

Wahrscheinlichkeit und Statistik

Janis Hutz

<https://janishutz.com>

5. Mai 2026

1 Basics

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 (Ω, \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 $\mathbb{P}[A] = \frac{\binom{n}{2}}{\binom{n}{2}} = \frac{2}{n-1}$

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

1.3 Eigenschaften/Interp. von Ereignissen

T \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
	$\forall \omega \in \Omega$
$\Omega = A_1 \cup A_2 \cup A_3$ mit	nur eines von A_1, A_2, A_3
A_1, A_2, A_3 paarw. disj.	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

T \mathbb{P} Wahrscheinlichkeitsmass auf (Ω, \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

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

T (*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

1.4.2 Anwendungen der Ungleichungen

Sie sind nützlich für schwer zu berechnende W.

T (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 $\mathbb{P}[B] > 0$:

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

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

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

T (*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:

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

T (*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

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

T 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:

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

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

2.2 Verteilungsfunktion

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

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

T \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 $\lim_{a \rightarrow \infty} F(a) = 1$

2.3 Unabhängigkeit

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

2.3.1 Gruppierung

T n ZV \mathcal{X}_i und $1 \leq i_1 < \dots < i_k \leq n$ sind indizes und ϕ_1, \dots, ϕ_k Abbildungen. Dann sind unabhängig:
 $Y_1 = \phi_1(\mathcal{X}_1, \dots, \mathcal{X}_{i_1}), \dots, Y_k = \phi_k(\mathcal{X}_{i_{k-1}+1}, \dots, \mathcal{X}_{i_k})$

2.3.2 Unabhängig identisch verteilte ZV

Def Eine Folge von ZV ist **(1)** unabh. falls \mathcal{X}_i unabh. sind und **(2)** uiv, falls unabh. und die ZV dieselbe Verteilungsf. haben, also: $\forall i, j \quad F_{\mathcal{X}_i} = F_{\mathcal{X}_j}$

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

Def (Bernoulli ZV) mit param. $p \in [0, 1]$ falls $\mathbb{P}[\mathcal{X} = 0] = 1 - p$ und $\mathbb{P}[\mathcal{X} = 1] = p$. Not.: $\mathcal{X} \sim \text{Ber}(p)$

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

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

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

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

T (Inversionsmethode) F erfüllt eig. v. T2.4, $U \sim \mathcal{U}(\dots)$, 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.

T 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

T (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 $\exists A' \subseteq A \mid \mathbb{P}[A'] = 1$

3.3 Diskrete Zufallsvariablen

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

Bem Falls der Grundraum Ω endlich oder abzählbar ist, dann ist jede Z.V. \mathcal{X} diskret.

Def (Verteilung) Für Z.V. \mathcal{X} mit W endl. oder abzählb. $(p(x))_{x \in W} \stackrel{\text{def}}{\Leftrightarrow} \forall x \in W \quad p(x) := \mathbb{P}[\mathcal{X} = x]$

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

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

T \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 $\mathcal{X} \sim \text{Ber}(p)$: $\mathbb{P}[\mathcal{X} = 0] = 1 - p$ und $\mathbb{P}[\mathcal{X} = 1] = p$

3.4.2 Binomialverteilung

Def $\mathcal{X} \sim \text{Bin}(n, p)$, falls

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

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

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

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 $W = \mathbb{N} \setminus \{0\}$ falls

$$\forall k \in W \quad \mathbb{P}[\mathcal{X} = k] = (1-p)^{k-1} \cdot p$$

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

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

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

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

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

T $T \sim \text{Geom}(p)$, 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

$$\forall k \in \{r, r+1, \dots\} \quad \mathbb{P}[\mathcal{X} = k] = \binom{k-1}{r-1} p^r (1-p)^{k-r}$$

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

$$\forall k \in \{0, \dots, \min(m, r)\} \quad \mathbb{P}[\mathcal{X} = k] = \frac{\binom{r}{k} \binom{n-r}{m-k}}{\binom{n}{m}}$$

