| Nachname | Vorname | Matrikelnummer |
|---|---|---|
| | | |

**Formelsammlung zur Vorlesung**

# Statistik II für Studierende der Soziologie und Nebenfachstudierende

**Prof. Dr. Thomas Augustin, Johanna Brandt, Julia Plaß**

**Sommersemester 2015**

Es gelten die Hinweise zu den Klausurmodalitäten und erlaubten Hilfsmitteln auf der Veranstaltungshomepage!

Insbesondere gilt: Für die Verwendung in der Klausur im Sommersemester 2015 darf diese Version auf den bedruckten Seiten (und nur dort!) mit zusätzlichen handschriftlichen Kommentaren versehen werden, falls

a) die Formelsammlung in Originalgröße ausgedruckt wurde und
b) die Eintragungen in der eigenen Handschrift im Original vorgenommen wurden (keine Kopie!).

# Inhaltsverzeichnis

# 1 Wahrscheinlichkeitsrechnung

## 1.1 Mengen und elementare Mengenoperationen

**Definition 1.1** Eine *Menge* ist eine Zusammenfassung verschiedener Objekte zu einem Ganzen. Die einzelnen Objekte einer Menge werden *Elemente* genannt.

### Grundlegende Begriffe der Mengenlehre

- **Standardmengen:**

| | |
|---|---|
| $\mathbb{N} = \{1, 2, 3, \ldots\}$ | Menge der natürlichen Zahlen |
| $\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$ | Menge der natürlichen Zahlen inklusive 0 |
| $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$ | Menge der ganzen Zahlen |
| $\mathbb{R} = (-\infty, \infty)$ | Menge der reellen Zahlen |
| $\emptyset$ | leere Menge |

- **Elementeigenschaft:**
  $x$ ist Element der Menge $M$: $x \in M$
  $x$ ist nicht Element der Menge $M$: $x \notin M$

- **Teilmengen:** $M_1$ ist Teilmenge von $M_2$, in Zeichen $M_1 \subset M_2$, wenn jedes Element von $M_1$ auch in $M_2$ ist.

- **Schnittmenge:** Die Schnittmenge $M_1 \cap M_2$ ist die Menge aller Elemente, die sowohl in $M_1$ als auch in $M_2$ enthalten sind:

$$M_1 \cap M_2 = \{x | x \in M_1 \text{ und } x \in M_2\}$$

  Eigenschaften:

  - Gilt $M_1 \subset M_2$, so ist $M_1 \cap M_2 = M_1$.
  - Für jede Menge $M_1$ gilt: $M_1 \cap M_1 = M_1$ und $M_1 \cap \emptyset = \emptyset$.
  - Zwei Mengen $M_1$ und $M_2$ mit $M_1 \cap M_2 = \emptyset$, d.h. zwei Mengen, die kein gemeinsames Element haben, heißen *disjunkt*.
  - Verallgemeinerung: Die Schnittmenge aus $n$ Mengen $M_1, \ldots, M_n$ enthält alle Elemente, die in jeder der Mengen $M_1, \ldots, M_n$ enthalten sind und wird bezeichnet mit

$$\bigcap_{i=1}^{n} M_i := M_1 \cap M_2 \cap \ldots \cap M_n.$$

- **Vereinigungsmenge:** Die Vereinigungsmenge $M_1 \cup M_2$ ist die Menge aller Elemente, die in $M_1$ oder $M_2$ enthalten sind:
$$M_1 \cup M_2 = \{x | x \in M_1 \text{ oder } x \in M_2\}$$
Verallgemeinerung: Die Vereinigungsmenge aus $n$ Mengen $M_1, \ldots, M_n$ enthält alle Elemente, die in mindestens einer der Mengen $M_1, \ldots, M_n$ enthalten sind und wird bezeichnet mit
$$\bigcup_{i=1}^{n} M_i := M_1 \cup M_2 \cup \ldots \cup M_n$$
- **Differenzmenge:** Die Differenzmenge $M_1 \setminus M_2$ ist die Menge aller Elemente, die in $M_1$, aber nicht in $M_2$ enthalten sind:
$$M_1 \setminus M_2 = \{x | x \in M_1 \text{ aber } x \notin M_2\}$$
- **Komplementärmenge:** Die Komplementärmenge $\overline{M} = A^C$ bezüglich einer Grundmenge $\Omega$ ist die Menge aller Elemente von $\Omega$, die nicht in $M$ sind:
$$\overline{M} = M^C = \{x \in \Omega | x \notin M\} = \{x : x \notin M\}$$
- **Potenzmenge:** Die Potenzmenge $\mathcal{P}(M)$ ist die Menge aller Teilmengen von $M$:
$$\mathcal{P}(M) = \{N | N \subset M\}.$$
- **Mächtigkeit:** Die Mächtigkeit $|M|$ einer Menge $M$ ist die Anzahl der Elemente von $M$

**Rechenregeln für Mengen**

a) Kommutativgesetze (Vertauschung):
$$M_1 \cap M_2 = M_2 \cap M_1, \quad M_1 \cup M_2 = M_2 \cup M_1.$$
b) Assoziativgesetze (Zusammenfassen):
$$(M_1 \cap M_2) \cap M_3 = M_1 \cap (M_2 \cap M_3).$$
$$(M_1 \cup M_2) \cup M_3 = M_1 \cup (M_2 \cup M_3).$$
c) Distributivgesetze (Ausklammern/Ausmultiplizieren):
$$(M_1 \cup M_2) \cap M_3 = (M_1 \cap M_3) \cup (M_2 \cap M_3).$$
$$(M_1 \cap M_2) \cup M_3 = (M_1 \cup M_3) \cap (M_2 \cup M_3).$$
d) De Morgansche Regeln:
$$(\overline{M_1 \cup M_2}) = \overline{M_1} \cap \overline{M_2}$$
$$(\overline{M_1 \cap M_2}) = \overline{M_1} \cup \overline{M_2}$$
e) Aus $M_1 \subset M_2$ folgt $\overline{M_2} \subset \overline{M_1}$.

f) Für die Differenzmenge gilt $M_1 \setminus M_2 = M_1 \cap \overline{M_2}$.

g) Für die Potenzmenge gilt $|\mathcal{P}(M_1)| = 2^{|M_1|}$.

### Das kartesische Produkt

Das kartesische Produkt zweier Mengen

$$M = \{m_1, m_2, m_3, \ldots, m_k\}$$
$$N = \{n_1, n_2, n_3, \ldots, n_l\}$$

ist die Menge

$$M \times N := \Big\{(m_i, n_j) \;\Big|\; i = 1, \ldots, k, j = 1, \ldots, l\Big\}$$

Sie besteht also aus allen möglichen Kombinationen, so dass

$$\begin{aligned} M \times N = \big\{&(m_1, n_1), (m_1, n_2), (m_1, n_3), \ldots, (m_1, n_l),\\ &(m_2, n_1), (m_2, n_2), (m_2, n_3), \ldots, (m_2, n_l),\\ &\vdots\\ &(m_k, n_1), (m_k, n_2), (m_k, n_3), \ldots, (m_k, n_l)\big\}. \end{aligned}$$

**Verallgemeinerungen:**

- Das kartesische Produkt der Mengen $\Omega_1, \Omega_2, \ldots, \Omega_n$ wird mit

  $$\bigtimes_{i=1}^{n} \Omega_i = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_n$$

  bezeichnet und besteht aus allen möglichen $n$-Tupeln, die sich (unter Beachtung der Reihenfolge) aus Elementen aus $\Omega_1, \Omega_2, \ldots, \Omega_n$ bilden lassen.

- Die Mengen $\Omega_1, \Omega_2, \ldots, \Omega_n$ müssen nicht endlich sein; für endliche Mengen gilt

  $$\left|\bigtimes_{i=1}^{n} \Omega\right| = |\Omega_1| \cdot |\Omega_2| \cdot \ldots \cdot |\Omega_n|$$

- Kartesische Produkte werden verwendet, um Ergebnisse komplexer Experimente aus Einzelexperimenten zusammenzusetzen.

# 1.2 Wahrscheinlichkeit – Ein komplexer Begriff und seine Formalisierung

## 1.2.1 Zufallsvorgänge

## 1.2.2 Laplace-Wahrscheinlichkeiten und Urnenmodelle

**Abzählregel**

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

**Laplace-Wahrscheinlichkeit**

In einem Laplace-Experiment gilt für $P(A)$ mit $|A| = M$ und $|\Omega| = N$:

$$P(A) = \frac{|A|}{|\Omega|} = \frac{M}{N}.$$

**Urnenmodell**

- Grundgesamtheit: Urne mit $N$ nummerierten Kugeln
- Stichprobe: Zufälliges Ziehen von $n$ Kugeln aus der Urne

**Ziehen mit Zurücklegen mit Berücksichtigung der Reihenfolge**

- Ziehe Stichprobe vom Umfang $n$ **mit** Zurücklegen.
- $\Omega = \{(\omega_1, \ldots, \omega_n) | \omega_j \in \{1, \ldots, N\}\}$
  (Das selbe Element kann mehrfach vorkommen.)
- Anzahl möglicher Stichproben:

$$|\Omega| = \underbrace{N \cdot N \cdot \ldots \cdot N}_{n\text{-Faktoren}} = N^n$$

**Ziehen ohne Zurücklegen mit Berücksichtigung der Reihenfolge**

- Ziehe Stichprobe vom Umfang $n$ **ohne** Zurücklegen.
- $\Omega = \{(\omega_1, \ldots, \omega_n) : \omega_j \in \{1, \ldots, N\},\ \omega_j \neq \omega_i \text{ für } i \neq j\}$
  (Jedes Element kann nur einmal vorkommen.)
- Anzahl möglicher Stichproben:

$$|\Omega| = N \cdot (N-1) \cdot \ldots \cdot N - n + 1 = \frac{N!}{(N-n)!}$$

**Wiederholung Fakultät:** Die *Fakultät* einer natürlichen Zahl $k$ ist definiert als

$$k! = k \cdot (k-1) \cdot (k-2) \cdot \ldots \cdot 2 \cdot 1.$$

Es gilt:

$$1! = 1, \quad 0! = 1.$$

**Ziehen ohne Zurücklegen ohne Berücksichtigung der Reihenfolge**

- Ziehe Stichprobe vom Umfang $n$ **ohne** Zurücklegen.
- $\Omega = \{\{\omega_1, \ldots, \omega_n\} : \omega_j \in \{1, \ldots, N\}, \omega_j \neq \omega_i \text{ für } j \neq i\}$
- Anzahl möglicher Stichproben:

$$|\Omega| = \frac{N!}{(N-n)!n!} = \binom{N}{n}$$

**Wiederholung Binomialkoeffizient:** Der *Binomialkoeffizient* $\binom{N}{n}$ ist definiert als

$$\binom{N}{n} = \frac{N!}{(N-n)! \cdot n!}\,.$$

Es gilt:

$$\binom{N}{0} = 1, \quad \binom{N}{1} = N, \quad \binom{N}{N} = 1, \quad \binom{N}{n} = 0, \quad \text{falls } N < n\,.$$

### 1.2.3 Die „induktive Brücke“ I

### 1.2.4 Das Axiomensystem von Kolmogoroff (1933) und wichtige Rechenregeln

**Definition 1.2** Eine Funktion $P$ ($P$ steht für Probability), die Ereignissen aus $\Omega$ reelle Zahlen zuordnet, heißt *Wahrscheinlichkeit*, wenn gilt

(K1) $P(A) \geq 0$ für alle Ereignisse $A \subset \Omega$.

