1 Diskrete Wahrscheinlichkeitstheorie

Definition 1.1. Sei Ω eine höchstens abzählbare Menge, $\mathfrak{A} := \mathfrak{p}(\Omega) := \{A | A \subseteq \Omega\}$ und $\mathbf{p}: \Omega \to [0,1]$ eine Abbildung mit

$$\sum_{\omega \in \Omega} \mathtt{p}(\omega) = 1.$$

Die durch

$$\mathbf{P}(A) = \sum_{\omega \in A} \mathbf{p}(\omega)$$

definierte Abbildung heißt (diskrete) <u>Wahrscheinlichkeitsverteilung</u> (WV) über Ω. Die Funktion $\mathbf{p}$ heißt <u>Zähldichte</u>. Das Tripel $(\Omega, \mathfrak{A}, \mathcal{P})$ heißt (diskreter) <u>Wahrscheinlichkeitsraum</u>, $(\Omega, \mathfrak{A})$ heißt diskreter <u>messbarer Raum.</u>

Bezeichnung 1.2. Ω heißt Grundmenge, Ergebnisraum, Stichprobenraum; $A \subset \Omega$ heißt Ereignis und speziell Elementarereignis falls $|A| = 1$; Kurzschreibweise: $P(\{\omega\}) \overset{Abk.}{=} P(\omega) = p(\omega), \omega \in \Omega$

Lemma 1.3.

1. Gegeben sei die Situation aus (Def. 1.1). Dann gilt:
   (a) $\mathbf{P}(A) \leq 1$ für alle $A \in \mathfrak{p}(\Omega)$
   (b) $\mathbf{P}(\Omega) = 1$
   (c) $\mathbf{P}$ ist σ-Additiv, d.h. für paarweise disjunkte $A_1, A_2, ... \in \mathfrak{p}(\Omega)$ (d.h. $A_i \cap A_j = \varnothing$ für $i \neq j$) gilt:
   $$\mathbf{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbf{P}(A_i)$$
   insbesondere ist $\mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbf{P}(A_i), \forall n \in \mathbb{N}$ und $\mathbf{P}(\Omega) = \sum_{\omega \in \Omega} \mathbf{p}(\omega) = 1$
2. Sei $\mathbf{P} : \mathfrak{Pot}(\Omega) \to \mathbb{R}$ eine Abbildung mit obigen Eigenschaften, dann gibt es genau eine Funktion $\mathbf{p} : \Omega \to [0,1]$ mit $\mathbf{P}(A) = \sum_{\omega \in A} \mathbf{p}(\omega), \forall A \in \Omega$

Bezeichnung 1.4 (Laplace-Verteilung)**.** Sei $\Omega = \{1, ..., n\}, p(\omega) = \frac{1}{n}, \forall \omega \in \Omega$, so heißt $\mathbf{P}$ diskrete Gleichverteilung oder Laplace-Verteilung. Es gilt $\mathbf{P}(A) = \frac{|A|}{|\Omega|} = \frac{1}{n}|A|, \forall A \subseteq \Omega$

Definition 1.5 (σ-Algebra)**.** *Sei $\Omega \neq \varnothing$ und $\mathfrak{A} \subseteq \mathfrak{Pot}(\Omega)$ (System von Teilmengen von Ω). $\mathfrak{A}$ heißt σ-Algebra (von Ereignissen) über Ω, falls gilt:*

1. $\Omega \in \mathfrak{A}$
2. $A \in \mathfrak{A} \to A^C \in \mathfrak{A}$ *für alle* $A \in \mathfrak{A}$
3. $(A_n)_{n \in \mathbb{N}} \subset \mathfrak{A} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathfrak{A}$

d.h. Eine σ-Algebra ist abgeschlossen gegenüber der Bildung von Komplementen und abzählbaren Vereinigungen.

Lemma 1.6.

1. Sei $\mathfrak{A}$ eine σ-Algebra. Dann gilt:
   (a) $\varnothing \in \mathfrak{A} (\varnothing = \Omega^C \in \mathfrak{A})$
   (b) $A, B \in \mathfrak{A} \Rightarrow A \cap B \in \Omega (A \cap B = \underbrace{(\underbrace{A^C}_{\in \mathfrak{A}} \cup \underbrace{B^C}_{\in \mathfrak{A}})}_{\in \mathfrak{A}}{}^C \in \mathfrak{A})$
   (c) $(A_n)_{n \in \mathbb{N}} \in \mathfrak{A} \Rightarrow \bigcap_{n=1}^{\infty} A_n \in \mathfrak{A}$
2. Seien $B \subset \Omega$ und $\mathfrak{A}$ eine σ-Algebra übern Ω, dann gilt: $B \cap \mathfrak{A} := \{B \cap A | A \in \mathfrak{A}\}$ ist eine σ-Algebra über B (Spur-σ-Algebra)
3. Sei $\Omega \neq \varnothing$, $\mathfrak{A} = \{A \subset \Omega | A$ oder A^C ist höchstens abzählbar über $\Omega\}$ ist eine σ-Algebra über Ω

Bemerkung 1.7. $\mathfrak{Pot}(\Omega)$ ist σ-Algebra über Ω; $\mathfrak{Pot}(\Omega)$ ist die feinste, $\mathfrak{A} = \{\varnothing, \Omega\}$ ist die gröbste σ-Algebra.

Bemerkung 1.8. Man kann zeigen: Es gibt stets eine kleinste σ-Algebra, die ein vorgegebenes System $\mathfrak{F}$ von „einfachen" Mengen enthält.

Satz 1.9. Seien $\Omega \neq \varnothing, \mathcal{F} \subset \mathfrak{Pot}(\Omega)$ die kleinste σ-Algebra, die $\mathcal{F}$ enthält, diese ist gegeben durch: $\mathfrak{A}(\mathcal{F}) := \{A \in \mathfrak{Pot}(\Omega) |$ für jede $\sigma - Algebra \mathfrak{A}$ mit $\mathcal{F} \subset \mathfrak{A}$ gilt: $A \in \mathfrak{A}\}$ $\mathfrak{A}(\mathcal{F})$ heißt die von $\mathcal{F}$ erzegte σ-Algebra.

Definition 1.10 (Wahrscheinlichkeitsverteilung)**.** Sei $\mathfrak{A}$ eine σ-Algebra über $\Omega \neq \varnothing$ eine Abbildung $\mathbf{P} : \mathfrak{A} \to [0,1]$ mit

1. $\mathbf{P}(A) \geq 0, \forall A \in \mathfrak{A}$
2. $\mathbf{P}(\Omega) = 1$
3. $\mathbf{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbf{P}(A_n)$ für alle paarweise disjunkten $A_1, A_2, ... \in \mathfrak{A}$($\sigma$-Additivität)

heißt Wahrscheinlichkeitsverteilung oder Wahrscheinlichkeitsmass auf $\mathfrak{A}$. $(\Omega, \mathfrak{A}, \mathbf{P})$ heißt Wahrscheinlichkeitsraum, $(\Omega, \mathfrak{A})$ heißt messbarer Raum.

Definition 1.11 (Dirac-Verteilung)**.** Sei $\Omega \neq \varnothing$ abzählbar, $\omega \in \Omega$ fest. Die durch $\varepsilon_\omega(A) = \begin{cases} 1, & \text{falls } \omega \in \text{A} \\ 0, & \text{sonst (falls } \omega \notin \text{A)} \end{cases}$ $(A \in \mathfrak{p}(\Omega)$ festgelegte Wahrscheinlichkeitsverteilung $\varepsilon_\omega : \mathfrak{p}(\Omega) \to \mathbb{R}$ heißt Dirac-Verteilung oder Einpunktverteilung im Punkt ω. ε_ω ist Wahrscheinlichkeitsverteilung.

Definition 1.12 (Träger von $\mathbf{P}$)**.** Sei $(\Omega, \mathfrak{A}, \mathbf{P})$ ein diskreter Wahrscheinlichkeitsraum.

$$T := supp(\mathbf{P}) := \{\omega \in \Omega | \mathbf{P}(\omega) > 0\}$$

heißt Träger von $\mathbf{P}$.

Lemma 1.13. Sei $(\Omega, \mathfrak{A}, \mathbf{P})$ ein diskreter Wahrscheinlichkeitsraum und T Träger von $\mathbf{P}$. Dann gilt:

$$\mathbf{P}(A) = \sum_{\omega \in T} \mathbf{P}(\omega) \varepsilon_\omega(A), A \in \mathfrak{A}$$

d.h. $\mathbf{P}$ ist darstellbar als gewichtete Summe von Einpunktverteilungen.

2 Grundformeln der Kombinatorik

Merkregel: Die Anzahl der k-elemetigen Teilmengen einer n-elementigen Menge ist $\binom{n}{k}$.

Satz 2.1.

	k-mal ziehen aus n Kugeln	mit Zurücklegen	ohne Zurücklegen	
Permutation	mit Reihenfolge	Laplace-Raum $\lvert\Omega_1\rvert = n^k$	Laplace-Raum $\lvert\Omega_2\rvert = (n)_k = \frac{n!}{(n-k)!}$	unterscheidbare Murmeln
Kombinationen	ohne Reihenfolge	<u>kein</u> Laplace-Raum $\lvert\Omega_4\rvert = \binom{n+k-1}{k}$	Laplace-Raum $\lvert\Omega_3\rvert = \binom{n}{k}$	ununterscheidbare Murmeln
		mit Mehrfachbelegung	ohne Mehrfachbelegung	k Murmeln verteilt auf n Urnen

Definition 2.2 (Hypergeometrische Verteilung)**.** Betrachte eine Urne mit S schwarzen und W weißen Kugeln, $n = W + S$, Ziehe $k(\leq n)$ Kugeln ohne Zurücklegen. Gesucht ist die Wahrscheinlichkeit, dass die Stichprobe genau s schwarze und $k - s = w$ weiße Kugeln enthält.

$$h(\underbrace{s}_{Variable}; \underbrace{k, n, S}_{Parameter}) = \frac{\binom{S}{s} \cdot \binom{n-S}{k-s}}{\binom{n}{k}}, 0 \leq s \leq k$$

Die durch h bestimmte Zähldichte ($h(0) + h(1) + ... + h(k) = 1$) definiert die hypergeometrische Verteilung.