3.4.6 Poisson-Verteilung

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

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

T Für $n \in \mathbb{N}$ Z.V. $\mathcal{X}_i \sim \text{Bin}(n, \frac{\lambda}{n})$ und $\mathcal{N} \sim \text{Poisson}(\lambda)$:
 $\forall k \in \mathbb{N} \quad \lim_{n \rightarrow \infty} \mathbb{P}[\mathcal{X}_i = k] = \mathbb{P}[\mathcal{N} = k]$

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) von \mathcal{X}

T Sei $F_{\mathcal{X}}$ st. stückw. diff. auf Partition $-\infty = x_0 < x_1 < \dots < x_n = \infty$. Dann \mathcal{X} stetig, mit a_k beliebig und

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

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

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 Geomemtrische 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 $f_{\mathcal{X}}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$,

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

T $\mathcal{X} \sim \mathcal{N}(\mu, \sigma^2)$, 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

T \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 $|\mathcal{X}| = \mathcal{X}_+ + \mathcal{X}_-$. 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

T 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. conv.

4.2.1 Beispiele

- $\mathcal{X} \sim \text{Ber}(p)$: $\mathbb{E}[\mathcal{X}] = p$
- $\mathcal{X} \sim \text{Bin}(n, p)$: $\mathbb{E}[\mathcal{X}] = np$
- $\mathcal{X} \sim \text{Poisson}(\lambda)$: $\mathbb{E}[\mathcal{X}] = \lambda$

4.2.2 Transformierte Zufallsvariablen

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

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

4.3.1 Beispiele

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

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

4.4 Eigenschaften des Erwartungswerts

T (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 : $\mathbb{E}[\sum_{k=1}^n \lambda_k \mathcal{X}_k] = \sum_{k=1}^n \lambda_k \mathbb{E}[\mathcal{X}_k]$

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

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

T \mathcal{X}_k alle unabhängig mit $\mathbb{E}[\mathcal{X}_k]$ endlich. Dann gilt

$$\mathbb{E}[\prod_{k=1}^n \mathcal{X}_k] = \prod_{k=1}^n \mathbb{E}[\mathcal{X}_k]$$

T $f : \mathbb{R} \rightarrow \mathbb{R}_+$ 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 $\mathbb{E}[\varphi(\mathcal{X})] = \int_{-\infty}^{\infty} \varphi(x) f(x) dx$

T äquivalent: **1** \mathcal{X}, \mathcal{Y} unabhängig, für alle $\varphi, \psi : \mathbb{R} \rightarrow \mathbb{R}$:

$$\mathbb{E}[\varphi(\mathcal{X})\psi(\mathcal{Y})] = \mathbb{E}[\varphi(\mathcal{X})]\mathbb{E}[\psi(\mathcal{Y})]$$

T äquivalent: **(1)** \mathcal{X}_i unabhängig,

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

4.5 Ungleichungen

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

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

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

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 $\mathbb{V}[\mathcal{X}] = \mathbb{E}[\mathcal{X}^2] - \mathbb{E}[\mathcal{X}]^2$

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. $\mathbb{V}[\sum_{k=1}^n \mathcal{X}_k] = \sum_{k=1}^n \mathbb{V}[\mathcal{X}_k]$

Bsp Varianz von bekannten Verteilungen

- $\mathcal{X} \sim \text{Ber}(p)$, $\mathbb{V}[\mathcal{X}] = p(1-p)$
- $\mathcal{X} \sim \text{Bin}(n, p)$, $\mathbb{V}[\mathcal{X}] = np(1-p)$
- $\mathcal{X} \sim \text{Poisson}(\lambda)$, $\mathbb{V}[\mathcal{X}] = \lambda = \mathbb{E}[\mathcal{X}]$
- $\mathcal{X} \sim \mathcal{U}([a, b])$, $\mathbb{V}[\mathcal{X}] = \frac{(b-a)^2}{12}$
- $\mathcal{X} \sim \mathcal{N}(\mu, \sigma^2)$, $\mathbb{V}[\mathcal{X}] = \sigma^2$

