	.99979	.99980	.99981	.99982	.99983	.99984	.99985	.99986

Chi-Quadrat-Verteilung

p -Quantile $\chi^2_{[p; fg]}$ der Chi-Quadrat-Verteilung mit fg Freiheitsgraden

[Approximation: $\chi^2_{[p; fg]} \approx \frac{1}{2} \cdot (z_{[p]} + \sqrt{2fg - 1})^2$ für $fg > 30$]

Beispiel: $\chi^2_{[0.95; 10]} = 18.31$

fg	p												
	0.005	0.010	0.025	0.05	0.1	0.2	0.5	0.8	0.9	0.95	0.975	0.990	0.995
1	0.000	0.000	0.001	0.004	0.016	0.064	0.455	1.642	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	0.446	1.386	3.219	4.605	5.991	7.378	9.210	10.60
3	0.072	0.115	0.216	0.352	0.584	1.005	2.366	4.642	6.251	7.815	9.348	11.34	12.84
4	0.207	0.297	0.484	0.711	1.064	1.649	3.357	5.989	7.779	9.488	11.14	13.28	14.86
5	0.412	0.554	0.831	1.145	1.610	2.343	4.351	7.289	9.236	11.07	12.83	15.09	16.75
6	0.676	0.872	1.237	1.635	2.204	3.070	5.348	8.558	10.64	12.59	14.45	16.81	18.55
7	0.989	1.239	1.690	2.167	2.833	3.822	6.346	9.803	12.02	14.07	16.01	18.48	20.28
8	1.345	1.647	2.180	2.733	3.490	4.594	7.344	11.03	13.36	15.51	17.53	20.09	21.95
9	1.735	2.088	2.700	3.325	4.168	5.380	8.343	12.24	14.68	16.92	19.02	21.67	23.59
10	2.156	2.558	3.247	3.940	4.865	6.179	9.342	13.44	15.99	18.31	20.48	23.21	25.19
11	2.603	3.054	3.816	4.575	5.578	6.989	10.34	14.63	17.28	19.68	21.92	24.72	26.76
12	3.074	3.571	4.404	5.226	6.304	7.807	11.34	15.81	18.55	21.03	23.34	26.22	28.30
13	3.565	4.107	5.009	5.892	7.042	8.634	12.34	16.98	19.81	22.36	24.74	27.69	29.82
14	4.075	4.660	5.629	6.571	7.790	9.467	13.34	18.15	21.06	23.68	26.12	29.14	31.32
15	4.601	5.229	6.262	7.261	8.547	10.31	14.34	19.31	22.31	25.00	27.49	30.58	32.80
16	5.142	5.812	6.908	7.962	9.312	11.15	15.34	20.47	23.54	26.30	28.85	32.00	34.27
17	5.697	6.408	7.564	8.672	10.09	12.00	16.34	21.61	24.77	27.59	30.19	33.41	35.72
18	6.265	7.015	8.231	9.390	10.86	12.86	17.34	22.76	25.99	28.87	31.53	34.81	37.16
19	6.844	7.633	8.907	10.12	11.65	13.72	18.34	23.90	27.20	30.14	32.85	36.19	38.58
20	7.434	8.260	9.591	10.85	12.44	14.58	19.34	25.04	28.41	31.41	34.17	37.57	40.00
21	8.034	8.897	10.28	11.59	13.24	15.44	20.34	26.17	29.62	32.67	35.48	38.93	41.40
22	8.643	9.542	10.98	12.34	14.04	16.31	21.34	27.30	30.81	33.92	36.78	40.29	42.80
23	9.260	10.20	11.69	13.09	14.85	17.19	22.34	28.43	32.01	35.17	38.08	41.64	44.18