Definition 2.3 (Binomialverteilung)**.** Gesucht ist die Anzahl der Stichproben, die s defekte Teile enthalten. $b\left(s; k, \frac{S}{n}\right) = \frac{1}{n^k} \cdot \binom{k}{s} S^s (n-k)^{k-s} = \binom{k}{s} \left(\frac{S}{n}\right)^a \cdot \left(1 - \frac{S}{n}\right)^{k-1} = \binom{k}{s} p^s (1-p)^{k-s}$ wobei $p := \frac{S}{n}$ (Schlechtanteil). Die durch b bestimmte Zähldichte (als Funktion von s) definiert die sogenannte Binomialverteilung; kurz $b(s; k, p)$ oder $b(k, p)$. Intuitiv folgt, dass für große n kaum ein Unterschied zwischen Ziehen mit und ohne Rücklegen besteht.

Lemma 2.4. Sei $(S_n)_{n \in \mathbb{N}}$ mit $\underset{n \to \infty}{lim} \frac{S_n}{n} = p \in (0,1)$. Dann gilt: $\lim_{n \to \infty} h(s; k, n, S_n) = b(s; k, p)$

Lemma 2.5. Gegeben sei Folge von Binomialverteilungen $b(s; k, p_k); k \in \mathbb{N}$ mit $k \cdot p_k = \lambda > 0$ für alle $k \in \mathbb{N}$. Dann gilt: $\lim_{k \to \infty} b(s; k, p_k) = \frac{\lambda^s}{s!} e^{-\lambda}$ für alle $s \in \mathbb{N}_0$.
Offensichtlich definiert $p(s) := \frac{\lambda^s}{s!} e^{-\lambda}$ eine Zähldichte auf $\Omega = \mathbb{N}_0$

Sei $\lambda > 0$ beliebig: $\sum_{s=1}^{\infty} p(s) = e^{-\lambda} \underbrace{\sum_{s=0}^{\infty} \frac{\lambda^s}{s!}}_{e^\lambda} =$

1

Bezeichnung 2.6. Die Wahrscheinlichkeitsverteilung auf $\mathbb{N}_0$ definiert durch die Zähldichte $p(s) = \frac{\lambda^s}{s!} e^{-\lambda}, s \in \mathbb{N}_0, \lambda > 0$ heißt Poissonverteilung mit Parameter λ. Kurz: $po(s; \lambda)$ bzw. $po(\lambda)$

3 Eigenschaften von Wahrscheinlichkeitsräumen

Lemma 3.1. Es gelten für $A, B \in \mathfrak{A}$:

i) $P(A \cup B) = P(A) + P(B)$, falls $A \cap B = \varnothing$

ii) $P(B \setminus A) = P(B) - P(A)$, falls $A \subset B$ (Subtraktivität von P)

iii) $P(A^C) = 1 - P(A)$

iv) $A \subset B \Rightarrow P(A) \leq P(B)$ (Monotonie von P)

v) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

vi) $P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i); A_i \in \mathfrak{A}$

Definition 3.2. Sei $(A_n)_{n\in\mathbb{N}} \subseteq \mathfrak{A}$, $\mathfrak{A}$ sei σ-Algebra über $\Omega \neq \varnothing$. $(A_n)_{n\in\mathbb{N}}$ heißt MONOTON $\begin{cases} \text{WACHSEND} & \text{, falls } A_n \subseteq A_{n+1} \forall n \in \mathbb{N} \\ \text{FALLEND} & \text{, falls } A_n \supseteq A_{n+1} \forall n \in \mathbb{N} \end{cases}$
(kurz: $(A_n)_{n\in\mathbb{N}} \uparrow$ bzw. $(A_n)_{n\in\mathbb{N}} \searrow$)
Für monoton wachsende bzw. fallende Ereignisfolgen heißt jeweils

$$\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n$$

bzw.

$$\lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n$$

der Limes von $(A_n)_{n\in\mathbb{N}}$. Für eine beliebige Ereignisfolge $(A_n)_{n\in\mathbb{N}}$ heißen

$$\limsup_{n\to\infty} A_n = \lim_{n\to\infty}\left(\bigcup_{k=n}^{\infty} A_k\right) = \bigcap_{n=1}^{\infty} \underbrace{\bigcup_{k=n}^{\infty} A_k}_{\text{fallend}}$$

der LIMES SUPERIOR und

$$\liminf_{n\to\infty} A_n = \lim_{n\to\infty}\left(\bigcap_{k=n}^{\infty} A_k\right) = \bigcup_{n=1}^{\infty} \underbrace{\bigcap_{k=n}^{\infty} A_k}_{\text{wachsend}}$$

der LIMES INFERIOR von $(A_n)_{n\in\mathbb{N}}$.

Bemerkung 3.3. Es ist mit $(A_n)_{n\in\mathbb{N}} \subset \mathfrak{A}$: $\limsup_{n\to\infty} A_n \in \mathfrak{A}, \liminf_{n\to\infty} A_n \in \mathfrak{A}$ nach Definition der σ-Algebra und $\limsup_{n\to\infty} A_n = \{\omega \in \Omega | \omega$ liegt in unendlich vielen der A_n's$\} \cong$ unendlich viele der A_n's treten ein; $\liminf_{n\to\infty} A_n = \{\omega \in \Omega | \omega$ liegt in allen der A_n's, bis auf endlich viele$\} \cong$ alle, bis auf endlich viele der $A_n's$ treten ein (fast alle A_i's treten ein).

Lemma 3.4. $(\Omega, \mathfrak{A}, P)$ Wahrscheinlichkeitsraum, $(A_n)_{n\in\mathbb{N}} \subseteq \mathfrak{A}$, dann gilt:

i) $P\left(\bigcup_{n=1}^{\infty} A_n\right) = P\left(\limsup_{n\to\infty} A_n\right) = \lim_{n\to\infty} P(A_n)$, falls $(A_n)_{n\in\mathbb{N}} \uparrow$ (Stetigkeit von P von unten)

ii) $P\left(\bigcap_{n=1}^{\infty} A_n\right) = P\left(\liminf_{n\to\infty} A_n\right) = \lim_{n\to\infty} P(A_n)$, falls $(A_n)_{n\in\mathbb{N}} \searrow$ (Stetigkeit von P von oben)

iii) $P\left(\limsup_{n\to\infty} A_n\right) = \lim_{n\to\infty} P\left(\bigcup_{k=n}^{\infty} A_k\right)$
$P\left(\liminf_{n\to\infty} A_n\right) = \lim_{n\to\infty} P\left(\bigcap_{k=n}^{\infty} A_k\right)$

iv) $P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n)$ (Sub-σ-Additivität)

Lemma 3.5. Für Ereignisse $(A_n)_{n\in\mathbb{N}}$ in einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ gilt: $P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k) - \sum_{1\leq i_1<i_2\leq n} P(A_{i_1} \cap A_{i_2}) + \sum_{1\leq i_1<i_2<i_3\leq n} P(A_{i_1} \cap A_{i_2} \cap A_{i_3}) \pm \ldots + (-1)^{n+1} P\left(\bigcap_{k=1}^n A_k\right)$

Korrolar 3.6. Aus der Sylvester-Pointcaré-Siebformel: Seien $A_1, \ldots, A_n$ Ereignisse aus $\mathfrak{A}$ im Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, \mathbf{P})$, dann gilt: $\sum_{k=1}^n P(A_k) - \sum_{1\leq i_1<i_2\leq n} P(A_{i_1} \cap A_{i_2}) \leq P\left(\bigcup_{k=1}^n A_k\right) \leq \sum_{k=1}^n P(A_k)$

Bemerkung 3.7 (Zusammenfassung)**.**

Mathematisches Objekt	Interpretation
Ω	Grundraum (Ereignisraum)
$\omega \in \Omega$	(mögliches) Ergebnis
$A \in \mathfrak{A}$	Ereignis
$\mathfrak{A}$	Menge der (möglichen) Ereignisse
Ω	sicheres Ereignis
$\varnothing$	unmögliches Ereignis
$\omega \in A$	Ereignis A tritt ein
$\omega \in A^C$	Ereignis A tritt *nicht* ein
$\omega \in A \cup B$	Ereignis A *oder* B tritt ein
$\omega \in A \cap B$	Ereignis A *und* B tritten ein
$A \subseteq B$	Eintreten von Ereignis A *impliziert* das Eintreten von Ereignis B
$A \cap B = \varnothing$	Ereignisse A und B schließen einander aus
$\omega \in \bigcup_{i\in\mathbb{I}} A_i$	*mindestens ein* Ereignis $A_i, i \in \mathbb{I}$ tritt ein
$\omega \in \bigcap_{i\in\mathbb{I}} A_i$	*alle* Ereignis $A_i, i \in \mathbb{I}$ tretten ein
$\omega \in \limsup_{i\in\mathbb{I}} A_i$	*unendlich viele* Ereignisse $A_i, i \in \mathbb{I}$ treten ein
$\omega \in \liminf_{i\in\mathbb{I}} A_i$	*alle bis auf endlich viele* Ereignisse $A_i, i \in \mathbb{I}$ treten ein (fast alle)
$P(A)$	Wahrscheinlichkeit, dass das Ereignis $A \in \mathfrak{A}$ eintritt (Wahrscheinlichkeit für A)
$P(A) = 1$	A tritt (fast) sicher ein
$P(A) = 0$	A tritt sicher *nicht* ein
$P(A) > P(B)$	A ist wahrscheinlicher als B

4 Bedingte Wahrscheinlichkeiten

Definition 4.1. Sei $(\Omega, \mathfrak{A}, P)$ ein Wahrscheinlichkeitsraum. Für jedes $B \in \mathfrak{A}$ mit $P(B) > 0$ wird durch

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, A \in \mathfrak{A}$$

eine Wahrscheinlichkeitsverteilung $P(\cdot|B)$ auf $\mathfrak{A}$ definiert, die sogenannte BEDINGTE VERTEILUNG unter (der Hypothese) B.
$P(A|B)$ heißt ELEMENTAR BEDINGTE WAHRSCHEINLICHKEIT von A unter B.

