# WuS - Cheatsheet

# Wahrscheinlichkeitsräume

## Grundraum

Terminologie: Die Menge $\Omega$ nennen wir Grundraum.

## Ereignisse

Ein Element $\omega \in \Omega$ nennen wir Elementarereignis (oder Ausgang des Experiments).
Definition 1.1: Ein Mengensystem $\mathcal{F} \subset \mathcal{P}(\Omega)$ heisst Sigma-Algebra ( $\sigma$-Algebra), falls es die folgenden Eigenschaften erfüllt

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
3. $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

Dabei nennen wir die Elemente der $\sigma$-Algebra Ereignisse.

## Wahrscheinlichkeitsmass

Definition 1.2: Sei $\Omega$ ein Grundraum und sei $\mathcal{F}$ eine $\sigma$-Algebra. Eine Abbildung

$$\begin{aligned} \mathbb{P} : \mathcal{F} &\to [0,1] \\ A &\mapsto \mathbb{P}[A] \end{aligned}$$

heisst Wahrscheinlichkeitsmass auf $(\Omega, \mathcal{F})$, falls folgende Eigenschaften gelten

1. $\mathbb{P}[\Omega] = 1.$
2. ( $\sigma$-Additivität) $\mathbb{P}[A] = \sum_{i=1}^{\infty} \mathbb{P}[A_i] \quad$ if $A = \bigcup_{i=1}^{\infty} A_i$ (disjunkte Vereinigung).

## Der Wahrscheinlichkeitsraum

Definition 1.3: Sei $\Omega$ ein Grundraum, $\mathcal{F}$ eine $\sigma$-Algebra, und $\mathbb{P}$ ein Wahrscheinlichkeitsmass. Wir nennen das Tripel $(\Omega, \mathcal{F}, \mathbb{P})$ Wahrscheinlichkeitsraum.
**Bemerkung 1.4:** Das Ereignis $A = \varnothing$ tritt niemals ein. Das Ereignis $A = \Omega$ tritt stets ein.
Definition 1.5: Sei $\Omega$ eine endlicher Grundraum. Das Laplace Modell auf $\Omega$ ist ein Tripel $(\Omega, \mathcal{F}, \mathbb{P})$, sodass

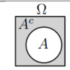
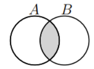
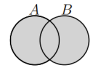
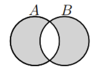
1. $\mathcal{F} = \mathcal{P}(\Omega)$,
2. $\mathbb{P} : \mathcal{F} \to [0,1]$ ist definiert durch

$$\forall A \in \mathcal{F} \quad \mathbb{P}[A] = \frac{|A|}{|\Omega|}.$$

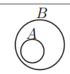
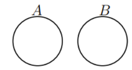
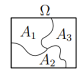
## Eigenschaften von Ereignissen

Satz 1.6 (Abgeschlossenheit der $\sigma$-Algebra bzgl. Operationen): Sei $\mathcal{F}$ eine $\sigma$-Algebra auf $\Omega$. Es gilt

1. $\varnothing \in \mathcal{F}$,
2. $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$,
3. $A, B \in \mathcal{F} \Rightarrow A \cup B \in \mathcal{F}$,
4. $A, B \in \mathcal{F} \Rightarrow A \cap B \in \mathcal{F}$.

| Ereignis | Grafische Darstellung | Probab. Interpretation |
|---|---|---|
| $A^c$ |   | $A$ tritt **nicht** ein |
| $A \cap B$ |   | $A$ **und** $B$ treten ein |
| $A \cup B$ |   | $A$ **oder** $B$ treten ein |
| $A \Delta B$ |   | **Entweder** $A$ oder $B$ tritt ein |

### Beziehung/Interpretationen zwischen Ereignissen

| Relation | Grafische Darstellung | Probab. Interpretation |
|---|---|---|
| $A \subset B$ |   | $B$ tritt ein, falls $A$ eintritt |
| $A \cap B = \varnothing$ |   | $A$ und $B$ können nicht gleichzeitig eintreten |
| $\Omega = A_1 \cup A_2 \cup A_3$ wobei $A_1, A_2, A_3$ paarweise disjunkt |   | für jedes Elementarereignis $\omega$, höchstens eins der Ereignisse $A_1$, $A_2$, $A_3$ kann eintreten. |

## Eigenschaften von Wahrscheinlichkeitsmassen

Satz 1.7: Sei $\mathbb{P}$ ein Wahrscheinlichkeitsmass auf $(\Omega, \mathcal{F})$. Es gilt:

1. $\mathbb{P}[\varnothing] = 0$
2. (Additivität) Sei $k \geq 1$, seien $A_1, \ldots, A_k$-viele paarweise disjunkte Ereignisse, dann gilt
$$\mathbb{P}[A_1 \cup \cdots \cup A_k] = \mathbb{P}[A_1] + \cdots + \mathbb{P}[A_k]$$
3. Sei $A$ ein Ereignis, dann gilt $\mathbb{P}[A^c] = 1 - \mathbb{P}[A]$
4. Falls $A$ und $B$ zwei (nicht notwendigerweise disjunkte) Ereignisse, dann gilt
$$\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B]$$

## Nützliche Ungleichungen

Satz 1.8 (Monotonie): Seien $A, B \in \mathcal{F}$, dann gilt $A \subset B \Rightarrow \mathbb{P}[A] \leq \mathbb{P}[B]$.
Satz 1.9 (Union Bound): Sei $A_1, A_2, \ldots$ eine Folge von (nicht notwendigerweise disjunkten) Ereignissen, dann gilt die folgende Ungleichung $\mathbb{P}\left[\bigcup_{i=1}^{\infty} A_i\right] \leq \sum_{i=1}^{\infty} \mathbb{P}[A_i]$.
**Bemerkung 1.10:** Die obrige Ungleichung gilt ebenfalls für eine Folge von endlich vielen nicht-leeren Ereignissen.

## Stetigkeit von Wahrscheinlichkeitsmassen

Satz 1.11: Sei $(A_n)$ eine monoton wachsende Folge von Ereignissen $(\forall n : A_n \subset A_{n+1})$. Dann gilt

$$\lim_{n \to \infty} P[A_n] = \mathbb{P}\left[\bigcup^{\infty} A_n\right]. \quad \text{monoton wachsender Grenzwert}$$

Sei $(B_n)$ eine monoton fallende Folge von Ereignissen $\forall n : (B_n \supset B_{n+1})$. Dann gilt

$$\lim_{n \to \infty} P[B_n] = \mathbb{P}\left[\bigcap_{n=1}^{\infty} B_n\right]. \quad \text{monoton fallender Grenzwert}$$

.
**Bemerkung 1.12:** Durch Monotonie erhalten wir $\mathbb{P}[A_n] \leq \mathbb{P}[A_{n+1}]$ und $\mathbb{P}[B_n] \geq \mathbb{P}[B_{n+1}]$ für jedes n. Daher sind die Grenzwerte in den obrigen Gleichungen wohldefiniert.

## Bedingte Wahrscheinlichkeit

Definition 1.13 (Bedingte Wahrscheinlichkeit): Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Seien $A, B$ zwei Ereignisse mit $\mathbb{P}[B] > 0$. Wir definieren die bedingte Wahrscheinlichkeit von $A$ gegeben $B$ wie folgt

$$\mathbb{P}[A \mid B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}.$$

**Bemerkung 1.14:** $\mathbb{P}[B \mid B] = 1$.
Satz 1.15: Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Sei $B$ ein Ereignis mit positiver Wahrscheinlichkeit. Dann ist $\mathbb{P}[. \mid B]$ ein $W$-Mass auf $\Omega$.
Satz 1.16 (Gesetz der totalen Wahrscheinlichkeit): Sei $B_1, \cdots, B_n$ eine Partition[a] des Grundraums $\Omega$, so dass $\mathbb{P}[B_i] > 0$ für jedes $1 \leq i \leq n$ gilt. Dann gilt

$$\forall A \in \mathcal{F} \quad \mathbb{P}[A] = \sum_{i=1}^{n} \mathbb{P}[A \mid B_i] \, \mathbb{P}[B_i]$$

[a] i.e. $\Omega = B_1 \cup \cdots \cup B_n$ mit paarweise disjunkten Ereignissen.
Satz 1.17 (Satz von Bayes): Sei $B_1, \ldots, B_n \in \mathcal{F}$ eine Partition von $\Omega$ sodass, $\mathbb{P}[B_i] > 0$ für jedes $i$ gilt. Für jedes Ereignis $A$ mit $\mathbb{P}[A] > 0$ gilt

$$\forall i = 1, \ldots, n \quad \mathbb{P}[B_i \mid A] = \frac{\mathbb{P}[A \mid B_i] \, \mathbb{P}[B_i]}{\sum_{j=1}^{n} \mathbb{P}[A \mid B_j] \, \mathbb{P}[B_j]}$$

**Bemerkung:** Eine Folgerung davon ist $\mathbb{P}[A|B] = \frac{\mathbb{P}[B|A] \cdot \mathbb{P}[A]}{\mathbb{P}[B]}$

## Unabhängigkeit von Ereignissen

Definition 1.18 (Unabhängigkeit): Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Zwei Ereignisse $A$ und $B$ heissen unabhängig falls $\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$.
**Bemerkung 1.19.** Falls $\mathbb{P}[A] \in \{0, 1\}$, dann ist A unabhängig von jedem Ereignis sodass,

$$\forall B \in \mathcal{F} \quad \mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B].$$

Falls ein Ereignis $A$ unabhängig von sich selbst ist, also $\mathbb{P}[A \cap A] = \mathbb{P}[A]^2$) gilt, dann muss $\mathbb{P}[A] \in \{0, 1\}$ gelten. $A$ ist unabhängig von $B$ genau dann wenn $A$ unabhängig von $B^c$ ist.
Satz 1.20: Seien $A, B \in \mathcal{F}$ zwei Ereignisse mit $\mathbb{P}[A], \mathbb{P}[B] > 0$. Dann sind folgende Aussagen äquivalent:

1. $\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$, $A$ und $B$ sind unabhängig
2. $\mathbb{P}[A \mid B] = \mathbb{P}[A]$, Eintreten von $B$ hat keinen Einfluss auf $A$
3. $\mathbb{P}[B \mid A] = \mathbb{P}[B]$. Eintreten von $A$ hat keinen Einfluss auf $B$

Definition 1.21: Sei $I$ eine beliebe Indexmenge. Eine Familie von Ereignissen $(A_i)_{i\in I}$ heisst unabhängig falls

$$\forall J \subset I \text{ endlich } \mathbb{P}\left[\bigcap_{j\in J} A_j\right] = \prod_{j\in J} \mathbb{P}\left[A_j\right].$$

**Bemerkung:** Drei Ereignisse $A, B$ und $C$ sind unabhängig falls alle 4 folgenden Gleichungen erfüllt sind (nicht nur die Letzte!):

$$\begin{aligned}\mathbb{P}[A\cap B] &= \mathbb{P}[A]\mathbb{P}[B],\\ \mathbb{P}[A\cap C] &= \mathbb{P}[A]\mathbb{P}[C],\\ \mathbb{P}[B\cap C] &= \mathbb{P}[B]\mathbb{P}[C],\\ \mathbb{P}[A\cap B\cap C] &= \mathbb{P}[A]\mathbb{P}[B]\mathbb{P}[C]\end{aligned}$$

# Zufallsvariablen und Verteilungsfunktionen

## Abstrakte Definition

Definition 2.1: Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Eine Zufallsvariable (Z.V.) ist eine Abbildung $X : \Omega \to \mathbb{R}$ sodass, für alle $a \in \mathbb{R}$ gilt,

$$\{\omega \in \Omega : X(\omega) \leq a\} \in \mathcal{F}.$$

Indikatorfunktion: Sei $A \in \mathcal{F}$. Wir definieren die Indikatorfunktion $\mathbb{1}_A$ auf $A$, durch

$$\forall \omega \in \Omega \quad \mathbb{1}_A(\omega) = \begin{cases} 0 & \text{if } \omega \notin A,\\ 1 & \text{if } \omega \in A.\end{cases}$$

.

