

Wahrscheinlichkeit und Statistik

Janis Hutz

<https://janishutz.com>

19. August 2026

1 Basics

IMPORTANT

Always simplify!!, i.e. insert $\mathbb{E}$, etc. and simplify

1.1 Wahrscheinlichkeitsräume

Begriff Ω Grundraum, $\omega \in \Omega$ Elementarereignis

Def Sigma-Algebra $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ ist σ -Algebra, falls:

- E1. $\Omega \in \mathcal{F}$
- E2. $A \in \mathcal{F} \Rightarrow A^C \in \mathcal{F}$ (A Ereignis $\Rightarrow$ nicht A auch)
- E3. $A_1, A_2, \dots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$
($A_1, \dots$ Ereignisse $\Rightarrow A_1$ oder A_2 oder $\dots$ ein Ereignis)

Bsp σ -Algebren bei 1x Würfeln ($\Omega = \{1, 2, 3, 4, 5, 6\}$)

- $\mathcal{F} = \{\emptyset, \{1, 2, 3, 4, 5, 6\}\}$
- $\mathcal{F} = \mathcal{P}(\Omega)$, dabei $|\mathcal{F}| = 64$
- $\mathcal{F} = \{\emptyset, \{1, 2\}, \{3, 4, 5, 6\}, \Omega\}$

Keine σ -Algebren sind bspw:

- $\mathcal{F} = \{\Omega\}$ (Komplementärereignis $\emptyset$ fehlt, E2 verletzt)
- $\mathcal{F} = \{\emptyset, \{1, 2, 3\}, \{4, 5, 6\}, \{1\}, \{2, 3, 4, 5, 6\}, \Omega\}$
(E3 verletzt, da bspw $\{4, 5, 6\} \cup \{1\} \notin \mathcal{F}$)

1.1.1 Wahrscheinlichkeitsmass

Def W.M $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ mit $A \mapsto \mathbb{P}[A]$, notiert $(\Omega, \mathcal{F})$, falls folgende Eigenschaften gelten

- E1. $\mathbb{P}[\Omega] = 1$
- E2. (σ -Additivität) $\mathbb{P}[A] = \sum_{i=1}^{\infty} \mathbb{P}[A_i]$, falls $A = \bigcup_{i=1}^{\infty} A_i$ (disjunkte Vereinigung)

Bsp Wieder mit Würfeln und $\mathcal{F} = \mathcal{P}(\Omega)$, sind W.M:

- Abbildung $\forall A \in \mathcal{F} \quad \mathbb{P}[A] = \frac{|A|}{6}$
- Abbildung $\forall A \in \mathcal{F} \quad \mathbb{P}[A] = \sum_{i \in A} p_i$ (p_i dabei prob. Zahl i würfeln; $p_i = \frac{1}{6} \forall i \in \Omega$ ist für fairen Würfel)

1.1.2 Wahrscheinlichkeitsraum

Def W.R ein Tripel $(\Omega, \mathcal{F}, \mathbb{P})$

Begriff A Ereignis, tritt (nicht) ein (für ω), if $\omega \in (\notin) A$

Bem $A = \emptyset$ tritt niemals ein, $A = \Omega$ immer.

1.2 Beispiele von Wahrscheinlichkeitsräumen

Def Laplace Modell $(\Omega, \mathcal{F}, \mathbb{P})$, sodass $\mathcal{F} = \mathcal{P}(\Omega)$ und

$\mathbb{P} : \mathcal{F} \rightarrow [0, 1] \stackrel{\text{def}}{=} \forall A \in \mathcal{F} \quad \mathbb{P}[A] = \frac{|A|}{|\Omega|}$, $\mathbb{P}$ ist W.M.

Bsp Auf Kreis mit $n \geq 3$ Punkten, Modell für Nachbarn

ist: $A = \{\{1, 2\}, \dots, \{n-1, n\}, \{n, 1\}\}$ für $\Omega = \{E \subseteq \{1, \dots, n\} \mid |E| = 2\}$, also $\mathbb{P}[A] = \frac{n}{\binom{n}{2}} = \frac{2}{n-1}$

Bsp W. 1. mal Kopf ist bei Wurf k : $p_k = p^{k-1}(1-p)$

1.3 Eigenschaften/Interp. von Ereignissen

Thm $\mathcal{F}$ σ -Algebra. Es gilt: E4. $\emptyset \in \mathcal{F}$

E5. $A_1, A_2, \dots \in \mathcal{F} \Rightarrow \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$

E6. $A, B \in \mathcal{F} \Rightarrow A \cup B \in \mathcal{F}$

E7. $A, B \in \mathcal{F} \Rightarrow A \cap B \in \mathcal{F}$

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
$A \subseteq B$	B tritt ein, falls A eintritt
$A \cap B = \emptyset$	A und B nicht gleichzeitig
$\Omega = A_1 \cup A_2 \cup A_3$ mit $\forall \omega \in \Omega$	nur eines von A_1, A_2, A_3 kann eintreten

Wir wählen nicht immer $\mathcal{F} = \mathcal{P}(\Omega)$, bspw. für mehrstufige Experimente ist dies nicht ideal (k. Filtern, Überabzählbarkeit)

1.4 Eigenschaften Wahrscheinlichkeitsmasse

Thm $\mathbb{P}$ Wahrscheinlichkeitsmass auf $(\Omega, \mathcal{F})$, A Ereignis:

E1. Es gilt $\mathbb{P}[\emptyset] = 0$

E2. **Additivität** $k \geq 1$, $A_1, \dots, A_k$ paarw. disj. Ereignisse:
 $\mathbb{P}[A_1 \cup \dots \cup A_k] = \mathbb{P}[A_1] + \dots + \mathbb{P}[A_k]$

E3. $\mathbb{P}[A^C] = 1 - \mathbb{P}[A]$

E4. B Ereignis, dann $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B]$

1.4.1 Nützliche Ungleichungen

Thm Monot. $A, B \in \mathcal{F}$, dann $A \subseteq B \Rightarrow \mathbb{P}[A] \leq \mathbb{P}[B]$

Thm Union Bound Für $A_1, A_2, \dots$ (mögl. disj.) gilt:
 $\mathbb{P}[\bigcup_{i=1}^{\infty} A_i] \leq \sum_{i=1}^{\infty} \mathbb{P}[A_i]$. Auch für end. n.-leere Ereignisse
Zudem: $\lim_{n \rightarrow \infty} \mathbb{P}[\bigcup_{i=1}^n A_i] = \mathbb{P}[\bigcup_{i=1}^{\infty} A_i]$

1.4.2 Anwendungen der Ungleichungen

Sie sind nützlich für schwer zu berechnende W.

Thm (A_n) mit $A_n \subseteq A_{n+1}$ (mon. wachsend). Dann:
 $\lim_{n \rightarrow \infty} \mathbb{P}[A_n] = \mathbb{P}[\bigcup_{n=1}^{\infty} A_n]$

und für (B_n) mit $B_n \supseteq B_{n+1}$ (mon. fallend) gilt:
 $\lim_{n \rightarrow \infty} \mathbb{P}[B_n] = \mathbb{P}[\bigcap_{n=1}^{\infty} B_n]$

Bem Mit Monotonie: $\mathbb{P}[A_n] \leq \mathbb{P}[A_{n+1}]$ und $\mathbb{P}[B_n] \geq \mathbb{P}[B_{n+1}]$. Grenzwerte oben wohldefiniert.

1.5 Bedingte Wahrscheinlichkeit

Def Für $(\Omega, \mathcal{F}, \mathbb{P})$ mit $A, B \in \mathcal{P}(\Omega)$ mit $\mathbb{P}[B] > 0$:

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} \quad (= \mathbb{P}[A] \text{ wenn } B = \Omega \text{ ist})$$

Bem $\mathbb{P}[B|B] = 1$

Thm $B \in \mathcal{P}(\Omega)$, dann ist $\mathbb{P}[\cdot|B]$ ein W-Mass auf Ω

Thm Totale W. $\Omega = B_1 \cup \dots \cup B_n$ mit B_i s eine Partition von Ω , mit B_i paarw. disj. und $\mathbb{P}[B_i] > 0 \forall 1 \leq i \leq n$. Dann:

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

Thm Bayes B_i wie oben, dann $\forall A$ mit $\mathbb{P}[A] > 0$:

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

1.6 Unabhängigkeit

Siehe zudem Abschnitt 2.3 (für Zufallsvariablen)

Def $A, B \in \mathcal{F}$ unabh. falls $\mathbb{P}[A \cap B] = \mathbb{P}[A] \cdot \mathbb{P}[B]$

Bem Falls $\mathbb{P}[A] \in \{0, 1\}$: $\forall B \in \mathcal{F} \quad \mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$.
Falls A unabh. von sich selbst ($\mathbb{P}[A \cap A] = \mathbb{P}[A]^2$), dann $\mathbb{P}[A] \in \{0, 1\}$.
Implik.: A unabh. v. $B \Leftrightarrow A$ unabh. v. B^C

Thm Sei $\mathbb{P}[A], \mathbb{P}[B] > 0$. Dann ist equivalent:

- (i) $\mathbb{P}[A \cap B] = \mathbb{P}[A] \cdot \mathbb{P}[B]$ (A und B unabhängig)
- (ii) $\mathbb{P}[A|B] = \mathbb{P}[A]$ (Eintreten von B beeinflusst A nicht)
- (iii) $\mathbb{P}[B|A] = \mathbb{P}[B]$ (Eintreten von A beeinflusst B nicht)

Def I eine beliebige Menge. $(A_i)_{i \in I}$ **unabhängig** falls:

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

Bem Für paarweise unabhängige Ereignisse A_i ist jede Kombination von Vereinigungen und Schnittmengen unabhängig von anderen Variablen.

2 Zufallsvar., Verteilungsfunktionen

2.1 Abstrakte Definition

Def **Zufallsvariable** kurz Z.V. ist $\mathcal{X} : \Omega \rightarrow \mathbb{R}$, sodass $\forall a \in \mathbb{R}$ gilt: $f = \{\omega \in \Omega \mid \mathcal{X}(\omega) \leq a\} \in \mathcal{F}$ (notwendige Bedingung für Wohldefiniertheit von $\mathbb{P}[f]$)

Notation Ohne ω : $\{X \leq a\} = \{\omega \in \Omega \mid X(\omega) \leq a\}$, etc

Def **Indikatorfunkt.** $A \in \mathcal{F}$, $1_A(\omega) = \begin{cases} 0 & \text{wenn } \omega \notin A \\ 1 & \text{wenn } \omega \in A \end{cases}$

Das A kann bspw $\mathcal{X} \leq a$ sein. Für $A \cap B$, $\mathbb{E}[1_A 1_B] = 0$ weil A und B nie gleichzeitig eintreten

2.2 Verteilungsfunktion

Die untenstehende Funktion ist die Cumulative Distribution Function (cdf)

Def $F_{\mathcal{X}} : \mathbb{R} \rightarrow [0, 1]$, def: $\forall a \in \mathbb{R} \quad F_{\mathcal{X}}(a) = \mathbb{P}[\mathcal{X} \leq a]$

Thm $a < b \in \mathbb{R}$. Dann: $\mathbb{P}[a < X \leq b] = F(b) - F(a)$

Nachfolgendes wird gebraucht, um zu Überprüfen, ob eine Funktion eine cdf ist

Thm $\mathcal{X}$ Z.V. auf $(\Omega, \mathcal{F}, \mathbb{P})$ und V.F. $F = F_{\mathcal{X}}$. Eig.:

- F ist monoton wachsend
- F ist rechtsseitig ($F(a) = \lim_{h \downarrow 0} F(a+h) \quad \forall a \in \mathbb{R}$)
- $\lim_{a \rightarrow -\infty} F(a) = 0$ und $\lim_{a \rightarrow \infty} F(a) = 1$

$$F_{\mathcal{X}^n}(a) = \mathbb{P}[\mathcal{X}^n \leq a] = \mathbb{P}[(-1)^{n-1} \sqrt[n]{a} \leq \mathcal{X} \leq \sqrt[n]{a}]$$

2.3 Unabhängigkeit

Siehe zudem Abschnitt 1.6

Def Z.V. $\mathcal{X}_1, \dots, \mathcal{X}_n$ sind **unabh.** falls $\forall x_1, \dots, x_n \in \mathbb{R}$ $\mathbb{P}[\mathcal{X}_1 \leq x_1, \dots, \mathcal{X}_n \leq x_n] = \mathbb{P}[\mathcal{X}_1 \leq x_1] \cdots \mathbb{P}[\mathcal{X}_n \leq x_n]$

Bem Alternativ: ZVs unabhängig, falls $\forall I_1 \subseteq \mathbb{R}, \dots, I_n \subseteq \mathbb{R}$ Intervalle $\{\mathcal{X}_1 \in I_1\}, \dots, \{\mathcal{X}_n \in I_n\}$ unabhängig

Def u.i.v. (i.i.d.) (unabhängig & identisch verteilt) Z.V.