Lemma 4.2. Sei $A, B, A_1, A_2, \ldots \in \mathfrak{A}$, $P(A) > 0, P(B) > 0, P\left(\bigcap_{i=1}^{n-1} A_i\right) > 0$. Dann:

i) $P(A|B) = P(B|A) \cdot \frac{P(A)}{P(B)}$

ii) $P\left(\bigcap_{i=1}^n A_i\right) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot \ldots \cdot P\left(A_n | \bigcap_{i=1}^{n-1} A_i\right)$

Lemma 4.3. Gegeben $A \in \mathfrak{A}, (B_n)_n \subset \mathfrak{A}, B_n$ paarweise disjunkt und $A \subset \bigcup_{n=1}^{\infty} B_n$ (häufig $\bigcup B_n = \Omega$, d.h. disjunkte Zerlegung von Ω). Dann gilt:

$$P(A) = \sum_{n=1}^{\infty} P(A|B_n) \cdot P(B_n)$$

(Ist $P(B_k) = 0$, so ist zunächst $P(A|B_k)$ nicht definiert; setze dann $P(B_k) \cdot P(A|B_k) = 0$)

Satz 4.4. Seien $A, (B_n)_n \subset \mathfrak{A}, B_n$ paarweise disjunkt, $A \subset \bigcup_{n=1}^{\infty} B_n$ und $P(A) > 0$. Dann gilt:

$$P(B_k|A) = \frac{P(B_k)P(A|B_k)}{\sum_{n=1}^{\infty} P(A|B_n)P(B_n)}, \forall k \in \mathbb{N}$$

5 Stochastische Unabhängigkeit

Definition 5.1. Sei $(\Omega, \mathfrak{A}, P)$ ein Wahrscheinlichkeitsraum, $A, B \in \mathfrak{A}$ heißen STOCHASTISCH UNABHÄNGIG, falls $P(A \cap B) = P(A) \cdot P(B)$

Lemma 5.2. Mit A, B sind auch A, B^C und A^C, B^C stochastisch unabhängig.

i) Ist $P(B) > 0$, so gilt: A, B stochastisch unabhängig $\Leftrightarrow P(A|B) = P(A)$

ii) Ist A eine Nullmenge, d.h. $P(A) = 0$, so sind A, B stochastisch unabhängig für alle $B \in \mathfrak{A}$.

Definition 5.3. Eine Familie $(A_i)_{i\in\mathbb{I}} \subset \mathfrak{A}, \mathbb{I}$ belibige Indexmenge, heißt PAARWEISE STOCHASTISCH UNABHÄNGIG, falls gilt:

$$P(A_i \cap A_j) = P(A_i) \cdot P(A_j), \forall i, j \in I, i \neq j.$$

Definition 5.4. Eine Familie $(A_I)_{i\in I}$ heißt (VOLLSTÄNDIG) STOCHASTISCH UNABHÄNGIG, falls für jede endliche Auswahl von Ereignissen gilt:

$$P\left(\bigcap_{j\in J} A_j\right) = \prod_{j\in J} P(A_j) \forall 0 \neq J \subseteq I, |J| < \infty.$$

Bemerkung 5.5.
1. $(A_i)_{i\in I}$ stochastisch unabhängig $\Rightarrow (A_I)_{i\in I}$ paarweise stochastisch unabhängig. Umkehrung gilt nicht!
2. jede Teilfamilie einer stochastisch unabhängigen Familie ist stochastisch unabhängig.
3. Beachte, dass die Definition ein System von Gleichungen leifert!

Satz 5.6.
1. Seien $(A_i)_{i\in\mathbb{I}}$ Familien stochastisch unabhängiger Ereignisse, $k \notin \mathbb{I}$ und $\mathbf{P}(A_k) \in \{0, 1\} \Rightarrow (A_i)_{i\in\mathbb{I}\cup\{k\}}$ stochastisch unabhängig.
2. $(A_i)_{i\in\mathbb{I}}$ stochastisch unabhängig und $B_i \in \{A_i, A_i^C, \varnothing, \Omega\} \forall i \in \mathbb{I} \Rightarrow (B_i)_{i\in\mathbb{I}}$ stochastisch unabhängig.
3. Sei $\mathbb{I} = \{1, \ldots, n\}, n \in \mathbb{N}$ fest gewählt: $(A_i)_{i\in\mathbb{I}}$ stochastisch unabhängig $\Leftrightarrow \mathbf{P}(\bigcap_{i=1}^n B_i) = \prod_{i=1}^n P(B_i), \forall B_i \in \{A_i, A_i^C\}, \forall i \in \mathbb{I}$

Bemerkung 5.7. Sei die Familie $(A_i)_{i\in\{1,\dots,n\}}$ stochastisch unabhängig oder kurz $A_1,...,A_n$ stochastisch unabhängig, dann gilt: $P\left(\bigcup_{i=1}^n A_i\right) = 1 - P\left(\bigcap_{i=1}^n A_i^C\right) = 1 - \prod_{i=1}^n P(A_i^C) = 1 - \prod_{i=1}^n (1 - P(A_i))$

Bemerkung 5.8. Relation „stochastische Unabhängigkeit" ist nicht transitiv, d.h. aus A_1, A_2 stochastisch unabhängig und A_2, A_3 stochastisch unabhängig folgt nicht notwendig A_1, A_3 stochastisch unabhängig.

Definition 5.9. Sei $(A_n)_{n\in\mathbb{N}} \subset \mathfrak{A}$. $(A_n)_n$ heißt KONVERGENT $\Leftrightarrow \limsup_{n\to\infty} A_n = \liminf_{n\to\infty} A_n$

Bemerkung 5.10. Es gilt stets: $\liminf_{n\to\infty} A_n \subset \limsup_{n\to\infty} A_n$.
Weiterhin: $(\limsup_{n\to\infty} A_n)^C = \left(\bigcap_{n=1}^\infty \bigcup_{k=n}^\infty A_k\right)^C = \bigcup_{n=1}^\infty \bigcap_{k=n}^\infty A_k^C = \liminf_{n\to\infty} A_n^C$

Lemma 5.11. Sei $\Omega, \mathfrak{A}, P)$ ein Wahrscheinlichkeitsraum und $(A_n)_{n\in\mathbb{N}} \subset \mathfrak{A}$, dann gilt:

i) $\sum_{n=1}^\infty P(A_n) < \infty \Rightarrow P\left(\limsup_{n\to\infty} A_n\right) = 0$

ii) Ist zusätzlich $(A_n)_{n\in\mathbb{N}}$ stochastisch unabhängig, $\sum_{n=1}^\infty P(A_n) = \infty \Rightarrow P\left(\limsup_{n\to\infty} A_n\right) = 1$

Bemerkung 5.12.

1. Analog gilt (mit deMorgan) $\sum_{n=1}^\infty P(A_n^C) < \infty \Rightarrow P\left(\liminf_{n\to\infty} A_n\right) = 1$; $(A_n)_n$ stochastisch unabhaengig, $\sum_{n=1}^\infty P(A_n^C) = \infty \Rightarrow P\left(\liminf_{n\to\infty} A_n\right) = 0$
2. Für stochastsich unabhängige Ereignisse $(A_i)_i$ gilt stets $P\left(\limsup_{n\to\infty} A_n\right)$, $P\left(\liminf_{n\to\infty} A_n\right) \in \{0, 1\}$
3. Sei $(A_n)_n \subset \mathfrak{A}$; falls eine unabhängige Teilfolge $(A_{n_k})_k$ von $(A_n)_n$ existiert mit $\sum_{k=1}^\infty P(A_{n_k})_k = \infty$, dann folgt $P\left(\limsup_{n\to\infty} A_n\right) = 1$

Definition 5.13. Für diskrete Wahrscheinlichkeitsräume $(\Omega_i, \mathfrak{A}_i, P_i), 1 \le i \le n$ heißt $(\Omega, \mathfrak{A}, P)$ mit $\Omega = \prod_{i=1}^n \Omega_i = \{\omega = (\omega_1, ..., \omega_n) | \omega_i \in \Omega, 1 \le i \le n\} := \times_{i=1}^n \Omega_i$ und P definiert durch $P(\omega) = \prod_{i=1}^n P_i(\omega_i)\,(P := \times_{i=1}^n P_i)$ $\mathfrak{A}$ ist Potenzmenge über Ω)
PRODUKT DER WAHRSCHEINLICHKEITSRÄUME $(\Omega_i, \mathfrak{A}_i, P_i), 1 \le i \le n$ (Produktraum)

Bezeichnung 5.14.

p	:	Erfolgswahrscheinlichkeit
$\omega_i = 1$	:	Erfolg im i-ten Teilexperiment

in n sogenannten Bernoulli-Experimenten mit Erfolgswahrscheinlichkeit p.
Die durch $E_k := \{\omega \in \Omega_i | \sum_{i=1}^n \omega_i = k\} \equiv$ Genau k Erfolge mit $k \le n$; $P(E_k) = \binom{n}{k} p^k (1-p)^{n-k} (E_j's$ sind disjunkt
definierte Wahrscheinlichkeitsverteilung auf $\{0, 1, ..., n\}$ heißt BINOMIALVERTEILUNG.

6 Zufallsvariablen

Zufallsvorgänge werden beschrieben durch einen Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, \mathbf{P})$. Dabei ist häufig nicht $\omega \in \Omega$ von Interesse, sondern eine Funktion X von ω. Diese bezeichnet man als ZUFALLSVARIABLE. Die zugehörige Wahrscheinlichkeitsverteilung wird durch $P^X(B) = P(\{\omega \in \Omega | X(\omega) \in B\}) = P(X^{-1}(B)), B \in \mathfrak{Pot}(\Omega)$ angegeben.