**Notation:** Für Ereignisse im Bezug auf Z.V. werden wir auf darauf verzichten sie mittels Beziehung zu $\omega$ darzustellen. Stattdessen schreiben wir für $a \leq b$

$$\begin{aligned}\{X \leq a\} &= \{\omega \in \Omega : X(\omega) \leq a\}\\ \{a < X \leq b\} &= \{\omega \in \Omega : a < X(\omega) < b\},\\ \{X \in \mathbb{Z}\} &= \{\omega \in \Omega : X(\omega) \in \mathbb{Z}\}\end{aligned}$$

Betrachten wir die Wahrscheinlichkeit nach obigen Beispiel. Dann lassen wir gerade die Klammern weg und schreiben einfach

$$\mathbb{P}[X \leq a] = \mathbb{P}[\{X \leq a\}] = \mathbb{P}[\{\omega \in \Omega : X(\omega) \leq a\}].$$

## Verteilungsfunktion

Definition 2.2: Sei $X$ eine Zufallsvariable auf einem $W$-Raum $(\Omega, \mathcal{F}, \mathbb{P})$. Die Verteilungsfunktion von $X$ ist eine Funktion $F_X : \mathbb{R} \to [0,1]$, definiert durch

$$\forall a \in \mathbb{R} \quad F_X(a) = \mathbb{P}[X \leq a].$$

**Example 2:** Indikatorfunktion eines Ereignisses Sei $A$ ein Ereignis. Sei $X = \mathbb{1}_A$ eine Indikatorfunktion auf einem Ereignis $A$. Dann gilt

$$F_X(a) = \begin{cases} 0 & \text{falls } a < 0,\\ 1 - \mathbb{P}[A] & \text{falls } 0 \leq a < 1\\ 1 & \text{falls } a \geq 1\end{cases}$$

Satz 2.3 (Einfache Identität): Seien $a < b$ zwei reelle Zahlen. Dann gilt $\mathbb{P}[a < X \leq b] = F(b) - F(a)$.

Theorem 2.4 (Eigenschaften der Verteilungsfunktion): Sei $X$ eine Z.V. auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$. Die Verteilungsfunktion $F = F_X : \mathbb{R} \to [0,1]$ von $X$ erfüllt folgende Eigenschaften:

1. $F$ ist monoton wachsend
2. $F$ ist rechtsstetig [a]
3. $\lim_{a\to-\infty} F(a) = 0$ und $\lim_{a\to\infty} F(a) = 1$.

[a] Formal: $F(a) = \lim_{h\downarrow 0} F(a+h)$ für jedes $a \in \mathbb{R}$.

## Unabhängigkeit von Zufallsvariablen

Definition 2.5: Seien $X_1, \ldots, X_n$ Zufallsvariablen auf einem $W$-Raum $(\Omega, \mathcal{F}, \mathbb{P})$. Dann heissen $X_1, \ldots, X_n$ unabhängig falls $\forall x_1, x_2, \ldots, x_n \in \mathbb{R} \quad \mathbb{P}\left[X_1 \leq x_1, \ldots, X_n \leq x_n\right] = \mathbb{P}\left[X_1 \leq x_1\right] \ldots \mathbb{P}\left[X_n \leq x_n\right]$

**Bemerkung 2.6:** Man kann zeigen, dass $X_1, \ldots, X_n$ genau dann unabhängig sind, wenn folgende Bedingung gilt $\forall I_1 \subset \mathbb{R}, \ldots, I_n \subset \mathbb{R}$ Intervalle $\{X_1 \in I_1\}, \ldots, \{X_n \in I_n\}$ sind unabhängig .

## Gruppierung von Zufallsvariablen

Satz 2.7 (Gruppieren von Zufallsvariablen): Seien $X_1, \ldots, X_n n$ unabhängige Zufallsvariablen. Seien $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ Indexes und $\phi_1, \ldots, \phi_k$ Abbildungen. Dann sind $Y_1 = \phi_1\left(X_1, \ldots, X_{i_1}\right), Y_2 = \phi_2\left(X_{i_1+1}, \ldots, X_{i_2}\right) \ldots, Y_k = \phi_k\left(X_{i_{k-1}+1}, \ldots, X_{i_k}\right)$ unabhängig.

## Folgen von u.i.v. Zufallsvariablen

Definition 2.8: Eine Folge von Zufallsvariablen $X_1, X_2, \ldots$ heißt

1. unabhängig falls $X_1, \ldots, X_n$ unabhängig sind, für alle $n \in \mathbb{N}$.
2. unabhängig und identisch verteilt (uiv) falls sie unabhängig ist und die Zufallsvariablen dieselbe Verteilungsfunktion haben d.h. $\forall i, j \quad F_{X_i} = F_{X_j}$.

## Transformation von Zufallsvariablen

Falls $X$ eine Zufallsvariable ist, und $\phi : \mathbb{R} \to \mathbb{R}$, so schreiben wir $\phi(X) := \phi \circ X$. Somit ist $\phi(X)$ eine neue Abbildung von $\Omega \to \mathbb{R}$, welche in dem nachfolgenden Diagram dargestellt ist:

$$\Omega \xrightarrow{X} \mathbb{R} \xrightarrow{\phi} \quad \mathbb{R}$$
$$\omega \longmapsto X(\omega) \longmapsto \phi(X(\omega)).$$

## Konstruktion von Zufallsvariablen

Definition 2.9: Sei $p \in [0,1]$. Eine Zufallsvariable $X$ heißt Bernoulli Zufallsvariable mit Parameter $p$ falls

$$\mathbb{P}[X = 0] = 1 - p \quad \text{ und } \quad \mathbb{P}[X = 1] = p.$$

Dabei schreiben wir stets $X \sim \text{Ber}(p)$.

Theorem 2.10 (Existenzsatz von Kolmogorov). Es existiert ein $W$-Raum $(\Omega, \mathcal{F}, \mathbb{P})$ und eine nicht endliche u.i.v. Folge von Bernoulli Zufallsvariablen $X_1, X_2, \ldots$ auf $(\Omega, \mathcal{F}, \mathbb{P})$ mit Parameter $\frac{1}{2}$.

Definition 2.11: Eine Zufallsvariable $U$ heißt gleichverteilt auf $[0,1]$ falls ihre Verteilungsfunktion gegeben ist durch

$$F_U(x) = \begin{cases} 0 & x < 0\\ x & 0 \leq x \leq 1\\ 1 & x > 1\end{cases}$$

Wir schreiben gerade $U \sim \mathcal{U}([0,1])$.

Satz 2.12 Seien $X_1, X_2, \ldots$ eine Folge von unabhängigen Bernoulli-Zufallsvariablen mit Parameter $1/2$. Für jedes festes $\omega$ haben wir $X_1(\omega), X_2(\omega) \cdots \in \{0,1\}$. Daraus folgt, dass die unendliche Reihe

$$Y(\omega) = \sum_{n=1}^{\infty} 2^{-n} X_n(\omega)$$

absolut konvergiert, wobei $Y(\omega) \in [0,1]$ ist. Die Abbildung $Y : \Omega \to [0,1]$ ist eine gleichverteilte Zufallsvariable auf $[0,1]$.

Definition 2.13 (Pseudoinverse): Die Pseudoinverse von $F$ ist eine Abbildung $F^{-1} : (0,1) \to \mathbb{R}$ definiert durch

$$\forall \alpha \in (0,1) \quad F^{-1}(\alpha) = \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}.$$

Theorem 2.14 (Inversionsmethode): Sei $F : \mathbb{R} \to [0,1]$ eine Abbildung, welche Eigenschaften (i)-(iii) erfüllt. Sei $U$ eine Gleichverteilte Zufallsvariable. Dann besitzt die Zufallsvariable

$$X = F^{-1}(U)$$

gerade die Verteilungsfunktion $F_X = F$.

**Bemerkung 2.15:** Wir wollen nochmals kurz erläutern, warum die Definition von $X$ nach 2.14 wohldefiniert ist. Sei $U : \Omega \to [0,1]$ und $F^{-1} : (0,1) \to \mathbb{R}$ analog zum obigen Theorem definiert. Dann gilt stets $P[U \in (0,1) = 1]$. Strenggenommen ist $X$ bis jetzt nur auf einer Menge mit Wahrscheinlichkeit 1 aber nicht auf ganz $\Omega$ definiert. Wir beheben das Problem mittels folgender Definition

$$X(\omega) = \begin{cases} F^{-1}(U(\omega)) & \text{falls } U(\omega) \in (0,1)\\ 0 & \text{sonst.}\end{cases}$$

Theorem 2.16: Seien $F_1, F_2 \ldots$ eine Folge von Funktionen $\mathbb{R}$ auf $[0,1]$, die die Eigenschaften $(i) - (\text{ iii })$ am Anfang des Abschnitts erfüllen. Dann existiert ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ und eine Folge von unabhängigen Zufallsvariablen $X_1, X_2, \ldots$ auf diesem Wahrscheinlichkeitsraum, sodass

1. für jedes $i$ gilt: $X_i$ hat Verteilungsfunktion $F_i$ (d.h. $\forall x \mathbb{P}\left[X_i \leq x\right] = F_i(x)$ ), und
2. $X_1, X_2, \ldots$ sind unabhängig.

# Diskrete und stetige Zufallsvariablen

## Unstetigkeit/Stetigkeit der Verteilungsfunktion $F$

Satz 3.1 (Wahrscheinlichkeit eines Punktes): Sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable mit Verteilungsfunktion $F$. Für jedes a in $\mathbb{R}$ gilt $\mathbb{P}[X = a] = F(a) - F(a-)$. Sei $a \in \mathbb{R}$ fixiert.

1. Wenn $F$ in einem Punkt $a \in \mathbb{R}$ nicht stetig ist, dann ist die "Sprunghöhe" $F(a) - F(a-)$ gleich der Wahrscheinlichkeit, dass $X = a$.
2. Falls $F$ stetig in einem Punkt $a \in \mathbb{R}$ ist, dann gilt $\mathbb{P}[X = a] = 0$.