Bem Paarweise unabh. $\not\Rightarrow$ unabh. von allen Z.V. zusammen

2.3.1 Gruppierung

Thm n ZV $\mathcal{X}_i$ und $1 \leq i_1 < \dots < i_k \leq n$ sind indizes und $\phi_1, \dots, \phi_k$ Abbildungen. Dann sind unabhängig:

$$Y_1 = \phi_1(\mathcal{X}_1, \dots, \mathcal{X}_{i_1}), \dots, Y_k = \phi_k(\mathcal{X}_{i_k+1}, \dots, \mathcal{X}_n)$$

2.4 Transformation von Zufallsvariablen

Da ZV Funk. $\Omega \rightarrow \mathbb{R}$ sind, mit komposition neue ZV aus and. ZV, bspw: $Z_1 = \exp(\mathcal{X}_1)$ oder $Z_2 = \mathcal{X}_1 + \mathcal{X}_2$

2.5 Konstruktion von Zufallsvariablen

Thm $\exists$ -T v. Kolmogorov $\exists$ W-Raum und endl. uiv Folge von $\mathcal{X}_i \sim \text{Ber}(\frac{1}{2})$ mit param 0.5

Def ZV U heisst **gleichverteilt auf** $[0, 1]$, falls

$$F_U(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x \leq 1. \text{ Wir schreiben } U \sim \mathcal{U}([0, 1]) \\ 1 & x > 1 \end{cases}$$

Thm $\mathcal{X}_i$ wie oben, da $\mathcal{X}_i(\omega) \in \{0, 1\}$ konvergiert $\mathcal{Y}(\omega) = \sum_{n=1}^{\infty} 2^{-n} \mathcal{X}_n(\omega)$ absolut, mit $\mathcal{Y}(\omega) \in [0, 1]$. $\mathcal{Y} \sim \mathcal{U}([0, 1])$

Def **Inverse Vert.** von F ist $F^{-1} : (0, 1) \rightarrow \mathbb{R}$. Def: $\forall \alpha \in (0, 1) \quad F^{-1}(\alpha) = \inf\{x \in \mathbb{R} \mid F(x) \geq \alpha\}$

Thm **Inversionsmethode** F erfüllt T2.4, u. $U \sim \mathcal{U}(\dots)$, dann hat ZV $X = F^{-1}(U)$ die Verteilfunkt. $F_X = F$

Bem X wohldefiniert mit $X(\omega) = F^{-1}(U(\omega))$ falls $U(\omega) \in (0, 1)$ und $X(\omega) = 0$ sonst.

Thm $F_1, \dots$ Folge von Funk. mit eig. v. T2.4. Dann $\exists$ W-Raum und Folge von unabhängigen ZV X_i , sodass:

- $\forall i \quad X_i$ has F_i (also $\forall x \quad \mathbb{P}[X_i \leq x] = F_i(x)$)
- $X_1, X_2, \dots$ sind unabhängig.

3 Diskrete und stetige ZV

3.1 Stetigkeit der Verteilungsfunktion

Thm **W. Punkt** Für Z.V $\mathcal{X}$ und V.F $F_{\mathcal{X}}$ gilt $\forall a \in \mathbb{R}$ $\mathbb{P}[\mathcal{X} = a] = F(a) - F(a-)$ mit $F(a-) := \lim_{h \downarrow 0} F(a-h)$.

- $\rightarrow F$ in a n. stetig, dann "Sprunghöhe" $F(a) - F(a-) = \mathbb{P}[\mathcal{X} = a]$
- $\rightarrow F$ stetig in a , dann $\mathbb{P}[\mathcal{X} = a] = 0$

3.2 Fast sichere Ereignisse

Def $A \in \mathcal{F}$ tritt **fast sicher** (f.s.) ein, falls $\mathbb{P}[A] = 1$.

Bem Für allg. Mengen: A f.s., falls $\exists A' \subseteq A : \mathbb{P}[A'] = 1$

3.3 Diskrete Zufallsvariablen

Def Z.V. $\mathcal{X}$ ist **diskret**, falls endl. oder abzählb. Menge $W \subsetneq \mathbb{R}$ existiert, s.d. $\mathbb{P}[\mathcal{X} \in W] = 1$ (Werte v. $\mathcal{X}$ f.s. in W)

Bem Falls Ω endl. / abzählb., dann ist jede Z.V. $\mathcal{X}$ diskret.

Def **Gewichtsf.** $\mathcal{X}$ mit W endl.: $p_{\mathcal{X}}(x) = \mathbb{P}[\mathcal{X} = x]$

Def **Verteilung** $(p(x))_{x \in W} \stackrel{\text{def}}{\iff} \forall x \in W \quad p(x)$

Thm $(p(x))_{x \in W} = \sum_{x \in W} p(x) = 1$

Bem $\forall (p(x))_{x \in W} \exists$ Z.V. mit dieser Verteilung. Können desh. schreiben: "Sei $\mathcal{X}$ disk. Z.V. mit Verteilung $(p(x))_{x \in W}$ "

3.3.1 Zusammenhang Verteilung, Verteilungsfunktion

Thm $\mathcal{X}$ disk. Z.V. wie oben, dann ist **Verteilungsfunktion**: $\forall x \in \mathbb{R} \quad F_{\mathcal{X}}(x) = \sum_{\substack{y \in W \\ y \leq x}} p(y)$. **Umgekehrt**: $p(x)$ ist die "Sprunghöhe" im Punkt $x \in W$, W pos. Sprünge in $F_{\mathcal{X}}$

3.4 Verteilungen

3.4.1 Bernoulli-Verteilung

Def $\mathcal{X} \sim \text{Ber}(p)$: $\mathbb{P}[\mathcal{X} = k] = p^k(1-p)^{1-k}$ mit $k \in \{0, 1\}$

3.4.2 Binomialverteilung

Def $\mathcal{X} \sim \text{Bin}(n, p)$, falls $\forall k \in \{0, \dots, n\}$ $\mathbb{P}[\mathcal{X} = k] = \binom{n}{k} p^k (1-p)^{n-k}$ ($F_{\mathcal{X}}$ mit Summe berechnen)

Bem $\sum_{k=0}^n p(k) = \sum_{k=0}^n \mathbb{P}[\mathcal{X} = k] = (p+1-p)^n = 1$

Thm $X_i \sim \text{Ber}(p_i)$ unab: $(S_n := \sum_{i=0}^n X_i) \sim \text{Ber}(n, p)$

Bem $\text{Bin}(1, p)$ ist $\text{Ber}(p)$ verteilt. Für $X, Y \sim \text{Bin}(n_i, p)$ mit X, Y unabhängig dann ist $X + Y \sim \text{Bin}(n_1 + n_2, p)$

3.4.3 Geometrische Verteilung

Warten auf den ersten Erfolg (in ∞ Folge von Bernoulli-Experimenten)

Def $\mathcal{X} \sim \text{Geom}(p)$ mit $W = \mathbb{N} \setminus \{0\}$ und $p \in [0, 1]$, falls $\forall k \in W \quad \mathbb{P}[\mathcal{X} = k] = (1-p)^{k-1} p$

Bem **Verteilung** $F_{\mathcal{X}}(k) = 1 - (1-p)^k$

Bem $\mathbb{P}[\mathcal{X} = 1] = p$, da wir Konvention $a^0 = 1$ verwenden.

Bem $\sum_{k=0}^{\infty} p(k) = p \cdot \sum_{k=0}^{\infty} (1-p)^{k-1} = p \cdot \frac{1}{p} = 1$

Thm $X_i \sim \text{Ber}(p)$ für $i \in \mathbb{N}$.

Dann $(T := \min\{n \geq 1 \mid X_n = 1\}) \sim \text{Geom}(p)$

Bem $T = \infty$ ist möglich, $\mathbb{P}[T = \infty] = 0$

Thm **Gedächtnislosigkeit** $T \sim \text{Geom}(p)$, dann $\forall n \geq 0 \quad \forall k \geq 1 \quad \mathbb{P}[T \geq n+k \mid T > n] = \mathbb{P}[T \geq k]$

3.4.4 Negativbinomiale Verteilung

Warten auf den r -ten Erfolg (in ∞ Folge von Bernoulli-Experimenten)

Def $\mathcal{X} \sim \text{NBin}(r, p)$, falls $\forall k \in \{r, r+1, \dots\}$

$$\mathbb{P}[\mathcal{X} = k] = \binom{k-1}{r-1} p^r (1-p)^{k-r} \quad (F_{\mathcal{X}} \text{ mit Summe})$$

Thm $\mathcal{X}_i \sim \text{Ber}(p)$, dann

$$T_r := \inf\{n \geq 1 \mid \sum_{i=1}^n \mathcal{X}_i = r\} \sim \text{NBin}(r, p)$$

Bem $\mathcal{X} := \sum_{i=1}^r \mathcal{X}_i \sim \text{NBin}(r, p)$ für $\mathcal{X}_i \sim \text{Geom}(p)$

3.4.5 Hypergeometrische Verteilung

r Elemente vom Typ I, $n-r$ El. vom Typ II, m davon gezogen, ohne Zurücklegen

Def $\mathcal{X} \sim \text{H}(n, r, m)$, falls $\forall k \in \{0, \dots, \min(m, r)\}$

$$\mathbb{P}[\mathcal{X} = k] = \frac{\binom{r}{k} \binom{n-r}{m-k}}{\binom{n}{m}} \quad (F_{\mathcal{X}} \text{ mit Summe berechnen})$$

3.4.6 Poisson-Verteilung

Approximation von Binomialverteilung

Def $\mathcal{X} \sim \text{Poisson}(\lambda)$ mit $\lambda > 0 \in \mathbb{R}$, falls

$$\forall k \in \mathbb{N} \quad \mathbb{P}[\mathcal{X} = k] = \frac{\lambda^k}{k!} e^{-\lambda} \quad F_{\mathcal{X}}(k) = e^{-\lambda} \sum_{j=0}^k \frac{\lambda^j}{j!}$$

Thm Für $n \in \mathbb{N}$ Z.V. $\mathcal{X}_i \sim \text{Bin}(n, \frac{\lambda}{n})$ und $\mathcal{N} \sim \text{Poisson}(\lambda)$:

$$\forall k \in \mathbb{N} \quad \lim_{n \rightarrow \infty} \mathbb{P}[\mathcal{X}_i = k] = \mathbb{P}[\mathcal{N} = k]$$

$\mathcal{X} + \mathcal{Y} \sim \text{Poisson}(\lambda + \mu)$, nicht f. $\mathcal{X} + 2\mathcal{Y}$ ($2\mathcal{Y}$ n. pois. dist)

3.5 Stetige Zufallsvariablen

Def Stetig verteilte Z.V. $\mathcal{X}$ stetig, falls $\exists f_{\mathcal{X}} : \mathbb{R} \rightarrow \mathbb{R}_+$, s.d. V.F. $F_{\mathcal{X}}(x) = \int_{-\infty}^x f_{\mathcal{X}}(t) dt$. $f_{\mathcal{X}}$ ist Dichte (pdf) v. $\mathcal{X}$. Falls Z.V. Sprungstelle hat, **weder stetig noch diskret**. Zudem wenn nicht $P[a < \mathcal{X} \leq b]$, dann Grenzwert an Seite mit Abweichung von diesem Schema

Bem $\mathbb{P}[\mathcal{X} = n] \approx 0$; $\mathbb{P}[m \leq \mathcal{X} \leq n] = \int_m^n f_{\mathcal{X}}(x) dx$

Thm $F_{\mathcal{X}}$ stetig stückw. diff. auf Partition $-\infty = x_0 < \dots < x_n = \infty$. Dann $\mathcal{X}$ stetig, mit a_k beliebig und

$$f_{\mathcal{X}} = \begin{cases} F'_{\mathcal{X}}(x) & \exists k \in \{0, \dots, n-1\} : x \in (x_k, x_{k+1}) \\ a_k & x \in \{x_1, \dots, x_{n-1}\} \end{cases}$$

Thm $\int_{-\infty}^{\infty} f_{\mathcal{X}}(x) dx = 1$ (nützl. f. berech. v. Param in f)

Bem $f_{\mathcal{X}}$ fast analog zu Gew.F $p_{\mathcal{X}}$. Also: $(\Sigma, p_{\mathcal{X}}) \mapsto (\int, f_{\mathcal{X}})$ vom diskreten zu stetigen Fall

3.6 Stetige Verteilungen

3.6.1 Gleichverteilung (Uniformmodell)

Def $\mathcal{X} \sim \mathcal{U}([a, b])$, falls $f_{\mathcal{X}} = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & \text{sonst} \end{cases}$

Bem $\mathbb{P}[\mathcal{X} \in [c, c+l]] = \frac{l}{b-a}$, $F_{\mathcal{X}}(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b \end{cases}$

3.6.2 Exponentialverteilung

Wie Geometrische Verteilung warten auf Erfolg

Def $\mathcal{X} \sim \text{Exp}(\lambda)$, falls $\forall x \in \mathbb{R} f_{\mathcal{X}}(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$

Bem Gedächtnisl. $\mathbb{P}[\mathcal{X} > t+s \mid \mathcal{X} > s] = \mathbb{P}[\mathcal{X} > t]$

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

3.6.3 Cauchy-Verteilung

Def $\mathcal{X} \sim \text{Cauchy}(x_0, \gamma)$, falls $f_{\mathcal{X}}(x) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (x - x_0)^2}$

Bem Verteilungsfunk. $F_{\mathcal{X}}(x) = \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{x - x_0}{\gamma}\right)$

Def Langschwänzige Verteilung für $|x| \rightarrow \infty$ nur sehr langsam gegen 0 (quadratisch vs. exponentiell bei Norm. V)

3.6.4 Normalverteilung

Def $\mathcal{X} \sim \mathcal{N}(\mu, \sigma^2)$ falls $f_{\mathcal{X}}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$,

mit σ Standardabweichung. Auch: Gauss'sche Verteilung

Def Standardnormalverteilung $\mathcal{X} \sim \mathcal{N}(0, 1)$:

$\Phi = \int_{-\infty}^x \varphi(t) dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{\frac{-t^2}{2}} dt$ mit $\varphi = f_{\mathcal{X}}$.