Definition 6.1. Sei $T : \Omega \to \Omega'$ eine Abbildung, so heißt $T^{-1} : \begin{cases} p(\Omega') & \to & p(\Omega) \\ A' & \mapsto & T^{-1}(A') : \{\omega \in \Omega | T(\omega) \in A'\} \end{cases}$ die zu T gehörige URBILDFUNKTION oder UMKEHRABBILDUNG/ PSEUDOINVERSE

Lemma 6.2 ((Eigenschafte der Urbildfunktion)). Seien $A', B', A'_i \in \mathfrak{Pot}(\Omega')$, so gilt:

i) $T^{-1}(\varnothing) = \varnothing, T^{-1}(\Omega') = \Omega$

ii) $T^{-1}(A' \backslash B') = T^{-1}(A') \backslash T^{-1}(B')$
speziell: $T^{-1}(B'^C) = (T^{-1}(B'))^C$

iii) $T^{-1}(\bigcap_{i\in\mathbb{I}} A'_i) = \bigcap_{i\in\mathbb{I}} T^{-1}(A'_i)$

iv) $T^{-1}(\bigcup_{i\in\mathbb{I}} A'_i) = \bigcup_{i\in\mathbb{I}} T^{-1}(A'_i)$
Spezielle disjunkte B'_i: (mit Symbol $\sum$ für disjunkte Vereinigung) $T^{-1}\left(\sum_{i\in\mathbb{I}} B'_i\right) = \sum_{i\in\mathbb{I}} T^{-1}(B'_i)$

v) $A' \subset B' \Rightarrow T^{-1}(A') \subset T^{-1}(B')$

vi) Ist $S : \Omega' \to \Omega''$ eine beliebige Abildung, dann gilt $(S \circ T)^{-1} = T^{-1} \circ S^{-1}$

Lemma 6.3. Sei $(\Omega, \mathfrak{A}, P)$ ein diskreter Wahrscheinlichkeitsraum und $X : \Omega \to \mathbb{R}$, dann ist $\mathbf{P}^X : \mathfrak{Pot}(\mathbb{R}) \to [0, 1]$ mit $P^X = \mathbf{P}(X^{-1}(A)) = \mathbf{P}(\{\omega \in \Omega | X(\omega) \in A\})$ eine diskrete Wahrscheinlichkeitsverteilung über (einer abzählbaren Teilmenge von) $\mathbb{R}$, bzw. über $X(\Omega)$. Kurz: Verteilung von X unter $\mathbf{P}$.

Definition 6.4. Eine Abbildung $X : \Omega \to \Omega'$ von einem messbaren Raum $(\Omega, \mathfrak{A})$ in einen anderen messbaren Raum $(\Omega', \mathfrak{A}')$ heißt MESSBAR, wenn $\forall A' \in \mathfrak{A}' : X^{-1}(A') \in \mathfrak{A}$.

Definition 6.5. Eine messbare Funktion im obigen Sinne von einem Wahrscheinlichkeitsraum in einen anderen heißt ZUFALLSVARIABLE (ZV), bzw. ZUFALLSVEKTOR, falls $\Omega' \subset \mathbb{R}^n$ (auch Zufallsgröße).

Bemerkung 6.6.

i) Ist $\mathfrak{A} = \mathfrak{Pot}(\Omega)$ wie beim diskreten Wahrscheinlichkeitsraum, so ist jede Abbildung $X : \Omega \to \Omega'$ messbar.

ii) Kurzschreibweise: $\{X \in A\} := X^{-1}(A) = \{\omega \in \Omega | X(\omega) \in A\}$, dann auch kurz $P(X \in A) = ...$

iii) Die Komposition messbarer Funktionen ist messbar.

Definition 6.7. Das Wahrscheinlichkeitsmaß P^X definiert durch $P^X(A) = P(X^{-1}(A)), A \subseteq \mathbb{R}$ heißt VERTEILUNG VON X UNTER P (kurz P^X oder P) oder X hat Verteilung P^X, X ist verteilt wie P^X, $X \sim P^X, X \sim P$.

Bezeichnung 6.8.

i) Zufallsvariable X ist binomialverteilt, falls $P^X(k) = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$

ii) Zufallsvariable X ist poisson-verteilt, falls für $\lambda > 0$ gilt $P^X(k) = P(X = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}, k \in \mathbb{N}_0$

iii) Seien $\Omega \neq \varnothing$, $\mathbf{P}$ eine diskrete Wahrscheinlichkeitsverteilung über $\Omega, A \subseteq \Omega$. Die Funktion $\mathcal{I}_A : \Omega \to \mathbb{R}$ definiert durch $\mathcal{I}_A(\omega) = \begin{cases} 1 & , \omega \in A \\ 0 & , \text{sonst} \end{cases}$ heißt INDIKATORFUNKTION von A und ist eine Zufallsvariable.
Dabei gilt $\mathcal{I}_A \sim b(1, p)$: $P(\mathcal{I}_A = 0) = 1 - p, P(\mathcal{I}_A = 1) = p$, zusätzlich gilt noch: $\mathcal{I}_{A\cup B} = max(\mathcal{I}_A, \mathcal{I}_B), \mathcal{I}_{A\cup B} = \mathcal{I}_A + \mathcal{I}_B$ (falls $A \cap B = \varnothing$), $\mathcal{I}_{A\cap B} = min(\mathcal{I}_A, \mathcal{I}_B) = \mathcal{I}_A \cdot \mathcal{I}_B, \mathcal{I}_{A^C} = 1 - \mathcal{I}_A$

Definition 6.9. Sei $X_1, .., X_m$ Zufallsverteilungen. Die Verteilung von $X = (X_1, ..., X_m)$ heißt die GEMEINSAME VERTEILUNG der Zufallsvariablen $X_1, ..., X_m$. Schreibweise: $P^X = P^{(X_1,...,X_m)}$ Die Verteilung von $(X_{i_1}, ..., X_{i_l}, 1 \le i_1 < ... < i_l \le m, l < m$ heißt L-DIMENSIONALE RANDVERTEILUNG (Marginalverteilung) zu $(i_1, ..., i_l)$. Die Verteilung von X_i heißt I-TE RANDVERTEILUNG (Marginalverteilung) $1 \le i \le m$.

Bemerkung 6.10. Die gemeinsame Verteilung ist durch die Angabe aller Wahrscheinlichkeiten $P(X_1 \in A_1, ..., X_m \in A_m), A_i \in \mathfrak{A}_i$ bestimmt.

Bemerkung 6.11. Die eindimensionale Randverteilungen legen die gemeinsame Verteilung nicht eindeutig fest.

Definition 6.12. Eine Familie von Zufallsvariablen $X_i : (\Omega, \mathfrak{A}, P) \to (\Omega_i, \mathfrak{A}_i, P^{X_i}), i \in \mathbb{I}$ heißt STOCHASTISCH UNABHÄNGIG (oder die Zufallsvariablen $X_i, i \in \mathbb{I}$ heißen stochastisch unabhängig), falls die Mengensysteme $X_i^{-1}(\mathfrak{A})$ stochastisch unabhängig sind, d.h. jedes Repräsentantensystem $B_i \in X_i^{-1}(\mathfrak{A}_i), i \in \mathbb{I}$ bildet eine unabhängige Familie von Ereignissen.

Bemerkung 6.13. Die Zufallsvariablen $X_i, i \in \mathbb{I}$ sind stochastisch unabhängig genau dann, wenn $P\left(\bigcap_{i\in\mathbb{J}} \underbrace{\{X_i \in A_i\}}_{X_i^{-1}(A_i)}\right) = \prod_{i\in\mathbb{J}} P(X_i \in A_i) \forall \mathbb{J} \subseteq \mathbb{I}, |\mathbb{J}| < \infty, \forall A_i \in \mathfrak{A}_i, i \in \mathbb{J}$

Satz 6.14. Sind die Zufallsvariablen $X_i : \Omega \to \Omega_i, i \in \mathbb{I}$, stochastisch unabhängig und sind die Abbildungen $f_i : \Omega_i \to \Omega'_i$ messbar, so sind die Zufallsvariablen $f_i \circ X_i, i \in \mathbb{I}$ stocastisch unabhängig.
Weiterhin: Seien $I_j \subset \mathbb{I}$ für $j \in \mathbb{J}$ disjunkte Teilmengen und $G_i : X_{i\in\mathbb{I}_j} \Omega_i \to \Omega''_i$ messbar $\Rightarrow g_j \circ (X_i, i \in \mathbb{I}_j) j \in \mathbb{J}$, stochastisch unabhängig (messbare Funktionen von Zufallsvariablen mit disjunkten Indexmengen)

Lemma 6.15. Sei $(\Omega, \mathfrak{A}, P)$ ein diskreter Wahrscheinlichkeitsraum. Dann gilt: $X_i, i \in \mathbb{I}$, stochastisch unabhängig $\Leftrightarrow P(X_j = x_j, j \in \mathbb{J}) = \prod_{j\in\mathbb{J}} P(X_j = x_j), \forall x_j \in X_j(\Omega), \forall j \in \mathbb{J}, \forall \mathbb{J} \subseteq \mathbb{I}, |\mathbb{J}| < \infty$

Satz 6.16. Seien X, Y seien stochastisch unabhängige Zufallsvariablen auf $\mathbb{Z}$ mit den Zähldichten f bzw. g (d.h. $P(X = n) = f(n), P(Y = m) = g(m); n, m \in \mathbb{Z}$) Dann hat $X + Y$ die Zähldichte h gegeben durch $P(X + Y = k) = h(k) = \sum_{j\in\mathbb{Z}} f(j) g(k - j) = \sum_{j\in\mathbb{Z}} f(k - j) g(j), h \in \mathbb{Z}$

Bezeichnung 6.17. h ist die FALTUNG der Dichten f und $g : h = f \star g$.

