

Stochastik I - Formelsammlung

von Julian Merkert, Wintersemester 2005/06, Prof. Bäuerle

Ereignisse und Wahrscheinlichkeiten

Ereignis

- Ergebnisraum: Ω
- Ereignis: $A \subset \Omega$
- Elementarereignis: $\{\omega\}, \omega \in \Omega$
- $A \cap B := AB := \{\omega \in \Omega | \omega \in A \text{ und } \omega \in B\}$
- $A \cup B := \{\omega \in \Omega | \omega \in A \text{ oder } \omega \in B\}$ (Für disjunkte Mengen: $A + B$)
- $A \setminus B = \{\omega \in \Omega | \omega \in A, \omega \notin B\}$
- $A^C = \Omega \setminus A$

Kartesisches Produkt: $\Omega_1 \times \Omega_2 \times \dots \times \Omega_n = \{(a_1, \dots, a_n) | a_i \in \Omega_i, i = 1 \dots n\}$

Potenzmenge: $\mathcal{P}(\Omega)$ (Menge aller Teilmengen von Ω)

DeMorgansche Regeln:

- $(A \cup B)^C = A^C \cap B^C$
- $(A \cap B)^C = A^C \cup B^C$

σ -Algebra über Ω : $\mathcal{A} \subset \mathcal{P}(\Omega)$ mit...

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

Messraum: (Ω, \mathcal{A})

Folge von Ereignissen: $\{A_n\}$

- Limes superior: $\overline{\lim}_{n \rightarrow \infty} A_n = \limsup A_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n$ („unendlich viele der A_n s treten ein“)
- Limes inferior: $\underline{\lim}_{n \rightarrow \infty} A_n = \liminf A_n = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n$ („alle bis auf endlich viele der A_n s treten ein“)
- Falls $\{A_n\}$ wachsend ($A_1 \subset A_2 \subset \dots = A_n \uparrow$) $\Rightarrow \overline{\lim}_{n \rightarrow \infty} A_n = \underline{\lim}_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} A_n$
- Falls $\{A_n\}$ fallend ($A_1 \supset A_2 \supset \dots = A_n \downarrow$) $\Rightarrow \overline{\lim}_{n \rightarrow \infty} A_n = \underline{\lim}_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} A_n$

Wahrscheinlichkeitsmaß auf \mathcal{A} : $P : \mathcal{A} \rightarrow [0, 1]$ im Messraum (Ω, \mathcal{A}) mit...

1. Normiertheit: $P(\Omega) = 1$
2. σ -Additivität: $P(\sum_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$ für alle paarweise disjunkten $A_1, A_2, \dots \in \mathcal{A}$

Wahrscheinlichkeitsraum: (Ω, \mathcal{A}, P)

- $P(A^C) = 1 - P(A)$
- $P(\emptyset) = 0$
- $A \subset B \Rightarrow P(A) \leq P(B)$
- $A \subset B \Rightarrow P(B \setminus A) = P(B) - P(A)$
- $P(B \setminus A) = P(B \setminus (B \cap A)) = P(B) - P(B \cap A)$

- Endliche Additivität: $P(\sum_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$ für alle paarweise disjunkten A_1, \dots, A_n
- Boolesche Ungleichung: $P(\bigcup_{i=1}^n A_i) \leq \sum_{i=1}^n P(A_i)$
- Siebformel: $P(\bigcup_{k=2}^n A_k) = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k})$
 - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
 - $P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3)$

Laplace'scher Wahrscheinlichkeitsraum: (Ω, \mathcal{A}, P) mit $\mathcal{A} = \mathcal{P}(\Omega)$ und $P(A) = \frac{|A|}{|\Omega|} \forall A \in \mathcal{A}$

- jedes Elementarereignis hat die gleiche Wahrscheinlichkeit $P(\omega) = \frac{1}{|\Omega|}$

Permutationen

- Die Anzahl der Permutationen von n verschiedenen Objekten ist $n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n$
- Die Anzahl der Permutationen von n Objekten mit jeweils n_1, n_2, \dots, n_k gleichen Elementen ist $\frac{n!}{n_1! \cdot n_2! \cdot \dots \cdot n_k!}$

Kombinatorik

# der Möglichkeiten bei Ziehung vom Umfang k aus $\{1 \dots n\}$	mit Zurücklegen	ohne Zurücklegen
mit Reihenfolge	n^k	$\frac{n!}{(n-k)!}$
ohne Reihenfolge	$\binom{n+k-1}{k}$	$\binom{n}{k} = \frac{n!}{k!(n-k)!}$

- $\binom{n+k-1}{k}$ kann auch als die Anzahl der Möglichkeiten, k (nicht unterscheidbare) Objekte auf n (unterscheidbare) Fächer aufzuteilen, angesehen werden (mit Mehrfachbelegungen).

