Cheat Sheet: Comp Sc BSc, WuS - Name, Surname - **WW-XXX-YY-ZZZ**

Wahrscheinlichkeit

Grundbegriffe

Def 1.2 - Grundraum: Der Ereignisraum oder Grundraum (sample space) $\Omega \neq \emptyset$ ist die Menge aller möglichen Ergebnisse des betrachteten Zufallsexperiments. Die Elemente $\omega \in \Omega$ heissen Elementarereignisse oder Ausgänge des Experiments (outcomes).
Def 1.4 - Potenzmenge und Ereignisse: Die Potenzmenge (power set) von $\Omega, \mathcal{P}(\Omega)$ oder 2^{Ω}, ist die Menge aller Teilmengen von Ω. Ein prinzipielles Ereignis (event) ist eine Teilmenge $\mathcal{A} \subseteq \Omega$, also eine Kollektion von Elementarereignissen. Die Klasse aller (beobachtbaren) Ereignisse bezeichnen wir mit $\mathcal{F}$. Das ist eine Teilmenge der Potenzmenge von Ω.
Note Berechnungen: Es gilt immer $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B]$ (Siebformel)
Def 1.5 -σ-Algebra : Ein Mengensystem $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ nennt man eine σ-Algebra (manchmal σ-field), wenn

1. $\Omega \in \mathcal{F}$,
2. für jedes $A \in \mathcal{F}$ auch das Komplement $A^{\complement} \in \mathcal{F}$ ist,
3. für jede Folge $(A_n)_{n \in \mathbb{N}}$ mit $A_n \in \mathcal{F}, n \in \mathbb{N}$, auch die Vereinigung $\cup_{n \in \mathbb{N}} A_n \in \mathcal{F}$ ist.

Def 1.9. (Wahrscheinlichkeitsmass): Sei Ω ein Grundraum und sei $\mathcal{F}$ eine σ-Algebra. Eine Abbildung

$$\mathbb{P} : \mathcal{F} \to [0,1], \text{ mit } A \mapsto \mathbb{P}[A]$$

heisst Wahrscheinlichkeitsmass (probability measure) auf $(\Omega, \mathcal{F})$, wenn die folgenden Axiome erfüllt sind,

1. Normiertheit: $\mathbb{P}[\Omega] = 1$,
2. σ-Additivität: $\mathbb{P}[\cup_{n \in \mathbb{N}} A_n] = \sum_{n=1}^{\infty} \mathbb{P}[A_n]$ für paarweise disjunkte Mengen A_n, d.h. $A_n \cap A_m = \emptyset$ für alle $n \neq m$.

Pro 1.10: Für ein Wahrscheinlichkeitsmass $\mathbb{P}$ auf $(\Omega, \mathcal{F})$ und Mengen $A, B \in \mathcal{F}$ gelten folgende Aussagen:

1. $\mathbb{P}\left[A^{\complement}\right] = 1 - \mathbb{P}[A]$, und insbesondere $\mathbb{P}[\emptyset] = 0$.
2. Monotonie: wenn $A \subseteq B$, dann $\mathbb{P}[A] \leq \mathbb{P}[B]$,
3. Additionsregel: $\mathbb{P}[A] + \mathbb{P}[B] = \mathbb{P}[A \cup B] - \mathbb{P}[A \cap B]$

Def Definition 1.12. (Wahrscheinlichkeitsraum): Sei Ω ein Grundraum, $\mathcal{F}$ eine σ-Algebra und $\mathbb{P}$ ein Wahrscheinlichkeitsmass auf $(\Omega, \mathcal{F})$. Das Tripel $(\Omega, \mathcal{F}, \mathbb{P})$ heisst Wahrscheinlichkeitsraum.
Def 1.14. (Laplace Modell) : Sei $\Omega = \{\omega_1, \ldots, \omega_N\}$ mit $|\Omega| = N$ ein endlicher Grundraum. $(\Omega, \mathcal{F}, \mathbb{P}$) heisst Laplace Modell auf Ω, wenn

- $\mathcal{F} = \mathcal{P}(\Omega)$
- $\mathbb{P}$ ist die diskrete Gleichverteilung auf Ω, d.h. alle Elementarereignisse sind gleich wahrscheinlich, $p_1 = p_2 = \ldots = p_N = \frac{1}{N}$. Insbesondere gilt für beliebige $A \subseteq \Omega$

$$\mathbb{P}[A] = \frac{|A|}{|\Omega|} = \frac{\text{Anzahl Elementarereignisse in } A}{\text{Anzahl Elementarereignisse in } \Omega}.$$

Def 1.22. (Bedingte Wahrscheinlichkeit) : Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Seien A, B zwei Ereignisse mit $\mathbb{P}[B] > 0$. Wir definieren die bedingte Wahrscheinlichkeit von A gegeben B (conditional probability) (d.h. unter der Bedingung, dass B eintritt) wie folgt,

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

S 1.25. : Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Sei B ein Ereignis mit positiver Wahrscheinlichkeit. Dann ist $\mathbb{P}^* : \mathcal{F} \to [0,1]$ definiert durch

$$A \mapsto \mathbb{P}^*[A] := \mathbb{P}[A \mid B]$$

wieder ein Wahrscheinlichkeitsmass auf $(\Omega, \mathcal{F})$.
S 1.29. (Satz von der totalen Wahrscheinlichkeit) : Sei $B_1, \ldots, B_N$ mit $\mathbb{P}[B_n] > 0$ für jedes $1 \leq n \leq N$ eine Partition des Grundraums Ω, d.h. $\bigcup_{n=1}^{N} B_n = \Omega$ mit $B_n \cap B_m = \emptyset$ für $n \neq m$. Dann gilt für alle $A \in \mathcal{F}$,

$$\mathbb{P}[A] = \sum_{n=1}^{N} \mathbb{P}[A \mid B_n] \mathbb{P}[B_n]$$

Das bedeutet insbesondere

$$\mathbb{P}[B] = \mathbb{P}[B \mid A] \cdot \mathbb{P}[A] + \mathbb{P}[B \mid A^{\complement}] \cdot \mathbb{P}[A^{\complement}]$$

S 1.32. (Satz von Bayes) : Sei $B_1, \ldots, B_N \in \mathcal{F}$ eine Partition von Ω mit $\mathbb{P}[B_n] > 0$ für alle n. Für jedes Ereignis A mit $\mathbb{P}[A] > 0$ und jedes $n \in \{1, \ldots, N\}$ gilt

$$\mathbb{P}[B_n \mid A] = \frac{\mathbb{P}[A \mid B_n] \mathbb{P}[B_n]}{\sum_{k=1}^{N} \mathbb{P}[A \mid B_k] \mathbb{P}[B_k]}.$$

Unabhängigkeit

Def 1.35. (Unabhängigkeit zweier Ereignisse) : Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Zwei Ereignisse A und B heissen (stochastisch) unabhängig, falls

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

Pro Proposition 1.37.: Seien $A, B \in \mathcal{F}$ zwei Ereignisse mit $\mathbb{P}[A], \mathbb{P}[B] > 0$ Dann sind die folgenden Aussagen äquivalent:

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

Def 1.40. (Unabhängigkeit) : Sei I eine beliebige Indexmenge. Eine Familie von Ereignissen $(A_i)_{i \in I}$ heisst (stochastisch) unabhängig, wenn für alle endlichen Teilmengen $J \subset I$ gilt:
$\mathbb{P}\left[\bigcap_{j \in J} A_j\right] = \prod_{j \in J} \mathbb{P}[A_j]$.
Note :

1. $\mathbb{P}(A) \in \{0,1\} \implies A$ zu jedem Ereignis unabhängig
2. A zu sich selbst unabhängig $\implies \mathbb{P}(A) \in \{0,1\}$
3. A, B unabhängig $\implies A, B^{C}$ unabhängig
4. $\mathbb{P}[B \mid A^{\complement}] = 1 - \mathbb{P}[B^{\complement} \mid A^{\complement}]$

Zuvallsvariablen

Allgemein

Def 3.4.: Sei $A \in \mathcal{F}$ ein Ereignis. Wir sagen A tritt $\mathbb{P}$-fast sicher ($\mathbb{P}$-f.s.) ein, falls $\mathbb{P}[A] = 1$. Im Englischen sagen wir $\mathbb{P}$-almost surely (P.a.s.). Wenn klar ist, welches Wahrscheinlichkeitsmass $\mathbb{P}$ gemeint ist, kürzen wir ab und schreiben einfach fast sicher (f.s.).
S 2.21. (Gruppierungen von Zufallsvariablen) : Seien $X_1, \ldots, X_n$ unabhängige Zufallsvariablen. Seien $1 \leq i_1 < i_2 < \ldots < i_k \leq n$ Indexe und $\varphi_1, \ldots, \varphi_k$ Abbildungen. Dann sind $Y_1 := \varphi_1(X_1, \ldots, X_{i_1}), Y_2 := \varphi_2(X_{i_1+1}, \ldots, X_{i_2}), \ldots, Y_k := \varphi_k(X_1, \ldots, X_{i_k})$ unabhängig.
Def 2.22. (Unabhängig und identisch verteilt) : Eine Folge von Zufallsvariablen $X_1, X_2, \ldots$ heisst

- unabhängig falls $X_1, \ldots, X_n$ für alle $n \in \mathbb{N}$ unabhängig sind,
- unabhängig und identisch verteilt (u.i.v.), falls sie unabhängig ist und die Zufallsvariablen dieselbe Verteilungsfunktion haben, d.h. für alle $k, \ell \in \mathbb{N}$ gilt $F_{X_k} = F_{X_\ell}$.
- Im Englischen sagt man independent and identically distributed und benutzt die Abkürzung i.i.d., die auch wir in dieser Vorlesung benutzen werden.

Def 3.7. (Diskrete Zufallsvariablen) : Eine Zufallsvariable $X : \Omega \to \mathbb{R}$ heisst diskret, falls eine endliche oder abzählbare Menge $W \subset \mathbb{R}$ existiert, sodass $\mathbb{P}[X \in W] = 1$, wenn also die Werte von X fast sicher in W liegen.
Def 2.1. (Zufallsvariable) : Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Eine (reellwertige) Zufallsvariable (Z.V.) ist eine Abbildung $X : \Omega \to \mathbb{R}$, sodass für alle $x \in \mathbb{R}$ gilt,

$$\{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{F}.$$

Gewichtsfunktion (pmf)

Def 3.9. (Gewichtsfunktion) : Für eine diskrete Zufallsvariable X mit Wertebereich $W(X) = \{x_1, x_2, \ldots\}$ und den dazugehörigen Wahrscheinlichkeiten $\{p_1, p_2, \ldots\}$ definieren wir die Gewichtsfunktion oder diskrete Dichte von X als

$$p_X : W(X) \to [0,1] \quad \text{mit} \quad p_X(x_k) := \mathbb{P}[X = x_k] =$$

Die Zahlenfolge $\{p_X(x_k)\}_{x_k \in W(X)}$ nennen wir auch Verteilung von X.
Pro 3.10. : Die Gewichtsfunktion p_X einer diskreten Zufallsvariablen X hat folgende Eigenschaften:

- Für alle $x_k \in W(X)$ gilt $p_X(x_k) \in [0,1]$.
- Die Wahrscheinlichkeiten addieren sich zu 1 , $\sum_{x_k \in W(X)} p_X(x_k) = \mathbb{P}[X \in W(X)] = 1$

Note Verhältnis cdf/pmf: Es gilt:

$$f_X(x) = \frac{d}{dx} F_X(x)$$

Umkehrung stimmt nicht!