24	9.886	10.86	12.40	13.85	15.66	18.06	23.34	29.55	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	18.94	24.34	30.68	34.38	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	19.82	25.34	31.79	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	20.70	26.34	32.91	36.74	40.11	43.19	46.96	49.64
28	12.46	13.56	15.31	16.93	18.94	21.59	27.34	34.03	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	22.48	28.34	35.14	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	23.36	29.34	36.25	40.26	43.77	46.98	50.89	53.67
31	14.46	15.66	17.54	19.28	21.43	24.26	30.34	37.36	41.42	44.99	48.23	52.19	55.00
32	15.13	16.36	18.29	20.07	22.27	25.15	31.34	38.47	42.58	46.19	49.48	53.49	56.33
33	15.82	17.07	19.05	20.87	23.11	26.04	32.34	39.57	43.75	47.40	50.73	54.78	57.65
34	16.50	17.79	19.81	21.66	23.95	26.94	33.34	40.68	44.90	48.60	51.97	56.06	58.96
35	17.19	18.51	20.57	22.47	24.80	27.84	34.34	41.78	46.06	49.80	53.20	57.34	60.27
36	17.89	19.23	21.34	23.27	25.64	28.73	35.34	42.88	47.21	51.00	54.44	58.62	61.58
37	18.59	19.96	22.11	24.07	26.49	29.64	36.34	43.98	48.36	52.19	55.67	59.89	62.88
38	19.29	20.69	22.88	24.88	27.34	30.54	37.34	45.08	49.51	53.38	56.90	61.16	64.18
39	20.00	21.43	23.65	25.70	28.20	31.44	38.34	46.17	50.66	54.57	58.12	62.43	65.48
40	20.71	22.16	24.43	26.51	29.05	32.34	39.34	47.27	51.81	55.76	59.34	63.69	66.77
41	21.42	22.91	25.21	27.33	29.91	33.25	40.34	48.36	52.95	56.94	60.56	64.95	68.05
42	22.14	23.65	26.00	28.14	30.77	34.16	41.34	49.46	54.09	58.12	61.78	66.21	69.34
43	22.86	24.40	26.79	28.96	31.63	35.07	42.34	50.55	55.23	59.30	62.99	67.46	70.62
44	23.58	25.15	27.57	29.79	32.49	35.97	43.34	51.64	56.37	60.48	64.20	68.71	71.89
45	24.31	25.90	28.37	30.61	33.35	36.88	44.34	52.73	57.51	61.66	65.41	69.96	73.17
46	25.04	26.66	29.16	31.44	34.22	37.80	45.34	53.82	58.64	62.83	66.62	71.20	74.44
47	25.77	27.42	29.96	32.27	35.08	38.71	46.34	54.91	59.77	64.00	67.82	72.44	75.70
48	26.51	28.18	30.75	33.10	35.95	39.62	47.34	55.99	60.91	65.17	69.02	73.68	76.97
49	27.25	28.94	31.55	33.93	36.82	40.53	48.33	57.08	62.04	66.34	70.22	74.92	78.23
50	27.99	29.71	32.36	34.76	37.69	41.45	49.33	58.16	63.17	67.50	71.42	76.15	79.49
100	67.33	70.06	74.22	77.93	82.36	87.95	99.33	111.7	118.5	124.3	129.6	135.8	140.2