(K2) $P(\Omega) = 1$.

(K3) Falls $A \cap B = \emptyset$, dann gilt $P(A \cup B) = P(A) + P(B)$

#### Rechenregeln für Wahrscheinlichkeiten

- $P(\overline{A}) = 1 - P(A)$
- Für beliebige Mengen $A, B$ gilt: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Falls $A_1, A_2, \ldots, A_n$ paarweise disjunkt sind, also $A_i \cap A_j = \emptyset$ für $i \neq j$, dann gilt:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)$$

#### Vollständige Zerlegung

Ist $A_1, A_2, \ldots, A_k$ eine *vollständige Zerlegung* von $\Omega$, d.h. gilt

$$\bigcup_{i=1}^{k} A_i = \Omega \text{ und } A_i \cap A_j = \emptyset \text{ für } i \neq j,$$

so gilt für jedes Ereignis $B$:

$$P(B) = \sum_{i=1}^{k} P(B \cap A_i).$$

### 1.2.5 Grundlegendes zum Begriff „Wahrscheinlichkeit“

# 1.3 Stochastische Unabhängigkeit und bedingte Wahrscheinlichkeiten

## 1.3.1 Stochastische Unabhängigkeit

**Definition 1.3** Zwei Ereignisse $A$ und $B$ heißen *(stochastisch) unabhängig*, wenn gilt

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

andernfalls heißen sie *stochastisch abhängig*.

**Definition 1.4** Ereignisse $A_1, A_2, \ldots, A_n$ heißen (vollständig) stochastisch unabhängig, wenn für alle $I \subseteq \{1, \ldots, n\}$ gilt

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i).$$

**Achtung:** Aus der paarweisen Unabhängigkeit

$$P(A_i \cap A_j) = P(A_i)P(A_j) \quad \text{für alle } i, j$$

folgt im Allgemeinen nicht die vollständige Unabhängigkeit.

## 1.3.2 Bedingte Wahrscheinlichkeiten

**Definition 1.5** Gegeben seien zwei Ereignisse $A$ und $B$ mit $P(B) > 0$. Dann heißt:

$$P(A|B) := \frac{P(A \cap B)}{P(B)}$$

*bedingte Wahrscheinlichkeit von A gegeben B* oder *bedingte Wahrscheinlichkeit von A unter der Bedingung B.*

## 1.3.3 Koppelung unabhängiger Experimente (unabhängige Wiederholungen)

**Unabhängige Koppelung mehrerer Experimente:** Gegeben sei eine Menge von Zufallsexperimenten, beschrieben durch die Ergebnisräume $\Omega_i$, $i = 1, \ldots, n$ und die Wahrscheinlichkeitsbewertungen $P_i$, $i = 1, \ldots, n$. Fasst man die Experimente zusammen, so ergibt sich der Ergebnisraum

$$\Omega := \Omega_1 \times \Omega_2 \times \ldots \times \Omega_n$$

mit den Elementen

$$\omega := (\omega_1, \omega_2, \ldots, \omega_n) .$$

Sind die Experimente unabhängig (Dies ist inhaltlich zu entscheiden!), so setzt man für beliebige $A_i \subset \Omega_i, i = 1, \ldots, n$,

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P_1(A_1) \cdot P_2(A_2) \cdot \ldots \cdot P_n(A_n).$$

Dies beschreibt ein Wahrscheinlichkeitsmaß auf $\Omega$, bei dem – per Konstruktion – beliebige Ereignisse aus den einzelnen $\Omega_i$ voneinander unabhängig sind.

### 1.3.4 Koppelung abhängiger Experimente

**Satz 1.6 (Satz von der totalen Wahrscheinlichkeit)** Gegeben sei eine vollständige Zerlegung $A_1, A_2, \ldots, A_k$ von $\Omega$. Dann gilt für jedes Ereignis $B$

$$P(B) = \sum_{j=1}^{k} P(B|A_j) \cdot P(A_j) = \sum_{j=1}^{k} P(B \cap A_j).$$

**Koppelung abhängiger Experimente:** Gegeben seien $n$ Experimente, beschrieben durch die Grundräume $\Omega_i = \{a_{i1}, \ldots, a_{ik_i}\}$ und die Wahrscheinlichkeitsbewertungen $P_i, i = 1, \ldots, n$. Bezeichnet man für beliebiges $i = 1, \ldots, n$ und $j = 1, \ldots, k_i$, mit $A_{ij}$ jeweils das zu $a_{ij}$ gehörige Elementarereignis (also das Ereignis „$a_{ij}$ tritt ein"), so gilt:

$$\begin{aligned} P\left(A_{1j_1} \cap A_{2j_2} \cap \ldots \cap A_{nj_n}\right) = & P_1(A_{1j_1}) \cdot P_2\left(A_{2j_2}|A_{1j_1}\right) \cdot P_3\left(A_{3j_3}|A_{1j_1} \cap A_{2j_2}\right) \\ & \cdot \ldots \cdot P_n\left(A_{nj_n}|A_{1j_1} \cap A_{2j_2} \cap \ldots \cap A_{n-1j_{n-1}}\right) \end{aligned}$$

Häufig werden die Indizes bei $P$ weggelassen.

**Korollar 1.7** Sei $A_1, A_2, \ldots, A_k$ eine vollständige Zerlegung. Dann gilt für beliebige Ereignisse $B$ und $C$ mit $P(C) > 0$

$$P(B|C) = \sum_{j=1}^{k} P(B|(A_j \cap C)) \cdot P(A_j|C)$$

#### Markovmodelle

Hier interpretiert man den Laufindex als Zeit. Gilt in der Koppelung abhängiger Experiment $\Omega_1 = \Omega_2 = \ldots = \Omega_n = \{a_1, \ldots, a_k\}$ und sind alle bedingten Wahrscheinlichkeiten nur vom jeweils unmittelbar vorhergehenden Zeitpunkt abhängig, d.h. gilt

$$P(A_{i+1,j_{i+1}}|A_{1j_1} \cap A_{2j_2} \cap \ldots \cap A_{ij_i}) = P(A_{i+1,j_{i+1}}|A_{ij_i}), \qquad (*)$$

so spricht man von einem *Markovmodell mit den Zuständen* $a_1, \ldots, a_k$.
Sind die sogenannten *Übergangswahrscheinlichkeiten* in (*) unabhängig von der Zeit, gilt also $P(A_{i+1,j}|A_{i\ell}) \equiv p_{j\ell}$ für alle $i, j, \ell$, so heißt das Markovmodell *homogen*.

### 1.3.5 Das Theorem von Bayes

**Satz 1.8** Sei $A_1, \ldots A_k$ eine vollständige Zerlegung von $\Omega$, wobei $P(A_i) > 0$, $P(B|A_i) > 0$, $i = 1, \ldots k$ und $P(B) > 0$ erfüllt seien. Dann gilt

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

# 1.4 Zufallsvariablen und ihre Verteilung

## 1.4.1 Diskrete Zufallsvariablen

**Definition 1.9** Gegeben seien ein diskreter, d.h. höchstens abzählbarer, Ergebnisraum $\Omega$ und die Wahrscheinlichkeit $P$ auf $\Omega$. Jede Abbildung

$$\begin{aligned} X : \Omega &\mapsto \Omega_X \\ \omega &\mapsto X(\omega) \end{aligned}$$

heißt Zufalls*element*. Setzt man für jede *Realisation* $x \in \Omega_X$

$$P_X(\{x\}) := P(\{X = x\}) := P(\{\omega | X(\omega) = x\})$$

so erhält man eine Wahrscheinlichkeit auf $\Omega_X$. (Oft wird auch $P(X = x)$ statt $P(\{X = x\})$ geschrieben.)

**Definition 1.10** Gegeben sei eine diskrete Zufallsvariable $X$ mit der Wahrscheinlichkeitsverteilung $P$. Die Menge

$$\mathcal{X} := \{x \in \mathbb{R} | P(\{x\}) > 0\}$$

heißt *Träger von X*.

**Definition 1.11** Die *Wahrscheinlichkeitsfunktion* $f(x)$ einer diskreten Zufallsvariable $X$ ist für $x \in \mathbb{R}$ definiert durch

$$f(x) = \begin{cases} P(X = x_i) = p_i, & x = x_i \in \mathcal{X} = \{x_1, x_2, \ldots, x_k, \ldots\} \\ 0, & \text{sonst.} \end{cases}$$

## 1.4.2 Verteilungsfunktion

**Definition 1.12** Sei $X$ eine Zufallsvariable. Die Funktion

$$\begin{aligned} F : \mathbb{R} &\mapsto [0; 1] \\ x &\mapsto F(x) \end{aligned}$$

$$F(x) := P(X \leq x)$$

heißt *Verteilungsfunktion*.

**Satz 1.13** Für beliebige reelle Zahlen $a$ und $b$ mit $a < b$ gilt: $P(a < X \leq b) = F(b) - F(a)$.

**Satz 1.14** Unter Regularitätsbedingungen gilt: Die Wahrscheinlichkeitsverteilung einer Zufallsvariablen $X$ kann man durch die Verteilungsfunktion eineindeutig erklären.

## 1.4.3 Stetige Zufallsvariablen

**Definition 1.15** Gegeben sei eine stetige Zufallsvariable $X$ mit differenzierbarer Verteilungsfunktion $F(x)$. Dann heißt die Ableitung von $F(x)$ nach $x$, also

$$f(x) = \frac{dF(x)}{dx},$$

*Dichte* der Zufallsvariablen $X$.

**Satz 1.16** In der Situation der obigen Definition gilt

$$F(x) = \int_{-\infty}^{x} f(u)\ du$$

und damit für beliebige reelle Zahlen $a$ und $b$ mit $a < b$

$$\begin{aligned} P(a \leq X \leq b) &= P(a < X \leq b) = P(a < X < b) \\ &= P(a \leq X < b) \\ &= \int_{a}^{b} f(x)\,dx. \end{aligned}$$

Jede Funktion $f$ auf $\mathbb{R}$ mit $f(x) \geq 0$ für alle $x$ und

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

kann als Dichte verwendet werden.

**Alternative Definition:** Eine Zufallsvariable $X$ heißt stetig, wenn es eine Funktion $f(x) \geq 0$ gibt, so dass für jedes Intervall $[a, b]$

$$P(a \leq X \leq b) = \int_{a}^{b} f(x)\,dx$$

gilt.

### 1.4.4 Lebensdauern; Hazardrate und Survivorfunktion

**Satz 1.17** Die Verteilung einer nicht negativen, stetigen Zufallsvariable $X$ wird eineindeutig durch sowohl die *Überlebensfunktion* (Survivorfunktion)

$$S(x) := P(X \geq x),$$

als auch durch die *Hazardrate*

$$\lambda(x) := \lim_{h \to 0} \frac{P(x \leq X \leq x + h | X \geq x)}{h}$$

beschrieben.