### Fast sichere Ereignisse

Definition 3.2: Sei $A \in \mathcal{F}$ ein Ereignis. Wir sagen $A$ tritt fast sicher (f.s.) ein, falls $\mathbb{P}[A] = 1$.
**Bemerkung 3.3:** Wir erweitern gerade diese Notation auf allgemeinere Mengen $A \subset \Omega$ (nicht zwangsweise ein Ereignis): Wir sagen dann, dass A fast sicher eintritt, falls ein Ereignis $A' \in \mathcal{F}$ existiert, sodass $A' \subset A$ und $\mathbb{P}[A'] = 1$.

### Diskrete Zufallsvariablen

Definition 3.4 (Diskrete Zufallsvariable): Eine Zufallsvariable $X : \Omega \to \mathbb{R}$ heisst diskret falls eine endliche oder abzählbare Menge $W \subset \mathbb{R}$ existiert, sodass $\mathbb{P}[X \in W] = 1$. "Die Werte von $X$ liegen in $W$ fast sicher."
**Bemerkung 3.5:** Wenn der Grundraum $\Omega$ endlich oder abzählbar ist, dann ist jede *Zu* fallsvariable $X : \Omega \to \mathbb{R}$ diskret. In der Tat ist das Bild $X(\Omega) = \{x \in \mathbb{R} : \exists \omega \in \Omega X(\omega) = x\}$ endlich oder abzählbar und wir haben $\mathbb{P}[X \in W] = 1$, mit $W = X(\Omega)$.
Definition 3.6: Sei $X$ eine diskrete Zufallsvariable mit Werten in einer endlichen oder abzählbaren Menge $W \subset \mathbb{R}$. Die Zahlenfolge $(p(x))_{x \in W}$ definiert durch

$$\forall x \in W \quad p(x) := \mathbb{P}[X = x]$$

heisst Verteilung von $X$.
Satz 3.7: Die Verteilung $(p(x))_{x \in W}$ einer diskreten Zufallsvariablen erfüllt $\sum_{x \in W} p(x) = 1$.
**Bemerkung 3.8:** Umgekehrt, wenn wir eine Folge von Zahlen $(p(x))_{x \in W}$ mit Werten in $[0, 1]$ gegeben haben, sodass $\sum_{x \in W} p(x) = 1$, dann gibt es einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ und eine Zufallsvariable $X$ mit zugehöriger Verteilung $(p(x))_{x \in W}$. Dies gilt nach dem Existenzsatz 2.16 in Kapitel 2. Diese Beobachtung ist in der Praxis wichtig, denn sie erlaubt uns zu schreiben: "Sei $X$ eine diskrete Zufallsvariable mit Verteilung $(p(x))_{x \in W}$."

### Verteilung $p$ vs. Verteilungsfunktion $F_X$

Satz 3.9: Sei $X$ eine diskrete Zufallsvariable, dessen Werte in einer endlichen oder abzählbaren Menge $W$ liegen, und deren Verteilung $p$ ist. Dann ist die Verteilungsfunktion von $X$ gegeben durch

$$\forall x \in \mathbb{R} \quad F_X(x) = \sum_{\substack{y \le x \\ y \in W}} p(y)$$

### Beispiele diskreter Zufallsvariablen

#### Bernoulli Verteilung

Definition 3.10 (Bernoulli Verteilung): Es sei $0 \le p \le 1$. Eine Zufallsvariable $X$ heisst Bernoulli Zufallsvariable mit Parameter $p$, wenn sie Werte in $W = \{0, 1\}$ annimmt und folgendes gilt

$$\mathbb{P}[X = 0] = 1 - p \quad \text{und} \quad \mathbb{P}[X = 1] = p.$$

In diesem Fall schreiben wir $X \sim \text{Ber}(p)$.

#### Binomialverteilung

Definition 3.11 (Binomialverteilung): Sei $0 \le p \le 1$, sei $n \in \mathbb{N}$. Eine Zufallsvariable $X$ heisst binomiale Zufallsvariable mit Parametern $n$ und p, wenn sie Werte in $W = \{0, \dots, n\}$ annimmt und folgendes gilt

$$\forall k \in \{0, \dots, n\} \quad \mathbb{P}[X = k] = \binom{n}{k} p^k (1-p)^{n-k}.$$

Wir schreiben dann $X \sim \text{Bin}(n, p)$.
**Bemerkung 3.12:** Wenn wir $p(k) = \binom{n}{k} p^k (1-p)^{n-k}$ definieren, haben wir

$$\sum_{k=0}^{n} p(k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p + 1 - p)^n = 1$$

**Binomialkoeffizient:** $\binom{n}{k} = \frac{n!}{(n-k)! \cdot k!}$
**Symmetrie der Binomialkoeffizienten:** $\binom{n}{k} = \binom{n}{n-k}$
Satz 3.13 (Summe von unabhängigen Bernoulli und Binomial Z.V.): Sei $0 \le p \le 1$, sei $n \in \mathbb{N}$. Seien $X_1, \dots, X_n$ unabhängige Bernoulli Z.V. mit Parameter $p$. Dann ist

$$S_n := X_1 + \cdots + X_n$$

eine binomialverteilte Z.V. mit Parametern $n$ und $p$.
**Bemerkung 3.14:** Insbesondere ist die Verteilung $\text{Bin}(1, p)$ gerade $\text{Ber}(p)$ verteilt. Es sei noch folgendes anzumerken: Falls $X \sim \text{Bin}(m, p), Y \sim \text{Bin}(n, p)$ und $X, Y$ unabhängig sind, dann ist $X + Y \sim \text{Bin}(m + n, p)$ verteilt.

#### Geometrische Verteilung

Definition 3.15 (Geometrische Verteilung): Es sei $0 < p \le 1$. Eine Zufallsvariable $X$ heisst geometrische Zufallsvariable mit Parameter $p$, falls sie Werte in $W = \mathbb{N} \backslash \{0\}$ annimmt und folgendes gilt

$$\forall k \in \mathbb{N} \backslash \{0\} \quad \mathbb{P}[X = k] = (1-p)^{k-1} \cdot p.$$

Wir schreiben dann $X \sim \text{Geom}(p)$.
**Bemerkung 3.16:** Für $p = 1$ und $k = 1$ erscheint in der obigen Gleichung ein Term $0^0$, wir verwenden die Konvention $0^0 = 1$ und damit gilt $\mathbb{P}[X = 1] = p$.
**Bemerkung 3.17:** Falls wir $p(k) = (1-p)^{k-1} p$ definieren, haben wir $\sum_{k=1}^{\infty} p(k) = p \sum_{k=1}^{\infty} (1-p)^{k-1} = p \cdot \frac{1}{p} = 1$.
Satz 3.18: Sei $X_1, X_2, \dots$ eine Folge von unendlich vielen unabhängigen BernoulliZ.V. mit Parameter $p$. Dann ist $T := \min\{n \ge 1 : X_n = 1\}$ eine geometrisch verteilte Zufallsvariable mit Parameter $p$.
**Bemerkung 3.19:** Falls wir sagen, dass $T$ eine geometrische Zufallsvariable ist, müssen wir folgendes präzisieren: Tatsächlich kann die Zufallsvariable $T$ den Wert $+\infty$ annehmen, wenn alle Zufallsvariablen $X_i$ gleich 0 sind. Dies ist jedoch kein Problem für den Bewei des Satzes. Man kann leicht überprüfen, dass $\mathbb{P}[T = \infty] = 0$ gilt.
Satz 3.20 (Gedächnislosigkeit der Geometrischen Verteilung): Sei $T \sim \text{Geom}(p)$ für $0 < p < 1$. Dann gilt
$\forall n \ge 0 \quad \forall k \ge 1 \quad \mathbb{P}[T \ge n + k \mid T > n] = \mathbb{P}[T \ge k]$.

#### Poisson Verteilung

Definition 3.21: Sei $\lambda > 0$ eine positive reelle Zahl. Eine Zufallsvariable $X$ heisst Poisson-Zufallsvariable mit Parameter $\lambda$, wenn sie Werte in $W = \mathbb{N}$ annimmt und folgendes gilt

$$\forall k \in \mathbb{N} \quad \mathbb{P}[X = k] = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Wir schreiben dann $X \sim \text{Poisson}(\lambda)$.
**Bemerkung 3.22:** Alternativ definieren wir $p(k) = \frac{\lambda^k}{k!} e^{-\lambda}$, haben wir $\sum_{k=0}^{\infty} p(k) = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} \cdot e^{\lambda} = 1$,
Satz 3.23 (Poisson-Approximation der Binomialverteilung): Sei $\lambda > 0$. Für jedes $n \ge 1$ seien $X_n \sim \text{Bin}\left(n, \frac{\lambda}{n}\right)$ Zufallsvariablen. Dann gilt

$$\forall k \in \mathbb{N} \quad \lim_{n \to \infty} \mathbb{P}[X_n = k] = \mathbb{P}[N = k],$$

wobei $N$ eine Poisson Zufallsvariable mit Parameter $\lambda$.
Reproduktivität: Die Poisson-Verteilung ist reproduktiv, d. h., die Summe $X_1 + X_2 + \cdots + X_n$ stochastisch unabhängiger Poisson-verteilter Zufallsvariablen $X_1, X_2, \dots, X_n$ mit den Parametern $\lambda_1, \lambda_2, \dots, \lambda_n$ ist wieder Poisson-verteilt mit dem Parameter $\lambda_1 + \lambda_2 + \cdots + \lambda_n$.

### Stetige Verteilungen

Definition 3.25 (Stetig verteilte Zufallsvariablen): Eine Zufallsvariable $X : \Omega \to \mathbb{R}$ heisst stetig, wenn ihre Verteilungsfunktion $F_X$ wie folgt geschrieben werden kann

$$F_X(a) = \int_{-\infty}^{a} f(x) dx \text{ für alle a in } \mathbb{R}.$$

wobei $f : \mathbb{R} \to \mathbb{R}_+$eine nicht-negative Funktion ist. Wir nennen dann $f$ Dichte von $X$.
**Intuition:** $f(x)dx$ ist die Wahrscheinlichkeit, dass $X$ Werte in $[x, x + dx]$ annimmt.
Theorem 3.26: Sei $X$ eine Zufallsvariable. Die Verteilungsfunktion $F_X$ sei stetig und stückweise $\mathcal{C}^1$, d.h. es gibt $x_0 = -\infty < x_1 < \cdots < x_{n-1} < x_n = +\infty$, sodass $F_X$ auf jedem Intervall $(x_i, x_{i+1})$ Element von $\mathcal{C}^1$ ist. Dann ist $X$ eine stetige Zufallsvariable und die Dichte $f$ kann konstruiert werden, indem man folgendes festlegt $\forall x \in (x_i, x_{i+1}) \quad f(x) = F_X'(x)$ mit beliebigen Werten in $x_1, \dots, x_{n-1}$.