Students *t*-Verteilung

p -Quantile $t_{[p;fg]}$ der t -Verteilung mit fg Freiheitsgraden

$$[t_{[p;fg]} = -t_{[1-p;fg]}]$$

Beispiel: $t_{[0.99;15]} = 2.6025$, $t_{[0.01;15]} = -2.6025$

<i>fg</i>	<i>p</i>							
	0.6	0.7	0.8	0.9	0.95	0.975	0.99	0.995
1	0.3249	0.7265	1.3764	3.0777	6.3138	12.706	31.821	63.657
2	0.2887	0.6172	1.0607	1.8856	2.9200	4.3027	6.9646	9.9248
3	0.2767	0.5844	0.9785	1.6377	2.3534	3.1824	4.5407	5.8409
4	0.2707	0.5687	0.9410	1.5332	2.1318	2.7764	3.7470	4.6041
5	0.2672	0.5594	0.9195	1.4759	2.0150	2.5706	3.3649	4.0322
6	0.2648	0.5534	0.9057	1.4398	1.9432	2.4469	3.1427	3.7074
7	0.2632	0.5491	0.8960	1.4149	1.8946	2.3646	2.9980	3.4995
8	0.2619	0.5459	0.8889	1.3968	1.8595	2.3060	2.8965	3.3554
9	0.2610	0.5435	0.8834	1.3830	1.8331	2.2622	2.8214	3.2498
10	0.2602	0.5415	0.8791	1.3722	1.8125	2.2281	2.7638	3.1693
11	0.2596	0.5399	0.8755	1.3634	1.7959	2.2010	2.7181	3.1058
12	0.2590	0.5386	0.8726	1.3562	1.7823	2.1788	2.6810	3.0545
13	0.2586	0.5375	0.8702	1.3502	1.7709	2.1604	2.6503	3.0123
14	0.2582	0.5366	0.8681	1.3450	1.7613	2.1448	2.6245	2.9768
15	0.2579	0.5357	0.8662	1.3406	1.7531	2.1314	2.6025	2.9467
16	0.2576	0.5350	0.8647	1.3368	1.7459	2.1199	2.5835	2.9208
17	0.2574	0.5344	0.8633	1.3334	1.7396	2.1098	2.5669	2.8982
18	0.2571	0.5338	0.8621	1.3304	1.7341	2.1009	2.5524	2.8784
19	0.2569	0.5333	0.8610	1.3277	1.7291	2.0930	2.5395	2.8609
20	0.2567	0.5329	0.8510	1.3253	1.7247	2.0860	2.5280	2.8453
21	0.2566	0.5325	0.8591	1.3232	1.7207	2.0796	2.5176	2.8314
22	0.2564	0.5321	0.8583	1.3212	1.7171	2.0739	2.5083	2.8188
23	0.2563	0.5318	0.8575	1.3195	1.7139	2.0687	2.4999	2.8073
24	0.2562	0.5314	0.8569	1.3178	1.7109	2.0639	2.4922	2.7969
25	0.2561	0.5312	0.8562	1.3163	1.7081	2.0595	2.4851	2.7874
26	0.2560	0.5309	0.8557	1.3150	1.7056	2.0555	2.4786	2.7787
27	0.2559	0.5307	0.8551	1.3137	1.7033	2.0518	2.4727	2.7707
28	0.2558	0.5304	0.8547	1.3125	1.7011	2.0484	2.4671	2.7633
29	0.2557	0.5302	0.8542	1.3114	1.6991	2.0452	2.4620	2.7564
30	0.2556	0.5300	0.8538	1.3104	1.6973	2.0423	2.4573	2.7500
31	0.2555	0.5298	0.8534	1.3095	1.6955	2.0395	2.4528	2.7440
32	0.2555	0.5297	0.8530	1.3086	1.6939	2.0369	2.4487	2.7385
33	0.2554	0.5295	0.8527	1.3077	1.6924	2.0345	2.4448	2.7333
34	0.2553	0.5294	0.8523	1.3070	1.6909	2.0322	2.4411	2.7284
35	0.2553	0.5292	0.8520	1.3062	1.6896	2.0301	2.4377	2.7238
36	0.2552	0.5291	0.8517	1.3055	1.6883	2.0281	2.4345	2.7195
37	0.2552	0.5290	0.8514	1.3049	1.6871	2.0262	2.4314	2.7154
38	0.2551	0.5288	0.8512	1.3042	1.6860	2.0244	2.4286	2.7116
39	0.2551	0.5287	0.8509	1.3036	1.6849	2.0227	2.4258	2.7079
40	0.2550	0.5286	0.8507	1.3031	1.6839	2.0211	2.4233	2.7045
41	0.2550	0.5285	0.8505	1.3025	1.6829	2.0195	2.4208	2.7012
42	0.2550	0.5284	0.8503	1.3020	1.6820	2.0181	2.4185	2.6981
43	0.2549	0.5283	0.8501	1.3016	1.6811	2.0167	2.4163	2.6951
44	0.2549	0.5282	0.8499	1.3011	1.6802	2.0154	2.4141	2.6923
45	0.2549	0.5281	0.8497	1.3006	1.6794	2.0141	2.4121	2.6896
46	0.2548	0.5281	0.8495	1.3002	1.6787	2.0129	2.4102	2.6870
47	0.2548	0.5280	0.8493	1.2998	1.6779	2.0117	2.4083	2.6846
48	0.2548	0.5279	0.8492	1.2994	1.6772	2.0106	2.4066	2.6822
49	0.2547	0.5278	0.8490	1.2991	1.6766	2.0096	2.4049	2.6800
50	0.2547	0.5278	0.8489	1.2987	1.6759	2.0086	2.4033	2.6778
∞	0.2533	0.5244	0.8416	1.2816	1.6449	1.9600	2.3263	2.5758