Verteilungsfunktion (cdf)

S 3.12. : Sei X eine diskrete Zufallsvariable mit Werten in W und Gewichtsfunktion p_X. Dann ist die Verteilungsfunktion von X gegeben durch

$$F_X(x) = \mathbb{P}[X \leq x] = \sum_{\substack{y \leq x \\ y \in W}} p_{X(y)}$$

Def 2.10. (Verteilungsfunktion): Sei X eine reellwertige Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$. Die Verteilungsfunktion von X ist die Funktion $F_X : \mathbb{R} \to [0,1]$, definiert durch

$$F_X(x) := \mathbb{P}[X \leq x]$$

Pro 2.12: Sei X eine reellwertige Zufallsvariable und seien $a < b$ zwei reelle Zahlen. Dann gilt

$$\mathbb{P}[a < X \leq b] = F_X(b) - F_X(a).$$

S 2.13. (Eigenschaften von Verteilungsfunktionen) : Sei X eine Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$. Die Verteilungsfunktion $F_X : \mathbb{R} \to [0,1]$ von X erfüllt folgende Eigenschaften:

1. F_X ist monoton wachsend.
2. F_X ist rechtsstetig, d.h. für alle $x \in \mathbb{R}$ gilt $F_X(x) = \lim_{h \searrow 0} F_X(x+h)$.
3. Es gelten die Grenzwerte $\lim_{x \to -\infty} F_X(x) = 0$ und $\lim_{x \to \infty} F_X(x) = 1$.

Note :

1. **Wenn F_X in einem Punkt $a \in \mathbb{R}$ nicht stetig ist, dann ist die "Sprunghöhe" $F_X(a) - F_X(a-)$ gleich der Wahrscheinlichkeit $\mathbb{P}(X = a)$.**

2. **Falls F_X stetig in einem Punkt $a \in \mathbb{R}$, dann gilt $\mathbb{P}(X = a) = 0$.**

Def 2.16. (Gemeinsame Verteilungsfunktion) : **Seien $X_1, \ldots, X_n$ Zufallsvariablen. Die gemeinsame Verteilungsfunktion von $X_1, \ldots, X_n$ (joint cumulative distribution function) ist die Abbildung $F : \mathbb{R}^n \to [0, 1]$ definiert durch**

$$(x_1, \ldots, x_n) \mapsto F(x_1, \ldots, x_n) = \mathbb{P}[X_1 \leq x_1, \ldots, X_n \leq x_n]$$

Def 2.18. (Unabhängigkeit): **Seien $X_1, \ldots, X_n$ Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$. Dann heissen $X_1, \ldots, X_n$ unabhängig, wenn für alle $x_1, \ldots, x_n \in \mathbb{R}$ gilt**

$$\mathbb{P}[X_1 \leq x_1, \ldots, X_n \leq x_n] = \mathbb{P}[X_1 \leq x_1] \cdot \ldots \cdot \mathbb{P}[X_n \leq x_n]$$

und somit auch

$$F_{X_1, X_2, \ldots}(X_1, X_2, \ldots) = F_{X_1}(x_1) F_{X_2}(x_2) \ldots$$

Def 3.37. (Stetig verteilte Zufallsvariablen): **Eine Zufallsvariable $X : \Omega \to \mathbb{R}$ heisst stetig, wenn eine nicht-negative Funktion $f_X : \mathbb{R} \to \mathbb{R}_+$ existiert, sodass die Verteilungsfunktion F_X dargestellt werden kann als**

$$F_X(x) = \int_{-\infty}^{x} f_X(t) \mathrm{d}t.$$

Wir nennen f_X die Dichte(-funktion) von X (probability density function (pdf)).

Def 3.38. (Stückweise stetig differenzierbare Funktionen): **Im Kontext von $\mathbb{R}$ sagt man oft, dass ein Objekt eine Eigenschaft stückweise (piecewise) erfüllt, wenn sie die Eigenschaft auf einer Partition des Definitionsbereichs erfüllt. Wir sagen, eine Funktion f ist stückweise stetig differenzierbar, wenn es eine Partition $-\infty = x_0 < x_1 < \ldots < x_{n-1} < x_n = \infty$ gibt, sodass f auf jedem Intervall (x_i, x_{i+1}) stetig differenzierbar ist.**

S 3.39: **Sei X eine Zufallsvariable und sei ihre Verteilungsfunktion F_X stetig und stückweise stetig differenzierbar auf einer Partition $-\infty = x_0 < x_1 < \ldots < x_{n-1} < x_n = \infty$. Dann ist X eine stetige Zufallsvariable und die Dichtefunktion f_X kann wie folgt konstruiert werden,**

$$f_X(x) = \begin{cases} F'_X(x) & \exists k \in \{0, \ldots, n-1\} : x \in (x_k, x_{k+1}) \\ a_k & x \in \{x_1, \ldots, x_{n-1}\}, \end{cases}$$

wobei die Werte a_k beliebig gewählt werden dürfen. In anderen Worten, es gilt $f_X(x) = F'_X(x)$ in allen Stetigkeitspunkten x von f_X.

Inverse

Def 2.30. (Verallgem einerte inverse Verteilungsfunktion): **Die verallgemeinerte inverse Verteilungsfunktion von F ist eine Abbildung $F^{-1} : (0, 1) \to \mathbb{R}$ definiert durch**

$$F^{-1}(\alpha) = \inf\{x \in \mathbb{R} \mid F(x) \geq \alpha\}.$$

Nach Definition des Infimums und unter Verwendung der Rechtsstetigkeit von F gilt für jedes $x \in \mathbb{R}$ und $\alpha \in (0, 1)$, dass

$$F^{-1}(\alpha) \leq x \iff \alpha \leq F(x).$$

S 2.31. (Inversionsmethode) : **Sei $F : \mathbb{R} \to [0, 1]$ eine Abbildung mit den Eigenschaften (i)-(iii) aus Satz 2.13. Sei $U \sim \mathcal{U}([0, 1])$. Dann hat die Zufallsvariable $X := F^{-1}(U)$ die Verteilungsfunktion F.**

Zusammengesetzte

Ex : **Sei X mit F_X gegeben und $Y \sim \phi(X)$. Dann kann $F_Y(y)$ berechnen mit:**

1. **Schreibe $F_Y(y) = \mathbb{P}[X \leq \phi^{-1}(y)]$ (Falls quadratisch: $\mathbb{P}[-\surd \leq X \leq \surd]$)**

2. **Finde Grenzen: $F_Y(y) \overset{!}{=} F_X(x) = 0, 1$**

3. **Schreibe $F_Y(y) = \begin{cases} 0 & x < c_0 \\ F_X(\phi^{-1}(y)) & 0 \leq x \leq 1 \\ 1 & x > c_1 \end{cases}$**

Verteilungen

Bernoulli ($X \sim \mathrm{Ber}(p)$)

Def 2.24. (Bernoulli) : **Sei $p \in [0, 1]$. Eine Zufallsvariable X heisst Bernoulli Zufallsvariable mit Parameter p, wenn gilt**

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

S 2.26. (Existenzsatz von Kolmogorov): **Es existiert ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ und eine unendliche Folge von unabhängig und gleichverteilte Bernoulli Zufallsvariablen $X_1, X_2, \ldots$ auf $(\Omega, \mathcal{F}, \mathbb{P})$ mit Parameter $\frac{1}{2}$.**

Note MLE:

1. $\ell = \sum_{i=1}^{n} (k_i \log(p) + (1 - k_i) \log(1 - p))$

2. $\frac{\partial \ell}{\partial p} = \frac{1}{p} \sum_{i=1}^{n} k_i - \frac{1}{1-p} \sum_{i=1}^{n} (1 - k_i)$

3. $T_{MLE} = \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$

4. $\mathrm{Bias}(T_{MLE}) = 0$ **somit Erwartungstreu**

5. **Konsistent**

Note Momenterzeugende:

$$M_X(t) = \mathbb{E}[e^{tX}]$$

Da X die Werte 0 und 1 annimmt,

$$M_X(t) = e^{t \cdot 0} \cdot P(X = 0) + e^{t \cdot 1} \cdot P(X = 1)$$

$$M_X(t) = e^{0} \cdot (1 - p) + e^{t} \cdot p$$

$$M_X(t) = (1 - p) + pe^{t}$$

Binomial ($X \sim \mathrm{Bin}(n, p)$)

Def Binomialverteilung: **Wiederholung von n unabhängigen Bernoulli-Experimenten mit gleichem Parameter p.**

$$p(k) := \mathbb{P}(X = k) = \binom{n}{k} \cdot p^{k} \cdot (1 - p)^{n-k}$$

mit $\forall k \in \{0, 1, \ldots, n\}$

Note Eigenschaften: **Es gilt**

1. **Für $X_n \sim \mathrm{Bin}(n, \frac{\lambda}{n})$ gilt $\lim_{n \to \infty} \mathbb{P}(X_n = k) = \mathbb{P}(Y = k)$ wobei $Y \sim \mathrm{Poisson}(\lambda)$.**

Proof :

1. $P[X_n = k] = \binom{n}{k} \left(1 - \frac{1}{n}\right)^{n-k} \frac{1}{n^k} = \frac{n!}{k!(n-k)!} \frac{1}{n^k} \left(1 - \frac{1}{n}\right)^{n} \left(1 - \frac{1}{n}\right)^{-k}$. **Nun gilt,** $\left(1 - \frac{1}{n}\right)^{n} \to e^{-1}$ **und** $\left(1 - \frac{1}{n}\right)^{-k} \to 1$ **wenn** $n \to \infty$, **und** $\frac{n!}{k!(n-k)!} \frac{1}{n^k} \to 1$, **da** $\frac{n!}{n^k (n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{n^k}$, **und** $\frac{(n-k+1)^k}{n^k} \leq \frac{n(n-1)\cdots(n-k+1)}{n^k} \leq \frac{n^k}{n^k}$, **wobei der linke und rechte Ausdruck jeweils Grenzwert 1 haben. Schliesslich,** $\lim_{n \to \infty} P[X_n = k] = \frac{1}{k!} e^{-1}$, **und $X \sim Poi(1)$ verteilt ist.**

Note MLE:

1. $\ell = k \log(p) + (N - k) \log(1 - p) + \log\left(\binom{N}{k}\right)$

2. $\frac{\partial \ell}{\partial p} = \frac{k}{p} - \frac{N-k}{1-p}$,
$\frac{\partial \ell}{\partial n} = k \log(p) + (N - k) \log(1 - p) + \frac{\partial}{\partial N} \log\left(\binom{N}{k}\right)$

3. $T_{MLE,p} = \frac{k}{N}$, $\quad T_{MLE,n}$ **hat keine geschlossene Form**

4. $\mathrm{Bias}(T_{MLE,p}) = 0$ **somit Erwartungstreu,** $\mathrm{Bias}(T_{MLE,n})$ **schwer zu berechnen :/**

5. **Konsistent**

Note Momenterzeugende:

$$\begin{aligned} M_X(t) &= \mathbb{E}[e^{tX}] \\ &= \sum_{k=0}^{n} e^{tk} \cdot P(X = k) \\ &= \sum_{k=0}^{n} e^{tk} \cdot \binom{n}{k} p^{k} (1 - p)^{n-k} \\ &= \sum_{k=0}^{n} \binom{n}{k} (pe^{t})^{k} (1 - p)^{n-k} \\ &= \left((1 - p) + pe^{t}\right)^{n} \end{aligned}$$

Geometrisch ($X \sim \mathrm{Geo}(p)$)

Def Geometrische Verteilung: **Warten auf den 1-ten Erfolg.**

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

Note Eigenschaften:

1. **Die Verteilung erfüllt Gedächtnislosigkeit:**

$$P[X > t + s \mid X > s] = P[X > t]$$

Proof :

1. $P[Z > n + k \mid Z > k] = \frac{P[Z > n+k, Z > k]}{P[Z > k]} = \frac{P[Z > n+k]}{P[Z > k]} = \frac{(1-q)^{n+k}}{(1-q)^{k}} = (1 - q)^{n} = P[Z > n]$

Note MLE:

1. $\ell = n \log(p) + \sum_{i=1}^{n} (x_i - 1) \log(1 - p)$

2. $\frac{\partial \ell}{\partial p} = \frac{n}{p} - \frac{\sum_{i=1}^{n} (x_i - 1)}{1-p}$