### Beispiele stetiger Zufallsvariablen

#### Gleichverteilung

Definition 3.27 (Gleichverteilung auf $[a, b], a < b$.): Eine stetige Zufallsvariable $X$ heisst gleichverteilt auf $[a, b]$ falls ihre Dichte gegeben ist durch

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

Wir schreiben dann stets $X \sim \mathcal{U}([a, b])$.
**Eigenschaften einer gleichverteilten Zufallsvariable $X$ auf $[a, b]$:**

1. Die Wahrscheinlichkeit in ein Intervall $[c, c + \ell] \subset [a, b]$ zu fallen ist lediglich abhängig von dessen Länge $\ell$ : $\mathbb{P}[X \in [c, c + \ell]] = \frac{\ell}{b-a}$.

2. Die Verteilungsfunktion $X$ ist gegeben durch

$$F_X(x) = \begin{cases} 0 & x < a, \\ \frac{x-a}{b-a} & a \leq x \leq b, \\ 1 & x > b. \end{cases}$$

### Exponentialverteilung

Definition 3.28 (Exponentialverteilung mit Parameter $\lambda > 0$ ): Eine stetige Zufallsvariable $T$ heisst exponentialverteilt mit Parameter $\lambda > 0$ falls ihre Dichte gegeben ist durch

$$f_\lambda(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Wir schreiben dann stets $T \sim \exp(\lambda)$.
**Eigenschaften einer exponentialverteilten Zufallsvariable $T$ mit Parameter $\lambda$:**

1. Die Wahrscheinlichkeit des Wartens ist exponentiell klein: $\forall t \geq 0 \quad \mathbb{P}[T > t] = e^{-\lambda t}$.
2. $T$ besitzt die Eigenschaft der Gedächnislosigkeit $\forall t, s \geq 0 \quad \mathbb{P}[T > t + s \mid T > t] = [T > s]$.

Die Verteilungsfunktion sieht folgendermassen aus

$$F_\lambda(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}.$$

### Normalverteilung

Definition 3.29: Eine stetige Zufallsvariable $X$ heisst normal verteilt mit Parametern $m$ und $\sigma^2 > 0$ falls ihre Dichte gegeben ist durch

$$f_{m,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

Wir schreiben dann stets $X \sim \mathcal{N}(m, \sigma^2)$.
**Eigenschaften der Normalverteilung:**

1. Seien $X_1, \ldots, X_n$ unabhängige normalverteilte Zufallsvariablen mit Parametern $(m_1, \sigma_1^2), \ldots, (m_n, \sigma_n^2)$, dann ist $Z = m_0 + \lambda_1 X_1 + \ldots + \lambda_n X_n$ eine normalverteilte Zufallsvariable mit Parametern $m = m_0 + \lambda_1 m_1 + \cdots + \lambda_n m_n$ und $\sigma^2 = \lambda_1^2 \sigma_1^2 + \cdots + \lambda_n^2 \sigma_n^2$
2. Wir sprechen im Fall von $X \sim \mathcal{N}(0, 1)$, gerade von einer standardnormalverteilten Zufallsvariable. Man merke sich dann folgende Beziehung $Z = m + \sigma \cdot X$ wobei $X$ eine normalverteilte Zufallsvariable mit Parametern $m$ und $\sigma^2$ ist.
3. Falls $X$ normalverteilt mit Parametern $m$ und $\sigma^2$ ist, dann liegt die "meiste" Wahrscheinlichkeitsmasse der Z.V. im Intervall $[m - 3\sigma, m + 3\sigma]$. Präzise gilt gerade $\mathbb{P}[|X - m| \geq 3\sigma] \leq 0.0027$

# Der Erwartungswert

## Der allgemeine Erwartungswert

Definition 4.1: Sei $X : \Omega \to \mathbb{R}_+$eine Zufallsvariable mit nicht-negativen Werten. Dann heisst

$$\mathbb{E}[X] = \int_0^\infty (1 - F_X(x))\, dx$$

der Erwartungswert von $X$.
**Bemerkung 4.2:** Der Erwartungswert kann sowohl endliche also auch nicht endliche Werte annehmen.
Satz 4.3: Sei $X$ eine nicht-negative Zufallsvariable. Dann gilt $\mathbb{E}[X] \geq 0$. Gleichheit gilt genau dann wenn $X = 0$ fast sicher hält.
Definition 4.4: Sei $X$ eine Zufallsvariable. Falls $\mathbb{E}[|X|] < \infty$, dann heisst

$$X_+(\omega) = \left\{ \begin{array}{ll} X(\omega) & \text{falls } X(\omega) \geq 0, \\ 0 & \text{falls } X(\omega) < 0, \end{array} \right. \quad \text{und}$$

$$X_-(\omega) = \begin{cases} -X(\omega) & \text{falls } X(\omega) \leq 0, \\ 0 & \text{falls } X(\omega) > 0. \end{cases}$$

$\mathbb{E}[X] = \mathbb{E}[X_+] - \mathbb{E}[X_-]$. Erwartungswert von $X$.

## Erwartungswert einer diskreten Zufallsvariable

Satz 4.6: Sei $X : \Omega \to \mathbb{R}$ eine diskrete Zufallsvariable dessen Werte in $W$ (endlich oder abzählbar) fast sicher liegen. Sei $\phi : \mathbb{R} \to \mathbb{R}$ eine Abbildung. Dann gilt

$$\mathbb{E}[\phi(X)] = \sum_{x \in W} \phi(x) \cdot \mathbb{P}[X = x],$$

solange der Erwartungswert wohldefiniert ist.

### Bernoulli Zufallsvariable

Sei $X$ eine Bernoulli Zufallsvariable mit Parameter $p$. Dann gilt $\mathbb{E}[X] = p$.

### Binomial Zufallsvariable

Sei $X \sim \text{Bin}(n, p)$, dann gilt $\mathbb{E}[X] = np$

### Poisson Zufallsvariable

Sei $X$ Poisson-verteilt mit Parameter $\lambda > 0$, dann gilt $\mathbb{E}[X] = \lambda$.

### Indikator Zufallsvariable

Sei $A \in \mathcal{F}$ ein Ereignis. Sei $\mathbb{1}_A$ die Indikator Funktion auf $A$, dann gilt $\mathbb{E}[\mathbb{1}_A] = \mathbb{P}[A]$.

### geometrische Zufallsvariable

Sei $X \sim \text{Geom}(p)$, dann gilt $\mathbb{E}[X] = \frac{1}{p}$

## Erwartungswert stetiger Zufallsvariablen

Satz 4.8: Sei $X$ eine stetige Zufallsvariable mit Dichte $f$. Dann gilt

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

solange das Integral wohldefiniert ist.
Theorem 4.9: Sei $X$ eine stetige Zufallsvariable mit Dichte $f$. Sei $\phi : \mathbb{R} \to \mathbb{R}$ eine Abbildung, sodass $\phi(X)$ eine Zufallsvariable ist. Dann gilt

$$E[\phi(X)] = \int_{-\infty}^{\infty} \phi(x) f(x) dx$$

solange das Integral wohldefiniert ist.

### Exponential Zufallsvariable

Sei $X \sim \exp(\lambda)$. Dann gilt $\mathbb{E}[X] = \frac{1}{\lambda}$.

### gleichverteilte Zufallsvariable

Sei $X \sim \mathcal{U}([a, b])$. Dann gilt $\mathbb{E}[X] = \frac{a+b}{2}$.

### Normalverteilung

Sei $X \sim \mathcal{N}(\mu, \sigma^2)$. Dann gilt $\mathbb{E}[X] = \mu$.

## Rechnen mit Zufallsvariable

Theorem 4.10 (Linearität des Erwartungswert): Seien $X, Y : \Omega \to \mathbb{R}$ Zufallsvariablen, sei $\lambda \in \mathbb{R}$. Falls die Erwartungswerte wohldefiniert sind gilt

1. $\mathbb{E}[\lambda \cdot X] = \lambda \cdot \mathbb{E}[X]$
2. $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$.

**Bemerkung 4.11:** Die Zufallsvariablen $X$ und $Y$ müssen dabei nicht unabhängig sein.
**Bemerkung 4.12:** Unter Anwendung der Induktion für $n \geq 1$ ergibt sich direkt $\mathbb{E}[\lambda_1 X_1 + \lambda_2 X_2 + \cdots + \lambda_n X_n] = \lambda_1 \mathbb{E}[X_1] + \lambda_2 \mathbb{E}[X_2] + \cdots + \lambda_n \mathbb{E}[X_n]$ für jede $Z.V. X_1, X_2, \ldots, X_n : \Omega \to E$, und für jedes $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{R}$, unter der Annahme, dass die Erwartungswerte wohldefiniert sind.
Theorem 4.13: Seien $X, Y$ zwei Zufallsvariablen. Falls $X$ und $Y$ unabhängig sind, dann ist $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

## Extremwert Formel

Satz 4.14 (Tailsum-Formel für nichtnegative Zufallsvariablen): Sei $X$ eine Zufallsvariable, sodass $X \geq 0$ fast sicher gilt. Dann folgt $\mathbb{E}[X] = \int_0^\infty \mathbb{P}[X > x] dx$.
Satz 4.15 (Extremwert Formel): Sei $X$ eine diskrete Zufallsvariable mit Werten in $\mathbb{N} = \{0, 1, 2, \ldots\}$. Dann gilt folgende Identität $\mathbb{E}[X] = \sum_{n=1}^\infty \mathbb{P}[X \geq n]$.

## Charakterisierung der Eigenschaften von Z.V.

Satz 4.16. Sei $X$ eine Zufallsvariable: Sei $f : \mathbb{R} \to \mathbb{R}_+$eine Abbildung, sodass $\int_{-\infty}^{+\infty} f(x) dx = 1$. Dann sind folgende Aussagen äquivalent

1. $X$ ist stetig mit Dichte $f$,
2. Für jede Abbildung stückweise stetige, beschränkte Abbildung $\phi : \mathbb{R} \to \mathbb{R}$ gilt $\mathbb{E}[\phi(X)] = \int_{-\infty}^\infty \phi(x) f(x) dx$.

Theorem 4.17: Seien $X, Y$ zwei diskrete Zufallsvariablen. Die folgenden Aussagen sind äquivalent

1. $X, Y$ sind unabhängig,
2. Für jedes $\phi : \mathbb{R} \to \mathbb{R}, \psi : \mathbb{R} \to \mathbb{R}$ (messbar) beschränkt und stückweise stetig gilt $\mathbb{E}[\phi(X)\psi(Y)] = \mathbb{E}[\phi(X)]\mathbb{E}[\psi(Y)]$.

Theorem 4.18: Seien $X_1, \ldots, X_n n$ Zufallsvariablen. Die folgenden Aussagen sind äquivalent

1. $X_1, \ldots, X_n$ sind unabhängig,
2. Für jedes $\phi_1 : \mathbb{R} \to \mathbb{R}, \ldots, \phi_n : \mathbb{R} \to \mathbb{R}$ (messbar) beschränkt gilt $\mathbb{E}[\phi_1(X_1) \cdots \phi_n(X_n)] = \mathbb{E}[\phi_1(X_1)] \cdots \mathbb{E}[\phi_n(X_n)]$

## Ungleichungen

Satz 4.19: Seien $X, Y$ zwei Zufallsvariablen, sodass $X \leq Y$ f.s. gilt. Falls beide Erwartungswerte wohldefiniert sind, folgt dann $\mathbb{E}[X] \leq \mathbb{E}[Y]$ fast sicher.