Fishers *F*-Verteilung

p -Quantile $f_{[p; fg1, fg2]}$ der F -Verteilung mit $fg1$ und $fg2$ Freiheitsgraden für $p = 0.95$ und $p = 0.99$

$$\left[f_{[p; fg1, fg2]} = \frac{1}{f_{[1-p; fg2, fg1]}} \right]$$

0.95-Quantile $f_{[0.95; fg1, fg2]}$

<i>fg2</i>	<i>fg1</i>											
	1	2	3	4	5	6	7	8	9	10	11	12
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	242.98	243.90
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.40	19.41
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.41	2.38
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.34	2.31
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	2.23
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.24	2.20
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.22	2.18
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.20	2.16
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.17	2.13
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.15	2.12
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.14	2.10
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09
31	4.16	3.30	2.91	2.68	2.52	2.41	2.32	2.25	2.20	2.15	2.11	2.08
32	4.15	3.29	2.90	2.67	2.51	2.40	2.31	2.24	2.19	2.14	2.10	2.07
33	4.14	3.28	2.89	2.66	2.50	2.39	2.30	2.23	2.18	2.13	2.09	2.06
34	4.13	3.28	2.88	2.65	2.49	2.38	2.29	2.23	2.17	2.12	2.08	2.05
35	4.12	3.27	2.87	2.64	2.49	2.37	2.29	2.22	2.16	2.11	2.07	2.04
36	4.11	3.26	2.87	2.63	2.48	2.36	2.28	2.21	2.15	2.11	2.07	2.03
37	4.11	3.25	2.86	2.63	2.47	2.36	2.27	2.20	2.14	2.10	2.06	2.02
38	4.10	3.24	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09	2.05	2.02
39	4.09	3.24	2.85	2.61	2.46	2.34	2.26	2.19	2.13	2.08	2.04	2.01
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.00
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03	1.99	1.95
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.02	1.97	1.93	1.89
80	3.96	3.11	2.72	2.49	2.33	2.21	2.13	2.06	2.00	1.95	1.91	1.88
90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94	1.90	1.86
100	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97	1.93	1.89	1.85
200	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93	1.88	1.84	1.80
500	3.86	3.01	2.62	2.39	2.23	2.12	2.03	1.96	1.90	1.85	1.81	1.77
1000	3.85	3.00	2.61	2.38	2.22	2.11	2.02	1.95	1.89	1.84	1.80	1.76

0.95-Quantile $f_{[0.95; fg1, fg2]}$ (Fortsetzung)