Bedingte Wahrscheinlichkeit: $P(A|B) = \frac{P(AB)}{P(B)}$

Multiplikationssatz: $P(A_1 \cap \dots \cap A_n) = \prod_{k=1}^n P(A_k | A_1 \cap \dots \cap A_{k-1})$ mit $A_0 := \Omega$

- $P(AB) = P(A|B) \cdot P(B)$

Ereignispartition von Ω : $B_1, B_2, \dots \in \mathcal{A}$ mit...

1. $B_i \cap B_j = \emptyset$ für $i \neq j$
2. $\sum_{i=1}^{\infty} B_i = \Omega$
3. $P(B_i) > 0 \forall i \in \mathbb{N}$

Satz von der totalen Wahrscheinlichkeit: $\forall A \in \mathcal{A} : P(A) = \sum_{j=1}^{\infty} P(B_j) \cdot P(A|B_j)$

Formel von Bayes: $P(B_k|A) = \frac{P(B_k) \cdot P(A|B_k)}{\sum_{j=1}^{\infty} P(B_j) \cdot P(A|B_j)}$

Unabhängigkeit zweier Ereignisse: $P(AB) = P(A) \cdot P(B)$

Unabhängigkeit der Ereignisse $A_1, \dots, A_n \in \mathcal{A}$: für alle $k = 1 \dots n$ und für alle k -Tupel $1 \leq i_1 < i_2 < \dots < i_k \leq n$:

$$P\left(\bigcap_{j=1}^k A_{i_j}\right) = \prod_{j=1}^k P(A_{i_j})$$

Zufallsvariablen und Verteilungen

Von \mathcal{E} erzeugte σ -Algebra: $\sigma(\mathcal{E}) := \bigcap_{\mathcal{A} \supset \mathcal{E}, \mathcal{A} \text{ ist } \sigma\text{-Algebra}} \mathcal{A}$

- $\mathcal{E} \subset \mathcal{P}(\Omega)$ heißt Erzeugendensystem

Borelsche σ -Algebra: $\mathcal{B} := \sigma(\mathcal{E})$ mit $\mathcal{E} = \{(a, b], -\infty < a < b < \infty\}$

Dynkin-System: $\mathcal{D} \subset \mathcal{P}(\Omega)$ mit..

1. $\Omega \in \mathcal{D}$
2. $A \in \mathcal{D} \Rightarrow A^C \in \mathcal{D}$
3. $A_1, A_2, \dots \in \mathcal{D}, A_i \cap A_j = \emptyset$ für $i \neq j \Rightarrow \sum_{i=1}^{\infty} A_i \in \mathcal{D}$

Durchschnittsstabiles (\cap -stabiles) Mengensystem \mathcal{E} : $A, B \in \mathcal{E} \Rightarrow A \cap B \in \mathcal{E}$

Zufallsvariable: $X : \Omega \rightarrow \mathbb{R}$

- $\Leftrightarrow X$ ist $(\mathcal{A}, \mathcal{B})$ -messbar
- $\Leftrightarrow X^{-1}(B) = \{\omega | X(\omega) \in B\} \in \mathcal{A} \forall B \in \mathcal{B}$
- $\Leftrightarrow X^{-1}((-\infty, a]) = \{\omega | X(\omega) \leq a\} \in \mathcal{A} \forall a \in \mathbb{R}$
- $\Leftrightarrow X$ stetig oder (schwach) monoton wachsend oder fallend für $(\mathbb{R}, \mathcal{B})$

Verkettung zweier Zufallsvariablen $X : \Omega \rightarrow \mathbb{R}$ und $Y : \mathbb{R} \rightarrow \mathbb{R} \Rightarrow Y \circ X : \Omega \rightarrow \mathbb{R}$ ist wieder ZV

Verteilung: $P_X : \mathcal{B} \rightarrow [0, 1]$ mit $P_X(B) = P(\{\omega \in \Omega | X(\omega) \in B\}) \forall B \in \mathcal{B}$

- Die Verteilung ist ein Wahrscheinlichkeitsmaß auf dem Messraum $(\mathbb{R}, \mathcal{B})$

Verteilungsfunktion: $F_X : \mathbb{R} \rightarrow [0, 1]$ mit $F_X(x) = P(X \leq x) = P(\{\omega \in \Omega | X(\omega) \leq x\}) = P_X((-\infty, x])$

- $\lim_{x \rightarrow -\infty} F_X(x) = 0, \lim_{x \rightarrow +\infty} F_X(x) = 1$
- F_X ist (schwach) monoton wachsend
- F_X ist rechtsseitig stetig

Quantilfunktion: $F^{-1} : [0, 1] \rightarrow \mathbb{R}$ mit $F^{-1}(y) := \inf \{x \in \mathbb{R} | F(x) \geq y\}$ für eine Verteilungsfunktion $F : \mathbb{R} \rightarrow [0, 1]$

- F stetig und streng monoton wachsend $\Rightarrow F^{-1}$ übliche Umkehrfunktion

Diskrete Verteilung

$X : \Omega \rightarrow \mathbb{R}$ heißt **diskret**, falls es eine endliche oder abzählbare Menge $C \subseteq \mathbb{R}$ gibt, so dass $P(X \in C) = p_X(C) = 1$ ist.