Definition 6.18. Seien $(\Omega, \mathfrak{A}, P)$ ein Wahrscheinlichkeitsraum und X eine Zufallsvariable mit Wahrscheinlichkeitsverteilung P^X. Die Funktion: $F^X \begin{cases} \mathbb{R}^1 \to \mathbb{R}^1 \\ x \mapsto P^X((-\infty, x]) \end{cases}$ heißt die zu P^X gehörige VERTEILUNGSFUNKTION (kurz: Verteilungsfunktion von X, F Verteilungsfunktion von X, X verteilt nach F,..., $X \sim F$)

$$P(X_i \le x) = F_i(x)$$

Bemerkung 6.19. i) Sei p^X die Zähldichte von P^X, dann ist $F^X(x) = P^X((-\infty, x]) = \sum_{\omega \leq X} p^X(\omega) = \sum_{\omega \leq x, \omega \in supp P^X} p^X(\omega), x \in \mathbb{R}$ (Träger)

ii) Für $P^X((-\infty, x]) = P(X \leq x)$ d.h. $F^X(x)$ ist die Wahrscheinlichkeit dafür, dass die Zufallvariable X Werte $\leq x$ annimmt.

Lemma 6.20. Sei F^X die zu P^X gehörige Verteilungsfunktion. Dann gilt:

1. (a) F^X ist monoton wachsend
   (b) F^X ist rechtsseitig stetig
   (c) $\lim_{x\to\infty} F^X(x) = 1$, $\lim_{x\to-\infty} F^X(x) = 0$
2. P^X ist durch F^X eindeutig betsimmt.

Definition 6.21. Sei $X = (X_1, ..., X_n)$ ein Zufallsvektor mit der Wahrscheinlichkeitsverteilung $\mathbf{P}^X$. Die durch $F^X := \mathbf{P}^X((-\infty, x_1] \times ... \times (-\infty, x_n]), (x_1, ..., x_n) \in \mathbb{R}^n$ definierte Funktion heißt MULTIPLIKATIVE VERTEILUNGSFUNKTION.

Bemerkung 6.22. Sind X_1, X_2 stochastischt unabhängig, so gilt: $F^{(X_1,X_2)}(x_1, x_2) = F^{X_1}(x_1) \cdot F^{X_2}(x_2), (x_1, x_2) \in \mathbb{R}^2$; Insbesondere ist bei stochastisch unabhängigen Zufallsvariablen die gemeinsame Verteilungsfunktion eindeutig durch die Verteilungsfunktionen der eindimensionalen Randwerte bestimmt.

7 Erwartungswerte

Definition 7.1. Sei $(\Omega, \mathbf{P})$ ein diskreter Wahrscheinlichkeitsraum.

1. Sei $X(\Omega) \subset \mathbb{R}^+$ oder $X(\Omega) \subset \mathbb{R}^-$: $EX := E(X) := \sum_{\omega\in\Omega} X(\omega) \cdot P(\{\omega\})$ heißt ERWARTUNGSWERT von X unter $\mathbf{P}$.
2. Sei X eine allgemeine Zufallsvariable mit $E(max(X, 0)) < \infty$ oder $E(max(X, 0)) > -\infty$, dann heißt $EX := E(X) := \sum_{\omega\in\Omega} X(\omega) \cdot P(\{\omega\})$ der ERWARTUNGSWERT von X unter $\mathbf{P}$. (Definition verhindert „$-\infty+\infty$“ in der Summe)

Bemerkung 7.2. 1. Für nicht-negative Zufallsvariablen ist E immer wohldefiniert. $E(X) = \infty$ ist erlaubt.

2. Für Zufallsvariablen mit positiven und negativen Werten muss die Wohldefiniertheit der Reihe $\sum_{\omega\in\Omega} X(\omega) \cdot \mathbf{P}(\omega)$ gewährleistet werden. Werden nur die endlichen Erwartungswerte betrachtet, kann absoluete Konvergenz $\sum_{\omega\in\Omega} |X(\omega)| \cdot \mathbf{P}(\omega) < \infty$ gefordert werden, d.h. die Änderung der Summationsreihenfolge ist erlaubt.
3. Im Folgenden wird stets die Wohldefiniertheit der auftretenden Erwartungswerte vorausgesetzt.
4. $E(X)$ hängt von der Verteilung von X ab: Sei $x_1, x_2, ...$ eine Abzählung von $X(\Omega)$, dann gilt: $E(X) = \sum_{i=1}^{\infty} x_i \cdot P^X(x_i), P^X(x) = 0$ falls $x \notin supp(P)$, d.h. $E(X)$ kann ebenso über die Summe $\sum x_i \cdot P^X(x_i)$ erklärt werden.
5. der Erwartungswert (als mögliche Kenngröße) dient dem Vergleich von Verteilungen.

Lemma 7.3. Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge aus $\mathbb{R}^+$, $b_j := \sum_{n=j}^{\infty} a_n$, dann gilt: $\sum_{j=1}^{\infty} b_j = \sum_{n=1}^{\infty} n \cdot a_n$

Korrolar 7.4. Sei $(\Omega, \mathfrak{A}, \mathbf{P})$ ein diskreter Wahrscheinlichkeitsraum, $X : \Omega \to \mathbb{N}_0$, dann gilt: $E(X) = \sum_{n=1}^{\infty} n \cdot P^X(n) = \sum_{n=1}^{\infty} P^X([n, \infty)) = \sum_{n=1}^{\infty} \mathbf{P}(X \geq n)$ $(\mathbf{P}(X \geq n) = 1 - F^X(n-1))$

Bezeichnung 7.5. Die Wahrscheinlichkeitsverteilung $P^X(n) = (1-p)^{n-1} \cdot p, n \in \mathbb{N}$ nennt man GEOMETRISCHE VERTEILUNG.

Satz 7.6. Sei $(\Omega, \mathfrak{A}, \mathbf{P})$ ein diskreter Wahrscheinlichkeitsraum, $X : \Omega \to \mathbb{R}^k$ ein Zufallsvektor und $f : \mathbb{R}^k \to \mathbb{R}$ eine mesbare Abbildung. Ferner existiert der Erwartungswert der Zufallsvariablen $f \circ X$. Dann gilt: $E(f \circ X) = \sum_{\omega\in\Omega} (f \circ X)(\omega) \cdot \mathbf{P}(\omega) = \sum_{x\in\mathbb{R}} x \cdot P^{f\circ X}(x) = \sum_{t\in\mathbb{R}^k} f(t) \cdot P^X(t) = E_{P^X}(f)$.

Lemma 7.7. Seien X, Y Zufallsvariablen mit endlichen Erwartungswerten, $a \in \mathbb{R}$, dann gilt:

(a) $E(a) = a$

(b) $E(aX) = a \cdot E(X)$

(c) $E(|X + Y|) \leq E(|X|) + E(|Y|)$

(d) $E(X + Y) = E(X) + E(Y)$ (zusammen mit (b) ergibt sich die Linearität)

(e) $X \leq Y \Rightarrow E(X) \leq E(Y)$ (speziell: $Y \geq 0 \Rightarrow E(Y) \geq 0, E(X) \leq E(|X|)$) (dabei $X < Y : X(\omega) \leq Y(\omega) \forall \omega \in \Omega$)

(f) $E(|X|) = 0 \Leftrightarrow P(X \neq 0) = 0$

Lemma 7.8. Seien $(X_i)_{i\in\mathbb{I}}$ Zufallsvariablen mit endlichen Erwartungswerten, dann gilt:

1. $E(\sup_{i\in\mathbb{I}} X_i) \geq \sup_{i\in\mathbb{I}} E(X_i)$
2. $E(\inf_{i\in\mathbb{I}} X_i) \leq \inf_{i\in\mathbb{I}} E(X_i)$

Satz 7.9. Seien X, Y stochastisch unabhängige Zufallsvariablen und $E(|X|) < \infty > E(|Y|)$, dann gilt: $E(X \cdot Y) < \infty$ und $E(X \cdot Y) = E(X) \cdot E(Y)$

Definition 7.10. Seien X, Y Zufallsvariablen und $c \in \mathbb{R}, k \in \mathbb{N}$:

1. $E((X - c)^k)$ heißtK-TES MOMENT von X und c (nichtzentrales Moment, $c = 0$ (zentrales) Moment $\Rightarrow$ 1.Moment $= E(X)$, $Var(X) =$ 2. Moment-(1.Moment)^2)
2. $E((X - EX)^2)$ heißt VARIANZ oder STREUUNG von X, kurz: $Var(X)$ oder $VarX = E(X^2) - E^2(X) = E(X(X - 1)) + E(X) - E^2(X)$
3. $E((X - EX)(Y - EY))$ heißt KOVARIANZ von X und Y, kurz: $Cov(X, Y) = E(XY) - E(X) \cdot E(Y)$ (Maß für linearen Zusammenhang zwische X und Y)

Satz 7.11. Sei X eine Zufallsvariable, $f : \mathbb{R} \to \mathbb{R}$ eine konvexe Funktion, so dass $E(f \circ X)$ und $E(X)$ existieren, dann gilt: $E(f \circ X) \geq f(EX)$. Speiziell $E(X^2) \geq (EX)^2$

Korrolar 7.12. Sei X eine reelwertige Zufallsvariable, $E(|X|^r) < \infty$ für ein $r \in (0, \infty)$. Dann existiert auch $E(|X|^s) \forall 0 < s \leq r$ und $(E(|X|^r))^{\frac{1}{r}} \geq (E(|X|^s))^{\frac{1}{s}}$

Lemma 7.13. Sei $X < \infty; a, b \in \mathbb{R}$

1. $Var(aX + b) = a^2 Var(X)$
2. $Var(X) = E(X^2) - (EX)^2$ (Häufig zur Berechnung von $Var(X)$)
3. $Var(X) = 0 \Leftrightarrow P(X \neq EX) = 0$
4. $Var(X) = \min_{a\in\mathbb{R}} E((X - a)^2)$

Bezeichnung 7.14. Eine Zufallsvariable X mit $EX = 0$ und $Var(X) = 1$ heißt STANDARDISIERT.
Weiterhin: Sei Y Zufallsvariable mit $EY = \mu < \infty$ und $0 < Var(Y) =: \sigma^2$, dann gilt: $X := \frac{Y - EY}{\sqrt{Var(Y)}} = \frac{Y-\mu}{\sigma}$ erfüllt gerade $EX = 0$ und $Var(X) = 1$.