## Markow Ungleichung

Theorem 4.20 (Markow-Ungleichung): Sei $X$ eine nicht-negative Zufallsvariable. Für jedes $a > 0$ gilt dann $\mathbb{P}[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$.

## Jensen Ungleichung

Theorem 4.21 (Jensen Ungleichung): Sei $X$ eine Zufallsvariable. Sei $\phi : \mathbb{R} \to \mathbb{R}$ eine konvexe Funktion. Falls $\mathbb{E}[\phi(X)]$ und $\mathbb{E}[X]$ wohldefiniert sind, gilt $\phi(\mathbb{E}[X]) \leq \mathbb{E}[\phi(X)]$.

## Cauchy-Schwarz Ungleichungen

Theorem 4.22: Seien $X, Y$ zwei Zufallsvariablen mit $\mathbb{E}[X^2], \mathbb{E}[Y^2] < \infty$. Dann gilt $\mathbb{E}[XY] \leq \sqrt{\mathbb{E}[X^2]} \cdot \sqrt{\mathbb{E}[Y^2]}$.

## Varianz

Definition 4.22: Sei $X$ eine Zufallsvariable, sodass $\mathbb{E}\left[X^2\right] < \infty$. Wir definieren die $Va$ rianz von $X$ durch
$\sigma_X^2 = \mathbb{E}\left[(X-m)^2\right], \quad$ wobei $m = \mathbb{E}[X]$ Die Wurzel aus $\sigma_X^2$ nennen wir gerade die Standardabweichung von $X$.
**Bemerkung 4.23:** Falls $\mathbb{E}\left[X^2\right] < \infty$, dann gilt gerade $\mathbb{E}[|X|] < \infty$ durch Gleichung. (4.7) und somit ist $m = \mathbb{E}[X]$ wohldefiniert.
Satz 4.24 (Grundlegende Eigenschaften der Varianz):

1. Sei $X$ eine Zufallsvariable mit $\mathbb{E}\left[X^2\right] < \infty$. Dann gilt $\sigma_X^2 = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2$
2. Sei $X$ eine Zufallsvariable mit $\mathbb{E}\left[X^2\right] < \infty$ und sei $\lambda \in \mathbb{R}$. Dann gilt $\sigma_{\lambda X}^2 = \lambda^2 \cdot \sigma_X^2$.
3. Seien $X_1, \ldots, X_n n$-viele paarweise unabhängigen $Z$ ufallsvariablen und $S = X_1 + \cdots + X_n$. Dann gilt $\sigma_S^2 = \sigma_{X_1}^2 + \cdots + \sigma_{X_n}^2$.

Theorem 4.24: Sei $X$ eine Zufallsvariable mit $\mathbb{E}\left[X^2\right] < \infty$. Dann gilt für alle $a \geq 0$ $\mathbb{P}[|X-m| \geq a] \leq \frac{\sigma_X^2}{a^2}, \quad$ where $m = \mathbb{E}[X]$.

## Beispiele Varianz von Zufallsvariablen

### Bernoulli Zufallsvariable

Sei $X$ eine Bernoulli Zufallsvariable mit Parameter $p$. Dann gilt $\mathrm{Var}[X] = p \cdot (1-p)$.

### Binomial Zufallsvariable

Sei $X \sim \mathrm{Bin}(n,p)$, dann gilt $\mathrm{Var}[X] = n \cdot p \cdot (1-p)$.

### Poisson Zufallsvariable

Sei $X$ Poisson-verteilt mit Parameter $\lambda > 0$, dann gilt $\mathrm{Var}[X] = \lambda$.

### Indikator Zufallsvariable

Sei $A \in \mathcal{F}$ ein Ereignis. Sei $\mathbb{1}_A$ die Indikator Funktion auf $A$, dann gilt $\mathrm{Var}[\mathbb{1}_A] = \mathbb{P}[A] - (\mathbb{P}[A])^2$.

### geometrische Zufallsvariable

Sei $X \sim \mathrm{Geom}(p)$, dann gilt $\mathrm{Var}[X] = \frac{1}{p^2} - \frac{1}{p}$

### Exponential Zufallsvariable

Sei $X \sim \exp(\lambda)$. Dann gilt $\mathrm{Var}[X] = \frac{1}{\lambda^2}$.

### gleichverteilte Zufallsvariable

Sei $X \sim \mathcal{U}([a,b])$. Dann gilt $\mathrm{Var}[X] = \frac{1}{12} \cdot (b-a)^2$.

### Normalverteilung

Sei $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$. Dann gilt $\mathrm{Var}[X] = \sigma^2$.

## Kovarianz

Definition 4.25: Seien $X, Y$ zwei Zufallsvariablen mit endlichen zweiten Momenten $\mathbb{E}\left[X^2\right] < \infty$ und $\mathbb{E}\left[Y^2\right] < \infty$. Wir definieren die Kovarianz zwischen $X$ und $Y$ durch $\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$.
**Bemerkung:** Die endlichen zweiten Momente von $X$ und $Y$ ermöglichen die Wohldefiniertheit der Kovarianz von $X$ und $Y$. Hierfür wende man die elementare Ungleichung $|XY| \leq \frac{1}{2}X^2 + \frac{1}{2}Y^2$ in Kombination mit Monotonie und Linearität des Erwartungswerts an, um folgende Ungleichung zu erhalten

$$\mathbb{E}[|XY|] \leq \frac{1}{2}\mathbb{E}\left[X^2\right] + \frac{1}{2}\mathbb{E}\left[Y^2\right] < \infty.$$

Analog zum Abschnitt 4, verschwindet die Kovarianz von unabhängigen Zufallsvariablen $X$ und $Y$, somit gilt

$$X, Y \text{ unabhängig } \implies \mathrm{Cov}(X,Y) = 0.$$

Die Umgekehrte Implikation ist falsch (Siehe Serie 6.6). Nichtsdestotrotz sehen wir in Abschnitt 6, dass wir eine Charakterisierung mittels beschränkten und stetigen Testfunktionen erhalten. Mittels Theorem 4.17, erhalten wir folgende Charakterisierung $X, Y$ Unabhängig $\iff \forall \phi, \psi$ stückweise stetig, beschränkt $\mathrm{Cov}(\phi(X), \psi(Y)) = 0$.

# Gemeinsame Verteilung

## Gemeinsame diskrete Verteilung

Definition 5.1: Seien $X_1, \ldots, X_n n$ diskrete Zufallsvariablen, sei $W_i \subset \mathbb{R}$ endlich oder abzählbar, wobei $X_i \in W_i$ fast sicher gilt. Die gemeinsame Verteilung von $(X_1, \ldots, X_n)$ ist eine Familie $p = \left(p\left(x_1, \ldots, x_n\right)\right)_{x_1 \in W_1, \ldots, x_n \in W_n}$, wobei jedes Mitglied definiert ist durch $p\left(x_1, \ldots, x_n\right) = \mathbb{P}\left[X_1 = x_1, \ldots, X_n = x_n\right]$.
Satz 5.2: Eine gemeinsame Verteilung von Zufallsvariablen $X_1, \ldots, X_n$ erfüllt $\sum_{x_1 \in W_1, \ldots, x_n \in W_n} p\left(x_1, \ldots, x_n\right) = 1$.
Satz 5.3: Sei $n \geq 1$ und seien $\phi : \mathbb{R}^n \to \mathbb{R}$ Abbildungen. Seien $X_1, \ldots, X_n n$ diskrete Zufallsvariablen in $(\Omega, \mathcal{F}, \mathbb{P})$, welche fast sicher Werten in endlichen oder abzählbaren Mengen $W_1, \ldots, W_n$ annehmen. Dann ist $Z = \phi\left(X_1, \ldots, X_n\right)$ eine diskrete Zufallsvariable, welche fast sicher Werte in $W = \phi\left(W_1 \times \cdots \times W_n\right)$ annimmt. Zudem ist die Verteilung von $Z$ gegeben durch

$$\forall z \in W \quad \mathbb{P}[Z = z] = \sum_{\substack{x_1 \in W_1, \ldots, x_n \in W_n \\ \phi(x_1, \ldots, x_n) = z}} \mathbb{P}\left[X_1 = x_1, \ldots X_n = x_n\right]$$

Satz 5.4 (Randverteilung): Seien $X_1, \ldots, X_n n$ diskrete Zufallsvariablen mit gemeinsamer Verteilung $p = \left(p\left(x_1, \ldots, x_n\right)\right)_{x_1 \in W_1, \ldots, x_n \in W_n}$. Für jedes $i$ gilt
$\forall z \in W_i \quad \mathbb{P}\left[X_i = z\right] = \sum_{x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n} p\left(x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_n\right).$
Satz 5.5 (Erwartungswert des Bildes): Seien $X_1, \ldots, X_n n$ diskrete Zufallsvariablen mit gemeinsamer Verteilung $p = \left(p\left(x_1, \ldots, x_n\right)\right)_{x_1 \in W_1, \ldots, x_n \in W_n}$. Sei $\phi : \mathbb{R}^n \to \mathbb{R}$, dann gilt

$$\mathbb{E}\left[\phi\left(X_1, \ldots, X_n\right)\right] = \sum_{x_1, \ldots, x_n} \phi\left(x_1, \ldots, x_n\right) p\left(x_1, \ldots, x_n\right),$$

solange die Summe wohldefiniert ist.
Satz 5.6: Seien $X_1, \ldots, X_n$ diskrete Zufallsvariablen mit gemeinsamer Verteilung $p = \left(p\left(x_1, \ldots, x_n\right)\right)_{x_1 \in W_1, \ldots, x_n \in W_n}$. Die folgenden Aussagen sind äquivalent

1. $X_1, \ldots, X_n$ sind unabhängig,
2. $p\left(x_1, \ldots, x_n\right) = \mathbb{P}\left[X_1 = x_1\right] \cdots \mathbb{P}\left[X_n = x_n\right]$ für jedes $x_1 \in W_1, \ldots, x_n \in W_n$.