OBdA: Sei $C = \{x_1, x_2, \dots\}$

Zähldichte von X : $p_X(k) = P(X = x_k)$

- $p_X(k) \geq 0 \forall k \in \mathbb{N}$
- $\sum_{k=1}^{\infty} p_X(k) = 1$

Stetige Verteilung

$X : \Omega \rightarrow \mathbb{R}$ heißt **absolutstetig**, falls die Verteilungsfunktion F_X von x folgende Darstellung besitzt:

$$F_X(x) = \int_{-\infty}^x f_X(y) dy \quad \forall x \in \mathbb{R}$$

Dichte von X : $f_X : \mathbb{R} \rightarrow [0, \infty)$

- $\int_{-\infty}^{\infty} f_X(y) dy = 1$

Diskrete Verteilungen:

- Binomialverteilung:

$$p_X(k) = P(X = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$
$$(n \in \mathbb{N}, p \in [0, 1])$$

- Hypergeometrische Verteilung:

$$P(X = k) = \frac{\binom{r}{k} \binom{n-r}{m-k}}{\binom{n}{m}}$$

- Geometrische Verteilung:

$$P(X = k) = (1-p) \cdot p^{k-1}$$
$$(p \in [0, 1])$$

- Poisson-Verteilung:

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$
$$(\lambda > 0)$$

- Diskrete Gleichverteilung auf $\{x_1, \dots, x_m\} \subset \mathbb{R}$:

$$P(X = x_i) = \frac{1}{m} \text{ für } i = 1 \dots m$$

Stetige Verteilungen:

- Gleichverteilung:

– Schreibweise: $X \sim U(a, b)$

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

$$F_X(x) = \int_a^x \frac{1}{b-a} dy = \frac{x-a}{b-a} \text{ für } a < x < b$$

- Exponentialverteilung:

– Schreibweise: $X \sim \exp(\lambda)$

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

$$F_X(x) = \int_0^x \lambda e^{-\lambda y} dy = 1 - e^{-\lambda x} \text{ für } x \geq 0$$

- Normalverteilung:

– Schreibweise: $X \sim N(\mu, \sigma^2)$

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

- Standardnormalverteilung ($\mu = 0, \sigma^2 = 1$):

$$F_X(x) = \Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} y^2\right) dy$$

Elementare Zufallsvariable: $X(\omega) = \sum_{i=1}^m \alpha_i 1_{A_i}(\omega)$

- $A_i \in \mathcal{A}, \alpha_i \in \mathbb{R}_+, m \in \mathbb{N}$
- Menge aller elementaren Zufallsvariablen im Messraum: M^E

Erwartungswert von X: $EX = \int X dP = \sum_{i=1}^m \alpha_i P(A_i)$

- Erwartungswert existiert $\Leftrightarrow EX < \infty$
- Linearität: $E(aX + bY) = aEX + bEY$
- Monotonie: $X \leq Y$ (d.h. $X(\omega) \leq Y(\omega) \forall \omega \in \Omega$) $\Rightarrow EX \leq EY$
- Mit $g: \mathbb{R} \rightarrow \mathbb{R}$ messbar gilt für diskreten und absolutstetigen Fall:

$$- Eg(x) \text{ existiert} \Leftrightarrow \sum_{k=0}^{\infty} |g(x_k)| \cdot p_X(k) < \infty$$

$$- Eg(x) = \sum_{k=0}^{\infty} g(x_k) \cdot p_X(k)$$

$$- EX = \sum_{x \in \mathbb{R}: P(X=x) > 0} x \cdot P(X = x)$$

$$- Eg(x) \text{ existiert} \Leftrightarrow \int_{-\infty}^{\infty} |g(x)| \cdot f_X(x) dx < \infty$$

$$- Eg(x) = \int_{-\infty}^{\infty} g(x) \cdot f_X(x) dx$$

$$- EX = \int_{-\infty}^{\infty} x \cdot f_X(x) dx$$

k-tes Moment von X: EX^k

k-tes zentriertes Moment von X: $E(X - EX)^k$

Varianz (2-tes zentriertes Moment): $\text{Var}(X) = E(X - EX)^2$

- $\text{Var}(X) = EX^2 - (EX)^2$
- $\text{Var}(aX + b) = a^2 \text{Var}(X) \forall a, b \in \mathbb{R}$
- $\text{Var}(X) \geq 0$ und $\text{Var}(X) = 0 \Leftrightarrow P(X = c) = 1$ für ein $c \in \mathbb{R}$

Tschebyscheff Ungleichung: $P(|X - EX| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} \text{Var}(X)$