Satz 7.15. Seien X, Y Zufallsvariablen mit $Var(X), Var(Y) < \infty$, da gilt $Var(X+Y) = Var(X) + Var(Y) + 2 \cdot Cov(X, Y)$ wobei $Cov(X, Y) = E((X - EX) \cdot (Y - EY)) = E(XY) - EX \cdot EY$

Bemerkung 7.16. Der Varianzoperator ist nicht lienar sondern benötigt den Korrekturterm $Cov(X, Y)$.

Satz 7.17. Seien X, Y Zufallsvariablen mit $EX^2, EY^2 < \infty$, dann gilt: $(E(X \cdot Y))^2 \leq EX^2 \cdot EY^2$, wobei Gleichheit genau dann gilt, wenn $\exists a \in \mathbb{R}$ $P(aX = Y) = 1$ gilt, d.h. X ein Vielfaches von Y ist.

Lemma 7.18. 1. $Cov(X, Y) = E(X \cdot Y) - EX \cdot EY$

2. $Cov(X, X) = Var(X)$
3. $Coc(X, Y) = Cov(Y, X)$
4. $Cov(aX + b, Y) = a \cdot Cov(X, Y)$
5. $Cov^2(X, Y) \leq Var(X) \cdot Var(Y)$
6. X, Y stochastisch unabhängig $\Rightarrow$ $Cov(X, Y) = 0$ Umkehrung gilt nicht!
7. $Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$

Bemerkung 7.19. 1. Die Kovarianz ist symmetrische Bilinearform

2. X, Y stochastisch unabhängig $\Rightarrow$ $Var(X + Y) = Var(X) + Var(Y)$
3. X, Y heißen UNKORRELIERT, wenn $Cov(X, Y) = 0$
4. Der KORRELATIONSKOEFFIZIENT ist definiert durch $Korr(X, Y) := \frac{Cov(X,Y)}{\sqrt{Var(X) \cdot Var(Y)}} \in [-1, 1]$
5. Aus der Unkorreliertheit folgt im Allgemeinen nicht die Unabhängigkeit!

8 Das schwache Gesetzt großer Zahlen

Definition 8.1. 1. Eine Folge $(X_n)_n$ von Zufallsvariablen über $(\Omega, \mathfrak{A}, \mathbf{P})$ heißt STOCHASTISCH KONVERGENT gegen 0, falls $\lim_{n\to\infty} P(|X_n| > \varepsilon) = 0, \forall \varepsilon > 0$ (Schreibweise $P - \lim_{n\to\infty} X_n = 0$)

2. Eine Folge $(X_n)_n$ heißt STOCHASTISCH KONVERGENT gegen $c \in \mathbb{R}$ bzw. gegen Zufallsvariable X, falls $P - \lim_{n\to\infty}(X_n - c) = 0$ bzw. $P - \lim_{n\to\infty}(X_n - X) = 0$

Satz 8.2. Der Grenzwert einer $P-$stochastisch konvergenten Folge ist im folgenden Sinne eindeutig definiert: $X_n \xrightarrow{P} X$ und $X_n \xrightarrow{P} Y \Rightarrow P(X = Y) = 1$ aber nicht zwangsläufig $X = Y$!

Satz 8.3. 1. $X_n \xrightarrow{P} X, Y_n \xrightarrow{P} Y, g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ stetig $\Rightarrow g(X_n, Y_n) \xrightarrow{P} g(X, Y)$

2. $X_n \to X$ (punktweise Konvergenz) $\Rightarrow X(\omega) = \lim_{n\to\infty} X_n(\omega), \forall \omega \in \Omega \Rightarrow X_n \xrightarrow{P} X$

Satz 8.4 (Markov'sche Ungleichung)**.** Sei X Zufallsvariable, $g : \mathbb{R}^+ \to \mathbb{R}^+$ monoton wachsend, dann gilt: $P(|X| > \varepsilon) \leq P(|X| \geq \varepsilon) \leq \frac{1}{g(\varepsilon)} \cdot E(g(|X|)), \forall \varepsilon > 0, g(\varepsilon) > 0$

Satz 8.5. Seien $(X_i)_{i\in\mathbb{N}}$ paarweise unkorrelierte Zufallsvariablen mit $EX_I =: \mu \forall i \in \mathbb{N}$ und $Var(X_i) \leq M \leq \infty$, dann gilt: $P\left(\left|\frac{1}{n}\sum_{i=1}^n X_i - \mu\right| > \varepsilon\right) \leq \frac{M}{n\cdot\varepsilon^2} \overset{n\to\infty}{\longrightarrow} 0 \Rightarrow P - \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n X_i = \mu$, d.h. das arithmetische Mittel von „Einzelversuchen“ (beschrieben durch $X_1, X_2, ...$) konvergiert stochastisch gegen den (unbekannten) Erwartungswert μ.

9 Borelmengen und Maße

Lemma 9.1. Ist $\mathcal{E} \subset \mathfrak{Pot}(\Omega)$, so existiert eine kleinste σ-Algebra $\mathfrak{A}(\mathcal{E})$, die $\mathcal{E}$ enthält, d.h.

i) $\mathfrak{A}(\mathcal{E})$ ist eine σ-Algebra

ii) $\mathcal{E} \subset \mathfrak{A}(\mathcal{E})$

iii) Ist $\mathfrak{A}'$ eine σ-Algebra mit $\mathcal{E} \subset \mathfrak{A}' \Rightarrow \mathfrak{A}(\mathcal{E}) \subseteq \mathfrak{A}'$

Bemerkung 9.2. i) $\mathfrak{A}(\mathcal{E})$ heißt die von $\mathcal{E}$ erzeugte σ-Algebra

ii) Ist $\mathcal{E}$ σ-Algebra $\Rightarrow \mathfrak{A}(\mathcal{E}) = \mathcal{E}$

iii) $\mathcal{E} = \{A\} \Rightarrow \mathfrak{A}(\mathcal{E}) = \{\varnothing, A, A^C, \Omega\}$

Bezeichnung 9.3. Seien $k \in \mathbb{N}$ und $\mathcal{E}^n := \{(a,b] | a, b \in \mathbb{R}\}$ (für $n = 1 : F^X(b) - F^X(a) = P(X \in (a,b]))$) so heißt $\mathfrak{B}^n := \mathfrak{A}(\mathcal{E}^n)(\mathfrak{B}^1 = \mathfrak{B})$ Borel'sche σ-Algebra oder σ-Algebra der Borelmenge über $\mathbb{R}^n$. Jedes $B \in \mathfrak{B}^n$ heißt Borelmenge

Definition 9.4. Sei $(\Omega, \mathfrak{A})$ ein messbarer Raum. Eine Abbildung $\mu : \mathfrak{A} \to [0, \infty]$ heißt Mass über $(\Omega, \mathfrak{A})$, falls gilt:

1. $\mu(\varnothing) = 0$
2. Für alle Familien $(A_i)_{i\in\mathbb{I}}$ von paarweise disjunkten $A_i \in \mathfrak{A}$ mit abzählhbarer Indexmenge $\mathbb{I}$ gilt: $\mu\left(\sum_{i\in\mathbb{I}} A_i\right) = \sum_{i\in\mathbb{I}} \mu(A_i)$

$(\Omega, \mathfrak{A}, \mu)$ heißt Massraum. Ein Maß μ mit $\mu(\Omega) = 1$ heißt Wahrscheinlichkeitsmass $\mathbf{P}$ und $(\Omega, \mathfrak{A}, \mathbf{P}$ heißt Wahrscheinlichkeitsraum.

Satz 9.5. Seien $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum und $A, B, A_1, A_2, ... \in \mathfrak{A}$. Das Maß μ bestitzt folgende Eigenschaften:

i) **Nulltreue:** $\mu(\varnothing) = 0$

ii) **Positivität:** $\mu(A) \geq 0, \forall A \in \mathfrak{A}$ ($\mu(A) = \infty$ ist möglich!)

iii) **Additivität:** Ist $A \cap B = \infty$, so gitl $\mu(A \cup B) = \mu(A) + \mu(B)$

iv) **Additivität:** Aus iii) folgt für paarweise disjunkte Mengen $A_1, ..., A_n$: $\mu\left(\sum_{i=1}^n A_I\right) = \sum_{i=1}^n \mu(A_i)$

v) **Isotonie:** Ist $A \subset B$, so gilt: $\mu(A) \leq \mu(B)$

vi) **Subtraktivität:** Sind $A \subset B$ und $\mu(\Omega) < \infty$, so gilt: $\mu(B \backslash A) = \mu(B) - \mu(A)$

vii) **Komplementarität:** Ist $\mu(\Omega) < \infty$, so gilt: $\mu(A^C) = \mu(\Omega) - \mu(A)$

viii) **Stetigkeit von unten:** Ist die Folge $(A_n)_{n\in\mathbb{N}}$ isoton (monoton wachsend), so gilt: $\mu\left(\bigcup_{n=1}^\infty A_n\right) = \lim_{n\to\infty} \mu(A_n)$

ix) **Stetigkeit von oben:** Ist die Folge $(A_n)_{n\in\mathbb{N}}$ antiton (monoton fallend), so gilt: $\mu\left(\bigcap_{n=1}^\infty A_n\right) = \lim_{n\to\infty} \mu(A_n)$

x) **Sub-Additivität:** Für Ereignisse $A_1, .., A_n$ gilt: $\mu\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mu(A_i)$

xi) **Subσ-Additivität:** Für eine Ereignisfolge $(A_i)_{i\in\mathbb{N}}$ gilt: $\mu\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mu(A_i)$

Satz 9.6. Sei $G : \mathbb{R} \to \mathbb{R}$ eine monoton wachsende, rechtsseitig stetige Funktion. Dann existiert genau ein Maß μ auf $(\mathbb{R}, \mathfrak{B})$ mit $\mu((a,b]) = G(b) - G(a), \forall a, b \in \mathbb{R}, a < b$