## Stetige Gemeinsame Verteilung

Definition 5.7: Sei $n \geq 1$. Wir sagen, dass die Zufallsvariablen $X_1, \ldots, X_n : \Omega \to \mathbb{R}$ eine stetige gemeinsame Verteilung besitzen, falls eine Abbildung $f : \mathbb{R}^n \to \mathbb{R}_+$ existiert, sodass

$$\mathbb{P}\left[X_1 \leq a_1, \ldots, X_n \leq b\right] = \int_{-\infty}^{a_1} \cdots \int_{-\infty}^{a_n} f\left(x_1, \ldots, x_n\right) dx_n \ldots dx_1$$

für jedes $a_1, \ldots, a_n \in \mathbb{R}$ gilt. Obige Abbildung $f$ nennen wir gerade gemeinsame Dichte von $(X_1, \ldots, X_n)$.
Satz 5.9: Sei $f$ die gemeinsame Dichte der Zufallsvariablen $(X_1, \ldots, X_n)$. Dann gilt $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f\left(x_1, \ldots, x_n\right) dx_n \ldots dx_1 = 1$.
Satz 5.10 (Erwartungswert): Sei $\phi : \mathbb{R}^n \to \mathbb{R}$ eine Abbildung. Falls $X_1, \ldots, X_n$ eine gemeinsame Dichte $f$ besitzen, dann lässt sich der Erwartungswert der $Z$ ufallsvariablen $Z = \phi\left(X_1, \ldots, X_n\right)$ mittels $\mathbb{E}[\phi(X,Y)] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \phi\left(x_1, \ldots, x_n\right) \cdot f\left(x_1, \ldots, x_n\right) dx_1 \ldots dx_n$, berechnen (solange das Integral wohldefiniert ist).
Randverteilung: Falls $X, Y$ eine gemeinsame Dichte $f_{X,Y}$ besitzt, dann gilt

$$\begin{aligned} \mathbb{P}[X \leq a] &= \mathbb{P}[X \in [-\infty, a], Y \in [-\infty, \infty]] \\ &= \int_{-\infty}^{a} \left( \int_{-\infty}^{\infty} f(x,y) dy \right) dx. \end{aligned}$$

Somit ist $X$ stetig mit folgender Dichte

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y) dy.$$

Analog ist $Y$ stetig mit folgender Dichte

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

Theorem 5.11: Seien $X_1, \ldots, X_n$ Zufallsvariablen mit Dichten $f_1, \ldots f_n$. Dann ist folgende Aussagen äquivalent

1. $X_1, \ldots, X_n$ sind unabhängig,
2. $X_1, \ldots, X_n$ sind insgesamt stetig mit gemeinsamer Dichte $f\left(x_1, \ldots, x_n\right) = f_1\left(x_1\right) \ldots f_n\left(x_n\right)$.

# Grenzwertsätze

**Vorbemerkung:** In diesem Kapitel fixieren wir einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ und eine Folge von u.i.v.-Zufallsvariablen $X_1, X_2, \ldots$. Mit anderen Worten, wir erhalten Zufallsvariablen $X_i : \Omega \to \mathbb{R}$, so dass $\forall i_1 < \cdots < i_k \forall x_1, \ldots x_k \in \mathbb{R} \quad \mathbb{P}\left[X_{i_1} \leq x_1, \ldots, X_{i_k} \leq x_k\right] = F(x_1) \cdots F(x_k)$. Zudem ist $S_n = X_1 + X_2 + \cdots + X_n$.

## Gesetz der grossen Zahlen (GGZ)

Theorem 6.1: Sei $\mathbb{E}\left[|X_1|\right] < \infty$. Setzte $m = \mathbb{E}\left[X_1\right]$ dann gilt

$$\lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} = m \text{ fast sicher.}$$

## Konvergenz in Verteilung

Definition 6.3: Seien $(X_n)_{n\in\mathbb{N}}$ und $X$ Zufallsvariablen. Wir schreiben $X_n \underset{\approx}{\text{Approx}} X$ as $n \to \infty$ falls für jedes $x \in \mathbb{R}$ $\lim_{n\to\infty} \mathbb{P}\left[X_n \leq x\right] = \mathbb{P}[X \leq x]$

## Zentraler Grenzwertsatz

Theorem 6.4 (Zentraler Grenzwertsatz (ZGWS)): Nehme an, dass der Erwartungswert $\mathbb{E}\left[X_1^2\right]$ wohldefiniert und endlich ist. Setzte $m = \mathbb{E}\left[X_1\right]$ und $\sigma^2 = \operatorname{Var}(X_1)$, dann gilt folgender Grenzwert

$$\mathbb{P}\left[\frac{S_n - n\cdot m}{\sqrt{\sigma^2 n}} \leq a\right] \underset{n\to\infty}{\longrightarrow} \Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2} dx$$

für jedes $a \in \mathbb{R}$.
**Bemerkung:** Beachte gerade, dass $\Phi$ gerade die Verteilungsfunktion einer Zufallsvariable $Z \sim \mathcal{N}(0,1)$ ist. Der Satz besagt somit, dass für grosse $n \in \mathbb{N}$ die Zufallsvariable

$$Z_n = \frac{S_n - n\cdot m}{\sqrt{\sigma^2 n}}$$

in Verteilung $Z \sim \mathcal{N}(0,1)$ ähnelt.

# Schätzer

## Grundbegriffe

Definition 1.1: Ein Schätzer ist eine Zufallsvariable $T : \Omega \to \mathbb{R}$ der Form

$$T = t(X_1, \ldots, X_n),$$

wobei $t : \mathbb{R}^n \to \mathbb{R}$.

## Bias

Definition 1.2: Ein Schätzer $T$ heisst erwartungstreu für $\theta$, falls für alle $\theta \in \Theta$ gilt $\mathbb{E}_\theta[T] = \theta$.
Definition 1.3: Sei $\theta \in \Theta$, und $T$ ein Schätzer. Der Bias (oder erwartete Schätzfehler) von $T$ im Modell $\mathbb{P}_\theta$ ist definiert als

$$\mathbb{E}_\theta[T] - \theta$$

Der mittlere quadratische Schätzfehler ("mean squared error", MSE) von $T$ im Modell $\mathbb{P}_\theta$ ist definiert als

$$\operatorname{MSE}_\theta[T] := \mathbb{E}_\theta\left[(T-\theta)^2\right]$$

**Bemerkung:** Man kann den MSE zerlegen als

$$\operatorname{MSE}_\theta[T] = \mathbb{E}_\theta\left[(T-\theta)^2\right] = \operatorname{Var}_\theta[T] + \left(\mathbb{E}_\theta[T] - \theta\right)^2,$$

also in die Summe aus der Varianz des Schätzers $T$ und dem Quadrat des Bias. Für erwartungstreue Schätzer sind Varianz und MSE dasselbe.

## Die Maximum-Likelihood-Methode (ML-Methode)

Definition 1.4: Die Likelihood-Funktion ist

$$L(x_1, \ldots, x_n; \theta) := \begin{cases} p_{\vec{X}}(x_1, \ldots, x_n; \theta) & \text{im diskreten Fall,} \\ f_{\vec{X}}(x_1, \ldots, x_n; \theta) & \text{im stetigen Fall.} \end{cases}$$

mit $p_{\vec{X}}(x_1, \ldots, x_n; \theta) = \prod_{i=1}^n p_X(x_i; \theta)$ und $f_{\vec{X}}(x_1, \ldots, x_n; \theta) = \prod_{i=1}^n f_X(x_i; \theta)$. Die Funktion $\log L(x_1, \ldots, x_n; \theta)$ heisst **log-Likelihood-Funktion**.
Definition 1.5: Für jedes $x_1, \ldots, x_n$, sei $t_{ML}(x_1, \ldots, x_n) \in \mathbb{R}$ der Wert, der $\theta \mapsto L(x_1, \ldots, x_n; \theta)$ als Funktion von $\theta$ maximiert. D.h.,

$$L(x_1, \ldots, x_n; t_{ML}(x_1, \ldots, x_n)) = \max_{\theta\in\Theta} L(x_1, \ldots, x_n; \theta)$$

Ein Maximum-Likelihood-Schätzer (ML-Schätzer) $T_{\mathrm{ML}}$ für $\theta$ wird definiert durch $T_{\mathrm{ML}} = t_{\mathrm{ML}}(X_1, \ldots, X_n)$.
**Bemerkung:** In den Rechnungen arbeitet man oft mit $L(x_1, \ldots, x_n; \theta)$, insbesondere beim Maximieren über $\theta$. Das optimale $\theta^*$ ist dann eine Funktion $t_{\mathrm{ML}}(x_1, \ldots, x_n)$ von $x_1, \ldots, x_n$. Damit der resultierende Schätzer $T_{\mathrm{ML}}$ von der Stichprobe $X_1, \ldots, X_n$ abhängt, muss dann aber $x_1, \ldots, x_n$ durch $X_1, \ldots, X_n$ ersetzt werden, d.h. der Maximum-Likelihood-Schätzer ist $T_{\mathrm{ML}} = t_{\mathrm{ML}}(X_1, \ldots, X_n)$.

# Konfidenzintervalle

## Definition

Definition 2.1: Sei $\alpha \in [0,1]$. Ein Konfidenzintervall für $\theta$ mit Niveau $1-\alpha$ ist ein Zufallsintervall $I = [A, B]$, sodass gilt

$$\forall \theta \in \Theta \quad \mathbb{P}_\theta[A \leq \theta \leq B] \geq 1 - \alpha$$

wobei $A, B$ Zufallsvariablen der Form $A = a(X_1, \ldots, X_n), B = b(X_1, \ldots, X_n)$ mittels $a, b : \mathbb{R}^n \to \mathbb{R}$ sind.
**Bemerkung 2.2:** In Eq. (2.1), ist der Paramter $\theta$ deterministisch und nicht zufällig. Die stochastischen Elemente sind gerade die Schranken $A = a(X_1, \ldots, X_n)$ und $B = b(X_1, \ldots, X_n)$.

## Verteilungsaussagen

Definition 2.3: Eine stetige Zufallsvariable $X$ heisst $\chi^2$-Verteilt mit $m$ Freiheitsgraden falls ihre Dichte gegeben ist durch

$$f_X(y) = \frac{1}{2^{\frac{m}{2}} \Gamma\left(\frac{m}{2}\right)} y^{\frac{m}{2}-1} e^{-\frac{1}{2}y} \text{ für } y \geq 0.$$

Wir schreiben dann $X \sim \chi^2_m$.
Erwartungswert von $X \sim \chi^2_m$: Sei $X \sim \chi^2_m$, dann ist $\mathbb{E}[X] = n$.
Varianz von $X \sim \chi^2_m$: Sei $X \sim \chi^2_m$, dann ist $\operatorname{Var}[X] = 2n$.