Erwartungswert und Varianz diskreter Verteilungen

Verteilung	Zähldichte $P(X = k)$	Erwartungsw.	Varianz	Anschauliche Interpretation
Binomialverteilung $X \sim B(n, p)$	$\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$	$n \cdot p$	$n \cdot p \cdot (1-p)$	Münze wird n-mal geworfen, P(i-ter Wurf=Kopf)=p. Die Anzahl X der Kopfwürfe ist binomialverteilt.
Hypergeometrische Verteilung $X \sim H(r, n, m)$	$\frac{\binom{r}{k} \binom{n-r}{m-k}}{\binom{n}{m}}$	$r \frac{m}{n}$	$r \frac{m}{n} \left(1 - \frac{m}{n}\right) \cdot \frac{\binom{n-r}{m-k}}{\binom{n-1}{m-k}}$	Urne mit r roten und s schwarzen Kugeln, $r + s = n$. Zieht man m Kugeln ohne Zurücklegen, ist die Anzahl X der gezogenen roten Kugeln hypergeometrisch verteilt.
Geometrische Verteilung	$p \cdot (1-p)^{k-1}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	Die Wahrscheinlichkeit, dass ein Ereignis zum Erfolg führt, sei p. Die Anzahl X der Versuche, die notwendig sind um einen ersten Erfolg zu haben, ist geometrisch verteilt.
Poisson-Verteilung	$e^{-\lambda} \cdot \frac{\lambda^k}{k!}$	λ	λ	Die Poisson-Verteilung ist eine Approximation der Binomialverteilung für große n und kleinem p.
Diskrete Gleichverteilung auf $\{x_1, \dots, x_m\}$	$\frac{1}{m}$	$\frac{1}{m} \sum_{i=1}^m x_i$	berechnen!	Ein Würfel wird einmal geworfen. Der geworfene Wert X ist gleichverteilt und hat die möglichen Ausprägungen $x_1 = 1, x_2 = 2, \dots, x_6 = 6$.

Erwartungswert und Varianz stetiger Verteilungen

Verteilung	Dichte $f(x)$	Erwartungsw.	Varianz	Verteilungsfunktion $F_X(x)$
Gleichverteilung $X \sim U(a, b)$	$\begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{sonst} \end{cases}$	$\frac{b+a}{2}$	$\frac{(b-a)^2}{12}$	$\int_a^x \frac{1}{b-a} dy = \frac{x-a}{b-a}$ für $a < x < b$
Exponentialverteilung $X \sim exp(\lambda)$	$\begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{sonst} \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\int_0^x \lambda e^{-\lambda y} dy = 1 - e^{-\lambda x}$ für $x \geq 0$
Normalverteilung $X \sim N(\mu, \sigma^2)$	$\varphi_{\mu, \sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}\right)$	μ	σ^2	Integral nicht lösbar
Standard-Normalverteilung	$\mu = 0, \sigma^2 = 1$	0	1	$\int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} y^2\right) dy = \Phi(x)$ [Tabelle]

Zufallsvektoren

Produkt- σ -Algebra: $\mathcal{A} = \mathcal{A}_1 \otimes \dots \otimes \mathcal{A}_n = \sigma(\{A_1 \times \dots \times A_n \mid A_i \in \mathcal{A}, i = 1 \dots n\})$

Randverteilung: $Q_i(A_i) = P(\Omega_1 \times \dots \times \Omega_{i-1} \times A_i \times \Omega_{i+1} \times \dots \times \Omega_n)$

Zufallsvektor: $X = (X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$

Verteilung: $P_X : \mathcal{B}(\mathbb{R}^n) \rightarrow [0, 1]$ mit $P_X(B) = P(\{\omega \in \Omega \mid X(\omega) \in B\})$, $B \in \mathcal{B}(\mathbb{R}^n)$

Verteilungsfunktion: $F_X(x_1, \dots, x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n)$

X heißt **diskret**, falls es eine endliche oder abzählbar unendliche Menge $C = \{x_1, x_2, \dots\} \subset \mathbb{R}^n$ gibt, so dass

$$P(X \in C) = 1$$

Zähldichte: $p_X(k) = P(X = x_k)$

Randzähldichte: $P(X_i = y_i)$

- $= P(\{\omega \mid X(\omega) \in C, X_i(\omega) = y_i\})$
- $= \sum_{x \in C, x_i = y_i} P(X = x)$

X heißt **absolutstetig**, falls es eine integrierbare Funktion $f_X : \mathbb{R}^n \rightarrow [0, \infty)$ gibt mit

$$F_X(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f_X(y_1, \dots, y_n) dy_1 \dots dy_n$$

Dichte: $f_X : \mathbb{R}^n \rightarrow [0, \infty)$

Randdichte: $f_{X_i}(x)$

$$\bullet = \underbrace{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_X(y_1, \dots, y_{i-1}, x_i, y_{i+1}, \dots, y_n)}_{n-1 \text{ mal}} dy_1 \dots dy_{i-1} dy_{i+1} \dots dy_n$$

Verteilung: $P_X(B) = \int_B f_X(y) dy \quad \forall B \in \mathcal{B}(\mathbb{R}^n)$

Multinomialverteilung:

- Experiment mit r möglichen Ausgängen E_1, \dots, E_r mit jeweiligen Wahrscheinlichkeiten p_1, \dots, p_r ($p_1 + \dots + p_r = 1$)
- Der Versuch wird n -mal unabhängig wiederholt
- $X_i(\omega)$ sei die Anzahl der E_i -Ausgänge
- $\Rightarrow P(X_1 = k_1, \dots, X_r = k_r) = p_1^{k_1} \cdot \dots \cdot p_r^{k_r} \cdot \frac{n!}{k_1! \cdot \dots \cdot k_r!}$