3. $T_{MLE} = \frac{1}{\bar{X}} = \frac{n}{\sum_{i=1}^{n} x_i}$

4. $\mathrm{Bias}(T_{MLE}) > 0$ **somit nicht Erwartungstreu**

5. **Konsistent**

Note Momenterzeugende:

$$\begin{aligned} M_X(t) &= \mathbb{E}[e^{tX}] \\ &= \sum_{k=0}^{\infty} e^{tk} \cdot P(X = k) \\ &= \sum_{k=0}^{\infty} e^{tk} \cdot (1 - p)^{k} \cdot p \\ &= p \sum_{k=0}^{\infty} (e^{t}(1 - p))^{k} \\ &= p \cdot \frac{1}{1 - e^{t}(1 - p)} \\ &= \frac{p}{1 - (1 - p)e^{t}} \end{aligned}$$

Poisson ($X \sim \mathrm{Poi}(\lambda)$)

Def Poisson-Verteilung:
Grenzwert der Binomialverteilung für grosse n und kleine p.

$$p(k) := \mathbb{P}(X = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \quad \forall k \in_0, \lambda > 0$$

Note Eigenschaften:

1. **Seien $X_1 \sim$ Poisson(λ_1) und $X_2 \sim$ Poisson(λ_2) unabhängig. Dann gilt $(X_1 + X_2) \sim$ Poisson$(\lambda_1 + \lambda_2)$.**
2. **Seien $X, Y \sim \mathrm{Poi}(\lambda), \mathrm{Poi}(\mu)$ Es gilt:** $P[X = k \mid X + Y = n] =$ $\binom{n}{k} \left(\frac{\lambda}{\lambda+\mu}\right)^k \left(1 - \frac{\lambda}{\lambda+\mu}\right)^{n-k} = \mathrm{bin}(n, \frac{\lambda}{\lambda+\mu})$

Proof :

1. **Für** $k \in \mathbb{N}_0$ $P[X + Y = k] = \sum_{l=0}^{\infty} P[X + Y = k, Y = l]$ **(Totale Wahrscheinlichkeit)**
$= \sum_{l=0}^{k} P[X + Y = k, Y = l] = \sum_{l=0}^{k} P[X = k - l, Y = l]$
$= \sum_{l=0}^{k} P[X = k - l] \cdot P[Y = l]$
$= \sum_{l=0}^{k} e^{-\lambda} \frac{\lambda^{k-l}}{(k-l)!} \cdot e^{-\mu} \frac{\mu^l}{l!}$
$= e^{-(\lambda+\mu)} \frac{1}{k!} \sum_{l=0}^{k} \frac{k!}{(k-l)!l!} \lambda^{k-l} \mu^l$
$= e^{-(\lambda+\mu)} \frac{1}{k!} \sum_{l=0}^{k} \binom{k}{l} \lambda^{k-l} \mu^l = e^{-(\lambda+\mu)} \frac{(\lambda+\mu)^k}{k!}$
(binomischer Satz)
2. **Sei** $k \in \{0, 1, \ldots, n\}$**:**
$P[X = k | X + Y = n] = \frac{P[X=k, X+Y=n]}{P[X+Y=n]} =$
$\frac{P[X=k, Y=n-k]}{P[X+Y=n]} = \frac{P[X=k]P[Y=n-k]}{P[X+Y=n]}$ **(Def. bedingte Wahrscheinlichkeit, Unabhängigkeit von X, Y**
$= \frac{e^{-\lambda} \frac{\lambda^k}{k!} \cdot e^{-\mu} \frac{\mu^{n-k}}{(n-k)!}}{e^{-(\lambda+\mu)} \frac{(\lambda+\mu)^n}{n!}} = \frac{e^{-\lambda} e^{-\mu}}{e^{-(\lambda+\mu)}} \cdot \frac{n!}{k!(n-k)!} \cdot$
$\frac{\lambda^k \mu^{n-k}}{(\lambda+\mu)^n} = \binom{n}{k} \cdot \frac{\lambda^k}{(\lambda+\mu)^k} \cdot \frac{\mu^{n-k}}{(\lambda+\mu)^{n-k}} = \binom{n}{k} \cdot (\frac{\lambda}{\lambda+\mu})^k \cdot$
$(\frac{\lambda+\mu-\lambda}{\lambda+\mu})^{n-k} = \binom{n}{k} \cdot (\frac{\lambda}{\lambda+\mu})^k \cdot (1 - \frac{\lambda}{\lambda+\mu})^{n-k}$
(Umformen)

Note MLE:

1. $\ell = \sum_{i=1}^{n} (x_i \log(\lambda) - \lambda - \log(x_i!))$
2. $\frac{\partial \ell}{\partial \lambda} = \frac{1}{\lambda} \sum_{i=1}^{n} x_i - n$
3. $T_{MLE} = \bar{X} = \frac{1}{n} \sum_{i=1}^{n} x_i$
4. $\mathrm{Bias}(T_{MLE}) = 0$ **somit Erwartungstreu**
5. **Konsistent**

Proof Log-Likelihood: **Die log-Likelihood-Funktion für die Poissonverteilung lautet**

$$\begin{aligned}\log L(k_1, \ldots, k_n; \lambda) &= \log\left(\prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{k_i}}{k_i!}\right) \\ &= \sum_{i=1}^{n} \log\left(e^{-\lambda} \frac{\lambda^{k_i}}{k_i!}\right) \\ &= -n\lambda + \log(\lambda) \sum_{i=1}^{n} k_i - \sum_{i=1}^{n} \log(k_i!) .\end{aligned}$$

Die Ableitung nach λ ist

$$\frac{\partial}{\partial \lambda} \log L(k_1, \ldots, k_n; \lambda) = -n + \frac{1}{\lambda} \sum_{i=1}^{n} k_i$$

und diese ist 0 für

$$\lambda = \frac{1}{n} \sum_{i=1}^{n} k_i .$$

Also ist der ML-Schätzer für λ gleich $T = \frac{1}{n} \sum_{i=1}^{n} X_i$.

Note Momenterzeugende:

$$\begin{aligned}M_X(t) &= \mathbb{E}[e^{tX}] \\ &= \sum_{k=0}^{\infty} e^{tk} \cdot P(X = k) \\ &= \sum_{k=0}^{\infty} e^{tk} \cdot \frac{\lambda^k e^{-\lambda}}{k!} \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{k!} \\ &= e^{-\lambda} \cdot e^{\lambda e^t} \\ &= e^{\lambda(e^t - 1)}\end{aligned}$$

Gleichverteilt ($X \sim \mathcal{U}([a, b])$)

Def 2.27. (Gleichverteilt): **Eine Zufallsvariable U heisst gleichverteilt auf $[0, 1]$, wir schreiben $U \sim \mathcal{U}([0, 1])$, falls ihre Verteilungsfunktion gegeben ist durch**

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

S 2.28.: **Die Abbildung $X : \Omega \to [0, 1]$ definiert in Gleichung $X(\omega) := \sum_{n=1}^{\infty} 2^{-n} X_n(\omega)$ ist eine gleichverteilte Zufallsvariable auf $[0, 1]$.**
2.29. (Binärdarstellung) **Jedes $x \in [0, 1)$ kann eindeutig in der Form**

$$x = \sum_{n=1}^{\infty} 2^{-n} x_n$$

dargestellt werden, wobei für alle $n \in \mathbb{N}$ gilt, $x_n \in \{0, 1\}$, und für jedes $N \in \mathbb{N}$ gibt es ein $k > N$, sodass $x_k = 0$ (also die Folge "endet" nicht in unendlichen vielen 1-en.) Die Folge $\{x_n\}_{n \in \mathbb{N}}$ heisst Binärdarstellung von x und wir schreiben $x = (.x_1 x_2 x_3 \ldots)_2$.
S 3.3. (Wahrscheinlichkeit eines Punktes):
Sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable mit Verteilungsfunktion F. Für jedes $x \in \mathbb{R}$ gilt $\mathbb{P}[X = x] = F(x) - F(x-)$
Ex : **Seien $X, Y, Z \sim \mathcal{U}(0, 1)$. Es gilt:**
$\mathbb{P}[X > Y] = \int\int_{(x,y) \in [0,1]^2} 1_{(x>y)} f_{X,Y}(x, y) dx dy =$
$\int_0^1 \int_0^x 1 dy dx = \int_0^1 x dx = \frac{1}{2}$
$\mathbb{P}[X > Y, X > Z] =$
$\int\int\int_{(x,y,z) \in [0,1]^3} 1_{(x>y \wedge x>z)} f_{X,Y,Z}(x, y, z) dz dy dx =$
$\int_0^1 \int_0^x \int_0^x 1 dz dy dx = \int_0^1 \int_0^x x dy dx = \int_0^1 x^2 dx = \frac{1}{3}$
$\mathbb{P}[X > Y \mid X > Z] = \frac{\mathbb{P}[X>Y, X>Z]}{\mathbb{P}[X>Y]} = \frac{2}{3}$
Ex : **Seien $X \sim \mathcal{U}(a, b)$ und $X \sim \mathcal{U}(c, d)$ Es gilt**
$\mathbb{P}[X > Y] = \frac{c-a}{b-a} + \frac{(b-c)^2}{2(b-a)(d-c)}$. **Des weiteren:**

$$f_{X+Y}(z) = \begin{cases} \frac{z-(a+c)}{(b-a)(d-c)} & \text{if } a + c \le z < a + d \\ \frac{1}{(b-a)(d-c)} & \text{if } a + d \le z < b + c \\ \frac{(b+a)-z}{(b-a)(d-c)} & \text{if } b + c \le z < b + d \end{cases}$$

Note MLE:

1. $\ell = -n \log(b - a)$ **für** $a \le \min(x_1, x_2, \ldots, x_n)$ **und** $b \ge \max(x_1, x_2, \ldots, x_n)$**, ansonsten** $\ell = -\infty$
2. $\frac{\partial \ell}{\partial a} = \frac{n}{a-b}$, $\quad \frac{\partial \ell}{\partial b} = \frac{n}{b-a}$
3. $T_{MLE,a} = \min(x_1, x_2, \ldots, x_n)$, $T_{MLE,b} = \max(x_1, x_2, \ldots, x_n)$
4. $\mathrm{Bias}(T_{MLE,a}) = \frac{a-b}{n+1}$, $\quad \mathrm{Bias}(T_{MLE,b}) = \frac{b-a}{n+1}$, **somit beide nicht Erwartungstreu**
5. **Konsistent**

Note Momenterzeugende:

$$\begin{aligned}M_X(t) &= \mathbb{E}[e^{tX}] \\ &= \int_a^b e^{tx} \cdot \frac{1}{b-a}, dx \\ &= \frac{1}{b-a} \int_a^b e^{tx}, dx \\ &= \frac{1}{b-a} \left[\frac{e^{tx}}{t}\right]_a^b \\ &= \frac{1}{b-a} \left(\frac{e^{tb} - e^{ta}}{t}\right) \\ &= \frac{e^{tb} - e^{ta}}{t(b-a)}\end{aligned}$$

Exponential ($X \sim \mathrm{Exp}(\lambda)$)

Note Eigenschaften:

1. **Die Exponential-Verteilung erfüllt Gedächtnislosigkeit:**
$$P[X > t + s \mid X > s] = P[X > t]$$
2. **Seien $X, Y \sim \mathrm{Exp}(\lambda_X), \mathrm{Exp}(\lambda_Y)$ Es gilt**
$$\mathbb{P}[X > Y] = \frac{\lambda_Y}{\lambda_X + \lambda_Y}$$

Proof :