Es gilt: $\Phi(-t) = 1 - \Phi(t)$

Thm $\mathcal{X} \sim \mathcal{N}(\mu, \sigma^2)$, dann $\frac{\mathcal{X}-\mu}{\sigma} \sim \mathcal{N}(0, 1)$, also:

$$F_{\mathcal{X}}(x) = \mathbb{P}[\mathcal{X} \leq x] = \mathbb{P}\left[\frac{\mathcal{X}-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right] = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

Bsp für Phänomene modellierbar mit Normalverteilung:

- Streuung von Messwerten um Mittelwert
- Grösse und Gewicht der Bevölkerung
- Renditen von Aktien

Bem Für $\mathcal{X}_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ unabhängig gilt:

$$\mathcal{Y} := \mu_0 + \sum_{k=1}^n a_k \mathcal{X}_k \sim \mathcal{N}\left(\mu_0 + \sum_{k=1}^n a_k \mu_k, \sum_{k=1}^n a_k^2 \sigma_k^2\right)$$

Bem Für $\mu \in \mathbb{R}, \sigma^2 > 0$ und $\mathcal{Z} \sim \mathcal{N}(0, 1)$ gilt $\mu + \sigma \mathcal{Z} \sim \mathcal{N}(\mu, \sigma^2)$ (nützlich für Simulation)

Bem $\mathbb{P}[|\mathcal{X} - \mu| \geq 3\sigma] \leq 0.0027$

4 Erwartungswert

4.1 Allgemeiner Erwartungswert

Def Für $\mathcal{X} : \Omega \rightarrow \mathbb{R}_+$, $\mathbb{E}[\mathcal{X}] = \int_0^{\infty} (1 - F_{\mathcal{X}}(x)) dx$

Bem $\mathbb{E}[\mathcal{X}]$ immer definiert und endlich oder unendlich

Thm $\mathcal{X}$ n.neg. Dann: $\mathbb{E}[\mathcal{X}] \geq 0$. =, wenn $\mathcal{X} = 0$ fast sicher

Def $\mathbb{E}[\mathcal{X}] = \mathbb{E}[\mathcal{X}_+] - \mathbb{E}[\mathcal{X}_-]$ mit $\mathcal{X}_-$ auch n.-neg.

Bem $|\mathcal{X}| = \mathcal{X}_+ + \mathcal{X}_-$; $\mathcal{X} = \mathcal{X}_+ - \mathcal{X}_-$. Für $\mathcal{X} \geq 0$ ist $\mathbb{E}[\mathcal{X}]$ immer definiert. Falls $\mathcal{X}$ kein konst. Vorzeichen, $\mathbb{E}[\mathcal{X}]$ undef.

Bem $\mathbb{E}[\mathcal{X}] = \int_0^{\infty} (1 - F_{\mathcal{X}}(x)) dx - \int_{-\infty}^0 F_{\mathcal{X}}(x) dx$

4.2 Diskrete Zufallsvariablen

Thm Für $\mathcal{X}$ mit Werten fast sicher in W :

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

Bem $\mathbb{E}[\mathcal{X}]$ wohldefiniert falls $(x \cdot p_{\mathcal{X}}(x))_{x \in W}$ abs. konv.

4.2.1 Beispiele

- $\mathcal{X} \sim \text{Ber}(p)$: $\mathbb{E}[\mathcal{X}] = p$ ($\mathbb{E}[1_A] = \mathbb{P}[A]$)
- $\mathcal{X} \sim \text{Bin}(n, p)$: $\mathbb{E}[\mathcal{X}] = np$
- $\mathcal{X} \sim \text{NBin}(r, p)$: $\mathbb{E}[\mathcal{X}] = \frac{r(1-p)}{p}$
- $\mathcal{X} \sim \text{Geom}(p)$: $\mathbb{E}[\mathcal{X}] = \frac{1}{p}$
- $\mathcal{X} \sim \text{H}(n, r, m)$: $\mathbb{E}[\mathcal{X}] = m \frac{n}{r}$
- $\mathcal{X} \sim \text{Poisson}(\lambda)$: $\mathbb{E}[\mathcal{X}] = \lambda$

4.2.2 Transformierte Zufallsvariablen

Thm Für $\varphi : \mathbb{R} \rightarrow \mathbb{R}$, $\mathbb{E}[\varphi(\mathcal{X})] = \sum_{x \in W} \varphi(x) \cdot \mathbb{P}[\mathcal{X} = x]$

4.3 Stetige Zufallsvariablen

Def $\mathcal{X}$ stetig $\mathbb{E}[\mathcal{X}] = \int_{-\infty}^{\infty} x f_{\mathcal{X}}(x) dx$

Thm $\mathbb{E}[\varphi(\mathcal{X})] = \int_{-\infty}^{\infty} \varphi(x) f_{\mathcal{X}}(x) dx$, falls int. wohldef.

Lem Int über gauss. Glockenk. $\int_{-\infty}^{\infty} e^{\frac{-x^2}{2\sigma^2}} dx = \sqrt{2\pi\sigma^2}$

4.3.1 Beispiele

- $\mathcal{X} \sim \mathcal{U}([a, b])$, $a < b$: $\mathbb{E}[\mathcal{X}] = \frac{a+b}{2}$
- $\mathcal{X} \sim \text{Exp}(\lambda)$, $\lambda > 0$: $\mathbb{E}[\mathcal{X}] = \frac{1}{\lambda}$
- $\mathcal{X} \sim \mathcal{N}(\mu, \sigma^2)$: $\mathbb{E}[\mathcal{X}] = \mu$
- $\mathcal{X} \sim \text{Cauchy}(x_0, \gamma)$: Existiert nicht (Int. ∞)
 $\mathbb{E}[\mathcal{X}_+] = \mathbb{E}[\mathcal{X}_-] = \infty$, Median: 0

4.4 Eigenschaften des Erwartungswerts

Thm Linearität Falls $\mathbb{E}[\mathcal{X}]$ und $\mathbb{E}[\mathcal{Y}]$ wohldefiniert:

$$\mathbb{E}[\lambda \mathcal{X}] = \lambda \mathbb{E}[\mathcal{X}] \text{ und } \mathbb{E}[\mathcal{X} + \mathcal{Y}] = \mathbb{E}[\mathcal{X}] + \mathbb{E}[\mathcal{Y}]$$

Bem Z.V $\mathcal{X}_k$ und λ_k : $\mathbb{E}[\sum_{k=1}^n \lambda_k \mathcal{X}_k] = \sum_{k=1}^n \lambda_k \mathbb{E}[\mathcal{X}_k]$

Thm Monotonie Sei $\mathcal{X} \leq \mathcal{Y}$ mit $\mathbb{E}$ wohldef.: $\mathbb{E}[\mathcal{X}] \leq \mathbb{E}[\mathcal{Y}]$

Thm $\mathcal{X}, \mathcal{Y}$ unabh., dann $\mathbb{E}[\mathcal{X}\mathcal{Y}] = \mathbb{E}[\mathcal{X}]\mathbb{E}[\mathcal{Y}]$

Thm $\mathcal{X}_k$ alle unabhängig mit $\mathbb{E}[\mathcal{X}_k]$ endlich. Dann gilt $\mathbb{E}[\prod_{k=1}^n \mathcal{X}_k] = \prod_{k=1}^n \mathbb{E}[\mathcal{X}_k]$

Thm $f : \mathbb{R} \rightarrow \mathbb{R}_+$ mit $\int_{-\infty}^{\infty} f(x) dx = 1$. Dann ist äquivalent:

(1) $\mathcal{X}$ stetig mit Dichte f und (2) für jede stückweise stetige Abb. $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ gilt $\mathbb{E}[\varphi(\mathcal{X})] = \int_{-\infty}^{\infty} \varphi(x) f(x) dx$

Thm äquivalent: (1) $\mathcal{X}_i$ unabhängig,

(2) $\forall \varphi_i$: $\mathbb{E}[\varphi_1(\mathcal{X}_1) \cdots \varphi_n(\mathcal{X}_n)] = \mathbb{E}[\varphi_1(\mathcal{X}_1)] \cdots \mathbb{E}[\varphi_n(\mathcal{X}_n)]$

Bem Bedingte W. $\mathbb{E}[I] = \mathbb{P}[K \cap R] \mathbb{E}[I \mid K, R] + \dots$

Dabei beeinflussen K und R den Wert nur durch Bedingte W.

4.5 Ungleichungen

Thm Markow $\mathcal{X}$ n.-neg., $g : \mathcal{X}(\Omega) \rightarrow [0, \infty)$. $\forall c \in \mathbb{R}$ mit $g(c) > 0$ gilt $\mathbb{P}[\mathcal{X} \geq c] \leq \frac{\mathbb{E}[g(\mathcal{X})]}{g(c)}$

Thm Jensensche $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ konvex, und falls $\mathbb{E}[\varphi(\mathcal{X})]$ und $\mathbb{E}[\mathcal{X}]$ wohldefiniert: $\varphi(\mathbb{E}[\mathcal{X}]) \leq \mathbb{E}[\varphi(\mathcal{X})]$

Thm Dreieck $\varphi(x) = |x|$, dann $|\mathbb{E}[\mathcal{X}]| \leq \mathbb{E}[|\mathcal{X}|]$. $\varphi(x) = x^2$, dann $\mathbb{E}[|\mathcal{X}|] \leq \sqrt{\mathbb{E}[\mathcal{X}^2]}$

Kor Cheb. $\mathbb{V}[\mathcal{Y}]$ end. $\forall c > 0$ gilt: $\mathbb{P}[|\mathcal{Y} - \mathbb{E}[\mathcal{Y}]| \geq c] \leq \frac{\mathbb{V}[\mathcal{Y}]}{c^2}$

4.6 Varianz

Def $\mathcal{X}$ mit $\mathbb{E}[\mathcal{X}^2] < \infty$, $\mathbb{V}[\mathcal{X}] = \mathbb{E}[(\mathcal{X} - \mathbb{E}[\mathcal{X}])^2]$

Def Standardabweichung $\sigma(\mathcal{X}) = \sqrt{\mathbb{V}[\mathcal{X}]}$

Bem $\mathbb{V}[\mathcal{X}] = \mathbb{E}[\mathcal{X}^2] - \mathbb{E}[\mathcal{X}]^2$ ($\mathbb{V}[\mathcal{X}] \geq 0$ ist wahr)

Bsp $\mathcal{X}$ determ. Z.V (= konst) mit Wert a , also $\mathcal{X} = a1_{\Omega}$.

Dann: $\mathbb{E}[\mathcal{X}] = a\mathbb{E}[1_{\Omega}] = a\mathbb{P}[\Omega] = a$ und $\mathbb{V}[\mathcal{X}] = \mathbb{E}[\mathcal{X}^2] - \mathbb{E}[\mathcal{X}]^2 = a^2\mathbb{E}[1_{\Omega}] - a^2 = 0$

Bem $\mathbb{E}[\mathcal{X}] < \infty$, dann $\mathbb{V}[\mathcal{X}] \geq 0$ mit = g.d.w. $\mathcal{X}$ konst; zudem $\mathbb{V}[a\mathcal{X}] = a^2\mathbb{V}[\mathcal{X}]$ und $\mathbb{V}[\mathcal{X} + a] = \mathbb{V}[\mathcal{X}]$

Prop $\mathcal{X}_k$ paarw. unabh. $\mathbb{V}[\sum_{k=1}^n \mathcal{X}_k] = \sum_{k=1}^n \mathbb{V}[\mathcal{X}_k]$. Falls $\mathcal{X}_k$ abhängig, dann gilt $\neq$. Falls **mermals** selbe Z.V. (e.g. $\mathcal{X} - \mathcal{Y} - \mathcal{Y}$, ist $\mathcal{X} - 2\mathcal{Y}$, dann $\mathbb{V} = \mathbb{V}[\mathcal{X}] + 2^2\mathbb{V}[\mathcal{Y}]$)

Bem Für Varianz von bspw. $Y = \sum_{i=1}^n \mathcal{X}_i$, $\mathcal{X}_i$ i.i.d: Unabh. nutzen (s. oben), dann ist Berechnung einfach

Bsp Varianz von bekannten Verteilungen

- $\mathcal{X} \sim \text{Ber}(p)$, $\mathbb{V}[\mathcal{X}] = p(1-p)$
- $\mathcal{X} \sim \text{Bin}(n, p)$, $\mathbb{V}[\mathcal{X}] = np(1-p)$
- $\mathcal{X} \sim \text{NBin}(r, p)$, $\mathbb{V}[\mathcal{X}] = \frac{r(1-p)}{p^2}$
- $\mathcal{X} \sim \text{Geom}(p)$, $\mathbb{V}[\mathcal{X}] = \frac{1-p}{p^2}$
- $\mathcal{X} \sim \text{H}(n, r, m)$, $\mathbb{V}[\mathcal{X}] = m \frac{rn-r^2}{n^2} \frac{n-m}{n-1}$
- $\mathcal{X} \sim \text{Poisson}(\lambda)$, $\mathbb{V}[\mathcal{X}] = \lambda = \mathbb{E}[\mathcal{X}]$
- $\mathcal{X} \sim \mathcal{U}([a, b])$, $\mathbb{V}[\mathcal{X}] = \frac{(b-a)^2}{12}$
- $\mathcal{X} \sim \text{Exp}(\lambda)$, $\mathbb{V}[\mathcal{X}] = \frac{1}{\lambda^2}$
- $\mathcal{X} \sim \mathcal{N}(\mu, \sigma^2)$, $\mathbb{V}[\mathcal{X}] = \sigma^2$
- $\mathcal{X} \sim \text{Cauchy}(x_0, \gamma)$: Existiert nicht

4.7 Kovarianz

Def $\text{cov}(\mathcal{X}, \mathcal{Y}) = \mathbb{E}[(\mathcal{X} - \mathbb{E}[\mathcal{X}])(\mathcal{Y} - \mathbb{E}[\mathcal{Y}])]$

Bem $\text{cov}(\mathcal{X}, \mathcal{Y}) = \mathbb{E}[\mathcal{X}\mathcal{Y}] - \mathbb{E}[\mathcal{X}]\mathbb{E}[\mathcal{Y}]$

Bem $\text{cov}(\mathcal{X}, \mathcal{X}) = \mathbb{V}[\mathcal{X}]$

Bem $\text{cov}(\mathcal{X}, \mathcal{Y}) = 0 \implies \mathcal{X}, \mathcal{Y}$ unabh. ($\Leftarrow$ impl. falsch!)