Es gelten folgende Zusammenhänge:

$$\begin{aligned} S(x) &= 1 - F(x) \\ S(x) &= \exp\left(-\int_{0}^{x} \lambda(u)du\right) \\ F(x) &= 1 - \exp\left(-\int_{0}^{x} \lambda(u)du\right) \\ f(x) &= \lambda(x) \cdot S(x) \end{aligned}$$

### 1.4.5 Unabhängigkeit von Zufallsvariablen

**Definition 1.18** Zwei Zufallsvariablen $X$ und $Y$ mit den Verteilungsfunktionen $F_X$ und $F_Y$ heißen *stochastisch unabhängig*, falls für alle $x$ und $y$ gilt

$$P(\{X \leq x\} \cap \{Y \leq y\}) = P(\{X \leq x\}) \cdot P(\{Y \leq y\}) = F_X(x) \cdot F_Y(y),$$

andernfalls heißen sie stochastisch abhängig.

## 1.5 Erwartungswert und Varianz

### 1.5.1 Diskrete Zufallsvariablen

**Definition 1.19** Gegeben sei eine diskrete Zufallsvariable $X$ mit Träger $\mathcal{X}$. Dann heißt

$$\mathrm{E}\,X := \mathrm{E}(X) := \mathbb{E}(X) := \sum_{x \in \mathcal{X}} x \cdot P(X = x)$$

*Erwartungswert* von $X$,

$$\begin{aligned}\mathrm{Var}\,X := \mathrm{Var}(X) := \mathbb{V}(X) &:= \mathrm{E}((X - \mathrm{E}(X))^2) \\ &= \sum_{x \in \mathcal{X}} (x - \mathrm{E}(X))^2 \cdot P(X = x)\end{aligned}$$

*Varianz* von $X$ und

$$\sigma_X := \sqrt{\mathrm{Var}(X)}$$

*Standardabweichung* von $X$.

Zur Berechnung der Varianz ist der sogenannte *Verschiebungssatz* sehr praktisch:

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}\,X)^2$$

### 1.5.2 Stetige Zufallsvariablen

**Definition 1.20** Sei $X$ eine stetige Zufallsvariable mit Dichte $f(x)$. Dann heißt

$$\mathrm{E}\,X := \mathrm{E}(X) := \mathbb{E}(X) := \int_{-\infty}^{\infty} x \cdot f(x)\, dx$$

*Erwartungswert* von $X$,

$$\begin{aligned}\mathrm{Var}\,X := \mathrm{Var}(X) := \mathbb{V}(X) &:= \mathrm{E}((X - \mathrm{E}(X))^2 \\ &= \int_{-\infty}^{\infty} (x - \mathrm{E}(X))^2 \cdot f(x)\, dx\end{aligned}$$

*Varianz* von $X$ und

$$\sigma_X := \sqrt{\mathrm{Var}(X)}$$

*Standardabweichung* von $X$.

### 1.5.3 Allgemeine Rechenregeln für Erwartungswert und Varianz

**Satz 1.21** Seien $X$ und $Y$ diskrete oder stetige Zufallsvariablen (mit existierenden Erwartungswerten und Varianzen). Dann gilt:

a) $\mathrm{E}(aX+bY) = a \cdot \mathrm{E}(X) + b \cdot \mathrm{E}(Y)$ und insbesondere auch

$$\begin{aligned} \mathrm{E}(a) &= a, \\ \mathrm{E}(aX) &= a \cdot \mathrm{E}(X) \\ \mathrm{E}(X+Y) &= \mathrm{E}(X) + \mathrm{E}(Y) \end{aligned}$$

b) $\mathrm{Var}(aX+b) = a^2 \cdot \mathrm{Var}(X)$.

Sind $X$ und $Y$ zusätzlich unabhängig, so gilt

$$\begin{aligned} \mathrm{E}(X \cdot Y) &= \mathrm{E}(X) \cdot \mathrm{E}(Y) \\ \mathrm{Var}(X+Y) &= \mathrm{Var}(X) + \mathrm{Var}(Y) \end{aligned}$$

**Definition 1.22** Die Zufallsvariable

$$Z := \frac{X - \mathrm{E}(X)}{\sqrt{\mathrm{Var}(X)}}$$

heißt *standardisierte Zufallsvariable*. Es gilt

$$\mathrm{E}(Z) = 0 \text{ und } \mathrm{Var}(Z) = 1.$$

### 1.5.4 Quantile und Modi

Gegeben sei eine Zufallsvariable X mit Wahrscheinlichkeitsverteilung $P_X$ und Verteilungsfunktion $F_X$.

a) Für $\alpha \in (0,1)$ heißt $x_\alpha$ zugehöriges *$\alpha$-Quantil* genau dann, wenn $F(x_\alpha) = \alpha$. Insbesondere ergeben sich für $\alpha$=0.5 der *Median* und für $\alpha$=0.25 bzw. 0.75 das *25%* bzw. *75% Quantil*.

b) Ist X diskret, so heißt jedes $x_{\mathrm{mod}}$ mit

$$P(X = x_{mod}) \geq P(X = x), \text{ für alle } x \in \mathcal{X}$$

*Modus (Modalwert) von $P_X$*.
Ist X stetig mit Dichtefunktion $f_X$, so heißt jedes $x_{\mathrm{mod}}$ mit $f_X(x_{\mathrm{mod}}) \geq f_X(x)$, für alle $x \in \mathcal{X}$ *Modus von $P_X$* (bezüglich der Dichte $f_X$).

**Modalbereiche** Sei $\gamma \in (0,1)$. Ein Intervall $A_\gamma \subset \mathbb{R}$, $A_\gamma = [a,b]$, heiße *Modalintervall zum Niveau $\gamma$*, wenn gilt

$$P(X \in A_\gamma) \geq \gamma$$

und $P(X \in C) < \gamma$ für alle Intervalle $C \subset \mathbb{R}$, $C = [c,d]$ mit $d - c < b - a$. Dann ist $A_\gamma \cap \mathcal{X}$ kleinster ("präzisester") zusammengehöriger Bereich, der mindestens Wahrscheinlichkeit $\gamma$ besitzt.

# 1.6 Wichtige Verteilungsmodelle

## 1.6.1 Binomialverteilung

**Definition 1.23** Eine Zufallsvariable $X$ heißt *binomialverteilt* mit dem Parametern $\pi$ bei $n$ Versuchen, kurz $X \sim B(n, \pi)$, wenn sie die Wahrscheinlichkeitsfunktion

$$f(x) = P(X = x) = \begin{cases} \binom{n}{x} \pi^x (1-\pi)^{n-x}, & x = 0, 1, \ldots, n \\ 0, & \text{sonst} \end{cases}$$

besitzt.

**Alternative Darstellung:** Die Zufallsvariable $X$ kann als

$$X = X_1 + \ldots + X_n$$

mit den binären Variablen

$$X_i = \begin{cases} 1 & \text{falls } A \text{ beim } i\text{-ten Versuch eintritt,} \\ 0 & \text{sonst.} \end{cases}$$

dargestellt werden. Die $X_i$ seien stochastisch unabhängig mit $X_i \sim B(1, \pi)$. $B(1, \pi)$ heißt *Bernoulliverteilung.*

### Erwartungswert und Varianz

- Erwartungswert der Binomialverteilung:

$$\mathrm{E}(X) = \mathrm{E}(X_1 + \ldots + X_n) = \mathrm{E}(X_1) + \ldots + \mathrm{E}(X_n) = n\pi$$

- Varianz der Binomialverteilung:

$$\mathrm{Var}(X) = \mathrm{Var}(X_1 + \ldots + X_n) = \mathrm{Var}(X_1) + \ldots + \mathrm{Var}(X_n) = n\pi(1-\pi)$$

### Eigenschaften

- Symmetrieeigenschaft:
  Sei $X \sim B(n, \pi)$ und $Y = n - X$. Dann gilt $Y \sim B(n, 1-\pi)$.

- Summeneigenschaft:
  Seien $X \sim B(n, \pi)$ und $Y \sim B(m, \pi)$. Sind $X$ und $Y$ unabhängig, so gilt

$$X + Y \sim B(n + m, \pi).$$

### 1.6.2 Poisson-Verteilung

**Definition 1.24** Eine Zufallsvariable $X$ mit der Wahrscheinlichkeitsfunktion

$$f(x) = P(X = x) = \begin{cases} \dfrac{\lambda^x}{x!}\,\mathrm{e}^{-\lambda}, & x \in \mathbb{N}_0 \\ 0, & \text{sonst} \end{cases}$$

heißt *Poisson-verteilt* mit Parameter (oder Rate) $\lambda > 0$, kurz $X \sim Po(\lambda)$.

**Erwartungswert und Varianz**

- Erwartungswert der Poissonverteilung:

$$\mathrm{E}(X) = \lambda$$

- Varianz der Poissonverteilung:

$$\mathrm{Var}(X) = \lambda$$

**Poisson-Verteilung als Näherungsmodell für die Binomialverteilung:**

$$X \sim B(n, \pi) \underset{\substack{n \text{ groß} \\ \pi \text{ klein}}}{\Longrightarrow} X \approxeq Po(n \cdot \pi)$$

**Satz 1.25 (Addition von Poisson-verteilten Zufallsvariablen)** Sind $X \sim Po(\lambda_X), Y \sim Po(\lambda_Y)$ voneinander unabhängig, so gilt

$$X + Y \sim Po(\lambda_X + \lambda_Y).$$

### 1.6.3 Normalverteilung

**Definition 1.26** Eine stetige Zufallsvariable $X$ heißt *normalverteilt* mit den Parametern $\mu$ und $\sigma^2$, kurz $X \sim \mathcal{N}(\mu, \sigma^2)$, wenn für ihre Dichte gilt :

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

und *standardnormalverteilt*, in Zeichen $X \sim \mathcal{N}(0; 1)$, falls $\mu = 0$ und $\sigma^2 = 1$ gilt ($\pi$ ist hier die Kreiszahl $\pi = 3.14\ldots$).

#### Bemerkungen

- $\mu$ und $\sigma^2$ sind genau der Erwartungswert und die Varianz, also, wenn $X \sim \mathcal{N}(\mu, \sigma^2)$, dann
$$\mathrm{E}(X) = \mu \quad \text{und} \quad \mathrm{Var}(X) = \sigma^2.$$
- Die Dichte ist symmetrisch um $\mu$, d.h.
$$f(\mu - x) = f(\mu + x)\,.$$
- Abgeschlossenheit gegenüber Linearkombinationen:
Seien $X_1$ und $X_2$ unabhängig und $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2), \ i = 1, 2$. Ferner seien $b, a_1, a_2$ feste reelle Zahlen. Dann gilt
$$\begin{aligned} Y_1 &:= a_1 X_1 + b \sim \mathcal{N}(a_1\mu_1 + b; a_1^2\sigma_1^2) \\ Y_2 &:= a_1 X_1 + a_2 X_2 \sim \mathcal{N}(a_1\mu_1 + a_2\mu_2; a_1^2\sigma_1^2 + a_2^2\sigma_2^2). \end{aligned}$$
- **Standardisierung:** Gilt $X \sim \mathcal{N}(\mu, \sigma^2)$, so ist die zugehörige standardisierte Zufallsvariable
$$Z = \frac{X - \mu}{\sigma}$$
standardnormalverteilt.