1. $\mathbb{P}[T > s + t \mid T > s] = \frac{\mathbb{P}[T>s+t \cap T>s]}{\mathbb{P}[T>s]} = \frac{\mathbb{P}[T>s+t]}{\mathbb{P}[T>s]} =$
$\frac{1-\mathbb{P}[T \le s+t]}{1-\mathbb{P}[T \le s]} = \frac{1-\left(1-e^{-\alpha(s+t)}\right)}{1-\left(1-e^{-\alpha s}\right)} = \frac{e^{-\alpha(s+t)}}{e^{-\alpha s}} =$
$e^{-\alpha t} = 1 - \mathbb{P}[T \le t] = \mathbb{P}[T > t]$
2. $\mathbb{P}[Z > Y] = \int_{-\infty}^{\infty} \underbrace{\mathbb{P}[Z > y]}_{=e^{-\mu y}} \cdot \underbrace{f_Y(y)}_{=\lambda e^{-\lambda y}} \, \mathrm{d}y =$
$\int_0^{\infty} e^{-\mu y} \lambda e^{-\lambda y} = \int_0^{\infty} \lambda e^{-\lambda y} e^{-\mu y} =$
$-\frac{\lambda}{\lambda+\mu} \int_{u(0)}^{u(\infty)} e^u \, \mathrm{d}u = -\frac{\lambda}{\lambda+\mu} \left[e^{-(\lambda+\mu)y}\right]_0^{\infty} =$
$-\frac{\lambda}{\lambda+\mu} (0 - 1) = \frac{\lambda}{\lambda+\mu}$

Note MLE:

1. $\ell = n \log(\lambda) - \lambda \sum_{i=1}^{n} x_i$
2. $\frac{\partial \ell}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i$
3. $T_{MLE} = \frac{1}{\bar{X}} = \frac{n}{\sum_{i=1}^{n} x_i}$
4. $\mathrm{Bias}(T_{MLE}) = 0$ **somit Erwartungstreu**
5. **Konsistent**

Note Momenterzeugende:

$$\begin{aligned}M_X(t) &= \mathbb{E}[e^{tX}] \\ &= \int_0^{\infty} e^{tx} \cdot \lambda e^{-\lambda x}, dx \\ &= \lambda \int_0^{\infty} e^{(t-\lambda)x}, dx \\ &= \lambda \left[\frac{e^{(t-\lambda)x}}{t - \lambda}\right]_0^{\infty} \\ &= \lambda \left(\frac{1}{t - \lambda}\right) \quad \textbf{(für } t < \lambda\textbf{)} \\ &= \frac{\lambda}{\lambda - t}\end{aligned}$$

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

Pro 3.49. (Standardnormalverteilung): **Für $X \sim \mathcal{N}(\mu, \sigma^2)$ gilt $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$, also**

$$F_X(x) = \mathbb{P}[X \le x] = \mathbb{P}\left[\frac{X - \mu}{\sigma} \le \frac{x - \mu}{\sigma}\right] = \Phi\left(\frac{x - \mu}{\sigma}\right.$$

Note : **Seien $X_1, \ldots, X_n$ unabhängige normalverteilte ZV mit Parameter $(m_1, \sigma_1^2), \ldots, (m_n, \sigma_n^2)$, dann ist**

$$Z = m_0 + \lambda_1 X_1 + \ldots + \lambda_n X_n$$

Lem 4.21.: **Für das uneigentliche Integral über die gaußsche Glockenkurve gilt**

$$\int_{-\infty}^{\infty} e^{\frac{-x^2}{2\sigma^2}} \, \mathrm{d}x = \sqrt{2\pi\sigma^2}$$

Note Eigenschaften:

1. $\frac{1}{n} (X_1 + \cdots + X_n)^2 \sim \chi_1^2$

Proof :

1. **Wir wissen, dass $X_1 + \cdots + X_n$ einer Normalverteilung folgt mit** $\mathbb{E}[S_n] = \mathbb{E}[X_1 + X_2 + \cdots + X_n] = n \cdot \mathbb{E}[X_i] = 0, \mathbb{V}(S_n) = \mathbb{V}(X_1 + X_2 + \cdots + X_n) = n \cdot \mathbb{V}(X_i) = n \cdot 1 = n$ **Somit ist $S_n \sim \mathcal{N}(\mathbb{E}[S_n], \mathbb{V}[S_n]) = \mathcal{N}(0, n)$. Wir standardisieren S_n und erhalten** $\frac{S_n - 0}{\sqrt{n}} \sim \mathcal{N}(0, 1)$
Dann $\chi_1^2 \sim Z^2 \sim \left(\frac{S_n}{\sqrt{n}}\right)^2$ **und somit**
$\left(\frac{S_n}{\sqrt{n}}\right)^2 \overset{\text{Def}}{=} \left(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\right)^2 = \frac{(X_1 + \cdots + X_n)^2}{n}$

Note MLE:

1. $\ell = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2$
2. $\frac{\partial \ell}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu)$, $\frac{\partial \ell}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (x_i - \mu)^2$
3. $T_{MLE,\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^{n} x_i$, $T_{MLE,\sigma^2} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{X})^2$

4. $\text{Bias}(T_{MLE,\mu}) = 0$, $\text{Bias}(T_{MLE,\sigma^2}) = -\frac{\sigma^2}{n}$ **somit μ ist Erwartungstreu, σ^2 ist nicht Erwartungstreu**

5. **Konsistent**

Note Momenterzeugende:

$$
\begin{aligned}
&M_X(t) \\
&= \mathbb{E}[e^{tX}] \\
&= \int_{-\infty}^{\infty} e^{tx} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{tx-\frac{(x-\mu)^2}{2\sigma^2}}, dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{(x^2-2\mu x+\mu^2)-2\sigma^2 tx+2\sigma^2 t\mu}{2\sigma^2}}, dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{(x-(\mu+\sigma^2 t))^2}{2\sigma^2}} e^{\frac{\mu^2+2\sigma^2 t\mu-2\sigma^2 t\mu}{2\sigma^2}}, dx \\
&= e^{\mu t+\frac{1}{2}\sigma^2 t^2}
\end{aligned}
$$

Chi-Quadrat-Verteilung ($\mathcal{X}^2$)

Def : **Es gilt $\mathcal{X}^2 = Z_1^2 + \cdots Z_n^2$ mit $Z_1, \ldots, Z_n \sim \mathcal{N}(0,1)$**

Note Eigenschaften:

1. $\chi_2^2 = X_1^2 + X_2^2 = \text{Exp}(\frac{1}{2})$ **mit** $X_1, X_2 \sim \mathcal{N}(0,1)$

Proof :

1. **Formel for χ^2 numerisch mit $n = 2$ ausrechnen.**

Extremwertverteilt ($\mathcal{X}^2$)

Def : **Es gilt:**

$$F_\alpha(x) = e^{-e^{-(x-\alpha)}}$$

Note Dichtefunktion:

$$f_\alpha(x) = F'_\alpha(x) = e^{-(x-\alpha)} e^{-e^{-(x-\alpha)}}$$

Note Erwartungswert:

$$E[X] = \gamma + \alpha$$

mit $\gamma = -\int_0^\infty \log x e^{-x} dx$

Proof : $E[X] = \int_{\mathbb{R}} x f_\alpha(x) dx$
$= \int_{\mathbb{R}} (x-\alpha) f_\alpha(x) dx + \alpha \int_{\mathbb{R}} f_\alpha(x) dx$
$= \int_{\mathbb{R}} (x-\alpha) e^{-(x-\alpha)} e^{-e^{-(x-\alpha)}} dx + \alpha$
$= \int_{\mathbb{R}} s e^{-s} e^{-e^{-s}} ds + \alpha$
Durch die Variablensubstitution $u = e^{-s}$ mit $s = -\log u$ und $\frac{ds}{du} = -\frac{1}{u}$ erhält man für das verbleibende Integral $\int_{\mathbb{R}} s e^{-s} e^{-e^{-s}} ds = \int_\infty^0 (-\log u) u e^{-u} \left(-\frac{1}{u}\right) du = -\int_0^\infty \log u e^{-u} du = \gamma$

Kompositionen

- $\mathbb{P}[X > Y] = \sum_{j=1}^n \mathbb{P}[X > Y \mid Y = j] \mathbb{P}[Y = j]$ **(diskret)**
- $\mathbb{P}[X > Y] = \int_{-\infty}^\infty \mathbb{P}[X > Y \mid Y = y] f_Y(y) dy$ **(stetig)**
- **Sei X mit f_x eine ZV und $Y \sim aX + b$. Dann**

Erwartungswert

Def 4.1. (Erwartungswert (nicht-negativ)) : **Sei $X : \Omega \to \mathbb{R}_+$ eine Zufallsvariable mit nicht-negativen Werten und Verteilungsfunktion F_X. Dann heisst**

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

der Erwartungswert von X (expected value).
S 4.3.: **Sei X eine nicht-negative Zufallsvariable. Dann gilt $\mathbb{E}[X] \geq 0$. Gleichheit gilt genau dann wenn $X = 0$ fast sicher gilt.**
Def 4.4. (Allgemeiner Erwartungswert): **Sei X eine reellwertige Zufallsvariable. Falls $\mathbb{E}[|X|] < \infty$, dann heisst**

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

der Erwartungswert von X.

Diskret

S 4.8 (Erwartungswert (diskret)) : **Sei $X : \Omega \to \mathbb{R}$ eine diskrete Zufallsvariable, deren Werte fast sicher in W (endlich oder abzählbar) liegen. Dann gilt**

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

solange der Erwartungswert wohldefiniert ist. Somit gilt auch

$$\mathbb{E}[a + b \cdot X^c] = a + \sum_{x \in W} b \cdot x^c p_X(x)$$

Stetig

S 4.17. (Erwartungswert (stetig)) : **Sei X eine stetige Zufallsvariable mit Dichte f_X. Dann gilt**

$$\mathbb{E}[X] = \int_{-\infty}^\infty x f_X(x) \mathrm{d}x$$

solange das Integral absolut konvergiert, d.h. falls $\int_{-\infty}^\infty |x| f_X(x) \mathrm{d}x < \infty$. Ist das Integral nicht absolut konvergent, so existiert der Erwartungswert nicht (zumindest nicht in $\mathbb{R}$). Das ist völlig analog zum diskreten Fall. Es gilt ebenfalls

$$\mathbb{E}\left[a + b \cdot X^c\right] = a + \int_{-\infty}^\infty b \cdot x^c \cdot f_X(x) dx$$

Eigenschaften

S 4.18.: **Sei X eine stetige Zufallsvariable mit Dichte f_X und sei $\varphi : \mathbb{R} \to \mathbb{R}$ eine Abbildung. Dann gilt**

$$\mathbb{E}[\varphi(X)] = \int_{-\infty}^\infty \varphi(x) f_X(x) \mathrm{d}x$$

solange das Integral wohldefiniert ist.
S 4.25. (Linearität des Erwartungswerts): **Seien $X, Y : \Omega \to \mathbb{R}$ Zufallsvariablen und sei $\lambda \in \mathbb{R}$. Falls die Erwartungswerte wohldefiniert sind, gilt**

1. $\mathbb{E}[\lambda X] = \lambda \mathbb{E}[X]$

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

S 4.29. (Monotonie des Erwartungswerts) : **Seien X, Y zwei Zufallsvariablen, sodass $X \leq Y$ f.s. gilt. Falls beide Erwartungswerte wohldefiniert sind, gilt**

$$\mathbb{E}[X] \leq \mathbb{E}[Y]$$

Unabhängigkeit

S 4.30 (Erwartungswerte bei Unabhängigkeit): **für n unabhängige Zufallsvariablen $X_1, \ldots, X_n$.**
Satz 4.31. Seien $X_1, \ldots, X_n$ unabhängige Zufallsvariablen mit endlichen Erwartungswerten $\mathbb{E}[X_1], \ldots, \mathbb{E}[X_n]$. Dann gilt

$$\mathbb{E}\left[\prod_{k=1}^n X_k\right] = \prod_{k=1}^n \mathbb{E}[X_k]$$

S 4.32.: **Sei X eine Zufallsvariable und sei $f : \mathbb{R} \to \mathbb{R}_+$ eine Abbildung, sodass $\int_{-\infty}^\infty f(x) \mathrm{d}x = 1$. Dann sind folgende Aussagen äquivalent**

1. **X ist stetig mit Dichte f,**

2. **für jede stückweise stetige, beschränkte Abbildung $\varphi : \mathbb{R} \to \mathbb{R}$ gilt**

$$\mathbb{E}[\varphi(X)] = \int_{-\infty}^\infty \varphi(x) f(x) \mathrm{d}x$$

S 4.33.: **Seien X, Y zwei Zufallsvariablen. Die folgenden Aussagen sind äquivalent:**

1. **X, Y sind unabhängig,**

2. **Für alle stückweise stetigen, beschränkten Abbildungen $\varphi, \psi : \mathbb{R} \to \mathbb{R}$ gilt**

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

S 4.34. : **Seien $X_1, \ldots, X_n$ Zufallsvariablen. Die folgenden Aussagen sind äquivalent:**

1. **$X_1, \ldots, X_n$ sind unabhängig,**

2. **Für alle stückweise stetigen, beschränkten Abbildungen $\varphi_1, \ldots, \varphi_n : \mathbb{R} \to \mathbb{R}$ gilt**

$\mathbb{E}[\varphi_1(X_1) \cdot \ldots \cdot \varphi_n(X_n)] = \mathbb{E}[\varphi_1(X_1)] \cdot \ldots \cdot \mathbb{E}[\varphi_n(X_n)]$.