Bem $\mathcal{X}, \mathcal{Y}$ unabh. $\iff \forall \varphi, \psi$ stückweise stetig, beschränkt gilt $\text{cov}(\varphi(\mathcal{X}), \psi(\mathcal{Y})) = 0$

Bem Folgende Terminologie (neg. korr. = antikorreliert):

- Wenn $\text{cov}(\mathcal{X}, \mathcal{Y}) > 0$, dann: $\mathcal{X}, \mathcal{Y}$ **positiv korreliert**
- Wenn $\text{cov}(\mathcal{X}, \mathcal{Y}) = 0$, dann: $\mathcal{X}, \mathcal{Y}$ **unkorreliert**
- Wenn $\text{cov}(\mathcal{X}, \mathcal{Y}) < 0$, dann: $\mathcal{X}, \mathcal{Y}$ **negativ korreliert**

Bsp $\mathcal{X}, \mathcal{Y}$ unkorreliert $\not\Rightarrow \mathcal{X}, \mathcal{Y}$ unabhängig

aber $\mathcal{X}, \mathcal{Y}$ unabh. $\implies \mathcal{X}, \mathcal{Y}$ unkorreliert

Bem Eigenschaften der Kovarianz (alle $a, \dots \in \mathbb{R}$):

- **Positive Semidefinitheit**: $\text{cov}(\mathcal{X}, \mathcal{X}) \geq 0$
- **Symmetrie**: $\text{cov}(\mathcal{X}, \mathcal{Y}) = \text{cov}(\mathcal{Y}, \mathcal{X})$
- **Bilin.**: $\text{cov}(a\mathcal{X} + b, c\mathcal{Y} + d) = \text{accov}(\mathcal{X}, \mathcal{Y})$ und $\text{cov}(\mathcal{X}, (e\mathcal{Y} + f) + (g\mathcal{Z} + h)) = e\text{cov}(\mathcal{X}, \mathcal{Y}) + g\text{cov}(\mathcal{X}, \mathcal{Z})$

Bem $\mathbb{V}\left[\sum_{k=1}^n \mathcal{X}_k\right] = \sum_{k=1}^n \mathbb{V}[\mathcal{X}]_k + 2 \sum_{k=1}^{n-1} \sum_{l=k+1}^n \text{cov}(\mathcal{X}_k, \mathcal{X}_l)$

Bem In Matrix-Notation für $\mathcal{X} = (\mathcal{X}_1, \dots, \mathcal{X}_n)^\top$

$$\Sigma = \begin{pmatrix} \mathbb{V}[\mathcal{X}_1] & \text{cov}(\mathcal{X}_1, \mathcal{X}_2) & \dots & \text{cov}(\mathcal{X}_1, \mathcal{X}_n) \\ \text{cov}(\mathcal{X}_2, \mathcal{X}_1) & \mathbb{V}[\mathcal{X}_2] & \dots & \text{cov}(\mathcal{X}_2, \mathcal{X}_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(\mathcal{X}_n, \mathcal{X}_1) & \text{cov}(\mathcal{X}_n, \mathcal{X}_2) & \dots & \mathbb{V}[\mathcal{X}_n] \end{pmatrix}$$

Für 2 Z.V (bsp): $\mathbb{V}[\mathcal{X} - \mathcal{Y}] = \mathbb{V}[\mathcal{X}] + \mathbb{V}[\mathcal{Y}] - \text{cov}(\mathcal{X}, \mathcal{Y})$

5 Gemeinsame Verteilungen

5.1 Gemeinsame Diskrete Verteilung

Def $\mathcal{X}_k$ mit Mengen $W_k \subseteq \mathbb{R}$ endlich oder abzählbar mit $\mathcal{X}_k \in W_k$ fast sicher. Die gemeinsame Verteilung (*Joint probability distribution*) von $(\mathcal{X}_1, \dots, \mathcal{X}_n)$ ist Familie von W.:

$$\{p(x_1, \dots, x_n)\}_{x_1 \in W_1, \dots, x_n \in W_n}$$

mit $p: \mathbb{R}^n \rightarrow [0, 1]$ die gemeinsame Gewichtsfunction (joint probability mass function) mit

$$p(x_1, \dots, x_n) = \mathbb{P}[\mathcal{X}_1 = x_1, \dots, \mathcal{X}_n = x_n]$$

Thm Gemeinsame Verteilung von $\mathcal{X}_i$ erfüllt stets

$$\sum_{x_1 \in W_1, \dots, x_n \in W_n} p(x_1, \dots, x_n) = 1$$

Bem Umgekehrt: Für endliche oder abzählbare W_i und Funktion $p: W_1 \times \dots \times W_n \rightarrow [0, 1]$, die obiges erfüllen, gibt es W-Raum mit Verteilung p

Prop Aus p Verteilungsfunktion (analog zu ZH $F_{\mathcal{X}}$ & $p_{\mathcal{X}}$):

$$\begin{aligned} F(x_1, \dots, x_n) &= \mathbb{P}[\mathcal{X}_1 \leq x_1, \dots, \mathcal{X}_n \leq x_n] \\ &= \sum_{y_1 \leq x_1, \dots, y_n \leq x_n} p(y_1, \dots, y_n) \end{aligned}$$

Thm Verteilung Bild Sei $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathcal{X}_i$ disk. Z.V. mit Werten jeweils f.s. in $W_1, \dots, W_n$. Dann $\mathcal{Z} = \varphi(\mathcal{X}_1, \dots, \mathcal{X}_n)$ disk. Z.V. mit Werten f.s. in $W = \varphi(W_1 \times \dots \times W_n)$. Verteilung von $\mathcal{Z}$ dann gegeben durch ($\forall z \in W$):

$$\mathbb{P}[\mathcal{Z} = z] = \sum_{\substack{x_1 \in W_1, \dots, x_n \in W_n \\ \varphi(x_1, \dots, x_n) = z}} \mathbb{P}[\mathcal{X}_1 = x_1, \dots, \mathcal{X}_n = x_n]$$

Thm Randverteilung $\mathcal{X}_i$ disk. Z.V. mit gem. p .

$\forall k \in \{1, \dots, n\}$ und $\forall x \in W_k$ gilt:

$$\mathbb{P}[\mathcal{X}_k = x] = \sum_{\substack{x_l \in W_l \\ l \in \{1, \dots, n\} \setminus \{k\}}} p(x_1, \dots, x_{k-1}, x, x_{k+1}, \dots, x_n)$$

Also: Elimination der anderen Variable(n) in p

Bem Verteilungsf. d. k -ten Randv. $F_{\mathcal{X}_k}(x) = \mathbb{P}[\mathcal{X}_k \leq x]$

$$F_{\mathcal{X}_k}(x) = \lim_{x_l \rightarrow \infty} F(x_1, \dots, x_{k-1}, x, x_{k+1}, \dots, x_n)$$

$$l \in \{1, \dots, n\} \setminus \{k\}$$

Thm Erwartungswert Bild (Solange Σ wohldef, summieren über $x_1 \in W_1, \dots, x_n \in W_n$), siehe rechts für details

$$\mathbb{E}[\varphi(\mathcal{X}_1, \dots, \mathcal{X}_n)] = \sum_{x_1, \dots, x_n} \varphi(x_1, \dots, x_n) p(x_1, \dots, x_n)$$

Thm Folgende Aussagen sind äquivalent (für $\mathcal{X}_i$ mit Verteilung $\{p(x_1, \dots, x_n)\}_{x_1 \in W_1, \dots, x_n \in W_n}$):

- $\mathcal{X}_1, \dots, \mathcal{X}_n$ sind unabhängig
- $\forall x_1 \in W_1, \dots, x_n \in W_n$ gilt (prod. Randv. von $\mathcal{X}_i$) $p(x_1, \dots, x_n) = \mathbb{P}[\mathcal{X}_1 = x_1] \cdot \dots \cdot \mathbb{P}[\mathcal{X}_n = x_n]$

5.2 Gemeinsame Stetige Verteilung

Def $\mathcal{X}_i$ haben gemeinsame stetige Dichte wenn Abbildung $f: \mathbb{R}^n \rightarrow \mathbb{R}_+$ existiert, so dass $\forall a_i \in \mathbb{R}$ gilt:

$$\mathbb{P}[\mathcal{X}_1 \leq a_1, \dots, \mathcal{X}_n \leq a_n] = \int_{-\infty}^{a_1} \dots \int_{-\infty}^{a_n} f(x_1, \dots) \, dx_n \dots dx_1$$

Thm f gem. Dichte

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots) \, dx_n \dots dx_1 = 1$$

Jeder Funk. f die obiges erfüllt ein W-Raum $(\Omega, \mathcal{F}, \mathbb{P})$ und n -Z.V. $\mathcal{X}_i$ zugeo. werden, s.d. f die gem. Dichte von $\mathcal{X}_i$ ist.

Intuition: $f(x_1, \dots) \, dx_1, \dots$ beschreibt die W., dass ein zufälliger Punkt $(\mathcal{X}_1, \dots)$ in $[x_1, x_1 + dx_1] \times \dots$ liegt

Bsp Gleichverteilungen:

- **Einheitsquadrat** $f(x, y) = \begin{cases} 1 & (x, y) \in [0, 1]^2 \\ 0 & \text{sonst} \end{cases}$
 - **Einheitskreisscheibe** $D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$
- $$f(x, y) = \begin{cases} \frac{1}{\pi} & (x, y) \in D \\ 0 & \text{sonst} \end{cases}$$

Thm Erwartungswert Bild

$$\mathbb{E}[\varphi(\mathcal{X}_1, \dots)] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \varphi(x_1, \dots) f(x_1, \dots) \, dx_n \dots$$

φ ist bspw. für $\mathbb{E}[\mathcal{X}\mathcal{Y}]$, $\varphi(\mathcal{X}, \mathcal{Y}) = \mathcal{X} \cdot \mathcal{Y}$, Integralgrenzen nach $f(x_1, \dots)$

Def Randverteilung Falls $\mathcal{X}, \mathcal{Y}$ gemeinsame Verteilung F haben, so ist die Verteilungsfunktion der **Randverteilung** von $\mathcal{X}$, $F_{\mathcal{X}}: \mathbb{R} \rightarrow [0, 1]$ gegeben durch:

$$x \mapsto F_{\mathcal{X}}(x) = \mathbb{P}[\mathcal{X} \leq x] = \mathbb{P}[\mathcal{X} \leq x, \mathcal{Y} < \infty] = \lim_{y \rightarrow \infty} F(x, y)$$

Analog für $\mathcal{Y}$ ist sie $F_{\mathcal{Y}}: \mathbb{R} \rightarrow [0, 1]$:

$$y \mapsto F_{\mathcal{Y}}(y) = \mathbb{P}[\mathcal{Y} \leq y] = \mathbb{P}[\mathcal{Y} < \infty, \mathcal{X} \leq y] = \lim_{x \rightarrow \infty} F(x, y)$$

Die Dichten der Randverteilungen sind:

$$f_{\mathcal{X}}(x) = \int_{-\infty}^{\infty} f(x, y) \, dy \quad f_{\mathcal{Y}}(y) = \int_{-\infty}^{\infty} f(x, y) \, dx$$

Bsp Randdichten mit Gleichverteilung

Einheitsquadrat $f(x, y) = 1_{(x, y) \in [0, 1]^2} = 1_{x \in [0, 1]} 1_{y \in [0, 1]}$

$$f_{\mathcal{X}}(x) = \int_0^1 1_{x \in [0, 1]} 1_{y \in [0, 1]} \, dy = 1_{x \in [0, 1]}$$

$$f_{\mathcal{Y}}(y) = \int_0^1 1_{x \in [0, 1]} 1_{y \in [0, 1]} \, dx = 1_{y \in [0, 1]}$$

Einheitskreisscheibe $f(x, y) = \frac{1}{\pi} \cdot 1_{x^2 + y^2 \leq 1}$

$$f_{\mathcal{X}}(x) = \int_{-\infty}^{\infty} \frac{1}{\pi} 1_{y^2 \leq 1-x^2} \, dy = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi} \, dy = \frac{2}{\pi} \sqrt{1-x^2}$$

$$f_{\mathcal{Y}}(y) = \int_{-\infty}^{\infty} \frac{1}{\pi} 1_{x^2 \leq 1-y^2} \, dx = \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \frac{1}{\pi} \, dx = \frac{2}{\pi} \sqrt{1-y^2}$$

Thm Unabhängigkeit $\mathcal{X}_i$ unabhängig g.d.w.

$$f_{\mathcal{X}_1, \dots, \mathcal{X}_n}(x_1, \dots, x_n) = f_{\mathcal{X}_1}(x_1) \cdot \dots \cdot f_{\mathcal{X}_n}(x_n)$$

Bsp Unabhängigkeit bei Gleichverteilungen

Einheitsquadrat Koordinaten sind unabhängig

Einheitskreisscheibe Koordinaten nicht unabhängig

5.3 Tips & Tricks

Bem Die **Dichten** können anstelle der Variablen in bspw. der Formel für bedingte W. verwendet werden! Dies ist besonders nützlich für die Berechnung von bedingten Dichten, etc. und Bestimmung Unabhängigkeit

How To: Randdichten berechnen

1. Intervalle zur Integration für Randdichten via Indikatorvariable 1_{cond} bestimmen. 1_x verschwindet bei Bestimmung von f_y in den Integral-grenzen
2. $cond$ nach entsprechender Var. auflösen, wenn nötig
3. Bei Integration Indikatorvariable als konstante betrachten (muss auch integriert werden!)
4. Übrig bleibende Indikatorvariablen wieder anfügen

Die Randdichten sind dann die Dichten von $\mathcal{X}$, etc.; Für weitere Berechnungen verwendbar, wie bspw. $\mathbb{E}[\mathcal{X} + \mathcal{Y}]$.