Unabhängigkeit der Zufallsvariablen $X_1, \dots, X_n : \Omega \rightarrow \mathbb{R} : F_{X_1, \dots, X_n}(x_1, \dots, x_n) = \prod_{i=1}^n F_{X_i}(x_i)$

- $\Leftrightarrow P(X_1 \leq x_1, \dots, X_n \leq x_n) = \prod_{i=1}^n P(X_i \leq x_i) \quad \forall (x_1, \dots, x_n) \in \mathbb{R}^n$
- $\Leftrightarrow P(X_1 \in B_1, \dots, X_n \in B_n) = \prod_{i=1}^n P(X_i \in B_i) \quad \forall B_1, \dots, B_n \in \mathcal{B}$
- $\Leftrightarrow P(B_1 \times \dots \times B_n) = \prod_{i=1}^n P_{X_i}(B_i) \quad \forall B_1, \dots, B_n \in \mathcal{B}$

Produkt-Maß: $P(A_1 \times \dots \times A_n) = P_1(A_1) \cdot \dots \cdot P_n(A_n)$

i -te Projektion: $X_i(\omega) = X_i((\omega_1, \dots, \omega_n) = \omega_i, i = 1 \dots n$

X_1, \dots, X_n **unabhängig**

- $\Leftrightarrow P(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^n P(X_i = x_i)$
 $\forall (x_1, \dots, x_n) \in \mathbb{R}^n$

X_1, \dots, X_n **unabhängig**

- $\Leftrightarrow f_X(x_1, \dots, x_n) = \prod_{i=1}^n f_{X_i}(x_i)$
 $\forall (x_1, \dots, x_n) \in \mathbb{R}^n \setminus B$
- (B vom Lebesguemaß 0)

Addition zweier absolutstetiger Zufallsvariablen $(X = (X_1, X_2) : \Omega \rightarrow \mathbb{R}^2)$ mit gemeinsamer Dichte f_X :

- $X_1 + X_2$ absolutstetig mit Dichte $f_{X_1+X_2}(x) = \int_{-\infty}^{\infty} f_X(y, x-y) dy \quad \forall x \in \mathbb{R}$

Faltungformel (Addition zweier unabhängiger absolutstetiger ZV):

- $f_{X_1+X_2}(x) = \int_{-\infty}^{\infty} f_{X_1}(y)f_{X_2}(x-y) dy \quad \forall x \in \mathbb{R}$

Faltung: $P_X * P_Y = P_{X+Y}$ für X,Y unabhängig

Erwartungswert zweier unabhängiger Zufallsvariablen: $EX \cdot Y = EX \cdot EY$

- sonst: $EX \cdot Y = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot y \cdot f_{X,Y}(x,y) d(x,y)$

Cauchy-Schwarzsche Ungleichung: $(EXY)^2 \leq EX^2 \cdot EY^2$

Kovarianz: $\text{Cov}(X, Y) := E[(X - EX)(Y - EY)]$

- $\text{Cov}(X, Y) = EXY - EX \cdot EY$
- $\text{Cov}(X, X) = \text{Var}(X)$
- $\text{Cov}(X, Y) = \text{Cov}(Y, X)$

X und Y **unkorreliert** $\Leftrightarrow \text{Cov}(X, Y) = 0$

- X,Y unabhängig \Rightarrow X,Y unkorreliert

Korrelationskoeffizient: $\varrho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}}$

- Ist $\varrho(X, Y)$ definiert, so gilt: $-1 \leq \varrho(X, Y) \leq 1$
- Der Korrelationskoeffizient ist ein Maß für die lineare Abhängigkeit der Zufallsvariable

Berechnung der Varianz: $\text{Var}(X_1 + \dots + X_n) = \sum_{i=1}^n \text{Var}(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j)$

- Für unkorrelierte X_i : $\text{Var}(X_1 + \dots + X_n) = \sum_{i=1}^n \text{Var}(X_i)$

Erwartungsvektor: $EX = (EX_1, \dots, EX_n)$

Kovarianzmatrix: $\text{Cov}(X) = (\text{Cov}(X_i, X_j))_{1 \leq i, j \leq n}$

Erzeugende Funktion: $g_X(s) = \sum_{k=0}^{\infty} p_X(k) \cdot s^k = Es^k$

- Nur für X diskret
- $g_X : [-1, 1] \rightarrow \mathbb{R}$ ist wohldefiniert für $|s| \leq 1$
- Zähldichte: $p_X(k) = \frac{g_X^{(k)}(0)}{k!}$
- Erwartungswert: $EX = g'_X(1-) = \lim_{s \uparrow 1} g'_X(s)$
- Varianz: $\text{Var}(X) = g''_X(1-) + g'_X(1-) - (g'_X(1-))^2$
- Eindeutigkeit: $p_X(k) = p_Y(k) \quad \forall k \in \mathbb{N} \Leftrightarrow g_X(s) = g_Y(s) \quad \forall s \in [-1, 1]$
- Addition von X, Y unabhängig: $g_{X+Y}(s) = g_X(s) \cdot g_Y(s) \quad \forall s \in [-1, 1]$
- X_1, \dots, X_n unabhängig $\Leftrightarrow g_{\sum_{i=1}^n X_i} = \prod_{i=1}^n g_{X_i}$