fg2	fg1										
	14	16	18	20	22	24	26	28	30	35	40
1	245.36	246.47	247.32	248.02	248.58	249.05	249.45	249.80	250.10	250.69	251.14
2	19.42	19.43	19.44	19.45	19.45	19.45	19.46	19.46	19.46	19.47	19.47
3	8.71	8.69	8.67	8.66	8.65	8.64	8.63	8.62	8.62	8.60	8.59
4	5.87	5.84	5.82	5.80	5.79	5.77	5.76	5.75	5.75	5.73	5.72
5	4.64	4.60	4.58	4.56	4.54	4.53	4.52	4.50	4.50	4.48	4.46
6	3.96	3.92	3.90	3.87	3.86	3.84	3.83	3.82	3.81	3.79	3.77
7	3.53	3.49	3.47	3.44	3.43	3.41	3.40	3.39	3.38	3.36	3.34
8	3.24	3.20	3.17	3.15	3.13	3.12	3.10	3.09	3.08	3.06	3.04
9	3.03	2.99	2.96	2.94	2.92	2.90	2.89	2.87	2.86	2.84	2.83
10	2.86	2.83	2.80	2.77	2.75	2.74	2.72	2.71	2.70	2.68	2.66
11	2.74	2.70	2.67	2.65	2.63	2.61	2.59	2.58	2.57	2.55	2.53
12	2.64	2.60	2.57	2.54	2.52	2.51	2.49	2.48	2.47	2.44	2.43
13	2.55	2.51	2.48	2.46	2.44	2.42	2.41	2.39	2.38	2.36	2.34
14	2.48	2.44	2.41	2.39	2.37	2.35	2.33	2.32	2.31	2.28	2.27
15	2.42	2.38	2.35	2.33	2.31	2.29	2.27	2.26	2.25	2.22	2.20
16	2.37	2.33	2.30	2.28	2.25	2.24	2.22	2.21	2.19	2.17	2.15
17	2.33	2.29	2.26	2.23	2.21	2.19	2.17	2.16	2.15	2.12	2.10
18	2.29	2.25	2.22	2.19	2.17	2.15	2.13	2.12	2.11	2.08	2.06
19	2.26	2.21	2.18	2.16	2.13	2.11	2.10	2.08	2.07	2.05	2.03
20	2.22	2.18	2.15	2.12	2.10	2.08	2.07	2.05	2.04	2.01	1.99
21	2.20	2.16	2.12	2.10	2.07	2.05	2.04	2.02	2.01	1.98	1.96
22	2.17	2.13	2.10	2.07	2.05	2.03	2.01	2.00	1.98	1.96	1.94
23	2.15	2.11	2.08	2.05	2.02	2.01	1.99	1.97	1.96	1.93	1.91
24	2.13	2.09	2.05	2.03	2.00	1.98	1.97	1.95	1.94	1.91	1.89
25	2.11	2.07	2.04	2.01	1.98	1.96	1.95	1.93	1.92	1.89	1.87
26	2.09	2.05	2.02	1.99	1.97	1.95	1.93	1.91	1.90	1.87	1.85
27	2.08	2.04	2.00	1.97	1.95	1.93	1.91	1.90	1.88	1.86	1.84
28	2.06	2.02	1.99	1.96	1.93	1.91	1.90	1.88	1.87	1.84	1.82
29	2.05	2.01	1.97	1.94	1.92	1.90	1.88	1.87	1.85	1.83	1.81
30	2.04	1.99	1.96	1.93	1.91	1.89	1.87	1.85	1.84	1.81	1.79
31	2.03	1.98	1.95	1.92	1.90	1.88	1.86	1.84	1.83	1.80	1.78
32	2.01	1.97	1.94	1.91	1.88	1.86	1.85	1.83	1.82	1.79	1.77
33	2.00	1.96	1.93	1.90	1.87	1.85	1.83	1.82	1.81	1.78	1.76
34	1.99	1.95	1.92	1.89	1.86	1.84	1.82	1.81	1.80	1.77	1.75
35	1.99	1.94	1.91	1.88	1.85	1.83	1.82	1.80	1.79	1.76	1.74
36	1.98	1.93	1.90	1.87	1.85	1.82	1.81	1.79	1.78	1.75	1.73
37	1.97	1.93	1.89	1.86	1.84	1.82	1.80	1.78	1.77	1.74	1.72
38	1.96	1.92	1.88	1.85	1.83	1.81	1.79	1.77	1.76	1.73	1.71
39	1.95	1.91	1.88	1.85	1.82	1.80	1.78	1.77	1.75	1.72	1.70
40	1.95	1.90	1.87	1.84	1.81	1.79	1.77	1.76	1.74	1.72	1.69
50	1.89	1.85	1.81	1.78	1.76	1.74	1.72	1.70	1.69	1.66	1.63
60	1.86	1.82	1.78	1.75	1.72	1.70	1.68	1.66	1.65	1.62	1.59
70	1.84	1.79	1.75	1.72	1.70	1.67	1.65	1.64	1.62	1.59	1.57
80	1.82	1.77	1.73	1.70	1.68	1.65	1.63	1.62	1.60	1.57	1.54
90	1.80	1.76	1.72	1.69	1.66	1.64	1.62	1.60	1.59	1.55	1.53
100	1.79	1.75	1.71	1.68	1.65	1.63	1.61	1.59	1.57	1.54	1.52
200	1.74	1.69	1.66	1.62	1.60	1.57	1.55	1.53	1.52	1.48	1.46
500	1.71	1.66	1.62	1.59	1.56	1.54	1.52	1.50	1.48	1.45	1.42
1000	1.70	1.65	1.61	1.58	1.55	1.53	1.51	1.49	1.47	1.43	1.41