Lemma 9.7. Sei $f : \mathbb{R} \to \mathbb{R}^+$ und $I = (a,b) \subset \mathbb{R}, -\infty \leq a < b \leq \infty$, mit $f(x) = 0 \forall x \in I^C$, f stetig auf I und $\int_{-\infty}^{\infty} f(x)dx = \int_a^b f(x)dx = 1$ (Riemann-Integral)

10 Wahrscheinlichkeitsmaße mit Riemann-Dichten über $\mathbb{R}$

Beispiel 10.1 (Rechteckverteilung (stetige Gleichverteilung))**.** auf $(a,b), a, b \in \mathbb{R}$ $f(x) = \frac{1}{b-a} \cdot \mathcal{I}_{[a,b]}(x) = \begin{cases} \frac{1}{b-a} & , x \in [a,b] \\ 0 & , \text{sonst} \end{cases}, x \in \mathbb{R}$ Die durch Dichtefunktion definierte Wahrscheinlichkeitsverteilung über (a,b) heißt stetige Gleichverteilung. Verteilungsfunktion: $F(x) = \begin{cases} 0 & , x < a \\ \frac{x-a}{b-a} & , a \leq x \leq b \\ 1 & , x > b \end{cases}$ Zum Berechnen der Wahrscheinlichkeiten: $P(X \in (c,d)) = \int_c^d f(x)dx = F(d) - F(c), (c,d) \subset (a,b)$

Beispiel 10.2 (Exponentialverteilung)**.** $f(x) = \begin{cases} \lambda \cdot e^{-\lambda x} & x > 0 \\ 0 & , x \leq 0 \end{cases}, x \in \mathbb{R}, \lambda > 0$ $F(X) = \begin{cases} 0 & x \leq 0 \\ 1 - e^{-\lambda x} & , x > 0 \end{cases}$ Schreibweise: $Exp(\lambda)$

Beispiel 10.3 (Weibull-Verteilung)**.** $f(x) = \begin{cases} \alpha\beta x^{\beta-1} e^{-\alpha x^\beta} & , x > 0 \\ 0 & , x \leq 0 \end{cases}, x \in \mathbb{R}, \alpha, \beta > 0$, u.a. für die Lebensdauer: $F(X) = \begin{cases} 0 & x \leq 0 \\ 1 - e^{-\alpha x^\beta} & , x > 0 \end{cases}$ Schreibweise: $Wei(\alpha, \beta)$

Beispiel 10.4 (Gammaverteilung)**.** $f(x) = \frac{p^b}{\Gamma(p)} \cdot e^{-bx} \cdot x^{p-1} \cdot \mathcal{I}_{(0,\infty)}(x), x \in \mathbb{R}, a, p > 0$ (Für $b = \lambda, p = 1$ ergibt sich $Exp(\lambda)$) $\Rightarrow$ Problem bei der Angabe von F, da dies geschlossen nur für $p \in \mathbb{N}$ möglich ist.

Beispiel 10.5 (Gauß'sche Normalverteilung)**.** $f(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \mu, x \in \mathbb{R}, \sigma > 0$ Schreibweise: $\mathcal{N}(\mu, \sigma)$

Definition 10.6. Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ heißt (Riemann-)Dichte auf $\mathbb{R}^n$, falls gilt:

- $f(x) \geq 0, \forall x \in \mathbb{R}^n$
- f ist Riemann-integrierbar mit
- $\int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} f(x_1, ...x_n)dx_1...dx_n = 1$

Satz 10.7. Ist f eine Dichte über $\mathbb{R}^n$, so definiert $F(x_1, ..., x_n) = \int_{-\infty}^{x_n} ... \int_{-\infty}^{x_1} f(y_1, ..., y_n)dy_1...dy_n, (x_1, ..., x_n) \in \mathbb{R}^n$ eine stetige Verteilungsfunktion über $\mathbb{R}^n$. Ist $\mathbf{P}$ das zugehörige Wahrscheinlichkeitsmaß über $(\mathbb{R}^n, \mathfrak{B}^n)$ heißt f eine Dichte von $\mathbf{P}$. Dann ist $P\left(\times_{i=1}^n [a_i, b_i]\right) = \int_{a_n}^{b_n} ... \int_{a_1}^{b_1} f(y_1, ..., y_n)dy_1...dy_n, \forall a_i \leq b_i \in \mathbb{R}$.

Bezeichnung 10.8. Sei f die Riemann-Dichte von X, so heißt $EX := \int_{-\infty}^{+\infty} x \cdot f(x)dx$ Erwartungswert von X (falls wohldefiniert).

Bemerkung 10.9. Es gelten (weiterhin) nach dem Ersetzen von Summen durch Integrale:

- $E(g \circ X) = \int_{-\infty}^{+\infty} g(x)f(x)dx$, z.B. $g(x) := (X - EX)^2 \Rightarrow E(g \circ X) = Var(X) = \int_{-\infty}^{+\infty} (X - EX)^2 f(x)dx$
- Eigenschaften der Erwartungswerte, Varianz, Kovarianz
- Ungleichungen von Jensen, Ljapunoff, Cauchy-Schwarz und Markov und Techebyscheff

Beispiel 10.10.

$X \sim$	$E(X)$	$E(X^2)$	Var(X)
$b(n,p)$	$n \cdot p$		$np(1-p)$
$po(\lambda)$	λ		λ
$R(a,b)$	$\frac{b+a}{2}$	$\frac{1}{3}(b^2 + ab + a^2)$	$\frac{(a-b)^2}{12}$
$Exp(\lambda)$	$\frac{1}{\lambda}$		$\frac{1}{\lambda^2}$
$Wei(\alpha, \beta)$	$\alpha^{-\frac{1}{\beta}} \cdot \Gamma\left(\frac{1}{\beta} + 1\right)$		$\alpha^{-\frac{2}{\beta}} \cdot \left(\Gamma\left(\frac{2}{\beta} + 1\right) - \Gamma^2\left(\frac{1}{\beta} + 1\right)\right)$
$\Gamma(b,p)$	$\frac{p}{b}$		$\frac{p}{b^2}$
$\mathcal{N}(\mu, \sigma^2)$	μ		σ^2

Bemerkung 10.11 (Ausfallrate)**.** $h(x) = \frac{f(x)}{1-F(x)}$ ($X \sim Exp(\lambda) \Rightarrow h(x) = \lambda; X \sim Wei(\alpha, \beta) \Rightarrow h(x) = \begin{cases} \uparrow & , \beta > 1 \equiv \text{Abnutzung} \\ \text{konst.} & , \beta = 1 \equiv \text{keine Alterung} \\ \downarrow & , \beta < 1 \equiv \text{System wird stabiler} \end{cases}$)

Bezeichnung 10.12. Verteilungsfunktion wurde eingeführt als Integral mit oberer Grenze als Argument. Die so eingeführte Verteilungsfunktion hat (sogar) die Eigenschaft der absoluten Stetigkeit, d.h. die Dichte f von X ist gegeben durch $F'(x) = f(x)$.

Bemerkung 10.13. Die Zufallsvariablen $X_1, ..., X_n : (\Omega, \mathfrak{A}, \mathbf{P}) \to (\mathbb{R}, \mathfrak{B})$ sind stochastisch unabhängig $\iff F^{(X_1,...,X_n)}(x_1, ..., x_n) = F^X(x) = F^{X_1}(x_1) \cdot ... \cdot F^{X_n}(x_n), \forall (x_1, ..., x_n) \in \mathbb{R}^n$, d.h. $\mathbf{P}(X_1 \leq x_1, ..., X_n \leq x_n) = \prod_{i=1}^n P(X_i \leq x_i), \forall x_i \in \mathbb{R}$ ist notwendig und hinreichend für die stochastische Unabhängigkeit von $X_1, ..., X_n$

Lemma 10.14. Sei $X = (X_1, ..., X_n)$ ein absolut stetiger Zufallsvektor mit Dichtefunktion f^X. Dann gilt für die i-te Randdichte $F^{X_i}(x) = f^{X_i}(x) = \underbrace{\int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty}}_{n-1 \text{ Integrale}} f^X(x_1, ..., x_{i-1}, x, x_{i+1}, ..., x_n) dx_1...dx_{i-1}dx_{i+1}...dx_n, x \in \mathbb{R}$.

Ferner gilt: $f^X(x_1,...,x_n) = \prod_{i=1}^n d^{X_i}(x_i), \forall x_i \in \mathbb{R} \iff X_1,...,X_n$ stochastisch unabhängig mit Dichten $f^{X_1},...,f^{X_n}$.

Korrolar 10.15. Sei (X_1, X_2) ein absolut stetiger Zufallsvektor mit der Dichtefunktion $f^{(X_1,X_2)}$ dann ist $Y = X_1 + X_2$ absolut stetig mit der Dichte $f^Y(y) = \int_{-\infty}^{+\infty} f^X(t, y-z)dt, y \in \mathbb{R}$

Bemerkung 10.16. Sind X_1 und X_2 stochastisch unabhängig, dann ist $f^{(X_1,X_2)}(x_1,x_2) = f^{X_1}(x_1) \cdot f^{X_2}(x), \forall x_1, x_2 \in \mathbb{R}$ und damit $f^{X_1+X_2}(y) = \int_{-\infty}^{+\infty} f^{X_1}(t) \cdot f^{X_2}(y-t)dt, y \in \mathbb{R}$

Bezeichnung 10.17. Die Verteilung $X_1 + X_2$ heißt Faltung von X_1 und X_2.

Bezeichnung 10.18. $X_1 + X_2$ besitzt eine Dreiecksverteilung.

Bezeichnung 10.19. Die Klasse der Γ-Verteilungen (bei festem Parameter) ist Faltungstabil, d.h. die Faltung führt nicht aus der Klasse der Verteilungen heraus.