#### Standardnormalverteilung

- Die *Dichte der Standardnormalverteilung* wird oft mit $\varphi(z)$ bezeichnet, also
$$\varphi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right)$$
und die zugehörige Verteilungsfunktion mit
$$\Phi(z) = \int_{-\infty}^{z} \varphi(u)\, du$$
- $\Phi(z)$ lässt sich nicht in geschlossener Form durch bekannte Funktionen beschreiben
$\Longrightarrow$ numerische Berechnung, Tabellierung.
- Funktionswerte für negative Argumente::
$$\Phi(-z) = 1 - \Phi(z)$$

# 1.7 Grenzwertsätze und Approximationen

## 1.7.1 Das i.i.d.-Modell

## 1.7.2 Das schwache Gesetz der großen Zahlen

**Satz 1.27 (Theorem von Bernoulli)** Seien $X_1, \ldots, X_n$, i.i.d. mit $X_i \in \{0, 1\}$ und $P(X_i = 1) = \pi$. Dann gilt für

$$H_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

und jedes $\epsilon > 0$

$$\lim_{n \to \infty} P(|H_n - \pi| \leq \epsilon) = 1$$

**Schwaches Gesetz der großen Zahl:** Gegeben seien $X_1, \ldots, X_n$ i.i.d. Zufallsvariablen mit (existierendem) Erwartungswert $\mu$ und (existierender) Varianz $\sigma^2$. Dann gilt für

$$\bar{X}_n := \frac{1}{n} \sum_{i=1}^{n} X_i$$

und beliebiges $\epsilon > 0$:

$$\lim_{n \to \infty} P(|\bar{X}_n - \mu| \leq \epsilon) = 1$$

Schreibweise:

$$\bar{X}_n \xrightarrow{P} \mu$$

(„Stochastische Konvergenz“, „$X_n$ konvergiert in Wahrscheinlichkeit gegen $\mu$“.)

## 1.7.3 Der Hauptsatz der Statistik

**Satz 1.28** Seien $X_1, \ldots, X_n$ i.i.d. mit Verteilungsfunktion $F$ und sei $F_n(x)$ die empirische Verteilungsfunktion der ersten $n$ Beobachtungen. Mit

$$D_n := \sup_x |F_n(x) - F(x)|,$$

gilt für jedes $c > 0$

$$\lim_{n \to \infty} P(D_n > c) = 0.$$

## 1.7.4 Der zentrale Grenzwertsatz

**Satz 1.29** Seien $X_1, \ldots, X_n$ i.i.d. mit (existierendem) $\mathrm{E}(X_i) = \mu$ und (existierender) $\mathrm{Var}(X_i) = \sigma^2 > 0$ sowie

$$Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left( \frac{X_i - \mathrm{E}(X_i)}{\sqrt{\mathrm{Var}(X_i)}} \right).$$

Dann gilt: $Z_n$ ist *asymptotisch standardnormalverteilt*, in Zeichen: $Z_n \overset{a}{\sim} \mathcal{N}(0; 1)$, d.h. es gilt für jedes $z$

$$\lim_{n \to \infty} P(Z_n \leq z) = \Phi(z).$$

**Wichtige Anwendung: Approximation der Binomialverteilung**

Sei $X \sim B(n, \pi)$. Falls $n$ groß genug ist, gilt:

$$P(X \leq x) \;\approx\; \Phi\left(\frac{x - n \cdot \pi}{\sqrt{n \cdot \pi(1-\pi)}}\right).$$

Stetigkeitskorrektur:

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

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

Es gibt verschiedene Faustregeln, ab wann diese Approximation gut ist, z.B.

$$n \cdot \pi \geq 5 \;\text{ und }\; n \cdot (1-\pi) \geq 5$$
$$n \cdot \pi(1-\pi) \geq 9$$

## 1.8 Mehrdimensionale Zufallsvariablen

**Definition 1.30** Betrachtet werden zwei eindimensionale diskrete Zufallselemente $X$ und $Y$ (zu demselben Zufallsexperiment). Die Wahrscheinlichkeit

$$P(X = x_i, Y = y_j) := P(\{X = x_i\} \cap \{Y = y_j\})$$

in Abhängigkeit von $x_i$ und $y_j$ $(i = 1, \ldots, k,\ j = 1, \ldots, m)$ heißt *gemeinsame Verteilung* der mehrdimensionalen Zufallsvariable $\binom{X}{Y}$ bzw. der Variablen $X$ und $Y$ $(i = 1, ..., k, j = 1, ..., m)$.

**Randwahrscheinlichkeiten:**

$$p_{i\bullet} = P(X = x_i) = \sum_{j=1}^{m} P(X = x_i, Y = y_j)$$

$$p_{\bullet j} = P(Y = y_j) = \sum_{i=1}^{k} P(X = x_i, Y = y_j)$$

**Bedingte Verteilungen:**

$$P(X = x_i | Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)}$$

$$P(Y = y_j | X = x_i) = \frac{P(X = x_i, Y = y_j)}{P(X = x_i)}$$

**Stetiger Fall**: Zufallsvariable mit zweidimensionaler Dichtefunktion $f(x, y)$:

$$P(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \left(\int_c^d f(x, y) dy\right) dx$$

**Definition 1.31** Seien $X$ und $Y$ zwei Zufallsvariablen. Dann heißt

$$\sigma_{X,Y} := \mathrm{Cov}(X,Y) = \mathrm{E}\big((X-\mathrm{E}(X))(Y-\mathrm{E}(Y))\big)$$

*Kovarianz* von $X$ und $Y$.

### Rechenregeln

a) $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$

b) $\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\cdot\mathrm{E}(Y)$ (Verschiebungssatz)

c) $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$

d) Mit $\tilde{X} = a_X X + b_X$ und $\tilde{Y} = a_Y Y + b_Y$ ist

$$\mathrm{Cov}(\tilde{X},\tilde{Y}) = a_X \cdot a_Y \cdot \mathrm{Cov}(X,Y)$$

e) $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\cdot\mathrm{Cov}(X,Y)$

**Definition 1.32** Zwei Zufallsvariablen $X$ und $Y$ mit $\mathrm{Cov}(X,Y) = 0$ heißen *unkorreliert.*

**Satz 1.33** Stochastisch unabhängige Zufallsvariablen sind unkorreliert. Die Umkehrung gilt jedoch im Allgemeinen nicht.

**Definition 1.34** Gegeben seien zwei Zufallsvariablen $X$ und $Y$. Dann heißt

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}$$

*Korrelationskoeffizient* von $X$ und $Y$.

### Eigenschaften des Korrelationskoeffizienten

- Mit $\tilde{X} = a_X X + b_X$ und $\tilde{Y} = a_Y Y + b_Y$ ist

$$|\rho(\tilde{X},\tilde{Y})| = |\rho(X,Y)|.$$

- $-1 \leq \rho(X,Y) \leq 1.$
- $|\rho(X,Y)| = 1 \iff Y = aX + b$
- Sind $\mathrm{Var}(X) > 0$ und $\mathrm{Var}(Y) > 0$, so gilt $\rho(X,Y) = 0$ genau dann, wenn $\mathrm{Cov}(X,Y) = 0$.

# 2 Induktive Statistik

## 2.1 Grundprinzipien der induktiven Statistik

## 2.2 Punktschätzung

### 2.2.1 Schätzfunktionen

**Definition 2.1** Sei $X_1, \dots, X_n$ i.i.d. Stichprobe. Eine Funktion

$$T = g(X_1, \dots, X_n)$$

heißt *Schätzer* oder *Schätzfunktion*.

**Andere Notation in der Literatur:** $\hat{\vartheta}$ Schätzer für $\vartheta$.

### 2.2.2 Gütekriterien

**Erwartungstreue, Bias**: Gegeben sei eine Stichprobe $X_1, \dots, X_n$ und eine Schätzfunktion $T = g(X_1, \dots, X_n)$ (mit existierendem Erwartungswert).

- $T$ heißt *erwartungstreu für den Parameter* $\vartheta$, falls gilt

  $$\mathrm{E}_\vartheta(T) = \vartheta$$

  für alle $\vartheta$.

- Die Größe

  $$\mathrm{Bias}_\vartheta(T) = \mathrm{E}_\vartheta(T) - \vartheta$$

  heißt *Bias* (oder *Verzerrung*) der Schätzfunktion. Erwartungstreue Schätzfunktionen haben per Definition einen Bias von 0.

- Die korrigierte Stichprobenvarianz

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

  ist erwartungstreu für $\sigma^2$.

### 2.2.3 Effizienz

#### Effizienz

- Gegeben seien zwei erwartungstreue Schätzfunktionen $T_1$ und $T_2$ für einen Parameter $\vartheta$. Gilt
$$\mathrm{Var}_\vartheta(T_1) \leq \mathrm{Var}_\vartheta(T_2) \text{ für alle } \vartheta$$
und
$$\mathrm{Var}_{\vartheta^*}(T_1) < \mathrm{Var}_{\vartheta^*}(T_2) \text{ für mindestens ein } \vartheta^*$$
so heißt $T_1$ *effizienter als* $T_2$.

- Eine für $\vartheta$ erwartungstreue Schätzfunktion $T^{\text{opt}}$ heißt *UMVU-Schätzfunktion* für $\vartheta$ (*u*niformly *m*inimum *v*ariance *u*nbiased), falls
$$\mathrm{Var}_\vartheta(T^{\text{opt}}) \leq \mathrm{Var}_\vartheta(T)$$
für alle $\vartheta$ und für alle erwartungstreuen Schätzfunktionen $T$.

#### MSE

$$\mathrm{MSE}_\vartheta(T) := \mathrm{E}_\vartheta(T - \vartheta)^2 = \mathrm{Var}_\vartheta(T) + (\mathrm{Bias}_\vartheta(T))^2$$

### 2.2.4 Asymptotische Gütekriterien

- Eine Schätzfunktion heißt *asymptotisch erwartungstreu*, falls
$$\lim_{n\to\infty} \mathrm{E}(\hat{\vartheta}) = \vartheta$$
bzw.
$$\lim_{n\to\infty} \mathrm{Bias}(\hat{\vartheta}) = 0$$
gelten.

- Ein Schätzer heißt *(MSE-)konsistent* oder *konsistent im quadratischen Mittel*, wenn gilt
$$\lim_{n\to\infty} (\mathrm{MSE}(T)) = 0.$$

### 2.2.5 Konstruktionsprinzipien guter Schätzer

#### Die Methode der kleinsten Quadrate

#### Das Maximum-Likelihood-Prinzip

**Definition 2.2** Gegeben sei die Realisation $x_1, \dots, x_n$ einer i.i.d. Stichprobe. Die Funktion in $\vartheta$

$$L(\vartheta || x_1, \dots, x_n) = \begin{cases} \prod_{i=1}^{n} P_\vartheta(X_i = x_i) & \text{falls } X_i \text{ diskret} \\ \prod_{i=1}^{n} f_\vartheta(x_i) & \text{falls } X_i \text{ stetig.} \end{cases}$$

heißt *Likelihood* des Parameters $\vartheta$ bei der Beobachtung $x_1, \dots, x_n$.
Derjenige Wert $\hat{\vartheta} = \hat{\vartheta}(x_1, \dots, x_n)$, der $L(\vartheta)$ maximiert, heißt *Maximum-Likelihood-Schätzwert*; die zugehörige Schätzfunktion $T(X_1, \dots, X_n)$ *Maximum-Likelihood-Schätzer*.