Konvergenz von Zufallsvariablen

p-fast-sichere Konvergenz: $P(\{\omega \in \Omega \mid \lim_{n \rightarrow \infty} X_n(\omega) - X(\omega)\}) = 1$

- $X_n \xrightarrow{f.s.} X \Leftrightarrow P(\lim_{n \rightarrow \infty} X_n = X) = 1$

Konvergenz in Wahrscheinlichkeit: $\lim_{n \rightarrow \infty} P(\{\omega \in \Omega \mid |X_n(\omega) - X(\omega)| \geq \varepsilon\}) = 0 \quad \forall \varepsilon > 0$

- $X_n \xrightarrow{p} X \Leftrightarrow \lim_{n \rightarrow \infty} P(|X_n - X| \geq \varepsilon) = 0 \quad \forall \varepsilon > 0$

Konvergenz in Verteilung: $\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x) \quad \forall x \in \mathbb{R}$ an denen $F_X(x)$ stetig ist

- $X_n \xrightarrow{d} X \Leftrightarrow \lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x) \quad \forall$ Stetigkeitsstellen x von X

Zusammenhang der Konvergenzen:

- $X_n \xrightarrow{f.s.} X \Rightarrow X_n \xrightarrow{p} X$
- $X_n \xrightarrow{p} X \Rightarrow X_n \xrightarrow{d} X$
- $X_n \xrightarrow{d} c, c \in \mathbb{R} \Rightarrow X_n \xrightarrow{p} c$

Charakteristische Funktion: $\varphi_X(t) := Ee^{itX} = E \cos(tX) + iE \sin(tX)$

- $\varphi_X : \mathbb{R} \rightarrow \mathbb{C}$ existiert immer $\quad \forall t \in \mathbb{R}$
- X diskret $\Rightarrow \varphi_X(t) = g_X(e^{it})$
- X absolutstetig $\Rightarrow \varphi_X(t) = \int_{-\infty}^{\infty} e^{itx} \cdot f_X(x) dx = \int_{-\infty}^{\infty} \cos(tx) \cdot f_X(x) dx + i \int_{-\infty}^{\infty} \sin(tx) \cdot f_X(x) dx$
- Eigenschaften:
 - $\varphi_X(0) = 1$
 - $|\varphi_X(t)| \leq 1 \quad \forall t \in \mathbb{R}$
 - $\varphi_{aX+b}(t) = e^{ibt} \cdot \varphi_X(at)$
- Eindeutigkeit: $P_X = P_Y \Leftrightarrow \varphi_X = \varphi_Y$
- X_1, \dots, X_n unabhängig $\Leftrightarrow \varphi_{\sum_{i=1}^n X_i}(t) = \prod_{i=1}^n \varphi_{X_i}(t)$
- n-tes Moment: $\varphi_X^{(n)}(0) = i^n EX^n$
- Stetigkeitssatz: $X_n \xrightarrow{d} X \Leftrightarrow \varphi_{X_n}(t) \xrightarrow{n \rightarrow \infty} \varphi(t) \quad \forall t \in \mathbb{R}$ und φ ist stetig in 0

Schwaches Gesetz der großen Zahlen: $\frac{X_1 + \dots + X_n}{n} \xrightarrow{p} \mu$

- nur für X_1, X_2, \dots u.i.v. (unabhängig und identisch verteilt)
- $\mu = EX_i$

Starkes Gesetz der großen Zahlen: $\frac{X_1 + \dots + X_n}{n} \xrightarrow{f.s.} \mu$

- nur für X_1, X_2, \dots u.i.v. (unabhängig und identisch verteilt)
- $\mu = EX_i$

Zentraler Grenzwertsatz: für X_1, X_2, \dots u.i.v. mit $EX_i = \mu$ und $0 < \text{Var}(X_i) = \sigma^2 < \infty$ gilt:

- $\frac{X_1 + \dots + X_n - n\mu}{\sqrt{n}\sigma} \xrightarrow{d} X \sim N(0, 1)$
- $P\left(\frac{X_1 + \dots + X_n - n\mu}{\sqrt{n}\sigma} \leq x\right) \xrightarrow{n \rightarrow \infty} \Phi(x) \quad \forall x \in \mathbb{R}$
- $P\left(\alpha \leq \frac{X_1 + \dots + X_n - n\mu}{\sqrt{n}\sigma} \leq \beta\right) \xrightarrow{n \rightarrow \infty} \Phi(\beta) - \Phi(\alpha)$

Parameterschätzung

Familie von Verteilungen: $\{P_\theta | \theta \in \Theta\}$ mit $\Theta \subset \mathbb{R}^m$