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

4.7 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

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.:** $\text{cov}(a\mathcal{X} + b, c\mathcal{Y} + d) = a\text{cov}(\mathcal{X}, \mathcal{Y}) + c\text{cov}(\mathcal{X}, d)$
 $\text{cov}(\mathcal{X}, (e\mathcal{Y} + f) + (g\mathcal{Z} + h)) = e\text{cov}(\mathcal{X}, \mathcal{Y}) + g\text{cov}(\mathcal{X}, \mathcal{Z})$

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 $\mathcal{X} = (\mathcal{X}_1, \dots, \mathcal{X}_n)^\top$

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

5 Gemeinsame Verteilungen

5.1 Gemeinsame Diskrete Verteilung

Def \mathcal{X}_i 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 Gewichtsfunktion (joint probability mass function) mit

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

T 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 $F_{\mathcal{X}}$ & $p_{\mathcal{X}}$):

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

T (*Verteilung Bild*) Sei $\phi: \mathbb{R}^n \rightarrow \mathbb{R}$, \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 ($\forall z \in W$):

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

T (*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. $F_{\mathcal{X}_k}(x) = \mathbb{P}[\mathcal{X}_k \leq x]$

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

T (*Erwartungswert Bild*) (Solange Summe wohldefiniert ist, summieren über $x_1 \in W_1, \dots, x_n \in W_n$)

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

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

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:

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

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

T (*Erwartungswert Bild*)

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

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:

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

Herleitung der Randdichte ("wegintegrieren"):

$$\begin{aligned} f_{\mathcal{X}}(x) &= \frac{d}{dx} F_{\mathcal{X}}(x) = \frac{d}{dx} \left(\int_{-\infty}^x \int_{-\infty}^{\infty} f(s, t) dt ds \right) \\ &= \int_{-\infty}^{\infty} f(x, t) dt \end{aligned}$$

(Hier wieder Umwandlung von Summe zu Integral von Diskret zu Stetig)

Bsp Beispiele von gemeinsamen stetigen Verteilungen

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

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

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

T (Unabhängigkeit) \mathcal{X}_i mit Dichten $f_{\mathcal{X}_i}$, dann ist äquiv.:

- $\mathcal{X}_1, \dots, \mathcal{X}_n$ sind unabhängig
- $\mathcal{X}_1, \dots, \mathcal{X}_n$ sind gem. stetig mit gem. Dichte (Prod. Randdichten): $f(x_1, \dots, x_n) = f_{\mathcal{X}_1}(x_1) \cdot \dots \cdot f_{\mathcal{X}_n}(x_n)$

Bsp (Gleichverteilungen)

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

$$f(x, y) = 1_{(x,y) \in [0,1]^2} = 1_{x \in [0,1]} 1_{y \in [0,1]} = f_{\mathcal{X}}(x) f_{\mathcal{Y}}(y)$$

Folglich sind die beiden Koordinaten unabhängig.

Einheitskreisscheibe $f(x, y) \neq f_{\mathcal{X}}(x) f_{\mathcal{Y}}(y)$, mit Dichten von oben. Also sind die beiden Koordinaten nicht unabhängig.

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

6.1 Schwaches Gesetz der grossen Zahlen

T (Schwach. Ges. der grossen Zahlen) 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

T (Starkes Ges. der grossen Zahlen) 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

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

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:

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

T (Zentraler Grenzwertsatz (ZGS))

i.i.d = independent and identically distributed (u.i.v in DE)

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

für grosse n , mit $\overset{\text{approx}}{\sim}$ gespr. "approx. gleichverteilt gemäss". Also ist für $\mathbb{E}[S_n^*] = 0$ und $\mathbb{V}[S_n^*] = 1$.