How To: $cond$ auflösen

In dem Integral sollte jeweils der äussere Integral der sein, für welchen wir volle Information haben. Zudem, bspw., für $0 \leq x \leq y^2 \leq 1$, wissen wir zwar für x und y die Grenzen, wählen aber x als inneren Integral, weil sich oft $\int_0^{y^2}$ schöner auflöst. Der äussere Integral ist dann auf dem ganzen Intervall (also äussere Beschränkungen).

Für einen einzelnen Integral wählen wir die Grenzen als die umliegenden Grenzen. Falls wir nur eine haben so müssen wir auflösen oder eine Grenze hinzufügen, siehe Randdichten von bspw. Einheitskreisscheibe.

How To: Gemeinsame Dichten finden

Nach Definition muss f die Ableitung von F sein. Oder umgekehrt, $\int f \, d\dots$ sollte F ergeben. Beispiel: $T \sim \mathcal{U}[1, 2]$, $\mathcal{X} \sim \text{Exp}(\lambda)$, $\mathcal{Y} \sim \text{Exp}(\gamma)$. Dann $f_{T, \mathcal{X}, \mathcal{Y}}(t, x, y) = \begin{cases} \lambda \gamma \exp(-\lambda x - \gamma y) & \text{falls } 1 \leq t \leq 2 \wedge x, y \geq 0 \\ 0 & \text{sonst} \end{cases}$

How To: Verteilungen / Dichten finden

Wir suchen $\mathbb{P}[\mathcal{Z} \leq a]$ (ausser für gem. V, dann s. oben). Dann überlegen, wie $\mathcal{Z} \leq a$ erfüllt werden kann. Beispiel: $\mathcal{Z} = \min(\mathcal{X}, \mathcal{Y})$, dann gesucht: $\mathbb{P}[\mathcal{X} \leq a \vee \mathcal{Y} \leq a]$ oder $1 - \mathbb{P}[\mathcal{Z} > a]$. Schliesslich F_Z aufstellen mit Bedingungen oder Indikatorvariablen.

Zudem kann Verteilung der einzelnen Variablen (und v.a. Unabhängigkeit) nützlich werden.

Falls Dichte, dann Verteilung ableiten.

How To: Multi-Variable Wahrscheinlichkeiten

Für W. wie $\mathbb{P}[\mathcal{X} \leq \mathcal{Y}^3]$ übersetzen wir dies in Integral $\int_B f(x, y) \, dx \, dy$ mit $B = \{(x, y) \mid x \leq y^3\}$, dann für $y \in [0, 1]$ gilt: $\int_0^1 \int_0^{y^3} f(x, y) \, dx \, dy$, wobei hier die untere Grenze durch gegebenes $x \geq 0$ entsteht.

Für unbekannte gemeinsame Dichte, die Eigenschaften ausfächern, Unabhängigkeit, etc ausnutzen.

Für W. im Exponent, dies beschreibt typ. Repetitionen, also über die Anzahl und W. nachdenken. Beispielsweise für $\mathbb{E}[(\frac{5}{6})^{\mathcal{X}-2} = k]$ berechnen wir $\mathbb{P}[\mathcal{X} = k] = (\frac{1}{6})^{k-1} \cdot \frac{5}{6}$, dann ist $\mathbb{E}[\dots] = \sum_{k=1}^{\infty} (\frac{5}{6})^{k-2} \cdot (\frac{1}{6})^{k-1} \cdot \frac{5}{6} = \sum_{k=1}^n (\frac{5}{36})^{k-1}$

6 Das Gesetz der grossen Zahlen

Def $\bar{\mathcal{X}}_n = \frac{1}{n} \sum_{k=1}^n \mathcal{X}_k$ ist das **arithmetische Mittel** der $\mathcal{X}_k$.

Begriff In Zusammenhang mit Zufallsvariablen auch **Sichprobenmittel** genannt.

Realisierung wird **empirischer Mittelwert** genannt

Bem Vereinfacht: Beide besagen, dass $\lim_{n \rightarrow \infty} \bar{\mathcal{X}}_n \approx \mathbb{E}[\mathcal{X}_k]$

Bem Erwartungswert und Varianz mit lin, etc berechnen

6.1 Schwaches Gesetz der grossen Zahlen

Thm Schwaches GGZ Sei $K = \{1, 2, \dots\}$ und $\forall k \in K : \mathcal{X}_k$ unabh. Z.V. mit $\mathbb{E}[\mathcal{X}_k] = \mu$; $\mathbb{V}[\mathcal{X}_k] = \sigma^2$:

$$\bar{\mathcal{X}}_n = \frac{1}{n} S_n = \frac{1}{n} \sum_{k=1}^n \mathcal{X}_k$$

Dann konvergiert $\bar{\mathcal{X}}_n$ für $n \rightarrow \infty$ in Wahrscheinlichkeit gegen $\mu = \mathbb{E}[\mathcal{X}_k]$, also $\forall \varepsilon > 0$ gilt $\mathbb{P}[|\bar{\mathcal{X}}_n - \mu| > \varepsilon] \xrightarrow{n \rightarrow \infty} 0$

6.2 Starkes Gesetz der grossen Zahlen

Thm Starkes GGZ Für $\mathcal{X}_1, \dots$ mit $\mathcal{X}_k$ unabhängig mit $\mathbb{E}[\mathcal{X}_k]$ endlich. Für

$$\bar{\mathcal{X}}_n = \frac{1}{n} S_n = \frac{1}{n} \sum_{k=1}^n \mathcal{X}_k$$

gilt $\bar{\mathcal{X}}_n \xrightarrow{n \rightarrow \infty} \mu$ $\mathbb{P}$ -fast sicher, also $\mathbb{P}\left[\left\{\omega \in \Omega \mid \bar{\mathcal{X}}_n(\omega) \xrightarrow{n \rightarrow \infty} \mu\right\}\right] = 1$

6.3 Zentraler Grenzwertsatz

Def Konvergenz in Verteilung $(\mathcal{X}_n)_{n \in \mathbb{N}}$, $\mathcal{X}$ mit V.F.

$(F_n)_{n \in \mathbb{N}}$, F . $(\mathcal{X}_n)_{n \in \mathbb{N}}$ **konvergiert in V.** gegen $\mathcal{X}$ ($\mathcal{X}_n \xrightarrow{d} \mathcal{X}$ für $n \rightarrow \infty$), falls $\forall$ Stetigkeitsp. $x \in \mathbb{R}$ von F gilt:

$$\lim_{n \rightarrow \infty} F_n(x) = \lim_{n \rightarrow \infty} \mathbb{P}[\mathcal{X}_n \leq x] = \mathbb{P}[\mathcal{X} \leq x] = F(x)$$

Thm Zentraler Grenzwertsatz (ZGS)

i.i.d = u.i.v in DE, siehe 2.3, $\mathcal{X}_k$ unten kann jegliche Verteilung sein

$\mathcal{X}_k$ i.i.d mit $\mathbb{E}[\mathcal{X}_k] = \mu$, $\mathbb{V}[\mathcal{X}_k] = \sigma^2$. Für Partialsummen $S_n = \sum_{k=1}^n \mathcal{X}_k$ gilt $\forall x \in \mathbb{R}$ (mit Φ V.F. von Std.-Norm.-V):

$$\lim_{n \rightarrow \infty} \mathbb{P}\left[\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right] = \Phi(x)$$

Bem $\mathbb{E}[S_n] = n\mu$, $\mathbb{V}[S_n] = n\sigma^2$; $S_n^* = \frac{S_n - n\mu}{\sigma\sqrt{n}} \overset{\text{approx}}{\sim} \mathcal{N}(0, 1)$

für grosse n , mit $\overset{\text{approx}}{\sim}$ gespr. "approximativ verteilt gemäss".

Also ist für $\mathbb{E}[S_n^*] = 0$ und $\mathbb{V}[S_n^*] = 1$.

Also: $S_n \overset{\text{approx}}{\sim} \mathcal{N}(n\mu, n\sigma^2)$, bzw. $\bar{\mathcal{X}}_n \overset{\text{approx}}{\sim} \mathcal{N}(\mu, \frac{1}{n}\sigma^2)$

Bem Für $S_n \sim \text{Bin}(n, p)$ ist $S_n^* \overset{\text{approx}}{\sim} \mathcal{N}(np, np(1-p))$

und $\mathbb{P}[a < S_n \leq b] \approx \Phi\left(\frac{b + \frac{1}{2} - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a + \frac{1}{2} - np}{\sqrt{np(1-p)}}\right)$

Bem Für Bestimmung von $\mathbb{P}[S_n \leq y]$, ZGS direkt verwenden, dann die Ungleichung in gewünschte Form umformen und $\mathbb{E}[\mathcal{X}_i]$ und $\mathbb{V}[\mathcal{X}_i]$ berechnen und einsetzen.

Bem Im Allgemeinen für ZGS, gegebenen Ausdruck umformen, und allfällige Werte bestimmen. **WICHTIG**: Muss genau in dieser Form sein.

Bsp $\lim_{n \rightarrow \infty} \mathbb{P}\left[\frac{S_n}{\sqrt{n}} > 1\right]$ mit S_n wie oben, $\mathbb{E}[S_n] = 0$, dann berechnen $1 - \mathbb{P}\left[\frac{S_n}{\sqrt{n}\sigma} \leq \frac{1}{\sigma}\right]$, durch Umformung der Ungl.

6.4 Chernoff-Schranken

Def Momenterzeugende Funktion von $\mathcal{X}$ ist für $t \in \mathbb{R}$ $M_{\mathcal{X}}(t) = \mathbb{E}[e^{t\mathcal{X}}]$

Bsp $\mathcal{X} \sim \text{Ber}(p)$, dann $M_{\mathcal{X}}(t) = 1 - p + pe^t$;

$\mathcal{X} \sim \text{Bin}(n, p)$, dann $M_{\mathcal{X}}(t) = (1 - p + pe^t)^n$

Thm Chernoff-Ungleichung $\mathcal{X}_k$ i.i.d. Z.V. mit jeweils $\forall t \in \mathbb{R}$ $M_{\mathcal{X}}(t)$ endl.; $\forall b \in \mathbb{R}$

$$\mathbb{P}[S_n \geq b] \leq \exp\left(\inf_{t \in \mathbb{R}} (n \log(M_{\mathcal{X}}(t) - tb))\right)$$

Thm Chernoff-Schranke $\mathcal{X}_k \sim \text{Ber}(p_k)$ unabhängig; $S_n = \sum_{k=1}^n \mathcal{X}_k$; $\mu_n = \mathbb{E}[S_n] = \sum_{k=1}^n p_k$ und $\delta > 0$. Dann:

$$\mathbb{P}[S_n \geq (1 + \delta)\mu_n] \leq \left(\frac{e^\delta}{(1 + \delta)^{1 + \delta}}\right)$$

Dabei ist δ der Faktor der Abweichung

7 Schätzer

7.1 Statistische Grundlagen

Begriff Stichprobe Die Gesamtheit der Beobachtungen $x_1, \dots, x_n$ oder Z.V. $\mathcal{X}_1, \dots, \mathcal{X}_n$. n ist **Stichprobenumfang**

Def Parameterraum Ein Datensatz $x_1, \dots, x_n$ aus einer Stichprobe $\mathcal{X}_1, \dots, \mathcal{X}_n$ kann durch (evtl. hochdim.) **Param.** ϑ aus **Param.-Raum** Θ dargestellt werden. ϑ ist dabei oft die Param der Verteilung (bspw. $\vartheta = (\mu, \sigma^2)$ bei Normalv.). Wir betrachten jeweils **Familie von W-Räumen**, $(\Omega, \mathcal{F}, \mathbb{P}_\vartheta)_{\vartheta \in \Theta}$.

Wir haben normalen Grundraum plus W-Mass $\mathbb{P}_\vartheta$

7.2 Schätzer

Def Schätzer ist eine Z.V. der Form $T_l = t_l(\mathcal{X}_1, \dots, \mathcal{X}_n)$, mit zu findender **Schätzfunktion** $t_l: \mathbb{R}^n \rightarrow \mathbb{R}$.