Stichprobenraum: \mathcal{X}

Zufallsstichprobe von u.i.v. Zufallsvariablen X_1, \dots, X_n mit Verteilung $P: (X_1, \dots, X_n)$

Stichprobe: Realisierung (x_1, \dots, x_n) einer Zufallsstichprobe

Schätzer von θ : messbare Abbildung $T: \mathcal{X}^n \rightarrow \tilde{\Theta}, \tilde{\Theta} \supset \Theta$

Stichprobenmittel: $T: \mathbb{R}^n \rightarrow \mathbb{R}$ mit $T(x) = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$

Stichprobenvarianz: $S^2: \mathbb{R}^n \rightarrow \mathbb{R}$ mit $S^2(x) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$

Maximum-Likelihood-Methode

- Likelihood-Funktion der Stichprobe $x = (x_1, \dots, x_n) \subset \mathcal{X}^n$:

$$\begin{aligned} L_x(\theta) &= p(x_1, \theta) \cdot \dots \cdot p(x_n, \theta) & \Bigg| & & L_x(\theta) = f(x_1, \theta) \cdot \dots \cdot f(x_n, \theta) \\ &= P_\theta(X_1 = x_1) \cdot \dots \cdot P_\theta(X_n = x_n) = P_\theta(X = x) \end{aligned}$$

- Maximum-Likelihood-Schätzer (MLS): $\hat{\theta}_{ML}(x)$ mit $L_x(\hat{\theta}_{ML}(x)) = \sup_{\theta \in \Theta} L_x(\theta)$
- Ggf. kann der MLS durch Nullsetzen der Ableitung der Log-Likelihoodfunktion bestimmt werden:
 $\frac{\partial}{\partial \theta} \log(L_x(\theta)) = 0$

Momentenmethode

- k-tes theoretisches Moment** von $X \sim P_\theta$: $\mu_k = \mu_k(\theta) = E_\theta X^k, k = 1, 2, \dots$
- k-tes empirisches Moment** einer Stichprobe $x = (x_1, \dots, x_n)$: $\bar{x}_k = \frac{1}{n} \sum_{i=1}^n x_i^k$
- Vorgehensweise:
 - Setze theoretische und empirische Momente gleich $\Rightarrow \mu_k(\theta) = \bar{x}_k, k = 1, \dots, m$
 - Aufgelöst nach θ ergibt sich dann der Momentenschätzer $\hat{\theta}_{MM}(x) \in \Theta$

Erwartungstreueit eines Schätzers T : $\forall \theta \in \Theta: E_\theta T(X_1, \dots, X_n) = \theta$

Verzerrung eines Schätzers T : $b_T(\theta) = E_\theta T(X_1, \dots, X_n) - \theta$

- Erwartungstreue Schätzer sind unverzerrt

Mittlerer quadratischer Fehler: $MSE(T) = E_\theta \left[(T(X_1, \dots, X_n) - \theta)^2 \right] = \text{Var}_\theta(T) + b_T(\theta)^2$

- Für erwartungstreue Schätzer gilt: $MSE(T) = \text{Var}_\theta(T(X))$

Ungleichung von Cramér-Rao: $\text{Var}_\theta(T(X)) \geq \frac{(1 + \frac{\partial}{\partial \theta} b_T(\theta))^2}{E_\theta \left[\left(\frac{\partial}{\partial \theta} \log L_x(\theta) \right)^2 \right]}$

- Fisher-Information: $I(\theta) = E_\theta \left[\left(\frac{\partial}{\partial \theta} \log L_x(\theta) \right)^2 \right]$
- Für erwartungstreue Schätzer gilt: $\text{Var}_\theta(T(X)) \geq \frac{1}{I(\theta)}$

$(1 - \alpha)$ -Konfidenzintervall: $[L(x), U(x)]$ mit...

- $L, U: \mathcal{X}^n \rightarrow \Theta \subset \mathbb{R}$ messbare Funktionen
- $L(x) \leq U(x) \forall x \in \mathcal{X}^n$
- $P_\theta(L(X) \leq \theta \leq U(X)) = 1 - \alpha$

Eigenschaften:

- Sowohl Lage als auch Länge des Konfidenzintervalls hängen von der konkreten Stichprobe ab
- Ist z.B. $\alpha = 0,05$, dann enthält das Konfidenzintervall den wahren Parameter in 95% der Fälle

Testtheorie

Sei $\theta_0 \in \Theta$

Einseitiges Testproblem: $H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$

Zweiseitiges Testproblem: $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$

Kritischer Bereich: $R \subset \mathbb{R}^n = \mathcal{X}^n$, so dass H_0 verworfen wird für $x \in R$

- $0 = H_0$ wird nicht verworfen
- $1 = H_0$ wird verworfen

Test / Testverfahren: $\varphi : \mathcal{X}^n \rightarrow \{0, 1\}$

- $R = \{x \in \mathcal{X}^n | \varphi(x) = 1\}$

Gütefunktion: $\beta(\theta) = P_\theta(X \in R) = P_\theta(\varphi(x) = 1)$

- $\beta : \Theta \rightarrow [0, 1]$
- Für $\theta \in \Theta_1$ heißt $\beta(\theta)$ die Macht des Tests