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

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

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

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 $M_{\mathcal{X}}(t) = (1 - p + pe^t)^n$

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

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

T (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:

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

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. hochdimensionalem) **Param.** ϑ aus **Param.-Raum** Θ dargestellt werden. 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}_ϑ

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}_ϑ) ist $\mathbb{E}_\vartheta[T] - \vartheta$

Def (MSE) Mittlerer quadratischer Schätzfehler (mean squared err.) von T mit \mathbb{P}_ϑ 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}_ϑ -W. gegen ϑ konvergiert, also $\forall \vartheta \in \Theta, \varepsilon > 0: \lim_{n \rightarrow \infty} \mathbb{P}_\vartheta[|T^{(n)} - \vartheta| > \varepsilon] = 0$

7.3 Maximum-Likelihood-Method

In Modell \mathbb{P}_ϑ 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 ind. Gew. $p_{\mathcal{X}}(x; \vartheta)$ bzw. Dichte $f_{\mathcal{X}}(x; \vartheta)$, 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 $\log(L(x_1, \dots, x_n; \vartheta))$. Ist im i.i.d-Fall eine Summe.

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

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

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

Bem Einfacher: Statt maximieren, Nullstellen von Ableitung nach ϑ .

Bsp Verteilungen

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 $x \in \{0, 1\}$. LH-Func:

$$L(x_1, \dots, x_n; \vartheta) = \vartheta^{\sum_{k=1}^n x_k} (1 - \vartheta)^{n - \sum_{k=1}^n x_k}$$

Log LH-Func: $\log(\vartheta) \sum_{k=1}^n x_k + \log(1 - \vartheta) (n - \sum_{k=1}^n x_k)$

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

Normalverteilung $\mathcal{X}_i \sim \mathcal{N}(\mu, \sigma^2)$ i.i.d. Dabei:

$$f_{\mathcal{X}}(x; \vartheta) = \frac{1}{\sqrt{2\pi v}} e^{-\frac{(x-\mu)^2}{2v}}$$

Weil i.i.d:

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

und somit:

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

Der Schätzer ist $T = (T_1, T_2)$ (**Momentschätzer**):

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

Für T gilt: $(\mathbb{E}_\vartheta[\mathcal{X}], \mathbb{V}_\vartheta[\mathcal{X}])$. Ist nicht Erwartungstreu. Es gilt $\mathbb{E}_\vartheta[\mathcal{X}_k \mathcal{X}_l] = \mathbb{E}_\vartheta[\mathcal{X}_k] \mathbb{E}_\vartheta[\mathcal{X}_l] = \mathbb{E}_\vartheta[\mathcal{X}]^2$ und folglich:

$$\mathbb{E}_\vartheta[(\bar{X}_n)^2] = \frac{1}{n} \mathbb{E}_\vartheta[\mathcal{X}^2] + \frac{n^2 - n}{n^2} (\mathbb{E}_\vartheta[\mathcal{X}])^2$$

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

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

T'_2 ist (korrigierte) empirische (Stichproben)varianz ((un)biased sample variance)

7.4 Verteilungsaussagen

Approx. von Verteilung von Schätzer unter \mathbb{P}_ϑ (wenn T Summe mit \mathcal{Y}_k im \mathbb{P}_ϑ). Solche sind approx. Normalverteilt unter \mathbb{P}_ϑ , mit Parametern $\mu = n \mathbb{E}_\vartheta[\mathcal{Y}_k]$ und $\sigma^2 = n \mathbb{V}_\vartheta[\mathcal{Y}_k]$.

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

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 **langschänziger**

T 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. $\frac{\bar{X}_n - \mu}{\frac{S}{\sqrt{n}}} = \frac{\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}}{\frac{S/\sigma}{\sqrt{\frac{n-1}{\sigma^2} S^2}}}$