Ungleichungen

S 4.35. (Markow-Ungleichung) : **Sei X eine nicht-negative Zufallsvariable und sei $g : X(\Omega) \to [0, \infty)$ eine wachsende Funktion. Für jedes $c \in \mathbb{R}$ mit $g(c) > 0$ gilt**

$$\mathbb{P}[X \geq c] \leq \frac{\mathbb{E}[g(X)]}{g(c)}$$

S 4.36. (Jensensche Ungleichung) : **Sei X eine Zufallsvariable und sei $\varphi : \mathbb{R} \to \mathbb{R}$ eine konvexe Funktion. Falls $\mathbb{E}[\varphi(X)]$ und $\mathbb{E}[X]$ wohldefiniert sind, gilt**

$$\varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)]$$

Cor 4.37. (Dreiecksungleichung) : **Anwendung der Jensenschen Ungleichung auf $\varphi(x) = |x|$ liefert die Dreiecksungleichung,**

$$|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$$

Varianz

Def 4.38. (Varianz Standardabweichung): **Sei X eine Zufallsvariable, sodass $\mathbb{E}\left[X^2\right] < \infty$. Wir definieren die Varianz von X (variance) durch**

$$\mathbb{V}[X] = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Die Wurzel der Varianz nennt man Standardabweichung von X (standard deviation) und sie wird oft mit $\sigma = \sigma_X = \sigma(X) = \sqrt{\mathbb{V}[X]}$ bezeichnet.

Eigenschaften

Pro 4.46. :
Seien $X_1, \ldots, X_n$ paarweise unabhängigen Zufallsvariablen, dann gilt

$$\mathbb{V}\left[\sum_{k=1}^n X_k\right] = \sum_{k=1}^n \mathbb{V}[X_k]$$

Note :

1. **Sei X ein ZV, sodass $(X^2) < \infty$ und $a, b \in$:**

$$\mathbb{V}(a \cdot X + b) = a^2 \cdot \mathbb{V}(X)$$

2. **Seien $X_1, \ldots, X_n$ paarweise unabhängig. Dann gilt**

$$\mathbb{V}(X_1 + \ldots + X_n) = \mathbb{V}(X_1) + \ldots + \mathbb{V}(X_n)$$

Cor 4.50. (Chebyshev-Ungleichung): **Sei Y eine Zufallsvariable mit endlicher Varianz. Für jedes $c > 0$ gilt dann**

$$\mathbb{P}[|Y - \mathbb{E}[Y]| \geq c] \leq \frac{\mathbb{V}[Y]}{c^2}$$

Note :

1. **$\textbf{Cov}(X, X) = \mathbb{V}(X)$**

2. **X, Y unabhängig $\implies$ $\textbf{Cov}(X, Y) = 0$ (Die Umkehrung ist falsch!)**

3. $\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2\textbf{Cov}(X, Y)$

4. $\mathbb{V}(X) \geq 0$

Kovarianz

Def 4.51. (Kovarianz): Seien X, Y zwei Zufallsvariablen mit endlichen zweiten Momenten, $\mathbb{E}\left[X^2\right], \mathbb{E}\left[Y^2\right] < \infty$. Die Kovarianz zwischen X und Y (covariance) ist definiert als

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

Note Kovarianz: Man kann die Kovarianz ebenfalls ausdrücken durch

$$\operatorname{cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Beweis:

$$\begin{aligned}
&\operatorname{cov}(X, Y) \\
&= \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \\
&= \mathbb{E}[(XY - \mathbb{E}[X]Y - \mathbb{E}[Y]X + \mathbb{E}[X]\mathbb{E}[Y]] \\
&\overset{4.25}{=} \mathbb{E}[XY] - \mathbb{E}[\mathbb{E}[X]Y] - \mathbb{E}[\mathbb{E}[Y]X] + \mathbb{E}[\mathbb{E}[X]\mathbb{E}[Y]] \\
&\overset{(1)}{=} \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Y] \\
&= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]
\end{aligned}$$

Note Kolleriertheit: Es gilt:
- Wenn $\operatorname{cov}(X, Y) > 0$, dann sind X und Y positiv korreliert.
- wenn $\operatorname{cov}(X, Y) = 0$, dann sind X und Y unkorreliert.
- Wenn $\operatorname{cov}(X, Y) < 0$, dann sind X und Y negativ korreliert oder antikorreliert.

Note Eigenschaften: Für die Kovarianz gilt:

1. Positive Semidefinitheit: $\operatorname{cov}(X, X) \geq 0$,
2. Symmetrie: $\operatorname{cov}(X, Y) = \operatorname{cov}(Y, X)$
3. Bilinearität: $\operatorname{cov}(aX + b, cY + d) = ac\operatorname{cov}(X, Y)$ und $\operatorname{cov}(X, (eY + f) + (gZ + h)) = e\operatorname{cov}(X, Y) + g\operatorname{cov}(X, Z)$.

Proof :

1.

$$\begin{aligned}
\operatorname{cov}(X, X) &\overset{4.52}{=} \mathbb{E}[XX] - \mathbb{E}[X]\mathbb{E}[X] \\
&= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&\overset{4.40}{=} \mathbb{V}[X]
\end{aligned}$$

2.

$$\begin{aligned}
\operatorname{cov}(X, Y) &\overset{4.52}{=} \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] && (1) \\
&= \mathbb{E}[YX] - \mathbb{E}[Y]\mathbb{E}[X] && (2) \\
&\overset{4.52}{=} \operatorname{cov}(Y, X) && (3)
\end{aligned}$$

3.

Zusammengesetzte

Note Verteilung einer Ungleichung: Es gilt für beliebige ZV X, Y:

$$\mathbb{P}[X > Y] = \int_{-\infty}^{\infty} \mathbb{P}[X > y] \cdot f_Y(y) \mathrm{d}y$$

Gemeinsame Verteilungen

S 5.3. : Eine gemeinsame Verteilung von Zufallsvariablen $X_1, \ldots, X_n$ erfüllt stets

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

Pro 5.5: Aus der gemeinsamen Gewichtsfunktion p bekommt man die gemeinsame Verteilungsfunktion via

$$\begin{aligned}
F\left(x_1, \ldots, x_n\right) &= \mathbb{P}\left[X_1 \leq x_1, \ldots, X_n \leq x_n\right] \\
&= \sum_{y_1 \leq x_1, \ldots, y_n \leq x_n} \mathbb{P}\left[x_1 = y_1, \ldots, X_n = y_n\right] \\
&= \sum_{y_1 \leq x_1, \ldots, y_n \leq x_n} p\left(y_1, \ldots, y_n\right).
\end{aligned}$$

Def 5.18. (Randverteilung) : Haben X und Y die gemeinsame Verteilungsfunktion F, so ist die Funktion $F_X : \mathbb{R} \to [0, 1]$, gegeben durch

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

die Verteilungsfunktion der Randverteilung von X. Analog ist

$$F_Y : \mathbb{R} \to [0, 1]$$

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

die Verteilungsfunktion der Randverteilung von Y.

Diskret

Def 5.1. (Gemeinsame diskrete Verteilung) : Seien $X_1, \ldots, X_n$ diskrete Zufallsvariablen, seien für $k \in \{1, \ldots, n\}$ Mengen $W_k \subset \mathbb{R}$ endlich oder abzählbar, sodass $X_k \in W_k$ fast sicher gilt. Die gemeinsame Verteilung von $(X_1, \ldots, X_n)$ ist eine Familie von Wahrscheinlichkeiten

$$\left\{p\left(x_1, \ldots, x_n\right)\right\}_{x_1 \in W_1, \ldots, x_n \in W_n}$$

wobei $p : \mathbb{R}^n \to [0, 1]$ die gemeinsame Gewichtsfunktion bezeichnet,

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

S 5.6. (Verteilung des Bildes) : Seien $n \geq 1$ und sei $\varphi : \mathbb{R}^n \to \mathbb{R}$ eine Abbildung. Seien $X_1, \ldots, X_n$ diskrete Zufallsvariablen, mit Werten jeweils in $W_1, \ldots, W_n$ (f.s.). Dann ist $Z = \varphi\left(X_1, \ldots, X_n\right)$ eine diskrete Zufallsvariable, die f.s. Werte in $W = \varphi\left(W_1 \times \ldots \times W_n\right)$ annimmt. Zudem ist die Verteilung von Z für alle $z \in W$ gegeben durch

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

S 5.7. (Randverteilung): Seien $X_1, \ldots, X_n$ diskrete Zufallsvariablen mit gemeinsamer Gewichtsfunktion p. Für jedes $k \in \{1, \ldots, n\}$ und jedes $x \in W_k$ gilt

$$\mathbb{P}\left[X_k = x\right] = \sum_{\substack{x_\ell \in W_\ell \\ \ell \in \{1, \ldots, n\} \setminus \{k\}}} p\left(x_1, \ldots, x_{k-1}, x, x_{k+1}, \ldots, x_n\right)$$

S 5.9. (Erwartungswert des Bildes): Seien $X_1, \ldots, X_n$ diskrete Zufallsvariablen mit gemeinsamer Verteilung $\left\{p\left(x_1, \ldots, x_n\right)\right\}_{x_1 \in W_1, \ldots, x_n \in W_n}$ und sei $\varphi : \mathbb{R}^n \to \mathbb{R}$ eine Abbildung. Es gilt

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

solange die Summe wohldefiniert ist, wobei wir hier über alle $x_1 \in W_1, \ldots, x_n \in W_n$ summieren.

S 5.10. : Seien $X_1, \ldots, X_n$ diskrete Zufallsvariablen mit gemeinsamer Verteilung $\left\{p\left(x_1, \ldots, x_n\right)\right\}_{x_1 \in W_1, \ldots, x_n \in W_n}$. Die folgenden Aussagen sind äquivalent

1. $X_1, \ldots, X_n$ sind unabhängig,
2. für alle $x_1 \in W_1, \ldots, x_n \in W_n$ gilt

$$p\left(x_1, \ldots, x_n\right) = \mathbb{P}\left[X_1 = x_1\right] \cdot \ldots \cdot \mathbb{P}\left[X_n = x_n\right]$$

d.h. die gemeinsame Gewichtsfunktion ist das Produkt der einzelnen Gewichtsfunktionen der Randverteilungen.

Stetig

Def 5.11. (Gemeinsame stetige Verteilung) : Sei $n \geq 1$. Wir sagen, dass die Zufallsvariablen $X_1, \ldots, X_n : \Omega \to \mathbb{R}$ eine stetige gemeinsame Verteilung besitzen, falls eine Abbildung $f : \mathbb{R}^n \to \mathbb{R}_+$existiert, sodass für alle $a_1, \ldots, a_n \in \mathbb{R}$ gilt,$\mathbb{P}\left[X_1 \leq a_1, \ldots, X_n \leq a_n\right] = \int_{-\infty}^{a_1} \cdots \int_{-\infty}^{a_n} f\left(x_1, \ldots, x_n\right) \mathrm{d}x_n \cdots \mathrm{d}x_1$
Die Funktion f heisst die gemeinsame Dichte von $\left(X_1, \ldots, X_n\right)$ (joint probability density function).