Niveau / Signifikanzniveau α eines Tests: $\beta(\theta) \leq \alpha$

- \Leftrightarrow Wahrscheinlichkeit für einen Fehler 1. Art ist gleich α

wahr Entscheidung	H_0	H_1
H_0	ok	Fehler 1. Art
H_1	Fehler 2. Art	ok

Gleichmäßig bester Test $\varphi^* \in D_\alpha : \beta^*(\theta) = P_\theta(\varphi^*(x) = 1) = \max_{\varphi \in D_\alpha} P_\theta(\varphi(x) = 1)$

- D_α Menge von Test zum Niveau α

Wahl der Nullhypothese: Möchte man sich für $\theta < \theta_0$ entscheiden, sollte man $H_0 : \theta \geq \theta_0$ wählen

X_0, X_1, \dots, X_r u.i.v. mit $X_i \sim N(0, 1), r \in \mathbb{N}$:

- χ^2 -Verteilung mit r Freiheitsgraden: $\sum_{i=1}^r X_i^2$
- **t-Verteilung** mit r Freiheitsgraden: $\frac{X_0}{\sqrt{\frac{1}{r} \sum_{i=1}^r X_i^2}}$
- Für Zufallsstichprobe $X = (X_1, \dots, X_n)$ zur $N(\mu, \sigma^2)$ -Verteilung gilt:
 - $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \bar{X} \sim N(\mu, \frac{\sigma^2}{n})$
 - $S^2(X) = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2, \frac{(n-1) \cdot S^2(X)}{\sigma^2} \sim \chi_{n-1}^2$
 - $\sqrt{n} \frac{\bar{X} - \mu}{S(X)} \sim t_{n-1}$

Randomisierter Test: $\varphi : \mathcal{X}^n \rightarrow [0, 1]$

- Gibt an, mit welcher Wahrscheinlichkeit $\varphi(x)$ H_0 abgelehnt wird
- Gütefunktion: $\beta(\theta) = E_\theta \varphi(X)$

Neyman-Pearson-Test: randomisierter Test $\varphi^* : \mathcal{X}^n \rightarrow [0, 1]$ mit $\varphi^*(x) = \begin{cases} 0 & \text{falls } L_x(\theta_1) < c^* \cdot L_x(\theta_0) \\ \gamma(x) & \text{falls } L_x(\theta_1) = c^* \cdot L_x(\theta_0) \\ 1 & \text{falls } L_x(\theta_1) > c^* \cdot L_x(\theta_0) \end{cases}$

- $c^* \in [0, \infty)$ konstant
- $\gamma : \mathcal{X}^n \rightarrow [0, 1]$ eine Funktion

Lemma von Neyman-Pearson:

- Ist φ^* ein NPT mit $\alpha = \beta_{\varphi^*}(\theta_0)$, dann ist φ^* gleichmässig bester Test unter allen Tests zum selben Niveau α
- Für jedes $\alpha \in (0, 1)$ existiert ein NPT φ^* zum Niveau α . Dabei kann $\gamma(x) \equiv \gamma$ gewählt werden.

Familie von (Zähl-)Dichten mit monotonem Dichtequotienten: $\{f(x, \theta) | \theta \in \Theta\}$ bzw. $\{p(x, \theta) | \theta \in \Theta\}$ mit...

1. \exists messbare Funktion $T : \mathcal{X}^n \rightarrow \mathbb{R}$ so dass $q(x) = \frac{L_x(\theta_1)}{L_x(\theta_0)} = q^*(T(x_1, \dots, x_n))$

2. q^* monoton in $T(x_1, \dots, x_n) = T(x)$ für alle $\theta_0 < \theta_1$

- Bei exp-Verteilungen: $T(x) = \bar{x}$ erfüllt die Bedingung
- $X \in \mathcal{X}^n$ Stichprobe zu einer Verteilung mit nicht monoton fallenden Dichtequotienten in $T(x)$.

Jeder Test der Form $\begin{cases} 0 & \text{falls } T(x) > t_0 \\ \gamma & \text{falls } T(x) = t_0 \\ 1 & \text{falls } T(x) < t_0 \end{cases}$ ist gleichmässig bester Test für das Testproblem

$H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$ zum Niveau $\alpha = E_{\theta_0} \varphi(X) = \sup_{\theta \leq \theta_0} E_{\theta} \varphi(X)$

Likelihood-Quotient: $q(x) = \frac{\sup_{\theta \in \Theta_0} L_x(\theta)}{\sup_{\theta \in \Theta_1} L_x(\theta)}$

Likelihood-Quotiententest: $\begin{cases} 0 & \text{falls } q(x) > c_0 \\ \gamma & \text{falls } q(x) = c_0 \text{ (nur im diskreten Fall)} \\ 1 & \text{falls } q(x) < c_0 \end{cases}$