Einsetzen von Daten liefert dann **Schätzwerte**. Oft schreiben wir $T = (T_1, \dots, T_m)$ und $\vartheta = (\vartheta_1, \dots, \vartheta_m)$

Bem Schätzwert ist Zahl $(T_l(\omega))$, **Schätzer** ist Z.V. (T_l)

Bsp Simple Schätzer (beide Erwartungstreu)

Letztes Ergebnis: $T = X_n$ mit X_n letztes Ergebnis

Durchschnitt: $T = \bar{X}_n = \frac{1}{n} S_n$ ($\frac{1}{n} \sum_{k=1}^n x_k$ für Daten x_i)

Def Erwartungstreue Schätzer T unbiased für ϑ , falls $\forall \vartheta \in \Theta$ gilt $\mathbb{E}_\vartheta[T] = \vartheta$. (T schätzt im Mittel richtig)

Def Bias Erwarteter Schätzf. von T ($\mathbb{P}_\vartheta$) ist $\mathbb{E}_\vartheta[T] - \vartheta$

Def MSE Mittlerer quadratischer Schätzfehler (mean squared err.) von T mit $\mathbb{P}_\vartheta$ ist $\text{MSE}_\vartheta[T] = \mathbb{E}_\vartheta[(T - \vartheta)^2]$

Bem $\text{MSE}_\vartheta[T] = \mathbb{V}_\vartheta[T] + (\mathbb{E}_\vartheta[T] - \vartheta)^2$

Def Folge von Schätzern $T^{(n)}$, $n \in \mathbb{N}$ ist **konsistent** für ϑ , falls $T^{(n)}$ für $n \rightarrow \infty$ in $\mathbb{P}_\vartheta$ -W. gegen ϑ konvergiert, also $\forall \vartheta \in \Theta, \varepsilon > 0: \lim_{n \rightarrow \infty} \mathbb{P}_\vartheta[|T^{(n)} - \vartheta| > \varepsilon] = 0$.

Beweise oft mit GGZ (s. 6.1), also T umformen, s.d. anwendbar (e.g. $T^{(n)} \div n$)

7.3 Maximum-Likelihood-Method

Funktion $p_{\mathcal{X}}(x_1, \dots, x_n; \vartheta) = \mathbb{P}_\vartheta[\mathcal{X}_1 = x_1, \dots]$ (gleich wie normale Gew.-F.)

In Modell $\mathbb{P}_\vartheta$ sind $\vec{\mathcal{X}} = (\mathcal{X}_1, \dots, \mathcal{X}_n)$ entweder

- **diskret** (gem. Gew. $p_{\vec{\mathcal{X}}}(x_1, \dots, x_n; \vartheta)$)
- **stetig** (gem. Dichte $f_{\vec{\mathcal{X}}}(x_1, \dots, x_n; \vartheta)$)

Da oft X_k i.i.d mit individueller Gew. $p_{\mathcal{X}}(x; \vartheta)$ bzw. Dichte $f_{\mathcal{X}}(x; \vartheta)$ sind, also (mit g ersetzt durch p oder f)

$$g_{\vec{\mathcal{X}}}(x_1, \dots, x_n; \vartheta) = \prod_{k=1}^n g_{\mathcal{X}}(x_k; \vartheta)$$

Def Likelihood-Funktion

$$L(x_1, \dots, x_n; \vartheta) = \begin{cases} p_{\vec{\mathcal{X}}}(x_1, \dots, x_n; \vartheta) & \text{im diskreten Fall} \\ f_{\vec{\mathcal{X}}}(x_1, \dots, x_n; \vartheta) & \text{im stetigen Fall} \end{cases}$$

log-Likelihood-Funktion $l(\vartheta) = \log(L(x_1, \dots, x_n; \vartheta))$. Ist im i.i.d-Fall eine Summe. **Immer in p durch x_k ersetzen!!**

Intuition: Likelihood-F. gibt W., dass in $\mathbb{P}_\vartheta$ die Stichprobe $\mathcal{X}_1, \dots, \mathcal{X}_n$ genau Werte $x_1, \dots, x_n$ liefert.

Def ML-Schätzer T_{ML} für ϑ maximiert die Funktion $\vartheta \mapsto L(\mathcal{X}_1, \dots, \mathcal{X}_n; \vartheta)$ für alle ϑ , also

$$T_{ML} = t_{TM}(\mathcal{X}_1, \dots, \mathcal{X}_n) \in \underset{\vartheta \in \Theta}{\operatorname{argmax}} L(\mathcal{X}_1, \dots, \mathcal{X}_n; \vartheta)$$

Wir suchen besten t_{TM} , durch geschickte Wahl von ϑ

Meistens sind $\mathcal{X}_k$ i.i.d. unter $\mathbb{P}_\vartheta$, dann L produkt, also besser $\log(L)$ maximieren (da Summe).

Def Momentschätzer (MM) $T_{MM}^k = \frac{1}{n} \sum_{i=1}^n X_i^k$, und die Lösung ist definiert als $T_{MM}^k = \mathbb{E}[\mathcal{X}^k]$. Für die Normalverteilung wird das Erste und Zweite Moment geschätzt.

How To: Finden von ML-Schätzern

1. log-LH-Funk. bestimmen (x_i statt $\mathcal{X}_i$ verwenden)
2. Nach ϑ ableiten (oder wodurch ϑ ersetzt wurde)
3. Nullstelle davon (für Max (via e.g. $f'' < 0$))
4. ML-Schätzer ist dann Nullstelle von Ableitung (mit x_i wieder durch $\mathcal{X}_i$ ersetzt)

Bsp ML-Schätzer von Verteilungen

n hier wenn nicht anders definiert, die Anzahl an Z.V. als arg von Dichte / Gewicht

Bernoulli $\mathcal{X}_i \sim \text{Ber}(p)$ i.i.d, hier $\vartheta = p$. Dabei: $p_{\mathcal{X}}(x; \vartheta) = \mathbb{P}_\vartheta[\mathcal{X} = x] = \vartheta^x (1 - \vartheta)^{1-x}$ mit $x \in \{0, 1\}$, $S = \sum_{k=1}^n x_k$
 $L(x_1, \dots, x_n; \vartheta) = \vartheta^S (1 - \vartheta)^{n-S}$

Log LH-Func: $l(\vartheta) = \log(\vartheta) \cdot S + \log(1 - \vartheta) \cdot (n - S)$

ML-Schätzer: $T = \frac{1}{n} \sum_{k=1}^n \mathcal{X}_k = \bar{X}_n$

Geometrisch dabei $\vartheta = p$ und **Log LH-Func**

$$l(\vartheta) = n \ln(\vartheta) + (S - n) \ln(1 - \vartheta) \quad S \text{ wie bei Ber}$$

ML-Schätzer $T = \frac{n}{S}$

Poisson dabei $\vartheta = \lambda$ und **Log LH-Func**

$$l(\vartheta) = -n\vartheta\lambda + \log(\lambda) \sum_{i=1}^n x_i - \sum_{i=1}^n \log(x_i!)$$

ML-Schätzer ist $T = \frac{1}{n} \sum_{i=1}^n X_i$

Normalverteilung

$\mathcal{X}_i \sim \mathcal{N}(\mu, \sigma^2)$ i.i.d. und $\vartheta = (\mu, \sigma^2) = (\mu, v)$. **Likelihood-Funktion** wie üblich und somit ist **Log LH-Func:**

$$l(\vartheta) = -\frac{1}{2} n (\log(2\pi) + \log(v)) - \sum_{k=1}^n \frac{(x_k - \mu)^2}{2v}$$

$$\partial_\mu l = 2 \sum_{k=1}^n \frac{x_k - \mu}{2v} \quad \partial_v l = -\frac{n}{2v} + 2v^{-2} \sum_{k=1}^n (x_k - \mu)^2$$

Der **Momentschätzer** ist $T = (T_1, T_2) = (E_\vartheta[\mathcal{X}], \mathbb{V}_\vartheta[\mathcal{X}])$:

$$T_1 = \bar{X}_n \quad T_2 = \frac{1}{n} \sum_{k=1}^n (\mathcal{X}_k - \bar{X}_n)^2 = \frac{1}{n} \sum_{k=1}^n \mathcal{X}_k^2 - (\bar{X}_n)^2$$

Erwartungstreuer Schätzer für $(\mathbb{E}_\vartheta[\mathcal{X}], \mathbb{V}_\vartheta[\mathcal{X}])$ (s. 7.2):

$$T'_1 = T_1 \quad T'_2 = \frac{n}{n-1} T_2$$

T'_2 ist die (korrigierte) **empirische (Stichproben)varianz** ((un)biased sample variance). Bei höherdimensionalen Parametern suchen wir den zugehörigen Nullvektor.

7.4 Verteilungsaussagen

Approx. V. von Schätzer unter $\mathbb{P}_\vartheta$ (wenn $T = \sum_{k=1}^n \mathcal{Y}_k$, mit $\mathcal{Y}_k$ i.i.d in $\mathbb{P}_\vartheta$). Dabei $T \sim \mathcal{N}(n\mathbb{E}_\vartheta[\mathcal{Y}_k], n\mathbb{V}_\vartheta[\mathcal{Y}_k])$ unter $\mathbb{P}_\vartheta$

Def χ^2 -Verteilung $\mathcal{X}$ ist χ^2 -verteilt mit m Freiheitsgraden, falls Dichte für $x \geq 0$ (wir schreiben $\mathcal{X} \sim \chi_m^2$):

$$f_{\mathcal{X}}(x) = \frac{1}{2^{\frac{m}{2}} \Gamma(\frac{m}{2})} x^{\frac{m}{2}-1} e^{-\frac{x}{2}}$$

Bem (Eulersche) Gammafunktion $\forall x \geq 0$:

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt \quad \Gamma(n+1) = n! \text{ mit } n \in \mathbb{N}_0$$

Bem $\mathcal{X}_k \sim \mathcal{N}(0, 1)$ i.i.d: $\left(\sum_{k=1}^m \mathcal{X}_k^2\right) \sim \chi_m^2$. $\chi_2^2 = \text{Exp}(\frac{1}{2})$

Bem Allgemein: $\mathcal{X}_k \sim \mathcal{N}(\mu, \sigma^2)$, dann $\frac{n-1}{\sigma^2} S^2 \sim \chi_{n-1}^2$, mit S eine Summe wie oberhalb

Def Studentsche t-Verteilung $\mathcal{X} \sim t_m$ falls Dichte

$$f_{\mathcal{X}}(x) = \frac{\Gamma(\frac{m+1}{2})}{\sqrt{m\pi} \Gamma(\frac{m}{2})} \left(1 + \frac{x^2}{m}\right)^{-\frac{m+1}{2}}$$

Bem $\mathcal{X} \sim \mathcal{N}(0, 1)$, $\mathcal{Y} \sim \chi_m^2$ unabh., dann: $\frac{\mathcal{X}}{\sqrt{\frac{1}{m}\mathcal{Y}}} \sim t_m$

Mit $m = 1$ Cauchy-V. mit $m \rightarrow \infty$ asympt. $\mathcal{N}(0, 1)$. t_m symm. um 0 wie $\mathcal{N}(0, 1)$, aber **langschwänziger** (see 3.6.3)

Untenstehendes sehr nützlich zur Bestimmung von V.

Thm Für $\mathcal{X}_k \sim \mathcal{N}(\mu, \sigma^2)$ i.i.d. und

$$\bar{X}_n = \sum_{k=1}^n \mathcal{X}_k, \quad S^2 = \frac{1}{n-1} \sum_{k=1}^n (\mathcal{X}_k - \bar{X}_n)^2$$

1. $\bar{X}_n \sim \mathcal{N}(\mu, \frac{1}{n}\sigma^2)$, also: $\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0, 1)$ mit $q = \frac{\sigma}{\sqrt{n}}$
2. $\frac{n-1}{\sigma^2} S^2 = \frac{1}{\sigma^2} \sum_{k=1}^n (\mathcal{X}_k - \bar{X}_n)^2 \sim \chi_{n-1}^2$
3. $\bar{X}_n$ und S^2 sind unabhängig
4. $\frac{\bar{X}_n - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1} \left(= \frac{\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}}{\sqrt{\frac{1}{n-1} \frac{n-1}{\sigma^2} S^2}} \right)$

8 Tests

8.1 Grundbegriffe

Wiederum Stichprobe $\mathcal{X}_1, \dots, \mathcal{X}_n$, mögliche Modelle durch Familie $(\mathbb{P}_\vartheta)_{\vartheta \in \Theta}$ beschrieben und ϑ ein-/mehrdimensional. Wir haben **Vermutung** wo in Θ richtiges ϑ liegt.

Wir entscheiden zwischen folgenden, mit $\Theta_0 \cap \Theta_A = \emptyset$:

- **Hypothese** $\Theta_0 \subseteq \Theta$ (oft: $H_0 : \vartheta \in \Theta_0$)
- **Alternative** $\Theta_A \subseteq \Theta$ (oft: $H_A : \vartheta \in \Theta_A$)

Falls keine Alternative spezifiziert, so gilt $\Theta_A = \Theta_0^C = \Theta \setminus \Theta_0$. Sie heissen **Einfach**, falls $\Theta_0 = \{\vartheta_0\}$, resp. $\Theta_A = \{\vartheta_A\}$, sonst **Zusammengesetzt** (e.g. $\Theta_A = (0.5, 1]$)

Def **Test** ist ein Paar (T, K) mit

- **Teststatistik** $T = t(\mathcal{X}_1, \dots, \mathcal{X}_n)$ eine Zufallsvariable mit messbarer Funktion $t : \mathbb{R}^n \rightarrow \mathbb{R}$ (oft LQ)
- **Verwerfungsbereich** $K \subseteq \mathbb{R}$ (auch **kritischer Bereich**)

Entscheidungsregel H_0 wird *verworfen*, falls $T(\omega) \in K$, sonst *angenommen*, bzw. nicht verworfen

Def **Fehler 1. Art** (Hypothese abgelehnt, ist aber richtig), falls $\vartheta \in \Theta_0$, aber $T \in K$, also $\mathbb{P}_\vartheta[T \in K]$ für $\vartheta \in \Theta_0$ ist W. für F. 1. Art

Def **Fehler 2. Art** (Hypothese angenommen, aber falsch), falls $\vartheta \in \Theta_A$, aber $T \notin K$, also $\mathbb{P}_\vartheta[T \notin K] = 1 - \mathbb{P}_\vartheta[T \in K]$ für $\vartheta \in \Theta_A$ ist W. für F. 2. Art

Bem Entscheidung hängt dabei von **Realisierung** ω ab. Da T eine Z.V. ist, können $\mathbb{P}_\vartheta[T \in K]$ in jedem $\mathbb{P}_\vartheta$ betrachten

Def **Signifikanzniveau** $\alpha \in (0, 1)$. Ziel: $\sup_{\vartheta \in \Theta_0} \mathbb{P}_\vartheta[T \in K] \leq \alpha$ (Fehler 1. Art so klein wie möglich). Typisch: $\alpha = 0.05$

Def **Macht** $\beta : \Theta_A \rightarrow [0, 1]$, mit $\beta(\vartheta) = \mathbb{P}_\vartheta[T \in K]$ soll möglichst gross werden (Fehler 2. Art so klein wie möglich)

Bem Seriöser Test nutzt immer **negation der Aussage** als Hypothese (da Hypothese eher angenommen als abgelehnt wird). Entscheidung ist nicht *Beweis*, sondern *Interpr.*

Bsp **Tea Tasting Lady** Lady behauptet, sie kann schmecken, ob zuerst Tee oder Milch eingegossen wurde. Hypothese: ist Raten ($H_0 : \vartheta = 0.5$, $H_A : \vartheta > 0.5$). Unter $\mathbb{P}_\vartheta$, $\mathcal{X}_i \sim \text{Ber}(\vartheta)$ i.i.d und $S_n = \sum_{k=1}^n \mathcal{X}_k \sim \text{Bin}(n, \vartheta)$. Grosser Wert von S_n eher H_A als H_0 .