Definition 10.20. Sei $(X_i)_{i\in\mathbb{N}}$ eine Folge von stochastisch unabhängig iid. nicht-negativen Zufallsvariablen. Die Folge der Partiasummen $\left(\underbrace{\sum_{i=1}^n X_i}_{=:S_n}\right)_{n\in\mathbb{N}}$ mit $S_0 = 0$ heißt Erneuerungsprozess. Für jedes $t \geq 0$ wird die Zufallsvariable N_t definiert durch $N_t := sup\{n \in \mathbb{N}, S_n \leq t\} (= \sum_{i=1}^\infty \mathcal{I}_{\{S_i \leq t\}})$. $(N_t)_{t\geq 0}$ heißt Erneuerungsprozess.

Lemma 10.21. Sei X_i die Lebensdauer von Komponenten und S_n die Gesamtlebensdauer, so ist N_t die Anzahl der ausgefallenen Komponenten bis zum Zeitpunkt t. Es gilt: $S_n \leq t \Leftrightarrow N_t \geq n \forall t > 0 \forall n \in \mathbb{N}$
Dann gilt: $X_i \sim Exp(\lambda) \forall i \in \mathbb{N}$ gilt $N_t \sim po(\lambda \cdot t) \forall t > 0$

11 Grundlagen der Simulation

Lemma 11.1. Für $[u,v] \subset [0,1]$ gilt: $|\mathbf{P}_m(\{\omega \in \Omega_m | u \leq \omega \leq v\}) - (v-u)| \leq \frac{1}{m}$ Wobei $(v-u)$ die Wahrscheinlichkeit nach der Rechteckverteilung für das Intervall $[u,v]$ ist.

Lemma 11.2. Sind $r_1,...,r_n, s_1,...,s_n \in [0,1]$ mit $r_j - s_j \leq \varepsilon, \forall j \in \{1,...,n\}$, so folgt $\left|\prod_{j=1}^n r_j - \prod_{j=1}^n s_j\right| \leq n \cdot \varepsilon$

Bezeichnung 11.3. Sei $m \in \mathbb{N}$ (Modul), $a \in \mathbb{N}_0$ (Faktor), $b \in \mathbb{N}_0$ (Inkrement), $z_0 \in \mathbb{N}_0$, $z_0 \leq m-1$ (Anfangsglied). Kongruenzschema: $z_{j+1} = (a \cdot z_j + b)(mod m), j \in \mathbb{N}_0$; Klar ist: $0 \leq z_j \leq m-1, \forall j \in \mathbb{N}_0$; Normierung: $x_j := \frac{z_k}{m}, j \in \mathbb{N}_0$ liefert die Folge $(x_n)_{n\in\mathbb{N}} \subset [0,1]$.

Satz 11.4. Für $b \geq 1$ wird die maximal mögliche Periodenlänge genau dann erreicht, wenn:

i) b ist teilerfremd zu m

ii) Jede Primzahl, die m teilt, teilt auch $a-1$

iii) Ist m durch 4 teilbar, so muss auch $a-1$ durch 4 teilbar sein

Bemerkung 11.5 (Regel)**.** Auch bei maximaler Periodenlänge dürfen nicht alle möglichen Pseudozufallszahlen für eine Simulation „verbraucht“ werden. Etwa wäre stets die letzte Zahl vorhersagbar.

12 Einführung in die Statistik

Definition 12.1. Die Menge aller möglichen Realisationen ist $\mathfrak{X}^n = (X(\Omega))^n$. Jede messbare Abbildung $\delta : \begin{cases} \mathfrak{X}^n & \to \mathcal{D} \\ x = (x_1,...,x_n) & \mapsto d \end{cases}$ heißt statistische Entscheidungsfunktion (SEF).

Bezeichnung 12.2. Eine Klasse $\mathcal{P}$ von Wahrscheinlichkeitsverteilungen heißt parametrisch, falls $\exists \Theta \subset \mathbb{R}^l$, Abbildung $h : \begin{cases} \Theta & \to \mathcal{P} \\ \theta & \mapsto P_\theta \end{cases}$, h bijektiv (Identifizierbarkeit).

Bezeichnung 12.3. Gegeben sei eine parametrische Klasse $\mathcal{P} = \{P_\theta | \theta \in \Theta\}$ von Wahrscheinlichkeiten. Eine SEF (statistische Entscheidungsfunktion) $\delta : \mathfrak{X}^n \to \Theta$ heißt Schätzfunktion. (Parameterschätzung, Punktschätzung)

Definition 12.4. Gegeben sei eine paramtrische Klasse $\mathcal{P}$ von Wahrscheinlichkeitsverteilungen, ein Stichprobenraum $\mathfrak{X}^n$, ein Parameterraum $\Theta = \mathcal{D} \subset \mathbb{R}$ und eine Menge $\Delta = \{\delta | \delta : \mathfrak{X}^n \to \Theta\}$ von Schätzfunktionen. $\delta \in \Delta$ nennt man erwartungstreu, falls gilt: $E_\theta(\delta(X)) = \theta \forall \theta \in \Theta = \left\{\begin{array}{ll} \int \delta(x) \cdot f_\theta(x)dx & \text{stetig} \\ \sum \delta(x_i) \cdot p_\theta(x_i) & \text{diskret} \end{array}\right\}$ (Im Mittel liefert die Schätzfunktion den wahren Wert)

Bezeichnung 12.5. Gegeben sei $\mathcal{P} = \{P_\theta | \theta \in \Theta\}$ auf $\mathfrak{X}^n$ mit Riemann- bzw. Lebesque-Dichten $f_\theta, \theta \in \Theta$ oder diskrete Wahrscheinlichkeitsverteilung, so heißt $L = \begin{cases} (\mathfrak{X}^n, \Theta) & \to [0,1] \\ (x,\theta) & \mapsto \begin{cases} f_\theta(x) & \text{kontinu. Fall} \\ P_\theta(X = x) & \text{diskreter Fall} \end{cases} \end{cases}$ Likelihood-Funktion zur Realisation X. (Funktion des Parameters θ bei festem x)

Bezeichnung 12.6. Fehler 1. Art: $p \in \mathcal{P}^0$ ablehnen, obwohl richtig; Fehler 2. Art: $p \in \mathcal{P}^0$ annehnen, obwohl falsch;

	$\mathcal{P}^0$	$\mathcal{P}^1$
$\mathcal{P}^0$	kein Fehler	Fehler 2. Art
$\mathcal{P}^1$	Fehler 1. Art	kein Fehler

Bemerkung 12.7. Im Allgemeinen ist es nicht möglich beide Fehlerwahrscheinlichkeiten simultan zu minimieren.

Bemerkung 12.8. Durch statistischen Test kann immer nur die Aussage in der Alternativen Hypothese belegt werden.

13 Formeln (aus anderen Bereichen der Mathematik

iid = stochastisch unabhängig, identisch verteilt

$$\sum_{n=0}^\infty \frac{(\lambda \cdot t)^k}{k!} = e^{\lambda t}$$

$$\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$$

$$\sum_{i=1}^n i = \frac{n(n+1)}{2}$$

$$\binom{n}{k} = \binom{n}{n-k}$$

$$\sum_{i=0}^\infty (1-p)^i = \frac{1}{1-(1-p)} = \frac{1}{p}$$

$$\sum_{i=0}^n a \cdot q^k = a \cdot \frac{q^{n+1}-1}{q-1}$$

$$\sum_{i=1}^n \binom{n}{i} a^i b^{n-i} = (a+b)^n$$

$$\sum_{k=0} \frac{\lambda^k}{k!} e^{-\lambda} = 1$$

$$\sum_{j=0}^\infty \frac{1}{2^j} = \frac{1}{1-\frac{1}{2}}$$

$$\lim_{n\to\infty} \frac{n!}{\left(\frac{n}{e}\right)^n \sqrt{2\pi n}} = 1$$

$$\binom{m}{m_1,...,m_n} = \frac{m!}{m_1! \cdot ... \cdot m_n!}$$

Lemma 13.1 (Regeln der Mengenlehre)**.**

1. $A \cup B = B \cup A$, $A \cap B = B \cap A$
2. $(A \cup B) \cup C = A \cup (B \cup C)$, $(A \cap B) \cap C = A \cap (B \cap C)$
3. $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$, $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$
4. $(A \cup B)^C = A^C \cap B^C$, $(A \cap B)^C = A^C \cup B^C$
5. $A \backslash B = A \cap B^C$

Lemma 13.2 (Bernoulli)**.** $(1+p)^n > 1 + np, (1-p)^n > 1 - np$ für $0 < p < 1, n > 1$

Lemma 13.3 (Logarithmen)**.** $log_a(b \cdot c) = log_a b + log_a c, log_a(b : c) = log_a b - lob_a c, log_a(b^r) = r \cdot log_a b$

Lemma 13.4 (Potenzgesetze)**.** $a^p \cdot a^q = a^{p+q}, a^p : a^r = a^{p-r}, (a^p)^r = a^{p \cdot r}$

Lemma 13.5 (Ableitungsregeln)**.** $(f \pm g)' = f' \pm g', (f \cdot g)' = f' \cdot g + f \cdot g', \left(\frac{f}{g}\right)' = \frac{f' \cdot g - f \cdot g'}{g^2}, \left(\frac{1}{g}\right)' = \frac{g'}{g^2}, (g \circ f)' = (g' \circ f) \cdot f', \int_a^b (f \circ g)(x) \cdot g'(x)dx = \int_{g(a)}^{g(b)} f(z)dz, \int_a^b f'(x) \cdot g(x)dx = f(x) \cdot g(x)|_a^b - \int_a^b f(x)\dot{g}'(x)dx$

Lemma 13.6 (Ableitungen)**.** $f(x) = \frac{1}{x} \Rightarrow F(x) = ln|x|, f(x) = \sqrt{x} \Rightarrow F(x) = \frac{2}{3}x^{\frac{3}{2}}, f(x) = ln(x) \Rightarrow F(x) = x(ln(x) - 1), f(x) = ln(a) \cdot a^x \Rightarrow F(x) = a^x$