0.99-Quantile $f_{[0.99; fg1, fg2]}$

fg2	fg1											
	1	2	3	4	5	6	7	8	9	10	11	12
1	4052.18	4999.34	5403.53	5624.26	5763.96	5858.95	5928.33	5980.95	6022.40	6055.93	6083.40	6106.68
2	98.50	99.00	99.16	99.25	99.30	99.33	99.36	99.38	99.39	99.40	99.41	99.42
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.34	27.23	27.13	27.05
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.45	14.37
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.96	9.89
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.79	7.72
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.54	6.47
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.73	5.67
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.18	5.11
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.77	4.71
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.46	4.40
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.22	4.16
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	4.02	3.96
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.86	3.80
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.73	3.67
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.62	3.55
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.52	3.46
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.43	3.37
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.36	3.30
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.29	3.23
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.24	3.17
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.18	3.12
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.14	3.07
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.09	3.03
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	3.06	2.99
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	3.02	2.96
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.99	2.93
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.96	2.90
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.93	2.87
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.91	2.84
31	7.53	5.36	4.48	3.99	3.67	3.45	3.28	3.15	3.04	2.96	2.88	2.82
32	7.50	5.34	4.46	3.97	3.65	3.43	3.26	3.13	3.02	2.93	2.86	2.80
33	7.47	5.31	4.44	3.95	3.63	3.41	3.24	3.11	3.00	2.91	2.84	2.78
34	7.44	5.29	4.42	3.93	3.61	3.39	3.22	3.09	2.98	2.89	2.82	2.76
35	7.42	5.27	4.40	3.91	3.59	3.37	3.20	3.07	2.96	2.88	2.80	2.74
36	7.40	5.25	4.38	3.89	3.57	3.35	3.18	3.05	2.95	2.86	2.79	2.72
37	7.37	5.23	4.36	3.87	3.56	3.33	3.17	3.04	2.93	2.84	2.77	2.71
38	7.35	5.21	4.34	3.86	3.54	3.32	3.15	3.02	2.92	2.83	2.75	2.69
39	7.33	5.19	4.33	3.84	3.53	3.30	3.14	3.01	2.90	2.81	2.74	2.68
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.73	2.66
50	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.78	2.70	2.63	2.56
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.56	2.50
70	7.01	4.92	4.07	3.60	3.29	3.07	2.91	2.78	2.67	2.59	2.51	2.45
80	6.96	4.88	4.04	3.56	3.26	3.04	2.87	2.74	2.64	2.55	2.48	2.42
90	6.93	4.85	4.01	3.53	3.23	3.01	2.84	2.72	2.61	2.52	2.45	2.39
100	6.90	4.82	3.98	3.51	3.21	2.99	2.82	2.69	2.59	2.50	2.43	2.37
200	6.76	4.71	3.88	3.41	3.11	2.89	2.73	2.60	2.50	2.41	2.34	2.27
500	6.69	4.65	3.82	3.36	3.05	2.84	2.68	2.55	2.44	2.36	2.28	2.22
1000	6.66	4.63	3.80	3.34	3.04	2.82	2.66	2.53	2.43	2.34	2.27	2.20

0.99-Quantile $f_{[0.99; fg1, fg2]}$ (Fortsetzung)