S 5.12.: Sei f die gemeinsame Dichte der Zufallsvariablen $\left(X_1, \ldots, X_n\right)$. Dann gilt

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f\left(x_1, \ldots, x_n\right) \mathrm{d}x_n \cdots \mathrm{d}x_1 = 1$$

Umgekehrt kann jeder Funktion $f : \mathbb{R}^n \to \mathbb{R}_+$, die Gleichung oben erfüllt, ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ und n Zufallsvariablen $X_1, \ldots, X_n$ zugeordnet werden, sodass f die gemeinsame Dichte von $X_1, \ldots, X_n$ ist.

S 5.15. (Erwartungswert des Bildes): Sei $\varphi : \mathbb{R}^n \to \mathbb{R}$ eine Abbildung. Falls $X_1, \ldots, X_n$ eine gemeinsame Dichte f besitzen, dann lässt sich der Erwartungswert der Zufallsvariable $\varphi\left(X_1, \ldots, X_n\right)$ berechnen als $\mathbb{E}\left[\varphi\left(X_1, \ldots, X_n\right)\right] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \varphi\left(x_1, \ldots, x_n\right) f\left(x_1, \ldots, x_n\right) \mathrm{d}x_n \cdots \mathrm{d}x_1$

S 5.21. (Unabhängigkeit von stetigen Zufallsvariablen): Seien $X_1, \ldots, X_n$ Zufallsvariablen mit Dichten $f_{X_1}, \ldots, f_{X_n}$. Dann sind die folgenden Aussagen äquivalent,

1. $X_1, \ldots, X_n$ sind unabhängig,
2. $X_1, \ldots, X_n$ sind gemeinsam stetig mit gemeinsamer Dichte $f : \mathbb{R}^n \to \mathbb{R}_+$,

$$f\left(x_1, \ldots, x_n\right) = f_{X_1}\left(x_1\right) \cdot \ldots \cdot f_{X_n}\left(x_n\right)$$

d.h. die gemeinsame Dichtefunktion f ist das Produkt der einzelnen Randdichten f_{X_k}.

Note Randdichten: Haben X und Y eine gemeinsame Dichte $f(x, y)$, so haben auch die Randverteilungen von X und Y Dichten $f_X : \mathbb{R} \to \mathbb{R}_+$bzw. $f_Y : \mathbb{R} \to \mathbb{R}_+$. $f_X(x) = \int_{-\infty}^{\infty} f(x, y) \mathrm{d}y$ und $f_Y(y) = \int_{-\infty}^{\infty} f(x, y) \mathrm{d}x$
Falls bei Verteilungen die Variablen eine Bedingung erfüllen müssen z.b $x^2 + y^2 \leq 1$, dann bekommt man die Grenzen des Integrals durch das umformen der jeweiligen Variabel. Z.b. Um $f_X(x)$ zu erhalten (nach y integrieren): $y^2 \leq 1 - x^2 \Rightarrow y = \pm\sqrt{1 - x^2}$ Obere Grenze $+\sqrt{1 - x^2}$ Untere Grenze $-\sqrt{1 - x^2}$ Falls Grenzen nicht beschränkt sind einfach nach ∞ integrieren z.b. mit $0 < x < y$ gilt $f_Y = \int_x^{\infty} f_{X,Y}(x, y) dy$

Grenzwertsätze

S 6.2. (Schwaches Gesetz der grossen Zahlen) : Sei $X_1, X_2, \ldots$ eine Folge von unabhängigen Zufallsvariablen mit gleichen Erwartungswerten $\mathbb{E}\left[X_k\right] = \mu$ und Varianzen $\mathbb{V}\left[X_k\right] = \sigma^2$. Sei

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

Dann konvergiert $\bar{X}_n$ für $n \to \infty$ in Wahrscheinlichkeit gegen $\mu = \mathbb{E}\left[X_k\right]$, d.h. für jedes $\varepsilon > 0$ gilt

$$\mathbb{P}\left[\left|\bar{X}_n - \mu\right| \varepsilon\right] \xrightarrow{n \to \infty} 0$$

Note :
- X_i, X_j unkorreliert $\Leftrightarrow \operatorname{Cov}\left(X_i, X_j\right) = 0$
- X_i, X_j unabhängig $\Longrightarrow X_i, X_j$ unkorreliert

S 6.5. (Starkes Gesetz der grossen Zahlen): Sei $X_1, X_2, \ldots$ eine Folge von unabhängigen Zufallsvariablen, die alle dieselbe Verteilung haben mit endlichem Erwartungswert $\mathbb{E}\left[X_k\right]$. Für

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

gilt dann

$$\bar{X}_n \xrightarrow{n \to \infty} \mu \quad \mathbb{P}\text{-fast sicher,}$$

das bedeutet,

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

Def 6.7. (Konvergenz in Verteilung): Seien $\left(X_n\right)_{n \in \mathbb{N}}$ und X Zufallsvariablen mit Verteilungsfunktionen $\left(F_n\right)_{n \in \mathbb{N}}$ und F. Wir sagen $\left(X_n\right)_{n \in \mathbb{N}}$ konvergiert in Verteilung gegen X und schreiben

$$X_n \xrightarrow{d} X \text{ für } n \to \infty$$

falls für jeden Stetigkeitspunkt $x \in \mathbb{R}$ von F gilt,

$$\lim_{n \to \infty} F_n(x) = \lim_{n \to \infty} \mathbb{P}\left[X_n \leq x\right] = \mathbb{P}[X \leq x] = F(x)$$

In der Literatur findet man unter anderem folgende Notation,

$$X_n \xrightarrow{w} X \text{ und } \quad X_n \xrightarrow{L} X.$$

Die Buchstaben d, w, L stehen dabei für "convergence in distribution", "weak convergence", bzw. "convergence in law".

S 6.10. (zentraler Grenzwertsatz, ZGS) : Sei $(X_k)_{k\in\mathbb{N}}$ eine Folge von i.i.d. Zufallsvariablen mit $\mathbb{E}[X_k]=\mu$ und $\mathbb{V}[X_k]=\sigma^2$. Für die Partialsummen $S_n=\sum_{k=1}^n X_k$ gilt dann für alle $x\in\mathbb{R}$,

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

wobei Φ die Verteilungsfunktion der $\mathcal{N}(0,1)$-Verteilung ist. und somit $\frac{\frac{1}{n}S_n-\mu}{\frac{\sigma}{\sqrt{n}}}=\frac{S_n-n\mu}{\sigma\sqrt{n}}\xrightarrow{d}\mathcal{N}(0,1)$

Def 6.16. (Momenterzeugende Funktion) : Die momenterzeugende Funktion einer Zufallsvariablen X ist für $t\in\mathbb{R}$ definiert durch

$$M_X(t)=\mathbb{E}\left[e^{tX}\right]\overset{S4.18}{=}\int_{-\infty}^{\infty}e^{tx}f_X(x)dx$$

S 6.19. (Chernoff-Ungleichung): Seien $X_1,\ldots,X_n$ i.i.d. Zufallsvariablen, für welche die momenterzeugende Funktion $M_X(t)$ für alle $t\in\mathbb{R}$ endlich ist. Für jedes $b\in\mathbb{R}$ gilt dann

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

S 6.20. (Chernoff-Schranke): Seien $X_1,\ldots,X_n$ unabhängig mit $X_k\sim\operatorname{Ber}(p_k)$ und sei $S_n=\sum_{k=1}^n X_k$. Sei $\mu_n=\mathbb{E}[S_n]=\sum_{k=1}^n p_k$ und $\delta>0$, dann gilt

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

Statistik

Def 7.2. (Schätzer) : Ein Schätzer ist eine Zufallsvariable der Form

$$T_\ell=t_\ell(X_1,\ldots,X_n)$$

Die Schätzfunktionen $t_\ell:\mathbb{R}^n\to\mathbb{R}$ müssen noch gewählt/gefunden werden. Einsetzen von Daten $x_k=X_k(\omega), k=1,\ldots,n$, liefert dann Schätzwerte $T_\ell(\omega)=t_\ell(x_1,\ldots,x_n)$ für $\vartheta_\ell,\ell=1,\ldots,m$. Der Kürze halber schreiben wir oft auch $T=(T_1,\ldots,T_m)$ und $\vartheta=(\vartheta_1,\ldots,\vartheta_m)$.

Note Fehler: Die Entscheidung bei einem Test kann auf zwei verschiedene Arten falsch herauskommen:

- Fehler 1. Art: die Hypothese wird abgelehnt, obwohl sie richtig ist. Das passiert für $\vartheta\in\Theta_0$ und $T\in K$. $\mathbb{P}_\vartheta[T\in K]$ heisst für $\vartheta\in\Theta_0$ die Wahrscheinlichkeit für einen Fehler 1. Art.
- Fehler 2. Art: die Hypothese wird angenommen, obwohl sie falsch ist. Das passiert für $\vartheta\in\Theta_A$ und $T\notin K$. $\mathbb{P}_\vartheta[T\notin K]=1-\mathbb{P}_\vartheta[T\in K]$ heisst für $\vartheta\in\Theta_A$ die Wahrscheinlichkeit für einen Fehler 2. Art.

Def 7.5. (Erwartungstreue): Ein Schätzer T heisst erwartungstreu für ϑ, falls für alle $\vartheta\in\Theta$ gilt

$$\mathbb{E}_\vartheta[T]=\vartheta$$

Def 7.6. (Bias und MSE) : Sei $\vartheta\in\Theta$ und T ein Schätzer. Der Bias (oder erwartete Schätzfehler) von T im Modell $\mathbb{P}_\vartheta$ ist definiert als

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

Erwartungstreu bedeutet also, dass der Bias gleich Null ist. Der mittlere quadratische Schätzfehler von T im Modell $\mathbb{P}_\vartheta$ ist definiert als

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

Note MSE: Man kann den MSE in folgender Form zerlegen.

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

Def Konsistenz: Eine Folge von Schätzern $T^{(n)},n\in\mathbb{N}$, heisst konsistent für ϑ, falls $T^{(n)}$ für $n\to\infty$ in $\mathbb{P}_\vartheta$-Wahrscheinlichkeit gegen ϑ konvergiert, d.h. für jedes $\vartheta\in\Theta$ und jedes $\varepsilon>0$ gilt

$$\lim_{n\to\infty}\mathbb{P}_\vartheta\left[\left|T^{(n)}-\vartheta\right|>\varepsilon\right]=0$$

Def 7.9. (Likelihood-Funktion) : Die Likelihood-Funktion ist

$$L(x_1,\ldots,x_n;\vartheta)=\begin{cases}p_{\vec{x}}(x_1,\ldots,x_n;\vartheta) & \text{diskret}\\ f_{\vec{x}}(x_1,\ldots,x_n;\vartheta) & \text{stetig}\end{cases}$$

Die Funktion $\log L(x_1,\ldots,x_n;\vartheta)$ heisst log-Likelihood-Funktion und hat im i.i.d.-Fall den Vorteil durch eine Summe gegeben zu sein.