Satz aus Aufgabe: Seien die Zufallsvariablen $X_1, \ldots, X_n$ unabhängig und je $\mathcal{N}\left(\mu, \sigma^2\right)$-verteilt unter $\mathbb{P}_\theta$, wobei $\theta = \left(\mu, \sigma^2\right) \in \mathbb{R} \times (0, \infty)$ ein 2-dimensionaler unbekannter Parameter ist. Dann gilt $\frac{1}{\sigma^2} \sum_{i=1}^n \left(X_i - \bar{X}_n\right)^2 \sim \chi^2_{n-1}$-verteilt ist unter $\mathbb{P}_\theta$.
**Bemerkung:** Die $\chi^2$ Verteilung mit $m$ Freiheitsgraden ist der Spezialfall einer $Ga(\alpha, \lambda)$ Verteilung mit $\alpha = \frac{m}{2}$ und $\lambda = \frac{1}{2}$. Für $m = 2$ ergibt das eine Exponentialverteilung mit Parameter $\frac{1}{2}$.
Satz 2.4: Sind die Zufallsvariablen $X_1, \ldots, X_m$ u.i.v. $\sim \mathcal{N}(0,1)$, so ist die Summe $Y := \sum_{i=1}^m X_i^2 \sim \chi^2_m$.
Definition 2.5: Eine stetige Zufallsvariable $X$ heisst t-verteilt mit m Freiheitsgraden falls ihre Dichte gegeben ist durch

$$f_X(x) = \frac{\Gamma\left(\frac{m+1}{2}\right)}{\sqrt{m\pi}\,\Gamma\left(\frac{m}{2}\right)} \left(1 + \frac{x^2}{m}\right)^{-\frac{m+1}{2}} \quad \text{für } x \in \mathbb{R}.$$

Wir schreiben dann $X \sim t_m$.
Satz 2.6: Sind $X$ und $Y$ unabhängig mit $X \sim \mathcal{N}(0,1)$ und $Y \sim \chi^2_m$, so ist der Quotient $Z := \frac{X}{\sqrt{\frac{1}{m}Y}}$ $t$-verteilt mit $m$ Freiheitsgraden.
**Schätzer für $\sigma$ und $\mu$:** Wir betrachten folgende Fälle:
Wir intressieren uns für $\mu$

1. $\sigma^2$ bekannt
2. $\sigma^2$ unbekannt

Wir intressieren uns für $\sigma^2$

3. $\mu$ bekannt
4. $\mu$ unbekannt

Für $\mathbb{P}[Z \in [q_-, q_+]] = 1 - \alpha$ gilt:

1. $Z = \sqrt{n}\frac{m-\mu}{\sigma} \sim \mathcal{N}(0,1)$ wobei $m = \overline{X_n} = \frac{1}{n}\sum_{i=1}^n X_i : Z \in \left[-q_{1-\frac{\alpha}{2}}, q_{1-\frac{\alpha}{2}}\right] \iff m - \frac{q_{1-\frac{\alpha}{2}}}{\sqrt{n}}\sigma \leq \mu \leq m + \frac{q_{1-\frac{\alpha}{2}}}{\sqrt{n}}\sigma$
2. $Z = \sqrt{n}\frac{m-\mu}{S} \sim t_{n-1}$ wobei $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X_n})^2 : Z \in \left[-q_{1-\frac{\alpha}{2}}, q_{1-\frac{\alpha}{2}}\right] \iff m - \frac{q_{1-\frac{\alpha}{2}}}{\sqrt{n}}|S| \leq \mu \leq m + \frac{q_{1-\frac{\alpha}{2}}}{\sqrt{n}}|S|$
3. $Z = \frac{n}{\sigma^2}\tilde{S}^2 \sim \chi^2_n$ wobei $\tilde{S}^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \mu)^2 : Z \in \left[q_{\frac{\alpha}{2}}, q_{1-\frac{\alpha}{2}}\right] \iff \frac{n\tilde{S}^2}{q_{1-\frac{\alpha}{2}}} \leq \sigma^2 \leq \frac{n\tilde{S}^2}{q_{\frac{\alpha}{2}}}$
4. $Z = \frac{n-1}{\sigma^2}S^2 \sim \chi^2_{n-1}$ wobei $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X_n})^2 : Z \in \left[q_{\frac{\alpha}{2}}, q_{1-\frac{\alpha}{2}}\right] \iff \frac{(n-1)S^2}{q_{1-\frac{\alpha}{2}}} \leq \sigma^2 \leq \frac{(n-1)S^2}{q_{\frac{\alpha}{2}}}$

Satz 2.7: Seien $X_1, \ldots, X_n$ i.i.d. $\sim \mathcal{N}\left(\mu, \sigma^2\right)$. Dann $\bar{X}_n$ und $S^2$ sind unabhängig.

## Approximative Konfidenzintervalle

Einen allgemeinen **approximativen Zugang** liefert der zentrale Grenzwertsatz. Oft ist ein Schätzer $T$ eine Funktion einer Summe $\sum_{i=1}^n Y_i$, wobei die $Y_i$ im Modell $P_\theta$ i.i.d. sind; das einfachste Beispiel ist $T = \bar{Y}_n = \frac{1}{n}\sum_{i=1}^n Y_i$. Nach dem zentralen Grenzwertsatz ist dann für grosse $n$

$$\sum_{i=1}^n Y_i \quad \text{approximativ normalverteilt unter } \mathbb{P}_\theta$$

mit Parametern $\mu = n\mathbb{E}_\theta[Y_i]$ und $\sigma^2 = n\operatorname{Var}_\theta[Y_i]$. Das kann man benutzen, um für die Verteilung von $T$ approximative Aussagen zu bekommen und damit gewisse Fragen zumindest approximativ zu beantworten. Wir erhalten also

$S_n^* := \frac{S_n - \mu}{\sqrt{n\sigma^2}} \overset{\text{Approx}}{\approx} \mathcal{N}(0,1)$ unter $\mathbb{P}_\theta$.

### Beispiel approximatives Intervall

Sei $S_n \sim \operatorname{Bin}(n,\theta)$. Dann folgt $S_n^* = \frac{S_n - n\theta}{\sqrt{n\theta(1-\theta)}}$. Also gilt $\mathbb{P}_\theta\left[\left|\frac{S_n - n\theta}{\sqrt{n\theta(1-\theta)}}\right| \leq z_{1-\frac{\alpha}{2}}\right] = \mathbb{P}_\theta\left[|S_n^*| \leq z_{1-\frac{\alpha}{2}}\right] \approx 1-\alpha$. Wir gehen davon aus, dass $\theta(1-\theta) \approx \frac{1}{4}$ ist und setzen das ein. Dann wollen wir also

$$|S_n - n\theta| \leq z_{1-\frac{\alpha}{2}}\sqrt{\frac{n}{4}},$$

und das approximative Konfidenzintervall für $\theta$ ergibt sich als

$$\left[\bar{S}_n - \frac{z_{1-\frac{\alpha}{2}}}{2\sqrt{n}}, \bar{S}_n + \frac{z_{1-\frac{\alpha}{2}}}{2\sqrt{n}}\right].$$

# Tests

## Null- und Alternativhypothese

1. Nullhypothese $H_0 : \theta \in \Theta_0$
2. Alternativhypothese $H_A : \theta \in \Theta_A$

mit $\Theta_0 \subset \Theta$ und $\Theta_A \subset \Theta$, wobei $\Theta_0 \cap \Theta_A = \emptyset$. Ist keine explizite Alternative spezifiziert, so hat man $\Theta_A = \Theta_0^c = \Theta \backslash \Theta_0$. Null- und/oder Alternativhypothese heissen **einfach**, falls $\Theta_0$ bzw. $\Theta_A$ aus einem einzelnen Wert, $\theta_0$ bzw. $\theta_A$, bestehen, also z.B. $\Theta_0 = \{\theta_0\}$ ist; sonst heissen sie **zusammengesetzt**.

## Test und Entscheidung

Definition 3.1: Ein Test ist ein Paar $(T,K)$, wobei

1. $T$ eine Zufallsvariable der Form $T = t(X_1,\ldots,X_n)$ ist, und
2. $K \subseteq \mathbb{R}$ eine (deterministische) Teilmenge von $\mathbb{R}$ ist.

Die Zufallsvariable $T = t(X_1,\ldots,X_n)$ heisst dann Teststatistik, und $K$ heisst kritischen Bereich oder Verwerfungsbereich.
**Entscheidungsregel:**

1. die Hypothese $H_0$ wird verworfen, falls $T(\omega) \in K$,
2. die Hypothese $H_0$ wird nicht verworfen bzw. angenommen, falls $T(\omega) \notin K$.

**Fehler:**

1. Bei einem Fehler 1. Art wird die Nullhypothese zu Unrecht verworfen, d.h. obwohl sie richtig ist. Das passiert für $\theta \in \Theta_0$ und $T \in K$; deshalb heisst $\mathbb{P}_\theta[T \in K]$ für $\theta \in \Theta_0$ die Wahrscheinlichkeit für einen Fehler 1. Art.
2. Bei einem Fehler 2. Art wird die Nullhypothese zu Unrecht nicht verworfen, d.h. man akzeptiert die Nullhypothese (verwirft sie nicht), obwohl sie falsch ist. Das passiert für $\theta \in \Theta_A$ und $T \notin K$, und deshalb heisst $\mathbb{P}_\theta[T \notin K] = 1 - \mathbb{P}_\theta[T \in K]$ für $\theta \in \Theta_A$ die Wahrscheinlichkeit für einen Fehler 2. Art.

## Signifikanzniveau und Macht

Definition 3.2: Sei $\alpha \in (0,1)$. Ein Test $(T,K)$ besitzt Signifikanzniveau $\alpha$, falls $\forall \theta \in \Theta_0 \quad \mathbb{P}_\theta[T \in K] \leq \alpha$.
Definition 3.3: Die Macht eines Tests $(T,K)$ wird definiert als folgende Funktion $\beta : \Theta_A \to [0,1], \quad \theta \mapsto \beta(\theta) := \mathbb{P}_\theta[T \in K]$.

## Konstruktion von Tests

Definition 3.5: Für jedes $x_1,\ldots,x_n$, definieren wir den Likelihood-Quotienten durch

$$R(x_1,\ldots,x_n) := \frac{L(x_1,\ldots,x_n;\theta_A)}{L(x_1,\ldots,x_n;\theta_0)}.$$

Als Konvention setzten wir $R(x_1,\ldots,x_n) = +\infty$, falls $L(x_1,\ldots,x_n;\theta_0) = 0$. Die Idee dieses Tests ist, dass man einen einfacheren Test finden kann, der sich änlich verhält wie der Likelihood-Quotienten und den man stattdessen nehmen kann. Ein Beispiel wäre $T := \sum_{i=1}^n X_i = S_n$ für $R(x_1,\ldots,x_n;\theta_0,\theta_A)$ wobei $X_1,\ldots,X_n$ i.i.d. $\sim \operatorname{Be}(\theta)$
Theorem 3.7 (Neyman-Pearson-Lemma): Sei $c \geq 0$. Sei $(T,K)$ ein LikelihoodQuotienten-Test mit Parameter c und Signifikanzniveau $\alpha^* := \mathbb{P}_{\theta_0}[T > c]$. Ist $(T',K')$ ein anderer Test mit Signifikanzniveau $\alpha \leq \alpha^*$, so gilt

$$\mathbb{P}_{\theta_A}[T' \in K'] \leq \mathbb{P}_{\theta_A}[T \in K]$$

**Bemerkung:** Falls $R \gg 1$, ist $\Theta_A$ wahrscheinlicher und falls $R \ll 1$, ist $\Theta_0$ wahrscheinlicher.
Definition 3.6: Sei $c \geq 0$. Der Likelihood-Quotienten-Test mit Parameter $c$ ist ein Test $(T,K)$, wobei Teststatistik und Verwerfungsbereich gegeben sind durch

$$T = R(X_1,\ldots,X_n) \quad \text{und} \quad K = (c,\infty]$$

.