#### Praktische Berechnung

Für die praktische Berechnung maximiert man statt der Likelihood typischerweise die Log-Likelihood

$$l(\vartheta || x_1, \dots, x_n) = \ln(L(\vartheta)) = \ln \prod_{i=1}^{n} P_\vartheta(X_i = x_i) = \sum_{i=1}^{n} \ln P_\vartheta(X_i = x_i)$$

bzw.

$$l(\vartheta || x_1, \dots, x_n) = \ln \prod_{i=1}^{n} f_\vartheta(x_i) = \sum_{i=1}^{n} \ln f_\vartheta(x_i).$$

## 2.3 Intervallschätzung

### 2.3.1 Motivation und Hinführung

### 2.3.2 Konfidenzintervalle: Definition

**Definition 2.3** Gegeben sei eine i.i.d. Stichprobe $X_1, \dots, X_n$ zur Schätzung eines Parameters $\vartheta$ und eine Zahl $\gamma \in (0;1)$. Ein zufälliges Intervall $\mathcal{C}(X_1, \dots, X_n)$ heißt *Konfidenzintervall* zum *Sicherheitsgrad* $\gamma$ (Konfidenzniveau $\gamma$), falls für jedes $\vartheta$ gilt:

$$P_\vartheta(\vartheta \in \underbrace{\mathcal{C}(X_1, \dots, X_n)}_{\text{zufälliges Intervall}}) \geq \gamma.$$

### 2.3.3 Konfidenzintervallen: Konstruktion

**Praktische Vorgehensweise:** Suche Zufallsvariable $Z_\vartheta$, die

- den gesuchten Parameter $\vartheta$ enthält und
- deren Verteilung aber nicht mehr von dem Parameter abhängt, („*Pivotgröße*“, dt. Angelpunkt).
- Dann wähle den Bereich $C_Z$ so, dass $P_\vartheta(Z_\vartheta \in C_Z) = \gamma$ und
- löse nach $\vartheta$ auf.

**Konfidenzintervall für den Mittelwert eines normalverteilten Merkmals**

- **bei bekannter Varianz:**

$$\left[\bar{X} - z_{\frac{1+\gamma}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\frac{1+\gamma}{2}} \frac{\sigma}{\sqrt{n}}\right] = \left[\bar{X} \pm z_{\frac{1+\gamma}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

- **bei unbekannter Varianz:**

$$\left[\bar{X} \pm t_{\frac{1+\gamma}{2}}(n-1) \frac{S}{\sqrt{n}}\right]$$

**Definition 2.4 (t-Verteilung)** Gegeben sei eine i.i.d. Stichprobe $X_1, \dots, X_n$ mit $X_i \sim \mathcal{N}(\mu, \sigma^2)$. Dann heißt die Verteilung von

$$Z = \frac{\bar{X} - \mu}{S} \cdot \sqrt{n}$$

*t-Verteilung* (oder Student-Verteilung) mit $\nu = n - 1$ *Freiheitsgraden*. In Zeichen: $Z \sim t(\nu)$.

### 2.3.4 Approximatives Konfidenzintervall für den Anteil $\pi$ eines binomialverteilten Merkmals

Seien $X_1, \dots, X_n$ i.i.d. mit $X_i = \begin{cases} 1 \\ 0 \end{cases}$ und $P(X_i = 1) = \pi$.

Dann gilt approximativ für großes $n$

$$\frac{\bar{X} - \pi}{\sqrt{\frac{\pi \cdot (1-\pi)}{n}}} \overset{a}{\sim} \mathcal{N}(0, 1)$$

und damit für das Konfidenzintervall:

$$\left[\bar{X} \pm z_{\frac{1+\gamma}{2}} \cdot \sqrt{\frac{\bar{X}(1-\bar{X})}{n}}\right].$$

### 2.3.5 Bestimmung des Stichprobenumfangs für die Anteilsschätzung

$\gamma$ Konfidenzniveau
$b_{\max}$ Schranke für die Genauigkeit

Löse $z_{\frac{1+\gamma}{2}} \cdot \sqrt{\frac{\bar{X}(1-\bar{X})}{n}} \leq b_{\max}$ nach $n$ auf: $\quad n \geq \frac{1}{b_{\max}^2} z_{\frac{1+\gamma}{2}}^2 \cdot \bar{X}(1-\bar{X})$

# 2.4 Hypothesentests

## 2.4.1 Grundprinzipien statistischer Hypothesentests

## 2.4.2 Konstruktion eines parametrischen statistischen Tests

## 2.4.3 Typische Tests I: Tests auf Lageparameter

### Gauß–Test

- Situation: $X_1, \ldots, X_n$ i.i.d. Stichprobe, wobei $X_i$ jeweils normalverteilt sei mit unbekanntem Mittelwert $\mu$ und bekannter Varianz $\sigma^2$.
- Statistische Hypothesen:

$$\begin{array}{llllll} \text{Fall 1:} & H_0: & \mu \leq \mu_0 & \text{gegen} & H_1: & \mu > \mu_0 \\ \text{Fall 2:} & H_0: & \mu \geq \mu_0 & \text{gegen} & H_1: & \mu < \mu_0 \\ \text{Fall 3:} & H_0: & \mu = \mu_0 & \text{gegen} & H_1: & \mu \neq \mu_0 \end{array}$$

- Testgröße:

$$T := \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}$$

  Falls $\mu = \mu_0$ gilt:

$$T = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n} \sim \mathcal{N}(0,1).$$

- Kritische Region:
  - Fall 1: $H_0$ ablehnen, falls:

$$T \geq z_{1-\alpha} \qquad \text{bzw.} \qquad \bar{X} \in \left[\mu_0 + z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}\, ;\, \infty\right)$$

  - Fall 2: $H_0$ ablehnen, falls:

$$T \leq -z_{1-\alpha} \qquad \text{bzw.} \qquad \bar{X} \in \left(-\infty; \mu_0 - z_{1-\alpha} \frac{\sigma}{\sqrt{n}}\right]$$

  - Fall 3: $H_0$ ablehnen, falls:

$$T \leq -z_{1-\frac{\alpha}{2}} \text{ oder } T \geq z_{1-\frac{\alpha}{2}}, \text{ also } |T| \geq z_{1-\frac{\alpha}{2}}.$$

### $t$–Test

- Situation: $X_1, \ldots, X_n$ i.i.d. Stichprobe, wobei $X_i$ jeweils normalverteilt sei mit unbekanntem Mittelwert $\mu$ und *unbekannter* Varianz $\sigma^2$
- Statistische Hypothesen:

$$\begin{array}{lllll} \text{Fall 1:} & H_0: & \mu \leq \mu_0 & \text{gegen} & H_1: & \mu > \mu_0 \\ \text{Fall 2:} & H_0: & \mu \geq \mu_0 & \text{gegen} & H_1: & \mu < \mu_0 \\ \text{Fall 3:} & H_0: & \mu = \mu_0 & \text{gegen} & H_1: & \mu \neq \mu_0 \end{array}$$

- Testgröße:

$$T = \frac{\bar{X} - \mu_0}{S} \sqrt{n}$$

  Falls $\mu = \mu_0$ gilt:

$$T = \frac{\bar{X} - \mu_0}{S} \sqrt{n} \sim t(n-1)$$

- Kritische Region:
  - Fall 1: $H_0$ ablehnen, falls $T \geq t_{1-\alpha}(n-1)$
  - Fall 2: $H_0$ ablehnen, falls $T \leq -t_{1-\alpha}(n-1)$
  - Fall 3: $H_0$ ablehnen, falls $T \leq -t_{1-\frac{\alpha}{2}}(n-1)$ oder $T \geq t_{1-\frac{\alpha}{2}}(n-1)$

### Approximative Tests für Hypothesen über Anteilswerte

- Situation: $X_1, \ldots, X_n$ i.i.d. Stichprobe, wobei $X_i = \begin{cases} 1 \\ 0 \end{cases}$ gelte und $P(X_i = 1) = \pi$ unbekannt sei.
- Statistische Hypothesen:

$$\begin{array}{lllll} \text{Fall 1:} & H_0: & \pi \leq \pi_0 & \text{gegen} & H_1: & \pi > \pi_0 \\ \text{Fall 2:} & H_0: & \pi \geq \pi_0 & \text{gegen} & H_1: & \pi < \pi_0 \\ \text{Fall 3:} & H_0: & \pi = \pi_0 & \text{gegen} & H_1: & \pi \neq \pi_0 \end{array}$$

- Testgröße:

$$\frac{\bar{X} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$$

  Falls $\pi = \pi_0$ gilt:

$$\frac{\bar{X} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} \overset{a}{\sim} \mathcal{N}(0,1),$$

- Kritische Region:
  - Fall 1: $H_0$ ablehnen, falls $T \geq z_{1-\alpha}$
  - Fall 2: $H_0$ ablehnen, falls $T \leq -z_{1-\alpha}$
  - Fall 3: $H_0$ ablehnen, falls $T \leq -z_{1-\frac{\alpha}{2}}$ oder $T \geq z_{1-\frac{\alpha}{2}}$

### 2.4.4 Typische Tests II: Lagevergleiche aus unabhängigen Stichproben

$X_1, \dots, X_n$ i.i.d. Stichprobe aus Gruppe $A$, $Y_1, \dots, Y_m$ i.i.d. Stichprobe aus Gruppe $B$

$$X_i \sim \mathcal{N}(\mu_X; \sigma_X^2) \qquad Y_i \sim \mathcal{N}(\mu_Y; \sigma_Y^2).$$

Statistische Hypothesen:

$$\begin{array}{llllll} \text{Fall 1:} & H_0: & \mu_X \le \mu_Y & \text{gegen} & H_1: & \mu_X > \mu_Y \\ \text{Fall 2:} & H_0: & \mu_X \ge \mu_Y & \text{gegen} & H_1: & \mu_X < \mu_Y \\ \text{Fall 3:} & H_0: & \mu_X = \mu_Y & \text{gegen} & H_1: & \mu_X \neq \mu_Y \end{array}$$

#### Zwei-Stichproben-Gauß–Test

Die Varianzen $\sigma_X^2$ und $\sigma_Y^2$ werden als bekannt angenommen.

- Testgröße:

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$$

  Falls $\mu_X = \mu_Y$ ist, gilt

$$T \sim \mathcal{N}(0, 1)$$

- Kritische Region:
  - Fall 1: $H_0$ ablehnen, falls $T \ge z_{1-\alpha}$
  - Fall 2: $H_0$ ablehnen, falls $T \le -z_{1-\alpha}$
  - Fall 3: $H_0$ ablehnen, falls $T \le -z_{1-\frac{\alpha}{2}}$ oder $T \ge z_{1-\frac{\alpha}{2}}$

#### Zwei-Stichproben-$t$-Test

Die Varianzen $\sigma_X^2$ und $\sigma_Y^2$ seien unbekannt.