Test: $T = S_n$, $K = (c, \infty)$. Für Bestimmung von α brauchen wir $\mathbb{P}_\vartheta[S_n > c]$ für $\vartheta = 0.5$. Berechnen Verteilung von T (nicht immer möglich, wenn nicht, nutzen Approximation). Für α muss $\mathbb{P}_\vartheta[S_n > c] \leq \alpha$ sein. Wählen c so, dass α ungefähr 5% ist.

Macht ersichtlich durch Berechnung von $\beta(0.6)$ hier. Fehler 2. Art ist gross ($1 - \beta(\vartheta)$ gross)

8.2 Konstruktion von Tests

$\mathcal{X}_i$ diskret oder gemeinsam stetig unter $\mathbb{P}_{\vartheta_0}$ und $\mathbb{P}_{\vartheta_A}$. Sei $L(x_1, \dots, x_n; \vartheta)$ die Likelihood-Funktion

Def **Likelihood-Quotient** Für $\vartheta_0 \in \Theta_0$, $\vartheta_A \in \Theta_A$ und $x_i \in \mathbb{R}$: $R(x_1, \dots, x_n; \vartheta_0, \vartheta_A) = \frac{L(x_1, \dots, x_n; \vartheta_A)}{L(x_1, \dots, x_n; \vartheta_0)}$. $R(x_1, \dots; \vartheta_0, \vartheta_A) = +\infty$ wenn $L(x_1, \dots; \vartheta_0) = 0$. Wenn R gross ist, sind Beobachtungen in $\mathbb{P}_{\vartheta_A}$ deutlich wahrscheinlicher als in $\mathbb{P}_{\vartheta_0}$. **Oft verwendeter Test**

Def **Likelihood-Quotienten-Test** Sei $c \geq 0$. Der **LQT** mit param c ist Test (T, K) mit $T = R(\mathcal{X}_1, \dots, \mathcal{X}_n; \vartheta_0, \vartheta_A)$ und $K = (c, \infty)$. H_0 wird verworfen wenn R gross wird.

Thm **Neyman-Pearson** Seien Θ_0, Θ_A einfach und (T, K) ein LQT mit c und sig.-niv. $\alpha^* := \mathbb{P}_{\vartheta_0}[T \in K]$. Anderer Test (T', K') mit $\mathbb{P}_{\vartheta_0}[T' \in K'] =: \alpha \leq \alpha^*$, so gilt auch $\mathbb{P}_{\vartheta_A}[T' \in K'] \leq \mathbb{P}_{\vartheta_A}[T \in K]$.

Intuition: Jeder Test mit kleinerem Signifikanzniveau hat auch kleinere Macht

Def **Verallgemeinerte LQ** ist gegeben durch

$$R(x_1, \dots, x_n) = \frac{\sup_{\vartheta \in \Theta_A} L(x_1, \dots, x_n; \vartheta)}{\sup_{\vartheta \in \Theta_0} L(x_1, \dots, x_n; \vartheta)}$$

$$\tilde{R}(x_1, \dots, x_n) = \frac{\sup_{\vartheta \in \Theta_A \cup \Theta_0} L(x_1, \dots, x_n; \vartheta)}{\sup_{\vartheta \in \Theta_0} L(x_1, \dots, x_n; \vartheta)}$$

8.3 Beispiele

How To: Tests

Was zu tun / erwähnen ist jeweils:

1. Modell: Verteilung der Stichproben $\mathcal{X}_k$ unter $\mathbb{P}_\vartheta$, mit $\vartheta =$ gesuchte Werte. Für $\mathcal{X}_k \sim \mathcal{N}(\mu, \sigma^2)$, ist $\vartheta = (\mu, \sigma^2)$, wenn beide unbekannt.
2. Hypothesen: $H_0 : \vartheta = \vartheta_0 =$ Vergleichswert, $H_A : \vartheta < \vartheta_0$ (oder grösser, $\neq$, etc). **Umgekehrt, s. links**
3. Teststatistik: Basierend auf Verteilung von $\mathcal{X}_k$ und Bekanntheit von Parametern (bspw. mit Theoremen von 3.6.4, wie Normalisierung, und bswp. $\sum_{i=1}^n \mathcal{X}_i \sim \mathcal{N}(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2)$)
4. Verteilung von T : Oft t_{n-1} , siehe 7.4 unten, oder bei entsprechendem Test
5. Verwerfungsbereich: Nach Anleitung unten
6. Wert der T : $T(\omega)$ berechnen, ω ist beobachteter Wert (bspw. Durchschnitt $\bar{\mathcal{X}}_n$)
7. Testentscheid: Basierend auf Resultat von vorherigem und K entscheiden.

Approximative Teststatistik Normalisierung unter H_0 , also $T = \frac{x_n - \mathbb{E}[S_n]}{\sqrt{\mathbb{V}[S_n]}}$, mit x_n Realisierung, S_n wie immer

Verwerfungsbereich $K_l = (c_>, \infty)$ (links), $K_r = (-\infty, c_<)$ (rechts), $K_b = (-\infty, -c_\neq) \cup (c_\neq, \infty)$ (beidseitig)

Bsp **Gauss-Test** Voraussetzung.: **(1)** $\mathcal{X}_i$ i.i.d. $\sim \mathcal{N}(\vartheta, \sigma^2)$ unter $\mathbb{P}_\vartheta$ **(2)** σ^2 bekannt. **Teststatistik:**

$$T = \frac{\bar{\mathcal{X}}_n - \vartheta_0}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0, 1) \quad \text{unter } \mathbb{P}_{\vartheta_0}$$

Bestimmung **cs:** $z_{1-\alpha} = \Phi^{-1}(1 - \alpha)$.

$$c_> = z_{1-\alpha}; c_< = -z_{1-\alpha}; c_\neq = z_{1-0.5\alpha}$$

Bsp **t-Test** $\mathcal{X}_i$ i.i.d. $\sim \mathcal{N}(\mu, \sigma^2)$ unter $\mathbb{P}_\vartheta$ mit $\vec{\vartheta} := (\mu, \sigma^2)$ (und insbesondere σ^2) unbekannt. Testen wieder $H_0 : \mu = \mu_0$. Teststatistik ist hier (see 7.4, t_m -V.):

$$T = \frac{\bar{\mathcal{X}}_n - \mu_0}{\frac{S}{\sqrt{n}}} \quad S^2 = \frac{1}{n-1} \sum_{k=1}^n (\mathcal{X}_k - \bar{\mathcal{X}}_n)^2$$

K : $c_> = t_{n-1, 1-\alpha}$; $c_< = t_{n-1, \alpha} = -c_>$; $c_\neq = t_{n-1, 1-0.5\alpha}$, mit $t_{m, \gamma}$ das γ -Quantil von t_m -Verteilung, welches Wert ist, so dass $\mathbb{P}[\mathcal{X} \leq t_{m, \gamma}] = \gamma$ für $\mathcal{X} \sim t_m$. Wert ist tabelliert!

Bem **Zweistichproben-Tests** $\mathcal{X}_k$ (i.i.d.) und jeweils unabhängig von allen $\mathcal{Y}_l$

Bsp **Gepaarter Zweist. m. N.V.** $\mathcal{X}_i$ i.i.d. $\sim \mathcal{N}(\mu_X, \sigma^2)$ und $\mathcal{Y}_j$ i.i.d. $\sim \mathcal{N}(\mu_Y, \sigma^2)$.

Differenzen $\mathcal{Z}_k = \mathcal{X}_k - \mathcal{Y}_k$ unter $\mathbb{P}_\vartheta$ i.i.d. $\sim \mathcal{N}(\mu_X - \mu_Y, 2\sigma^2)$. Wenn nicht unabhängig: $\sim \mathcal{N}(\mu_X - \mu_Y, 2(1 - \rho)\sigma^2)$ mit $\rho \in (-1, 1)$ Korrelation und $\rho = 0$ Unabhängigkeit.

Bsp **Ungepaart** $\mathcal{X}_i, \mathcal{Y}_j$ wie zuvor (n und m je).

z-Test (falls σ^2 bekannt):

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

$\mu_X - \mu_Y$ muss aus H_0 bekannt sein. Der kritische Bereich ist wieder Quantile der $\mathcal{N}(0, 1)$ -Verteilung.

t-Test (falls σ^2 unbekannt):

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

mit S_Y^2 analog zu S_X^2 (n durch m ersetzt)

$$S_X^2 = \frac{1}{n-1} \sum_{k=1}^n (\mathcal{X}_k - \bar{\mathcal{X}}_n)^2$$

Dann:

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

8.4 p-Wert

Für Stichprobe $\mathcal{X}_i$ (n el.) testen $H_0 : \vartheta = \vartheta_0$ gegen eine $H_A : \vartheta \in \Theta_A$.

p -Wert ist W. unter H_0 einen mindestens so extremen Wert für T zu bekommen wie $T(\omega) = t(x_1, \dots, x_n)$, angenommen, dass die Nullhypothese korrekt ist.

Def Geordnete Testsammlung T Teststatistik, Familie von Tests $(T, K_t)_{t \geq 0}$ **geordnet bezgl.** T , falls $\forall t$ gilt $K_t \subseteq \mathbb{R}$ und $\forall s \leq t$ gilt $K_s \supset K_t$. Typische Beispiele:

- $K_t = (t, \infty)$ (rechtsseitiger Test)
- $K_t = (-\infty, t)$ (linksseitiger Test)
- $K_t = (-\infty, t) \cup (t, \infty)$ (beidseitiger Test)

Def p-Wert p -Wert = $G(T) : \Omega \rightarrow [0, 1]$, mit $G(t) = \mathbb{P}_{\vartheta_0}[T \in K_t]$. Er ist eine Zufallsvariable, hängt direkt von x_i ab. Wiederholen des Tests generiert neuen p -Wert.

Wenn T stetig und $K_t = (t, \infty)$, dann gleichverteilt auf $[0, 1]$.

Intuition: Besagt, welche Tests H_0 ablehnen würden (alle Tests mit $\alpha > p$ lehnen H_0 ab). H_A spielt *keine* Rolle. Es heisst *NICHT*, dass der p -Wert die W. *IST*, dass die Hypothese richtig ist.

How To: Berechnung p -Wert

- Für rechtsseitigen Test: $p_r = \mathbb{P}[T \geq t]$
- Für linksseitigen Test: $p_l = \mathbb{P}[T \leq t]$
- Für beidseitigen Test: $p = 2 \min\{p_l, p_r\}$, falls Verteilung symmetrisch um 0, dann $p = \mathbb{P}[|T| \geq |t|]$

wobei für alle t eine Realisierung der Teststatistik T ist.

9 Konfidenzbereiche

Def Konfidenzbereich $C(x_1, \dots, x_n) \subseteq \Theta$. Mit genrierenden Z.V. ist $\tilde{C} := C(\mathcal{X}_1, \dots, \mathcal{X}_n)$ eine **zufällige Teilmenge** von Θ , (*Realisierung*) $\tilde{C}(\omega) = C(\mathcal{X}_1(\omega), \dots, \mathcal{X}_n(\omega))$
Für $\alpha \in [0, 1]$ ist C **Konfidenzbereich zum Niveau** $1 - \alpha$, falls $\mathbb{P}_{\vartheta}[C(\mathcal{X}_1, \dots, \mathcal{X}_n) \ni \vartheta] \geq 1 - \alpha \quad \forall \vartheta \in \Theta$

Man erwischt also mit grosser W. den Param in jedem Modell

Bsp $C(\mathcal{X}_1, \dots) = [\bar{X}_n - c, \bar{X}_n + c]$, mit c zu bestimmen.