### z-Test

Normalverteilung, Test für Erwartungswert bei bekannter Varianz: Dieser Test ist unter dem Namen $z$-Test bekannt. Hier sind $X_1,\ldots,X_n$ i.i.d. $\sim \mathcal{N}(\theta,\sigma^2)$ unter $\mathbb{P}_\theta$ mit bekannter Varianz $\sigma^2$, und wir wollen die Hypothese $H_0 : \theta = \theta_0$ testen. Mögliche Alternativen $H_A$ sind $\theta > \theta_0$ oder $\theta < \theta_0$ (einseitig), oder $\theta \neq \theta_0$ (zweiseitig). Die Teststatistik hier ist in jedem Fall

$$T := \frac{\bar{X}_n - \theta_0}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1) \text{ unter } \mathbb{P}_{\theta_0}.$$

Wir können die Intervalle folgender massen wählen:

1. $\alpha = \mathbb{P}_{\theta_0}[T \in K_>] = \mathbb{P}_{\theta_0}[T > c_>] = 1 - \mathbb{P}_{\theta_0}[T \leq c_>] = 1 - \Phi(c_>)$
2. $\alpha = \mathbb{P}_{\theta_0}[T < c_<] = \mathbb{P}_{\theta_0}[T > -c_<] = 1 - \mathbb{P}_{\theta_0}[T \leq -c_<] = \Phi(c_<)$
3. $\alpha = \mathbb{P}_{\theta_0}[T \in K_{\neq}] = \mathbb{P}_{\theta_0}[T < -c_{\neq}] + \mathbb{P}_{\theta_0}[T > c_{\neq}] = \Phi(-c_{\neq}) + 1 - \Phi(c_{\neq}) = 2(1 - \Phi(c_{\neq}))$

## Der p-Wert

Definition 3.9 (Geordnete Testsammlung): Sei $T$ eine Teststatistik. Eine Familie von Tests $(T,(K_t)_{t\geq 0})$ heisst geordert bzgl. $\boldsymbol{T}$ falls $K_t \subset \mathbb{R}$ und

$$s \leq t \Longrightarrow K_s \supset K_t$$

gilt.
Definition 3.10: Sei $H_0 : \theta = \theta_0$ eine einfache Nullhypothese. Sei $(T,K_t)_{t\geq 0}$ eine geordnete Familie von Test. Der $\boldsymbol{p}$ - Wert ist definiert als Zufallsvariable

$$\text{p-Wert } = G(T),$$

wobei $G : \mathbb{R}_+ \to [0,1]$ mittels $G(t) = \mathbb{P}_{\theta_0}[T \in K_t]$ definiert ist.
**Anmerkungen:**

1. Der P-Wert ist als Funktion einer Teststatistik $T$ selbst eine Zufallsvariable.
2. Der P-Wert hängt direkt von den anfänglichen Beobachtungen $X_1,\ldots,X_n$ ab. Somit wird das Wiederholen des Test auch einen neuen (zufälligen) P-Wert generieren.
3. Der p-Wert liegt stets in $[0,1]$. Sei $T$ stetig ist und $K_t = (t,\infty)$, dann kann gezeigt werden, dass der P-Wert unter $\mathbb{P}_{\theta_0}$ auf $[0,1]$ gleichverteilt ist.
4. Der P-Wert liefert uns die Information, welche Tests in unserer Familie $(T,K_t)$, $t \geq 0$ die Nullhypothese $H_0$ ablehnen würden.
5. Für einen P-Wert mit Wert $p$ gilt, dass alle Tests mit Signifikanzniveau $\alpha > p$ die Nullhypothese $H_0$ verwerfen würden und alle Tests mit Signifikanzniveau $\alpha \leq p$ die Nullhypothese $H_0$ nicht verwerfen würden.
6. Der P-Wert ist nur von der Nullhypothese abhängig. Die Alternativhypothese spielt keine Rolle in der Definition des P-Werts.

**Intuition:** p-Wert ist klein $\Longrightarrow H_0$ wird wahrscheinlich verworfen.

### Beispiel eines realisierten approximativen p-Wert

p-Wert $(\omega) =$

$$\mathbb{P}_{\theta_0}[|T| > t_0]\Big|_{t_0=T(\omega)} \approx 2\mathbb{P}_{\theta_0}[T > t_0]\Big|_{t_0=T(\omega)} \approx 2(1-\Phi(t_0))\Big|_{t_0=T(\omega)}$$

für $T \overset{\text{Approx}}{\approx} \mathcal{N}(0,1)$ unter $\mathbb{P}_{\Theta_0}$ und $K := (-\infty,-c) \cup (+c,+\infty)$.

# Zusätzliches

### t-Test

Normalverteilung, Test für Erwartungswert bei unbekannter Varianz. Hier sind $X_1,\ldots,X_n$ i.i.d. $\sim \mathcal{N}(\mu,\sigma^2)$ unter $P_{\vec{\theta}}$, wobei $\vec{\theta} = (\mu,\sigma^2)$. Wir wollen die Hypothese $\mu = \mu_0$ testen. Die

Teststatistik ist: $T := \frac{\bar{X}_n - \mu_0}{S/\sqrt{n}} \sim t_{n-1}$ unter $P_{\theta_0}$
$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$ Und die Verwerfungsbereiche:

$$
\begin{array}{lll}
c_< = t_{n-1,\alpha} & (-\infty, c_<) & \mu < \mu_0 \\
c_> = t_{n-1,1-\alpha} & (c_>, \infty) & \mu > \mu_0 \\
c_{\neq} = t_{n-1,1-\frac{\alpha}{2}} & (-\infty, c_<) \cup (c_>, \infty) & \mu \neq \mu_0
\end{array}
$$

Wobei gilt $t_{m,\alpha} = -t_{m,1-\alpha}$.

## Gepaarter Zweiproben-Test

Hier sind $X_1, \ldots, X_n$ i.i.d. $\sim \mathcal{N}(\mu_X, \sigma^2)$ und $Y_1, \ldots, Y_n$ i.i.d. $\sim \mathcal{N}(\mu_Y, \sigma^2)$ unter $P_\theta$. Insbesondere ist $m = n$ und die Varianz beider Stichproben dieselbe. Differenzen $Z_i := X_I - Y_i$ sind unter $P_\theta$ i.i.d. $\mathcal{N}(\mu_X - \mu_Y, 2\sigma^2)$. Dann analog $z$ und $t$-Test. (Setzt natürliche Paarung von Daten voraus!)

## Ungepaarter Zweiproben-Test

Hier sind unter $P_\theta$ die Zufallsvariablen $X_1, \ldots, X_n$ i.i.d. $\sim \mathcal{N}(\mu_X, \sigma^2)$ und $Y_1, \ldots, Y_m$ i.i.d. $\sim \mathcal{N}(\mu_Y, \sigma^2)$, wobei die Varianz in beiden Fällen dieselbe ist.

1. Bei bekannter Varianz:

$$H_0 : \mu_X - \mu_Y = \mu_0 \quad (z.B. \mu_0 = 0)$$

$$T = \frac{\bar{X}_n - \bar{Y}_m - (\mu_X - \mu_Y)}{\sigma\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim \mathcal{N}(0,1)$$

   Die kritischen Werte für den Verwerfungsbereich sind wie oben geeignete Quantile der $\mathcal{N}(0,1)$ Verteilung, je nach Alternative. Das ist der ungepaarte Zweistichproben-$z$-Test.

2. Bei unbekannter Varianz:

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

$$S_Y^2 := \frac{1}{m-1}\sum_{j=1}^{m}(Y_j - \bar{Y}_m)^2$$

$$S^2 := \frac{1}{m+n-2}\left((n-1)\cdot S_X^2 + (m-1)\cdot S_Y^2\right)$$

$$T = \frac{\bar{X}_n - \bar{Y}_m - (\mu_X - \mu_Y)}{S\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{n+m-2}$$

   unter jedem $P_\theta$. Dieser Test heisst ungepaarter Zweistichproben- $t$-Test.

## Integral

Prob. 3.36: Let be $A = \{(x,y) | a \leq x \leq b, \rho_1(x) \leq y \leq \rho_2(x)\}$. Let $f : A \to \mathbb{R}$ be continuous. Then $f$ is integrable on $A$ and the following is true

$$\int_A f(x,y)d(x,y) = \int_a^b dx \int_{\rho_1(x)}^{\rho_2(x)} f(x,y)dy.$$

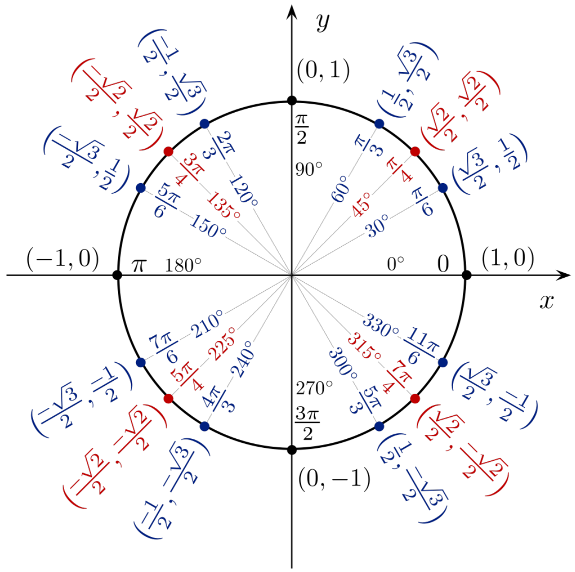
Satz 5.30 (Partielle Integration): Seien $a < b$ relle Zahlen und $f, g : [a,b] \longrightarrow \mathbb{R}$ stetig differenzierbar. Dann gilt:
$\int_a^b f(x)g'(x)dx = f(b)g(b) - f(a)g(a) - \int_a^b f'(x)g(x)dx$.

Satz 5.31 (Substitution): Sei $a < b, \phi : [a,b] \longrightarrow \mathbb{R}$ stetig differenzierbar, $I \subseteq \mathbb{R}$ ein Intervall mit $\phi([a,b]) \subseteq I$ und $f : I \longrightarrow \mathbb{R}$ eine stetige Funktion. Dann gilt:
$\int_{\phi(a)}^{\phi(b)} f(x)dx = \int_a^b f(\phi(t))\phi'(t)dt$.

## Allgemeines

**Mitternachtsformel:** $ax^2 + bx + c = 0 \Longleftrightarrow x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

**pq-Formel:** $x^2 + px + q = 0 \Longleftrightarrow x_{1,2} = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}$




## Reihen

**Geometrische Reihe:** Sei $q \in \mathbb{C}$ mit $|q| < 1$. Dann konvergiert $\sum_{k=0}^{\infty} q^k$ und dessen Wert ist: $\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$.
**allgemeine Formel für geometrische Reihe:**
$\sum_{i=0}^{k} p^i = \frac{p^{k+1}}{p-1} - \frac{1}{p-1}$