- **Variante I**
  Ist bekannt, dass die Varianzen gleich sind, so schätzt man sie mittels $S_X^2$ und $S_Y^2$.
  Testgröße:

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\left(\frac{1}{n} + \frac{1}{m}\right) \frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}}}$$

  Falls $\mu_X = \mu_Y$ gehorcht $T$ einer $t$-Verteilung mit $(n + m - 2)$ Freiheitsgraden.

  Kritische Region:
  - Fall 1: $H_0$ ablehnen, falls $T \ge t_{1-\alpha}(n + m - 2)$
  - Fall 2: $H_0$ ablehnen, falls $T \le -t_{1-\alpha}(n + m - 2)$
  - Fall 3: $H_0$ ablehnen, falls $T \le -t_{1-\frac{\alpha}{2}}(n + m - 2)$ oder $T \ge t_{1-\frac{\alpha}{2}}(n + m - 2)$

- **Variante II**

  Können die Varianzen nicht als gleich angenommen werden, so kann für großes $n$ und großes $m$ mit folgender Testgröße gerechnet werden:

  Testgröße:

  $$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}}}$$

  $T$ ist für $\mu_X = \mu_Y$ approximativ standardnormalverteilt und kann auch angewendet werden, wenn keine Normalverteilung vorliegt.

  Kritische Region:

  - Fall 1: $H_0$ ablehnen, falls $T \geq z_{1-\alpha}$
  - Fall 2: $H_0$ ablehnen, falls $T \leq -z_{1-\alpha}$
  - Fall 3: $H_0$ ablehnen, falls $T \leq -z_{1-\frac{\alpha}{2}}$ oder $T \geq z_{1-\frac{\alpha}{2}}$

### 2.4.5 Gauß–Test und $t$-Test für verbundene Stichproben

*Verbundene Stichproben:* Vergleich zweier Merkmale $X$ und $Y$, die jetzt an *denselben* Einheiten erhoben werden.

$$\begin{aligned} X_1, \ldots, X_n &\quad \text{i.i.d.} \quad \mathcal{N}(\mu_X, \sigma_X^2) \\ Y_1, \ldots, Y_n &\quad \text{i.i.d.} \quad \mathcal{N}(\mu_Y, \sigma_Y^2) \end{aligned}$$

Zum Testen von Hypothesen der Form

$$\begin{array}{llll} \text{Fall 1:} & H_0: \ \mu_X \leq \mu_Y & \text{gegen} & H_1: \ \mu_X > \mu_Y \\ \text{Fall 2:} & H_0: \ \mu_X \geq \mu_Y & \text{gegen} & H_1: \ \mu_X < \mu_Y \\ \text{Fall 3:} & H_0: \ \mu_X = \mu_Y & \text{gegen} & H_1: \ \mu_X \neq \mu_Y \end{array}$$

betrachtet man die Differenz $D_i = X_i - Y_i$. Für den Erwartungswert $\mu_D$ gilt

$$\mu_D = \mathrm{E}(D_i) = \mu_X - \mu_Y$$

und für die Varianz $\sigma_D^2$

$$\sigma_D^2 = \sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY} \quad \text{mit } \sigma_{XY} = \mathrm{Cov}(X, Y)$$

Wegen $D_i \sim \mathcal{N}(\mu_D, \sigma_D^2)$ mit $\mu_D = \mu_X - \mu_Y$ und $\sigma_D^2 = \sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY}$ sind obige Hypothesen äquivalent zu den Hypothesen

$$\begin{array}{llll} \text{Fall 1':} & H_0: \ \mu_D \leq 0 & \text{gegen} & H_1: \ \mu_D > 0 \\ \text{Fall 2':} & H_0: \ \mu_D \geq 0 & \text{gegen} & H_1: \ \mu_D < 0 \\ \text{Fall 3':} & H_0: \ \mu_D = 0 & \text{gegen} & H_1: \ \mu_D \neq 0, \end{array}$$

und man kann unmittelbar die Tests aus 2.4.3 anwenden. Sind die Varianzen unbekannt, so kann man $\sigma_D$ aus den Differenzen $D_i$, $i = 1, \ldots, n$ schätzen. Zur Prüfung ist dann die $t$-Verteilung heranzuziehen.

### 2.4.6 $\chi 2$-Tests am Beispiel des $\chi 2$-Unabhängigkeitstests

#### $\chi 2$-Unabhängigkeitstest

$(X_1, Y_1), \ldots, (X_n, Y_n)$ i.i.d. Stichprobe des zwei-dimensionalen (diskreten) Merkmals $(X, Y)$.

- Statistische Hypothesen:
  $H_0$: Es herrscht Unabhängigkeit.
  $H_1$: Es herrscht keine Unabhängigkeit.

  d.h.

  $$\begin{array}{rll} H_0: & P(X = x_i, Y = y_j) = P(X = x_i) \cdot P(Y = y_j) & \text{für alle Paare } i, j \\ \text{gegen} \quad H_1: & P(X = x_i^*, Y = y_j^*) \neq P(X = x_i^*) \cdot P(Y = y_j^*) & \text{für mindestens } \textit{ein} \text{ Paar } i^*, j^* \end{array}$$

- Teststatistik:

  $$\begin{aligned} T &= \sum_{i=1}^{k} \sum_{j=1}^{m} \frac{\left(h_{ij} - \frac{h_{i\bullet} h_{\bullet j}}{n}\right)^2}{\frac{h_{i\bullet} h_{\bullet j}}{n}} \\ &= \sum_{i=1}^{k} \sum_{j=1}^{m} n \cdot \frac{\left(\frac{h_j}{n} - \frac{h_{i\bullet} h_{\bullet j}}{n^2}\right)^2}{\frac{h_{i\bullet} h_{\bullet j}}{n^2}} \\ &= \sum_{i=1}^{k} \sum_{j=1}^{m} n \cdot \frac{(f_{ij} - f_{i\bullet} f_{\bullet j})^2}{f_{i\bullet} f_{\bullet j}} \end{aligned}$$

  Unter $H_0$ gehorcht $T$ approximativ einer sogenannten $\chi^2$-*Verteilung mit* $(k-1) \cdot (m-1)$ *Freiheitsgraden.*

- Kritische Region:
  $H_0$ ablehnen, falls

  $$T \geq z,$$

  wobei $z = \chi^2_{1-\alpha}((k-1) \cdot (m-1))$ das $(1-\alpha)$-Quantil der $\chi^2$-Verteilung mit $(k-1)(m-1)$ Freiheitsgraden bezeichnet. Dabei ist $\alpha$ das Signifikanzniveau.

### 2.4.7 Zur praktischen Anwendung statistischer Tests

**Testentscheidungen und Statistik-Software, $p$-Wert**

**Dualität von Test und Konfidenzintervall**

**Signifikanz versus Relevanz**

**Multiple Testprobleme**

**Nichtparametrische Tests**

## 2.5 Lineare Regressionsmodelle

### 2.5.1 Wiederholung aus Statistik I

### 2.5.2 Lineare Einfachregression

#### Statistische Sichtweise

- Wahres Modell: $y_i = \beta_0 + \beta_1 x_i$
- Gestört durch zufällige Fehler $\epsilon_i$: Beobachtung von Datenpaaren $(X_i, Y_i)$, $i = 1, \ldots, n$ mit
$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i, \quad \text{wobei}$$
  - $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$
  - $\sigma^2$ für alle $i$ gleich
  - $\epsilon_{i_1}, \epsilon_{i_2}$ stochastisch unabhängig für $i_1 \neq i_2$

#### Maximum Likelihood Schätzer

$$\hat{\beta}_1 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2},$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X},$$

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^n \hat{\varepsilon}_i^2 \quad \text{mit den geschätzten Residuen } \hat{\varepsilon}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i$$

#### Konstruktion von Konfidenzintervallen und Tests

Mit

$$\hat{\sigma}_{\hat{\beta}_0} := \frac{\hat{\sigma}\sqrt{\sum_{i=1}^n X_i^2}}{\sqrt{n \sum_{i=1}^n (X_i - \bar{X})^2}} \qquad \text{gilt} \qquad \frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}_{\hat{\beta}_0}} \sim t(n-2)$$

und analog mit

$$\hat{\sigma}_{\hat{\beta}_1} := \frac{\hat{\sigma}}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2}} \qquad \text{gilt} \qquad \frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}} \sim t(n-2).$$

- Konfidenzintervalle zum Sicherheitsgrad $\gamma$:
$$\text{für } \beta_0\text{:} \quad \left[\hat{\beta}_0 \pm \hat{\sigma}_{\hat{\beta}_0} \cdot t_{\frac{1+\gamma}{2}}(n-2)\right]$$
$$\text{für } \beta_1\text{:} \quad \left[\hat{\beta}_1 \pm \hat{\sigma}_{\hat{\beta}_1} \cdot t_{\frac{1+\gamma}{2}}(n-2)\right]$$

- Hypothesen und kritischen Region:

| | Hypothesen | | | kritische Region |
|---|---|---|---|---|
| Fall 1: | $H_0\colon\ \beta_1 \leq \beta_1^*$ | gegen | $\beta_1 > \beta_1^*$ | $T_{\beta_1^*} \geq t_{1-\alpha}(n-2)$ |
| Fall 2: | $H_0\colon\ \beta_1 \geq \beta_1^*$ | gegen | $\beta_1 < \beta_1^*$ | $T_{\beta_1^*} \leq t_{1-\alpha}(n-2)$ |
| Fall 3: | $H_0\colon\ \beta_1 = \beta_1^*$ | gegen | $\beta_1 \neq \beta_1^*$ | $\lvert T_{\beta_1^*}\rvert \geq t_{1-\frac{\alpha}{2}}(n-2)$ |

mit der Teststatistik

$$T_{\beta_1^*} = \frac{\hat{\beta}_1 - \beta_1^*}{\hat{\sigma}_{\hat{\beta}_1}}$$

(analog für $\hat{\beta}_0$).

**Typischer SPSS-Output**

Koeffizienten[a]

| | | | Standardisierte Koeffizienten | | |
|---|---|---|---|---|---|
| | $\beta$ | Standardfehler | Beta | $T$ | Signifikanz |
| Konstante | $\hat{\beta}_0$ | $\hat{\sigma}_{\hat{\beta}_0}$ | 5) | 1) | 3) |
| Unabhängige Variable | $\hat{\beta}_1$ | $\hat{\sigma}_{\hat{\beta}_1}$ | 6) | 2) | 4) |

[a] abhängige Variable

1) Wert der Teststatistik
$$T_{\beta_0^*} = \frac{\hat{\beta}_0}{\hat{\sigma}_{\hat{\beta}_0}}.$$
zum Testen von $H_0$: $\beta_0 = 0$ gegen $H_1$: $\beta_0 \neq 0$.

2) Analog: Wert von
$$T_{\beta_1^*} = \frac{\hat{\beta}_1}{\hat{\sigma}_{\hat{\beta}_1}}$$
zum Testen von $H_0$: $\beta_1 = 0$ gegen $H_1$: $\beta_1 \neq 0$.