Def 7.10. (ML-Schätzer): Der Maximum-Likelihood-Schätzer T_{ML} für ϑ wird dadurch definiert, dass er die Funktion

$$\vartheta\mapsto L(X_1,\ldots,X_n;\vartheta)$$

über alle ϑ maximiert, d.h.

$$T_{\mathrm{ML}}=t_{\mathrm{TM}}(X_1,\ldots,X_n)\in\underset{\vartheta\in\Theta}{\arg\max}L(X_1,\ldots,X_n;\vartheta)$$

RC Log-MLE:
Gegeben seien $X_1\ldots X_n$ unter P_θ i.i.d. Für ein fixes n gilt: (falls n bekannt, gleich einsetzen)

1. Gemeinsame Dichte finden $g(x_1,\ldots,x_n;)=\prod_{i=1}^n\mathbb{P}[X_i=x_i]$
2. Bestimme $f(\theta)=\log(L(x_1,\ldots,x_n;\theta))$ (θ einsetzen für respective Variable in allen Verteilungen)
3. Maximum von $f(\theta)$ finden ($f(\theta)'=0$ und $f(\theta)''<0$, Ränder Überprüfen)

RC Erwartungstreue eines Schätzers:
Berechne $\mathbb{E}_\lambda[T]-\lambda$. Falls 0, dann ist Schätzer Erwartungstreu. Da X_i i.i.d sind kann man Linearität des Erwargungswertes anwenden.

RC Konsistenz eines Schätzers:
Benutze Chebyshev

$$\mathbb{P}_p[|T_n-\lambda|>\varepsilon]\leq\frac{\operatorname{Var}(T_n)}{\varepsilon^2}$$

Varianz des Schätzers durch Proposition 4.46 ausrechenbar, da i.i.d.
Falls $\lim n\to\infty\frac{\operatorname{Var}(T_n)}{\varepsilon^2}=0$ dann konsistent

---

Note MLE-Schätzer:

- $X_1,\ldots,X_n\sim\operatorname{Ber}(\theta)$ iid.: $T=\frac{1}{n}\sum_{i=1}^n X_i$
- $X_1,\ldots,X_n\sim\operatorname{Exp}(\theta)$ iid.: $T=\frac{\sum_{i=1}^n X_i}{n}=\bar{X}_n$
- $X_1,\ldots,X_n\sim\operatorname{Geo}(\theta)$ iid.: $T=\frac{n}{\sum_{i=1}^n X_i}=\frac{1}{\bar{X}_n}$
- $X_1,\ldots,X_n\sim\operatorname{Bin}(N,\theta)$ iid.: $T=\frac{1}{N}\frac{\sum_{i=1}^n X_i}{n}$
- $X_1,\ldots,X_n\sim Poi(\theta)$ iid.: $T=\frac{\sum_{i=1}^n X_i}{n}=\bar{X}_n$
- $X_1,\ldots,X_n\sim\mathcal{U}([\theta_1,\theta_2])$ iid.: $T_{\theta_1}=\max(X_i),T_{\theta_2}=\min(X_i)$
- $X_1,\ldots,X_n\sim\mathcal{N}(\theta_1,\theta_2)$ iid. : $T_{\theta_1}=\bar{X}_n,T_{\theta_2}=S^2$

Def 7.18. (Studentsche t-Verteilung): Eine stetige Zufallsvariable X heisst t-verteilt mit m Freiheitsgraden falls ihre Dichte für $x\in\mathbb{R}$ gegeben ist durch

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

Wir schreiben dann $X\sim t_m$.

S 7.20.:
Seien $X_1,\ldots,X_n$ i.i.d. $\sim\mathcal{N}(\mu,\sigma^2)$, und

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

Es gelten folgende Aussagen:

1. $\bar{X}_n\sim\mathcal{N}(\mu,\frac{1}{n}\sigma^2)$, also $\frac{\bar{X}_n-\mu}{\frac{\sigma}{\sqrt{n}}}\sim\mathcal{N}(0,1)$.
2. $\frac{n-1}{\sigma^2}S^2=\frac{1}{\sigma^2}\sum_{k=1}^n(X_k-\bar{X}_n)^2\sim\chi^2_{n-1}$.
3. $\bar{X}_n$ und S^2 sind unabhängig.
4. $\frac{\bar{X}_n-\mu}{\frac{S}{\sqrt{n}}}=\frac{\frac{\bar{X}_n-\mu}{\sigma/\sqrt{n}}}{S/\sigma}=\frac{\frac{\bar{X}_n-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{1}{n-1}\frac{n-1}{\sigma^2}S^2}}\sim t_{n-1}$.

Testing

Def Entscheidungsregel: Die Hypothese H_0 wird verworfen, wenn $T(\omega)\in K$, Die Hypothese H_0 wird nicht verworfen bzw. angenommen, wenn $T(\omega)\notin K$

Def Signifikanzniveau: Für einen Test wählt zuerst ein Signifikanzniveau $\alpha\in(0,1)$ und verlangt

$$\sup_{\vartheta\in\Theta_0}\mathbb{P}_\vartheta[T\in K]\leq\alpha.$$

Man kontrolliert also die Wahrscheinlichkeit für einen Fehler 1. Art durch α. Für $\alpha\downarrow$ gilt

- Prob für Fehler 1. Art wird kleiner
- Verwerfungsbereich muss kleiner gewählt werden
- Prob für Fehler 2. Art wird grösser
- Macht des Tests wird kleiner

Def 8.7. (Likelihood-Quotient): Für $\vartheta_0\in\Theta_0,\vartheta_A\in\Theta_A$ und $x_1,\ldots,x_n\in\mathbb{R}$ definieren wir den Likelihood-Quotienten durch

$$R(x_1,\ldots,x_n;\vartheta_0,\vartheta_A)=\frac{L(x_1,\ldots,x_n;\vartheta_A)}{L(x_1,\ldots,x_n;\vartheta_0)}.$$

Als Konvention setzen wir $R(x_1,\ldots,x_n;\vartheta_0,\vartheta_A)=+\infty$, wenn $L(x_1,\ldots,x_n;\vartheta_0)=0$

Def 8.8. (Likelihood-Quotienten-Test): Sei $c\geq0$. Der Likelihood-Quotienten-Test mit Parameter c ist ein Test (T,K) mit Teststatistik $T=R(X_1,\ldots,X_n;\vartheta_0,\vartheta_A)$ und Verwerfungsbereich $K=(c,\infty]$.

Lem 8.9. (Neyman-Pearson-Lemma): Sei $\Theta_0=\{\vartheta_0\}$ und $\Theta_A=\{\vartheta_A\}$. Sei (T,K) ein Likelihood-Quotienten-Test mit Parameter c und Signifikanzniveau $\alpha^*:=\mathbb{P}_{\vartheta_0}[T\in K]$. Ist (T',K') ein anderer Test mit Signifikanzniveau $\mathbb{P}_{\vartheta_0}[T'\in K']=:\alpha\leq\alpha^*$, so gilt auch

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

Das bedeutet, jeder andere Test mit kleinerem Signifikanzniveau hat auch kleinere Macht bzw. grössere Wahrscheinlichkeit für einen Fehler 2. Art.

Def 8.10. (Verallgemeinerung des Likelihood-Quotient): Der verallgemeinerte Likelihood-Quotient ist gegeben durch

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

oder auch

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

RC Z-Test:
Gegeben:

- Daten: x_i und $\bar{x}=\frac{\sum^{i=1}{}_n x_i}{n}$
- Nullhypothese : μ_0
- Alternativhypothese :
  - $\mu>\mu_0$ (rechtsseitig)
  - $\mu<\mu_0$ (linksseitig)
  - $\mu\neq\mu_0$ (beidseitig)
- Signifikanzniveau : α

1. Berechne Z-Score

$$Z=\frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$$

2. Berechne z
   - $z=z_{1-\alpha}=\Phi^{-1}(1-\alpha)$ (rechtsseitig)
   - $z=z_\alpha=\Phi^{-1}(\alpha)$ (linksseitig)
   - $z_{+,-}=\pm z_{\frac{\alpha}{2}}=\pm\Phi^{-1}(\frac{\alpha}{2})$ (beidseitig)

3. Entscheide:
   - $Z>z\implies$ verwerfe μ_0 (rechtsseitig)
   - $Z<z\implies$ verwerfe μ_0 (linksseitig)
   - $Z>z_+\vee Z<z_-\implies$ verwerfe μ_0 (beid-seitig)

Aufgaben

Ex : **Die Zufallsvariable X habe eine Verteilungsfunktion F_X mit zugehöriger Dichte f_X. Sei $Y = aX + b$, mit $a > 0$ und $b \in \mathbb{R}$. Dann gilt für die Dichte f_Y von Y**

Für die Verteilungsfunktion F_Y von Y gilt, dass $F_Y(y) = P[Y \leq y] = P[aX + b \leq y] = P\left[X \leq \frac{y-b}{a}\right] = F_X\left(\frac{y-b}{a}\right)$.

Somit folgt mit der Kettenregel
$f_y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X\left(\frac{y-b}{a}\right) = f_x\left(\frac{y-b}{a}\right)\frac{1}{a}$.

Analysis

Note Gamma Funktion: **Die Funktion Γ nennt man (Eulersche) Gammafunktion und sie ist für $x \geq 0$ definiert durch**

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} \, \mathrm{d}t$$

Γ hat eine grundlegende Verbindung zur Fakultätsfunktion, denn

$$\Gamma(n+1) = n\,! \text{ für } n \in \mathbb{N}_0.$$

Note Partielle Integration:

$$\int f'(x)g(x)\,dx = f(x)g(x) - \int f(x)g'(x)\,dx$$

oder

$$\int_a^b f'(x)g(x)\,dx = |f(x)g(x)|_a^b - \int_a^b f(x)g'(x)\,dx$$

Note Substitution: **Um $\int_a^b f(g(x))\,dx$ zu berechnen: Ersetze $g(x)$ durch u und integriere $\int_{g(a)}^{g(b)} f(u) \frac{\mathbf{d}u}{g'(x)}$.**

- **$g'(x)$ muss sich herauskürzen, sonst nutzlos.**
- **Grenzen substituieren nicht vergessen.**
- **Alternativ: unbestimmtes Integral berechnet werden und dann u wieder durch x substituieren.**
- **Man kann auch das Theorem in die andere Richtung anwenden:**

$$\int_a^b f(u)\,du = \int_{g^{-1}(a)}^{g^{-1}(b)} f(g(x))g'(x)$$

- **Sei $,Y$ kompakt, $f : Y \subset^n \to$ stetig.** **Sei $\gamma :\to Y$ mit $=_0 \cup B, Y = Y_0 \cup C$ (B, C Rand von $,Y$).** **Wenn $\gamma :_0 \to Y_0$ bijektiv und C^1 mit $\mathbf{det}(J_\gamma(x)) \neq 0, \forall x \in_0$, dann gilt**

$$\int_Y f(y)\,dy = \int f(\gamma(x))|\mathbf{det}(J_\gamma(x))|$$

Note Binomialsatz:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k \quad \textbf{mit:} \quad \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Des weiteren gilt:

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

Note Trigo-Werte:

deg	0	30	45	60	90	180
rad	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	π
cos	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0	-1
sin	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1	0
tan	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$	$+\infty$	0

Note Log/Exp Regeln:

Exponential	Logarithm
$a^0 = 1$	$\log_a 1 = 0$
$a^1 = a$	$\log_a a = 1$
$a^{m+n} = a^m \cdot a^n$	$\log_a(xy) = \log_a x + \log_a y$
$a^{m-n} = \frac{a^m}{a^n}$	$\log_a\left(\frac{x}{y}\right) = \log_a x - \log_a y$
$(a^m)^n = a^{mn}$	$\log_a(x^r) = r \log_a x$
$(ab)^m = a^m \cdot b^m$	$\log_a x = \frac{\log_b x}{\log_b a}$
$\left(\frac{a}{b}\right)^m = \frac{a^m}{b^m}$	$\log_a(a^x) = x$
$e^{\ln x} = x$	$\ln(e^x) = x$
	$e^{\ln x} = x$ (**for** $x > 0$)