Def Konfidenzintervall (K.I.) mit Niveau $1 - \alpha$ ist Zufallsintervall $I = [A, B]$ s.d. $\forall \vartheta \in \Theta$ gilt:

$$\mathbb{P}_{\vartheta}[[A, B] \ni \vartheta] = \mathbb{P}_{\vartheta}[A \leq \vartheta \leq B] \geq 1 - \alpha$$

mit A, B Z.V. wie $A = a(\mathcal{X}_1, \dots, \mathcal{X}_n)$, mit $a, b : \mathbb{R}^n \rightarrow \mathbb{R}$. Dabei ist ϑ deterministisch. (**Berechnung:** s. 2.2)

Bem Approx. K.I. Nur wenn genaue Verteilungsaussagen zur Verfügung stehen, sind exakte K.I. möglich, sonst mittels zentralem Grenzwertsatz approx. berechnen ($z = z_{1-\frac{\alpha}{2}}$)

$$P_{\vartheta}[|S_n^*| \leq z] \approx 1 - \alpha$$

Auflösen von untenstehend nach ϑ (S_n wie in 6.1)

$$|S_n - n\vartheta| \leq z\sqrt{n\vartheta(1-\vartheta)} \text{ bzw. } (S_n - n\vartheta)^2 \leq z^2 n\vartheta(1-\vartheta)$$

Ist kompliziert, ersetzen der Ugl. durch Gl. (quad. Gl.)

$$\hat{\vartheta}_{\pm} = \frac{2S_n + z^2 \pm \sqrt{(2S_n + z^2)^2 - 4S_n^2(1 + \frac{z^2}{n})}}{2n(1 + \frac{z^2}{n})}$$

Konfidenzintervall ist $[\hat{\vartheta}_-, \hat{\vartheta}_+]$.

Alternative Methoden (Annahme $\vartheta(1-\vartheta) \approx \frac{1}{4}$):

$$\left[\bar{S}_n - \frac{z}{2\sqrt{n}}, \bar{S}_n + \frac{z}{2\sqrt{n}} \right]$$

(Ersetzen von ϑ durch $\bar{S}_n$):

$$\left[\bar{S}_n - \frac{z}{2\sqrt{n}} \sqrt{\bar{S}_n(1-\bar{S}_n)}, \bar{S}_n + \frac{z}{2\sqrt{n}} \sqrt{\bar{S}_n(1-\bar{S}_n)} \right]$$

Alle Methoden liefern unterschiedliche Resultate

Normalverteilung: $\hat{\vartheta} = \bar{X}_n \pm z \frac{\sigma}{\sqrt{n}}$, $I = [\hat{\vartheta}_-, \hat{\vartheta}_+]$

Def $z_{1-\alpha} = \Phi^{-1}(1 - \alpha)$. α ist Verwerfungsbereich, nicht "Akzeptanzbereich"

10 Kombinatorik

10.1 Permutationen

Auf wieviele Arten kann man n Objekte anordnen

ohne Wiederholung $n!$. Zudem $\binom{n}{0} = 1$ nach Konvention

10.2 Kombinationen

Auf wieviele Arten kann man $k \leq n$ von n Objekten auswählen?

Ohne Zurücklegen $\binom{n}{k} = \frac{n!}{k!(n-k)!}$; **Mit** $\binom{n+k-1}{k}$

10.3 Variationen

Wie viele Sequenzen der Länge m kann man mit m Elementen bilden?

Ohne Zurücklegen $\frac{n!}{(n-m)!}$; **Mit** n^m

10.4 Binomialsatz

Def:

$$(x + y)^n = \sum_{i=0}^n \binom{n}{i} x^i y^{n-i} = \sum_{i=0}^n \binom{n}{i} x^{n-i} y^i$$

11 Tabellen

11.1 Standard-Normalverteilung

Tabelle mit Werten der $\Phi(z)$ -Funktion der Verteilung

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.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

11.2 Verschiedene Funktionen

Bem Ungleichungen Zur Erinnerung: Richtung wechselt bei Multiplikation mit (Division durch) negative Zahl.

11.2.1 Logarithmen

Hier ist wieder $\log = \ln$ (wenn keine Basis gegeben)

- (Basiswechsel) $\log_a(x) = \frac{\ln(x)}{\ln(a)}$
- (Potenzen) $\log_a(x^y) = y \log_a(x)$
- (Div, Mul) $\log_a(x \cdot (\div)y) = \log_a(x) + (-) \log_a(y)$
- $\log_a(1) = 0 \quad \forall a \in \mathbb{N}$

11.2.2 Trigonometrie

$$\cot(\xi) = \frac{\cos(\xi)}{\sin(\xi)}, \tan(\xi) = \frac{\sin(\xi)}{\cos(\xi)}$$

$$\sinh(x) := \frac{e^x - e^{-x}}{2} : \mathbb{R} \rightarrow \mathbb{R}, \cosh(x) := \frac{e^x + e^{-x}}{2} : \mathbb{R} \rightarrow$$

$$[1, \infty], \cosh(x) := \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}} : \mathbb{R} \rightarrow [-1, 1]$$

1. $\cos(x) = \cos(-x)$ und $\sin(-x) = -\sin(x)$
2. $\cos(\pi - x) = -\cos(x)$ und $\sin(\pi - x) = \sin(x)$
3. $\sin(x + w) = \sin(x) \cos(w) + \cos(x) \sin(w)$
4. $\cos(x + w) = \cos(x) \cos(w) - \sin(x) \sin(w)$
5. $\cos(x)^2 + \sin(x)^2 = 1$
6. $\sin(2x) = 2 \sin(x) \cos(x)$
7. $\cos(2x) = \cos(x)^2 - \sin(x)^2$

Werte der trigonometrischen Funktionen

°	rad	$\sin(\xi)$	$\cos(\xi)$	$\tan(\xi)$
0°	0	0	1	1
30°	$\frac{\pi}{6}$	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{3}}{2}$
45°	$\frac{\pi}{4}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	1
60°	$\frac{\pi}{3}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\sqrt{3}$
90°	$\frac{\pi}{2}$	1	0	$\emptyset$
120°	$\frac{2\pi}{3}$	$\frac{\sqrt{3}}{2}$	$-\frac{1}{2}$	$-\sqrt{3}$
135°	$\frac{3\pi}{4}$	$\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2}}{2}$	-1
150°	$\frac{5\pi}{6}$	$\frac{1}{2}$	$-\frac{\sqrt{3}}{2}$	$-\frac{\sqrt{3}}{2}$
180°	π	0	-1	0

11.2.3 Reihen

$$\sum_{k=1}^n \vartheta(1 - \vartheta)^{x_k-1} = \vartheta^n(1 - \vartheta)^{S-n} \text{ mit } S = \sum_{k=1}^n x_k$$

$$\sum_{k=1}^n k = \frac{n \cdot (n+1)}{2}; \sum_{k=1}^n k^2 = \frac{n \cdot (n+1) \cdot (2n+1)}{6};$$

$$\sum_{k=1}^n k^3 = \left(\frac{n \cdot (n+1)}{2} \right)^2; \sum_{k=1}^n \frac{1}{k^2} = \frac{\pi^2}{6}$$

$$\text{Arithmetic series: } s_n = \frac{n \cdot (2a_1 + (n-1) \cdot d)}{2}, \text{ mit } d \text{ Differenz}$$

$$\text{Geometric series: } s_n = a_1 \frac{1-q^n}{1-q}, \text{ mit } q \text{ Quotient}$$

11.3 Tabelle von Auf- und Ableitungen

Für Integrale von stückweisen Funktionen, stückweise integrieren und aufsummieren. Für ungerade Funktionen (mit $f(-x) = -f(x)$, bspw $f(x) = x$ und $f(x) = \cos(x)$) ist Integral $\int_{-C}^C f(x) dx = 0$, auch bei Multiplikation von mehreren Funktionen.

Bem Integration by parts

$$\int_a^b f(x)g'(x) dx = [f(x)g(x)]_a^b - \int_a^b f'(x)g(x) dx$$

$f(x)$ einfach ableitbar, $g'(x)$ einfach aufleitbar wählen.

Bem Integration by substitution

$$\int_a^b f(\varphi(t))\varphi'(t) dt = \int_{\varphi(a)}^{\varphi(b)} f(x) dx$$

dabei die innere Funktion gut wählen, ableiten und zurück substituieren. Grenzen des Integrals anpassen!

Bem Multiplication rule $f'(x)g(x) + f(x)g'(x)$

Bem Quotient rule $\frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}$

Thm Fundamental Theorem $\int_a^b f(x) dx = F(a) - F(b)$

Stammfunktion	Funktion	Ableitung
$\frac{x^{n+1}}{n+1}$	x^n	$n \cdot x^{n-1}$
$\ln x $	$\frac{1}{x} = x^{-1}$	$-x^{-2} = -\frac{1}{x^2}$
$\frac{2}{3}x^{\frac{3}{2}}$	$\sqrt{x} = x^{\frac{1}{2}}$	$\frac{1}{2 \cdot \sqrt{x}}$
$\frac{n}{n+1}x^{\frac{1}{n}+1}$	$\sqrt[n]{x} = x^{\frac{1}{n}}$	$\frac{1}{n}x^{\frac{1}{n}-1}$
e^x	e^x	e^x
$\exp(x)$	$\exp(x)$	$\exp(x)$
$\frac{1}{a \cdot (n+1)}(ax+b)^{n+1}$	$(ax+b)^n$	$n \cdot (ax+b)^{n-1} \cdot a$
$x \cdot (\ln x - 1)$	$\ln(x)$	$\frac{1}{x} = x^{-1}$
$\frac{1}{\ln(a)} \cdot a^x$	a^x	$a^x \cdot \ln(a)$
$\frac{x}{\ln(a)} \cdot (\ln x - 1)$	$\log_a x $	$\frac{1}{x \cdot \ln(a)}$
$-\cos(x)$	$\sin(x)$	$\cos(x)$
$\sin(x)$	$\cos(x)$	$-\sin(x)$
$-\ln \cos(x) $	$\tan(x)$	$\frac{1}{\cos^2(x)}$
$x \cdot \arcsin(x) + \sqrt{1-x^2}$	$\arcsin(x)$	$\frac{1}{\sqrt{1-x^2}}$
$x \cdot \arccos(x) - \sqrt{1-x^2}$	$\arccos(x)$	$-\frac{1}{\sqrt{1-x^2}}$
$x \cdot \arctan(x) - \frac{\ln(x^2+1)}{2}$	$\arctan(x)$	$\frac{1}{x^2+1}$

11.4 Limits

$\lim_{x \rightarrow \infty} \frac{1}{x} = 0$	$\lim_{x \rightarrow \infty} 1 + \frac{1}{x} = 1$
$\lim_{x \rightarrow \infty} e^x = \infty$	$\lim_{x \rightarrow -\infty} e^x = 0$
$\lim_{x \rightarrow \infty} e^{-x} = 0$	$\lim_{x \rightarrow -\infty} e^{-x} = \infty$
$\lim_{x \rightarrow \infty} \frac{e^x}{x^m} = \infty$	$\lim_{x \rightarrow -\infty} x e^x = 0$
$\lim_{x \rightarrow \infty} \ln(x) = \infty$	$\lim_{x \rightarrow 0} \ln(x) = -\infty$
$\lim_{x \rightarrow \infty} (1+x)^{\frac{1}{x}} = 1$	$\lim_{x \rightarrow 0} (1+x)^{\frac{1}{x}} = e$
$\lim_{x \rightarrow \infty} (1+\frac{1}{x})^b = 1$	$\lim_{x \rightarrow \infty} (1+\frac{1}{x})^b = 1$
$\lim_{x \rightarrow \infty} x^a q^x = 0, \forall 0 \leq q < 1$	$\lim_{x \rightarrow \infty} n^{\frac{1}{n}} = 1$
$\lim_{x \rightarrow \pm \infty} (1+\frac{1}{x})^x = e$	$\lim_{x \rightarrow \infty} (1-\frac{1}{x})^x = \frac{1}{e}$
$\lim_{x \rightarrow \pm \infty} (1+\frac{k}{x})^{mx} = e^{km}$	$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$
$\lim_{x \rightarrow 0} \frac{1}{\cos(x)} = 1$	$\lim_{x \rightarrow 0} \frac{\cos x - 1}{x} = 0$
$\lim_{x \rightarrow 0} \frac{\log \frac{1-x}{x}}{x} = -1$	$\lim_{x \rightarrow 0} x \log x = 0$
$\lim_{x \rightarrow 0} \frac{1 - \cos x}{x^2} = \frac{1}{2}$	$\lim_{x \rightarrow 0} \frac{e^x - 1}{x} = 1$
$\lim_{x \rightarrow 0} \frac{x}{\arctan x} = 1$	$\lim_{x \rightarrow \infty} \arctan x = \frac{\pi}{2}$
$\lim_{x \rightarrow \infty} \left(\frac{x}{x+k} \right)^x = e^{-k}$	$\lim_{x \rightarrow 0} \frac{e^x - 1}{x} = 1$
$\lim_{x \rightarrow 0} \frac{a^x - 1}{x} = \ln(a) \quad \forall a > 0$	$\lim_{x \rightarrow 0} \frac{e^{ax} - 1}{x} = a$
$\lim_{x \rightarrow 0} \frac{\ln(x+1)}{x} = 1$	$\lim_{x \rightarrow 1} \frac{\ln(x)}{x-1} = 1$
$\lim_{x \rightarrow \infty} \frac{\ln(x)}{x} = 0$	$\lim_{x \rightarrow \infty} \frac{\log(x)}{x^a} = 0$
$\lim_{x \rightarrow \infty} \sqrt[x]{x} = 1$	$\lim_{x \rightarrow \infty} \frac{2x}{2^x} = 0$
$\lim_{x \rightarrow \frac{\pi}{2}^-} \tan x = +\infty$	$\lim_{x \rightarrow \frac{\pi}{2}^+} \tan x = -\infty$
$\lim_{x \rightarrow \infty} \frac{\sin x}{x} = 0$	$\lim_{x \rightarrow 0^+} x \ln x = 0$