3) p-Wert zu 1)

4) p-Wert zu 2)

5), 6) hier nicht von Interesse.

### 2.5.3 Multiple lineare Regression

- Modellierungsansatz:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_p X_{ip} + \epsilon_i$$

- Es gilt für jedes $j = 0, \ldots, p$

$$\frac{\hat{\beta}_j - \beta_j}{\hat{\sigma}_{\hat{\beta}_j}} \sim t(n-p-1)$$

und man erhält wieder Konfidenzintervalle für $\beta_j$:

$$[\hat{\beta}_j \pm \hat{\sigma}_{\hat{\beta}_j} \cdot t_{\frac{1+\gamma}{2}}(n-p-1)]$$

sowie entsprechende Tests.

### 2.5.4 Varianzanalyse (Analysis of Variance, ANOVA)

Beobachtungen $Y_{ij}$, wobei: $j = 1, \ldots, J$ Faktorstufen
$i = 1, \ldots, n_j$ Personenindex in der $j$-ten Faktorstufe

**Zwei äquivalente Modellformulierungen:**

- Modell in Mittelwertdarstellung:

$$Y_{ij} = \mu_j + \epsilon_{ij} \qquad j = 1, \ldots, J, \; i = 1, \ldots, n_j,$$

mit: $\mu_j$ faktorspezifischer Mittelwert
$\epsilon_{ij}$ zufällige Störgröße
$\epsilon_{ij}$ $\sim \mathcal{N}(0, \sigma^2)$, $\epsilon_{11}, \epsilon_{12}, \ldots, \epsilon_{Jn_J}$ unabhängig.

**Testproblem:** $H_0$: $\mu_1 = \mu_2 = \ldots \mu_J$
gegen $H_1$: $\mu_l \neq \mu_q$ für mindestens ein Paar $(l, q)$

- Modell in Effektdarstellung:

$$Y_{ij} = \mu + \alpha_j + \epsilon_{ij}$$

wobei $\alpha_j$ so, dass

$$\sum_{j=1}^{J} n_j \alpha_j = 0.$$

mit: $\mu$ globaler Erwartungswert
$\alpha_j$ Effekt in der $j$-ten Faktorstufe, faktorspezifische systematische Abweichung vom gemeinsamen Mittelwert $\mu$

**Testproblem:** $H_0$: $\alpha_1 = \alpha_2 = \ldots \alpha_J = 0$
gegen $H_1$: $\alpha_j \neq 0$ für mindestens ein $j$

Die beiden Modelle sind äquivalent: setze $\mu_j := \mu + \alpha_j$!

**Streuungszerlegung**

Mittelwerte: $\bar{Y}_{\bullet\bullet}$ Gesamtmittelwert in der Stichprobe
$\bar{Y}_{\bullet j}$ Mittelwert in der $j$-ten Faktorstufe

Es gilt (vgl. Statistik I) die Streuungszerlegung:

$$\sum_{j=1}^{J}\sum_{i=1}^{n_j}(Y_{ij}-\bar{Y}_{\bullet\bullet})^2 = \underbrace{\sum_{j=1}^{J} n_j(\bar{Y}_{\bullet j}-\bar{Y}_{\bullet\bullet})^2}_{SQE} + \underbrace{\sum_{j=1}^{J}\sum_{i=1}^{n_j}(Y_{ij}-\bar{Y}_{\bullet j})^2}_{SQR}$$

- Hypothesen:

$$\begin{aligned} H_0:&\quad \mu_1 = \mu_2 = \ldots \mu_J \\ \text{gegen } H_1:&\quad \mu_l \neq \mu_q \quad \text{für mindestens ein Paar } (l,q) \end{aligned}$$

  beziehungsweise

$$\begin{aligned} H_0:&\quad \alpha_1 = \alpha_2 = \ldots \alpha_J = 0 \\ \text{gegen } H_1:&\quad \alpha_j \neq 0 \text{ für mindestens ein } j \end{aligned}$$

- Testgröße:

$$F = \frac{SQE/(J-1)}{SQR/(n-J)}$$

  Falls $H_0$ wahr ist, ist $T$ $F$-verteilt mit $(J-1)$ und $(n-J)$ Freiheitsgraden.

- Kritische Region:
  $H_0$ ablehnen, falls

$$T > F_{1-\alpha}(J-1, n-J),$$

  mit dem entsprechenden $(1-\alpha)$-Quantil der $F$-Verteilung mit $(J-1)$ und $(n-J)$ Freiheitsgraden.

# Standardnormalverteilung

Tabelliert sind die Werte der Verteilungsfunktion $\Phi(z) = P(Z \leq z)$ für $z \geq 0$.
Ablesebeispiel: $\Phi(1.75) = 0.9599$
Funktionswerte für negative Argumente: $\Phi(-z) = 1 - \Phi(z)$
Die $z$-Quantile ergeben sich über die Umkehrfunktion, z.B.: $z_{0.9599} = 1.75$ und $z_{0.9750} = 1.96$

| | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| 0.1 | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5753 |
| 0.2 | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| 0.3 | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| 0.4 | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| 0.5 | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |
| 0.6 | 0.7257 | 0.7291 | 0.7324 | 0.7357 | 0.7389 | 0.7422 | 0.7454 | 0.7486 | 0.7517 | 0.7549 |
| 0.7 | 0.7580 | 0.7611 | 0.7642 | 0.7673 | 0.7704 | 0.7734 | 0.7764 | 0.7794 | 0.7823 | 0.7852 |
| 0.8 | 0.7881 | 0.7910 | 0.7939 | 0.7967 | 0.7995 | 0.8023 | 0.8051 | 0.8078 | 0.8106 | 0.8133 |
| 0.9 | 0.8159 | 0.8186 | 0.8212 | 0.8238 | 0.8264 | 0.8289 | 0.8315 | 0.8340 | 0.8365 | 0.8389 |
| 1.0 | 0.8413 | 0.8438 | 0.8461 | 0.8485 | 0.8508 | 0.8531 | 0.8554 | 0.8577 | 0.8599 | 0.8621 |
| 1.1 | 0.8643 | 0.8665 | 0.8686 | 0.8708 | 0.8729 | 0.8749 | 0.8770 | 0.8790 | 0.8810 | 0.8830 |
| 1.2 | 0.8849 | 0.8869 | 0.8888 | 0.8907 | 0.8925 | 0.8944 | 0.8962 | 0.8980 | 0.8997 | 0.9015 |
| 1.3 | 0.9032 | 0.9049 | 0.9066 | 0.9082 | 0.9099 | 0.9115 | 0.9131 | 0.9147 | 0.9162 | 0.9177 |
| 1.4 | 0.9192 | 0.9207 | 0.9222 | 0.9236 | 0.9251 | 0.9265 | 0.9279 | 0.9292 | 0.9306 | 0.9319 |
| 1.5 | 0.9332 | 0.9345 | 0.9357 | 0.9370 | 0.9382 | 0.9394 | 0.9406 | 0.9418 | 0.9429 | 0.9441 |
| 1.6 | 0.9452 | 0.9463 | 0.9474 | 0.9484 | 0.9495 | 0.9505 | 0.9515 | 0.9525 | 0.9535 | 0.9545 |
| 1.7 | 0.9554 | 0.9564 | 0.9573 | 0.9582 | 0.9591 | 0.9599 | 0.9608 | 0.9616 | 0.9625 | 0.9633 |
| 1.8 | 0.9641 | 0.9649 | 0.9656 | 0.9664 | 0.9671 | 0.9678 | 0.9686 | 0.9693 | 0.9699 | 0.9706 |
| 1.9 | 0.9713 | 0.9719 | 0.9726 | 0.9732 | 0.9738 | 0.9744 | 0.9750 | 0.9756 | 0.9761 | 0.9767 |
| 2.0 | 0.9772 | 0.9778 | 0.9783 | 0.9788 | 0.9793 | 0.9798 | 0.9803 | 0.9808 | 0.9812 | 0.9817 |
| 2.1 | 0.9821 | 0.9826 | 0.9830 | 0.9834 | 0.9838 | 0.9842 | 0.9846 | 0.9850 | 0.9854 | 0.9857 |
| 2.2 | 0.9861 | 0.9864 | 0.9868 | 0.9871 | 0.9875 | 0.9878 | 0.9881 | 0.9884 | 0.9887 | 0.9890 |
| 2.3 | 0.9893 | 0.9896 | 0.9898 | 0.9901 | 0.9904 | 0.9906 | 0.9909 | 0.9911 | 0.9913 | 0.9916 |
| 2.4 | 0.9918 | 0.9920 | 0.9922 | 0.9925 | 0.9927 | 0.9929 | 0.9931 | 0.9932 | 0.9934 | 0.9936 |
| 2.5 | 0.9938 | 0.9940 | 0.9941 | 0.9943 | 0.9945 | 0.9946 | 0.9948 | 0.9949 | 0.9951 | 0.9952 |
| 2.6 | 0.9953 | 0.9955 | 0.9956 | 0.9957 | 0.9959 | 0.9960 | 0.9961 | 0.9962 | 0.9963 | 0.9964 |
| 2.7 | 0.9965 | 0.9966 | 0.9967 | 0.9968 | 0.9969 | 0.9970 | 0.9971 | 0.9972 | 0.9973 | 0.9974 |
| 2.8 | 0.9974 | 0.9975 | 0.9976 | 0.9977 | 0.9977 | 0.9978 | 0.9979 | 0.9979 | 0.9980 | 0.9981 |
| 2.9 | 0.9981 | 0.9982 | 0.9982 | 0.9983 | 0.9984 | 0.9984 | 0.9985 | 0.9985 | 0.9986 | 0.9986 |
| 3.0 | 0.9987 | 0.9987 | 0.9987 | 0.9988 | 0.9988 | 0.9989 | 0.9989 | 0.9989 | 0.9990 | 0.9990 |
| 3.1 | 0.9990 | 0.9991 | 0.9991 | 0.9991 | 0.9992 | 0.9992 | 0.9992 | 0.9992 | 0.9993 | 0.9993 |
| 3.2 | 0.9993 | 0.9993 | 0.9994 | 0.9994 | 0.9994 | 0.9994 | 0.9994 | 0.9995 | 0.9995 | 0.9995 |
| 3.3 | 0.9995 | 0.9995 | 0.9995 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9997 |
| 3.4 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9998 |
| 3.5 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 |
| 3.6 | 0.9998 | 0.9998 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 |
| 3.7 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 |

## Students $t$-Verteilung

Tabelliert sind die Quantile für $n$ Freiheitsgrade.
Für das Quantil $t_{1-\alpha}(\nu)$ gilt $F(t_{1-\alpha}(\nu)) = 1 - \alpha$.
Links vom Quantil $t_{1-\alpha}(\nu)$ liegt die Wahrscheinlichkeitsmasse $1 - \alpha$.
Ablesebeispiel: $t_{0.99}(20) = 2.528$
Die Quantile für $0 < 1 - \alpha < 0.5$ erhält man aus $t_\alpha(\nu) = -t_{1-\alpha}(\nu)$