Note Quadratic Form: **Für $a^2 + bx + c = 0$ gilt:**

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Ableitungen

$\mathbf{F(x)}$	$\mathbf{f(x)}$	$\mathbf{f'(x)}$
$\frac{x^{-a+1}}{-a+1}$	$\frac{1}{x^a}$	$\frac{a}{x^{a+1}}$
$\frac{x^{a+1}}{a+1}$	x^a $(a \neq -1)$	$a \cdot x^{a-1}$
$\frac{1}{k\ln(a)} a^{kx}$	a^{kx}	$ka^{kx}\ln(a)$
$\ln\lvert x\rvert$	$\frac{1}{x}$	$-\frac{1}{x^2}$
$\frac{2}{3}x^{3/2}$	$\sqrt{x}$	$\frac{1}{2\sqrt{x}}$
$\frac{n}{n+1}x^{\frac{1}{n}+1}$	$\sqrt[n]{x}$	$\frac{1}{n}x^{\frac{1}{n}-1}$
$-\cos(x)$	$\sin(x)$	$\cos(x)$
$\sin(x)$	$\cos(x)$	$-\sin(x)$
$\frac{1}{2}(x - \frac{1}{2}\sin(2x))$	$\sin^2(x)$	$2\sin(x)\cos(x)$
$\frac{1}{2}(x + \frac{1}{2}\sin(2x))$	$\cos^2(x)$	$-2\sin(x)\cos(x)$
$-\ln\lvert\cos(x)\rvert$	$\tan(x) = \frac{\sin}{\cos}$	$\frac{1}{\cos^2(x)}$ $1 + \tan^2(x)$
$\cosh(x)$	$\sinh(x)$	$\cosh(x)$
$\log(\cosh(x))$	$\tanh(x)$	$\frac{1}{\cosh^2(x)}$
$\ln\lvert\sin(x)\rvert$	$\cot(x)$	$-\frac{1}{\sin^2(x)}$
$\frac{1}{c} \cdot e^{cx}$	e^{cx}	$c \cdot e^{cx}$
$x(\ln\lvert x\rvert - 1)$	$\ln\lvert x\rvert$	$\frac{1}{x}$
$\frac{1}{2}(\ln(x))^2$	$\frac{\ln(x)}{x}$	$\frac{1-\ln(x)}{x^2}$
$\frac{x}{\ln(a)}(\ln\lvert x\rvert - 1)$	$\log_a\lvert x\rvert$	$\frac{1}{\ln(a)x}$

$\mathbf{F(x)}$	$\mathbf{f(x)}$
$\frac{1}{a\cdot(n+1)}(ax+b)^{n+1}$	$(ax+b)^n$
$\arcsin(x)$	$\frac{1}{\sqrt{1-x^2}}$
$\arccos(x)$	$\frac{-1}{\sqrt{1-x^2}}$
$\arctan(x)$	$\frac{1}{1+x^2}$
$\mathbf{arcsinh}(x)$	$\frac{1}{\sqrt{1+x^2}}$
$\mathbf{arccosh}(x)$	$\frac{1}{\sqrt{x^2-1}}$
$\mathbf{arctanh}(x)$	$\frac{1}{1-x^2}$
x^x $(x > 0)$	$x^x \cdot (1 + \ln x)$
$\log_a\lvert x\rvert$	$\frac{1}{x\ln a} = \log_a(e)\frac{1}{x}$

Integrale

$\mathbf{f(x)}$	$\mathbf{F(x)}$
$\int f'(x)f(x)$	$\frac{1}{2}(f(x))^2$
$\int \frac{f'(x)}{f(x)}$	$\ln\lvert f(x)\rvert$
$\int_{-\infty}^{\infty} e^{-x^2}$	$\sqrt{\pi}$
$\int(ax+b)^n$	$\frac{1}{a(n+1)}(ax+b)^{n+1}$
$\int x(ax+b)^n$	$\frac{(ax+b)^{n+2}}{(n+2)a^2} - \frac{b(ax+b)^{n+1}}{(n+1)a^2}$
$\int(ax^p+b)^n x^{p-1}$	$\frac{(ax^p+b)^{n+1}}{ap(n+1)}$
$\int(ax^p+b)^{-1}x^{p-1}$	$\frac{1}{ap}\ln\lvert ax^p+b\rvert$
$\int \frac{ax+b}{cx+d}$	$\frac{ax}{c} - \frac{ad-bc}{c^2}\ln\lvert cx+d\rvert$
$\int \frac{1}{x^2+a^2}$	$\frac{1}{a}\arctan\frac{x}{a}$
$\int \frac{1}{x^2-a^2}$	$\frac{1}{2a}\ln\left\lvert\frac{x-a}{x+a}\right\rvert$
$\int\sqrt{a^2+x^2}$	$\frac{x}{2}f(x) + \frac{a^2}{2}\ln(x+f(x))$
$\int \frac{1}{(x+a)^2}dx$	$-\frac{1}{a+x}$
$\int \frac{1}{(x+a)^3}dx$	$-\frac{1}{2(a+x)^2}$
$\int \frac{1}{(x+a)^t}dx$	$\frac{1}{(1-t)(x+a)^{t+1}}$
$\int \frac{x}{(x+a)^2}dx$	$\frac{a}{a+x} + \log\lvert a+x\rvert$
$\int \frac{x}{(x+a)^3}dx$	$-\frac{a+2x}{2(a+x)^2}$
$\int \frac{1}{ax^2+bx+c}dx$	$\frac{2\arctan\left(\frac{2ax+b}{\sqrt{4ac-b^2}}\right)}{\sqrt{4ac-b^2}}$
$\int\sin(kx)\cdot\cos(kx)dx$	$-\frac{1}{4k}\cos(2kx) = \frac{-(\cos(x))^2}{2}$
$\int\cos^n(x)dx$	$\frac{n-1}{n}\int\cos^{n-2}(x)dx + \frac{\cos^{n-1}(x)\sin(x)}{n}$
$\int\sin^n(x)dx$	$\frac{n-1}{n}\int\sin^{n-2}(x)dx - \frac{\cos(x)\sin^{n-1}(x)}{n}$
$\int\sin\cos dx$	$\frac{-1}{2}\cos^2$
$\int\frac{\cos}{\sin}dx$	$\log(\cos(x))$

Aufgabe 1

Seien U_1, U_2, U_3 i.i.d. $\mathrm{Uni}([0,1])$ Zufallsvariablen. Wir definieren $L = \min(U_1, U_2, U_3)$ und $M = \max(U_1, U_2, U_3)$.

Berechne die Wahrscheinlichkeitsdichte von M.

$$F_M(m) = \mathbb{P}[U_1 \leq m, U_2 \leq m, U_3 \leq m] = \Pi_{i=1}^3 \mathbb{P}[U_i \leq m$$

Berechne die gemeinsame Wahrscheinlichkeitsdichte von L und M.

$$\begin{aligned} &P[M < m, L \leq l] \\ =&P[M < m] - P[M < m, L > l] \\ &= m^3 - P[l < U_1 < m, l < U_2 < m, l < U_3 < m] \\ &= m^3 - (P[l < U_1 < m])^3 = m^3 - (m-l)^3 \\ &\textbf{für } 0 \leq l \leq m \leq 1 \end{aligned}$$

So $f_{M,L}(m, L) = 6(m-l)\not\Vdash_{\{0 \leq l \leq m \leq 1\}}$.

Generell gilt:

$$F_{\mathrm{Max}(X,Y)}(z) = P(X \leq z) \cdot P(Y \leq z) = F_X(z) \cdot F_Y(z)$$

Oder

$$F_{\mathrm{Min}(X,Y)}(z) = P(X \leq z) + P(Y \leq z) - P(X \leq z, Y \leq z$$

und dann

$$f_{\mathrm{Max}(X,Y)}(z) = \frac{d}{dz} F_{\mathrm{Max}(X,Y)}(z)$$

Verteilung	Parameter	$\mathbb{E}[X]$	$\mathbb{V}(X)$	$p_X(t)$	$F_X(t) = \mathbb{P}[X \leq t]$
Bernoulli $\sim \mathrm{Ber}(p)$	p: ErfolgsWK	p	$p \cdot (1-p)$	$p^t(1-p)^{1-t}$	$1-p$ für $0 \leq t < 1$
Binomial $\sim \mathrm{Bin}(n,p)$	n: Anzahl Versuche, p: ErfolgsWK	np	$np(1-p)$	$\binom{n}{t}p^t(1-p)^{n-t}$	$\sum_{k=0}^{t}\binom{n}{k}p^k(1-p)^{n-k}$
Geometrisch $\sim \mathrm{Geo}(p)$	p: ErfolgsWK, t: Anzahl Versuche	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$p(1-p)^{t-1}$	$1-(1-p)^t$
Negativbinomial $\sim \mathrm{NegBin}(r,p)$	r: Anzahl der Erfolge, p: Erfolgswahrscheinlichkeit	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$	$\binom{t+r-1}{t}p^t(1-p)^r$	$I_{1-p}(r,t+1)$
Hypergeometrisch $\sim \mathrm{H}(N,M,n)$	N: Gesamtzahl, M: Anzahl der Erfolge, n: Zufalsstichprobe	$n\frac{M}{N}$	$n\frac{M}{N}\frac{N-M}{N}\frac{N-n}{N-1}$	$\frac{\binom{M}{t}\binom{N-M}{n-t}}{\binom{N}{n}}$	$\frac{\sum_{k=0}^{t}\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$
Poisson $\sim \mathrm{Poi}(\lambda)$	λ: Erwartungswert λ: Varianz	λ	λ	$\frac{\lambda^t}{t!}e^{-\lambda}$	$e^{-\lambda}\sum_{k=0}^{t}\frac{\lambda^k}{k!}$
Gleichverteilung $\sim \mathcal{U}([a,b])$	n: Anzahl Ereignisse x_i : Ereignisse	$\frac{1}{n}\sum_{i=1}^{n}x_i$	$\frac{1}{n}\sum_{i=1}^{n}x_i^2 -$ $\frac{1}{n^2}\left(\sum_{i=1}^{n}x_i\right)^2$	$\frac{1}{n}$	$\frac{\lvert\{k : x_k \leq t\}\rvert}{n}$
Gleichverteilung $\sim \mathcal{U}([a,b])$(Intervall)	$[a,b]$: Intervall	$\frac{a+b}{2}$	$\frac{1}{12}(b-a)^2$	$\begin{cases}\frac{1}{b-a} & a \leq x \leq b\\ 0 & \text{sonst}\end{cases}$	$\begin{cases}0 & x \leq a\\ \frac{t-a}{b-a} & a < x < b\\ 1 & x \geq b\end{cases}$
Exponentialverteilung $\sim \mathrm{Exp}(\lambda)$	$\lambda : \frac{1}{\mathbb{E}[X]}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\begin{cases}\lambda e^{-\lambda t} & t \geq 0\\ 0 & t < 0\end{cases}$	$\begin{cases}1-e^{-\lambda t} & t > 0\\ 0 & t \leq 0\end{cases}$
Cauchyverteilung $\sim \mathrm{Cauchy}(x_0,\gamma)$	x_0: Lageparameter, γ: Skalenparameter	Undefiniert	Undefiniert	$\frac{1}{\pi\gamma\left[1+\left(\frac{t-x_0}{\gamma}\right)^2\right]}$	$\frac{1}{\pi}\arctan\left(\frac{t-x_0}{\gamma}\right)+\frac{1}{2}$
Normalverteilung $\sim \mathcal{N}(\mu,\sigma^2)$	σ^2: Varianz, $\mu : \mathbb{E}[X]$	μ	σ^2	$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(t-\mu)^2}{2\sigma^2}}$	$\frac{1}{2}\left[1+\mathrm{erf}\left(\frac{t-\mu}{\sigma\sqrt{2}}\right)\right]$
χ^2-Verteilung $\sim \chi_n^2$	n: Freiheitsgrad	n	$2n$	$\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}t^{\frac{n}{2}-1}e^{-\frac{t}{2}}$ für $t > 0$	$\mathrm{Poi}\left(\frac{n}{2},\frac{t}{2}\right)$
t-Verteilung $\sim t_n$	n: Freiheitsgrad	$\begin{cases}0 & n > 1\\ \text{undef.} & \text{sonst}\end{cases}$	$\begin{cases}\frac{n}{n-2} & n > 2\\ \infty & 1 < n \leq 2\\ \text{undef.} & \text{sonst}\end{cases}$	$\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\cdot\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}$	Undefiniert