fg2	fg1										
	14	16	18	20	22	24	26	28	30	35	40
1	6143.00	6170.01	6191.43	6208.66	6223.10	6234.27	6244.52	6252.90	6260.35	6275.25	6286.43
2	99.43	99.44	99.44	99.45	99.46	99.46	99.46	99.46	99.47	99.47	99.48
3	26.92	26.83	26.75	26.69	26.64	26.60	26.56	26.53	26.50	26.45	26.41
4	14.25	14.15	14.08	14.02	13.97	13.93	13.89	13.86	13.84	13.79	13.75
5	9.77	9.68	9.61	9.55	9.51	9.47	9.43	9.40	9.38	9.33	9.29
6	7.60	7.52	7.45	7.40	7.35	7.31	7.28	7.25	7.23	7.18	7.14
7	6.36	6.28	6.21	6.16	6.11	6.07	6.04	6.02	5.99	5.94	5.91
8	5.56	5.48	5.41	5.36	5.32	5.28	5.25	5.22	5.20	5.15	5.12
9	5.01	4.92	4.86	4.81	4.77	4.73	4.70	4.67	4.65	4.60	4.57
10	4.60	4.52	4.46	4.41	4.36	4.33	4.30	4.27	4.25	4.20	4.17
11	4.29	4.21	4.15	4.10	4.06	4.02	3.99	3.96	3.94	3.89	3.86
12	4.05	3.97	3.91	3.86	3.82	3.78	3.75	3.72	3.70	3.65	3.62
13	3.86	3.78	3.72	3.66	3.62	3.59	3.56	3.53	3.51	3.46	3.43
14	3.70	3.62	3.56	3.51	3.46	3.43	3.40	3.37	3.35	3.30	3.27
15	3.56	3.49	3.42	3.37	3.33	3.29	3.26	3.24	3.21	3.17	3.13
16	3.45	3.37	3.31	3.26	3.22	3.18	3.15	3.12	3.10	3.05	3.02
17	3.35	3.27	3.21	3.16	3.12	3.08	3.05	3.03	3.00	2.96	2.92
18	3.27	3.19	3.13	3.08	3.03	3.00	2.97	2.94	2.92	2.87	2.84
19	3.19	3.12	3.05	3.00	2.96	2.92	2.89	2.87	2.84	2.80	2.76
20	3.13	3.05	2.99	2.94	2.90	2.86	2.83	2.80	2.78	2.73	2.69
21	3.07	2.99	2.93	2.88	2.84	2.80	2.77	2.74	2.72	2.67	2.64
22	3.02	2.94	2.88	2.83	2.78	2.75	2.72	2.69	2.67	2.62	2.58
23	2.97	2.89	2.83	2.78	2.74	2.70	2.67	2.64	2.62	2.57	2.54
24	2.93	2.85	2.79	2.74	2.70	2.66	2.63	2.60	2.58	2.53	2.49
25	2.89	2.81	2.75	2.70	2.66	2.62	2.59	2.56	2.54	2.49	2.45
26	2.86	2.78	2.72	2.66	2.62	2.58	2.55	2.53	2.50	2.45	2.42
27	2.82	2.75	2.68	2.63	2.59	2.55	2.52	2.49	2.47	2.42	2.38
28	2.79	2.72	2.65	2.60	2.56	2.52	2.49	2.46	2.44	2.39	2.35
29	2.77	2.69	2.63	2.57	2.53	2.49	2.46	2.44	2.41	2.36	2.33
30	2.74	2.66	2.60	2.55	2.51	2.47	2.44	2.41	2.39	2.34	2.30
31	2.72	2.64	2.58	2.52	2.48	2.45	2.41	2.39	2.36	2.31	2.27
32	2.70	2.62	2.55	2.50	2.46	2.42	2.39	2.36	2.34	2.29	2.25
33	2.68	2.60	2.53	2.48	2.44	2.40	2.37	2.34	2.32	2.27	2.23
34	2.66	2.58	2.51	2.46	2.42	2.38	2.35	2.32	2.30	2.25	2.21
35	2.64	2.56	2.50	2.44	2.40	2.36	2.33	2.30	2.28	2.23	2.19
36	2.62	2.54	2.48	2.43	2.38	2.35	2.32	2.29	2.26	2.21	2.18
37	2.61	2.53	2.46	2.41	2.37	2.33	2.30	2.27	2.25	2.20	2.16
38	2.59	2.51	2.45	2.40	2.35	2.32	2.28	2.26	2.23	2.18	2.14
39	2.58	2.50	2.43	2.38	2.34	2.30	2.27	2.24	2.22	2.17	2.13
40	2.56	2.48	2.42	2.37	2.33	2.29	2.26	2.23	2.20	2.15	2.11
50	2.46	2.38	2.32	2.27	2.22	2.18	2.15	2.12	2.10	2.05	2.01
60	2.39	2.31	2.25	2.20	2.15	2.12	2.08	2.05	2.03	1.98	1.94
70	2.35	2.27	2.20	2.15	2.11	2.07	2.03	2.01	1.98	1.93	1.89
80	2.31	2.23	2.17	2.12	2.07	2.03	2.00	1.97	1.94	1.89	1.85
90	2.29	2.21	2.14	2.09	2.04	2.00	1.97	1.94	1.92	1.86	1.82
100	2.27	2.19	2.12	2.07	2.02	1.98	1.95	1.92	1.89	1.84	1.80
200	2.17	2.09	2.03	1.97	1.93	1.89	1.85	1.82	1.79	1.74	1.69
500	2.12	2.04	1.97	1.92	1.87	1.83	1.79	1.76	1.74	1.68	1.63
1000	2.10	2.02	1.95	1.90	1.85	1.81	1.77	1.74	1.72	1.66	1.61

Für großes $fg2$ können die Quantile der F -Verteilung mit $fg1$ und $fg2$ Freiheitsgraden durch die Quantile der χ^2 -Verteilung mit $fg1$ Freiheitsgraden approximiert werden.