**Approximation für $\nu > 30$:**

$$t_\alpha(\nu) \approx z_\alpha \quad (z_\alpha \text{ ist das } \alpha\text{-Quantil der Standardnormalverteilung})$$

| $\nu$ | 0.6 | 0.8 | 0.9 | 0.95 | 0.975 | 0.99 | 0.995 | 0.999 | 0.9995 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.3249 | 1.3764 | 3.0777 | 6.3138 | 12.706 | 31.821 | 63.657 | 318.31 | 636.62 |
| 2 | 0.2887 | 1.0607 | 1.8856 | 2.9200 | 4.3027 | 6.9646 | 9.9248 | 22.327 | 31.599 |
| 3 | 0.2767 | 0.9785 | 1.6377 | 2.3534 | 3.1824 | 4.5407 | 5.8409 | 10.215 | 12.924 |
| 4 | 0.2707 | 0.9410 | 1.5332 | 2.1318 | 2.7764 | 3.7469 | 4.6041 | 7.1732 | 8.6103 |
| 5 | 0.2672 | 0.9195 | 1.4759 | 2.0150 | 2.5706 | 3.3649 | 4.0321 | 5.8934 | 6.8688 |
| 6 | 0.2648 | 0.9057 | 1.4398 | 1.9432 | 2.4469 | 3.1427 | 3.7074 | 5.2076 | 5.9588 |
| 7 | 0.2632 | 0.8960 | 1.4149 | 1.8946 | 2.3646 | 2.9980 | 3.4995 | 4.7853 | 5.4079 |
| 8 | 0.2619 | 0.8889 | 1.3968 | 1.8595 | 2.3060 | 2.8965 | 3.3554 | 4.5008 | 5.0413 |
| 9 | 0.2610 | 0.8834 | 1.3830 | 1.8331 | 2.2622 | 2.8214 | 3.2498 | 4.2968 | 4.7809 |
| 10 | 0.2602 | 0.8791 | 1.3722 | 1.8125 | 2.2281 | 2.7638 | 3.1693 | 4.1437 | 4.5869 |
| 11 | 0.2596 | 0.8755 | 1.3634 | 1.7959 | 2.2010 | 2.7181 | 3.1058 | 4.0247 | 4.4370 |
| 12 | 0.2590 | 0.8726 | 1.3562 | 1.7823 | 2.1788 | 2.6810 | 3.0545 | 3.9296 | 4.3178 |
| 13 | 0.2586 | 0.8702 | 1.3502 | 1.7709 | 2.1604 | 2.6503 | 3.0123 | 3.8520 | 4.2208 |
| 14 | 0.2582 | 0.8681 | 1.3450 | 1.7613 | 2.1448 | 2.6245 | 2.9768 | 3.7874 | 4.1405 |
| 15 | 0.2579 | 0.8662 | 1.3406 | 1.7531 | 2.1314 | 2.6025 | 2.9467 | 3.7328 | 4.0728 |
| 16 | 0.2576 | 0.8647 | 1.3368 | 1.7459 | 2.1199 | 2.5835 | 2.9208 | 3.6862 | 4.0150 |
| 17 | 0.2573 | 0.8633 | 1.3334 | 1.7396 | 2.1098 | 2.5669 | 2.8982 | 3.6458 | 3.9651 |
| 18 | 0.2571 | 0.8620 | 1.3304 | 1.7341 | 2.1009 | 2.5524 | 2.8784 | 3.6105 | 3.9216 |
| 19 | 0.2569 | 0.8610 | 1.3277 | 1.7291 | 2.0930 | 2.5395 | 2.8609 | 3.5794 | 3.8834 |
| 20 | 0.2567 | 0.8600 | 1.3253 | 1.7247 | 2.0860 | 2.5280 | 2.8453 | 3.5518 | 3.8495 |
| 21 | 0.2566 | 0.8591 | 1.3232 | 1.7207 | 2.0796 | 2.5176 | 2.8314 | 3.5272 | 3.8193 |
| 22 | 0.2564 | 0.8583 | 1.3212 | 1.7171 | 2.0739 | 2.5083 | 2.8188 | 3.5050 | 3.7921 |
| 23 | 0.2563 | 0.8575 | 1.3195 | 1.7139 | 2.0687 | 2.4999 | 2.8073 | 3.4850 | 3.7676 |
| 24 | 0.2562 | 0.8569 | 1.3178 | 1.7109 | 2.0639 | 2.4922 | 2.7969 | 3.4668 | 3.7454 |
| 25 | 0.2561 | 0.8562 | 1.3163 | 1.7081 | 2.0595 | 2.4851 | 2.7874 | 3.4502 | 3.7251 |
| 26 | 0.2560 | 0.8557 | 1.3150 | 1.7056 | 2.0555 | 2.4786 | 2.7787 | 3.4350 | 3.7066 |
| 27 | 0.2559 | 0.8551 | 1.3137 | 1.7033 | 2.0518 | 2.4727 | 2.7707 | 3.4210 | 3.6896 |
| 28 | 0.2558 | 0.8546 | 1.3125 | 1.7011 | 2.0484 | 2.4671 | 2.7633 | 3.4082 | 3.6739 |
| 29 | 0.2557 | 0.8542 | 1.3114 | 1.6991 | 2.0452 | 2.4620 | 2.7564 | 3.3962 | 3.6594 |
| 30 | 0.2556 | 0.8538 | 1.3104 | 1.6973 | 2.0423 | 2.4573 | 2.7500 | 3.3852 | 3.6460 |
| $\infty$ | 0.2533 | 0.8416 | 1.2816 | 1.6449 | 1.9600 | 2.3263 | 2.5758 | 3.0903 | 3.2906 |

# $\chi^2$-Verteilung

Tabelliert sind die Quantile für $n$ Freiheitsgrade.
Für das Quantil $\chi^2_{1-\alpha,n}$ gilt $F(\chi^2_{1-\alpha,n}) = 1 - \alpha$.
Links vom Quantil $\chi^2_{1-\alpha,n}$ liegt die Wahrscheinlichkeitsmasse $1 - \alpha$.
Ablesebeispiel: $\chi^2_{0.95,\,10} = 18.307$

**Approximation für $n > 30$:**

$$\chi^2_\alpha(n) \approx \frac{1}{2}(z_\alpha + \sqrt{2n-1})^2 \quad (z_\alpha \text{ ist das } 1-\alpha\text{-Quantil der Standardnormalverteilung})$$

| $n$ | 0.01 | 0.025 | 0.05 | 0.1 | 0.5 | 0.9 | 0.95 | 0.975 | 0.99 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.0002 | 0.0010 | 0.0039 | 0.0158 | 0.4549 | 2.7055 | 3.8415 | 5.0239 | 6.6349 |
| 2 | 0.0201 | 0.0506 | 0.1026 | 0.2107 | 1.3863 | 4.6052 | 5.9915 | 7.3778 | 9.2103 |
| 3 | 0.1148 | 0.2158 | 0.3518 | 0.5844 | 2.3660 | 6.2514 | 7.8147 | 9.3484 | 11.345 |
| 4 | 0.2971 | 0.4844 | 0.7107 | 1.0636 | 3.3567 | 7.7794 | 9.4877 | 11.143 | 13.277 |
| 5 | 0.5543 | 0.8312 | 1.1455 | 1.6103 | 4.3515 | 9.2364 | 11.070 | 12.833 | 15.086 |
| 6 | 0.8721 | 1.2373 | 1.6354 | 2.2041 | 5.3481 | 10.645 | 12.592 | 14.449 | 16.812 |
| 7 | 1.2390 | 1.6899 | 2.1674 | 2.8331 | 6.3458 | 12.017 | 14.067 | 16.013 | 18.475 |
| 8 | 1.6465 | 2.1797 | 2.7326 | 3.4895 | 7.3441 | 13.362 | 15.507 | 17.535 | 20.090 |
| 9 | 2.0879 | 2.7004 | 3.3251 | 4.1682 | 8.3428 | 14.684 | 16.919 | 19.023 | 21.666 |
| 10 | 2.5582 | 3.2470 | 3.9403 | 4.8652 | 9.3418 | 15.987 | 18.307 | 20.483 | 23.209 |
| 11 | 3.0535 | 3.8157 | 4.5748 | 5.5778 | 10.341 | 17.275 | 19.675 | 21.920 | 24.725 |
| 12 | 3.5706 | 4.4038 | 5.2260 | 6.3038 | 11.340 | 18.549 | 21.026 | 23.337 | 26.217 |
| 13 | 4.1069 | 5.0088 | 5.8919 | 7.0415 | 12.340 | 19.812 | 22.362 | 24.736 | 27.688 |
| 14 | 4.6604 | 5.6287 | 6.5706 | 7.7895 | 13.339 | 21.064 | 23.685 | 26.119 | 29.141 |
| 15 | 5.2293 | 6.2621 | 7.2609 | 8.5468 | 14.339 | 22.307 | 24.996 | 27.488 | 30.578 |
| 16 | 5.8122 | 6.9077 | 7.9616 | 9.3122 | 15.338 | 23.542 | 26.296 | 28.845 | 32.000 |
| 17 | 6.4078 | 7.5642 | 8.6718 | 10.085 | 16.338 | 24.769 | 27.587 | 30.191 | 33.409 |
| 18 | 7.0149 | 8.2307 | 9.3905 | 10.865 | 17.338 | 25.989 | 28.869 | 31.526 | 34.805 |
| 19 | 7.6327 | 8.9065 | 10.117 | 11.651 | 18.338 | 27.204 | 30.144 | 32.852 | 36.191 |
| 20 | 8.2604 | 9.5908 | 10.851 | 12.443 | 19.337 | 28.412 | 31.410 | 34.170 | 37.566 |
| 21 | 8.8972 | 10.283 | 11.591 | 13.240 | 20.337 | 29.615 | 32.671 | 35.479 | 38.932 |
| 22 | 9.5425 | 10.982 | 12.338 | 14.041 | 21.337 | 30.813 | 33.924 | 36.781 | 40.289 |
| 23 | 10.196 | 11.689 | 13.091 | 14.848 | 22.337 | 32.007 | 35.172 | 38.076 | 41.638 |
| 24 | 10.856 | 12.401 | 13.848 | 15.659 | 23.337 | 33.196 | 36.415 | 39.364 | 42.980 |
| 25 | 11.524 | 13.120 | 14.611 | 16.473 | 24.337 | 34.382 | 37.652 | 40.646 | 44.314 |
| 26 | 12.198 | 13.844 | 15.379 | 17.292 | 25.336 | 35.563 | 38.885 | 41.923 | 45.642 |
| 27 | 12.879 | 14.573 | 16.151 | 18.114 | 26.336 | 36.741 | 40.113 | 43.195 | 46.963 |
| 28 | 13.565 | 15.308 | 16.928 | 18.939 | 27.336 | 37.916 | 41.337 | 44.461 | 48.278 |
| 29 | 14.256 | 16.047 | 17.708 | 19.768 | 28.336 | 39.087 | 42.557 | 45.722 | 49.588 |
| 30 | 14.953 | 16.791 | 18.493 | 20.599 | 29.336 | 40.256 | 43.773 | 46.979 | 50.892 |