Zusammenfassung Stochastik I

Tim Baumann, http://timbaumann.info/uni-spicker

Der abstrakte Maßbegriff

Definition. Eine **Ereignisalgebra** oder **Boolesche Algebra** ist eine Menge $\mathfrak{A}$ mit zweistelligen Verknüpfungen $\wedge$ („und") und $\vee$ („oder"), einer einstelligen Verknüpfung $\overline{\cdot}$ (Komplement) und ausgezeichneten Elementen $U \in \mathfrak{A}$ (unmögliches Ereignis) und $S \in \mathfrak{A}$ (sicheres Ereignis), sodass für $A, B, C \in \mathfrak{A}$ gilt:

i. $A \wedge A = A$
ii. $A \wedge B = B \wedge A$
iii. $A \wedge S = A$
iv. $A \wedge U = U$
v. $A \wedge \overline{A} = U$
vi. $A \wedge (B \wedge C) = (A \wedge B) \wedge C$
vii. $A \vee A = A$
viii. $A \vee S = S$
ix. $A \vee U = A$
x. $A \vee \overline{A} = S$
xi. $A \vee (B \vee C) = (A \vee B) \vee C$
xii. $A \wedge (B \vee C) = (A \wedge B) \vee (A \wedge C)$

Definition. Sei $\mathfrak{A}$ eine Boolesche Algebra. Dann definiert

$$A \leq B :\iff A \wedge B = B$$

eine Partialordnung auf $\mathfrak{A}$, gesprochen A impliziert B.

Definition. Eine **Algebra** (auch Mengenalgebra) $\mathfrak{A} \subset \mathcal{P}(\Omega)$ ist ein System von Teilmengen einer Menge Ω mit $\emptyset \in \mathfrak{A}$, das unter folgenden Operationen stabil ist:

- Vereinigung: $A, B \in \mathfrak{A} \implies A \cup B \in \mathfrak{A}$
- Durchschnitt: $A, B \in \mathfrak{A} \implies A \cap B \in \mathfrak{A}$
- Komplementbildung: $A \in \mathfrak{A} \implies A^c := \Omega \backslash A \in \mathfrak{A}$

Satz (Isomorphiesatz von Stone)**.** Zu jeder Booleschen Algebra $\mathfrak{A}$ gibt es eine Menge Ω derart, dass $\mathfrak{A}$ isomorph zu einer Mengenalgebra $\mathfrak{A}$ in $\mathcal{P}(\Omega)$ ist.

Definition. Eine **σ-Algebra** ist eine Algebra $\mathfrak{A} \subset \mathcal{P}(\Omega)$, die nicht nur unter endlichen, sondern sogar unter abzählbaren Vereinigungen stabil ist, d. h.

$$(A_n)_{n\in\mathbb{N}} \text{ Folge in } \mathfrak{A} \implies \bigcup_{n=0}^{\infty} A_n \in \mathfrak{A}.$$

Bemerkung. Es gilt damit:

- $\Omega = \emptyset^c \in \mathfrak{A}$
- Abgeschlossenheit unter abzählbaren Schnitten:

$$(A_n)_{n\in\mathbb{N}} \text{ Folge in } \mathfrak{A} \implies \bigcap_{n=0}^{\infty} A_n = \left(\bigcup_{n=0}^{\infty} (A_n)^c\right)^c \in \mathfrak{A}.$$

Definition. Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge in einer σ-Algebra $\mathfrak{A}$. Setze

$$\limsup_{n\to\infty} A_n := \bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_n \in \mathfrak{A}, \quad \liminf_{n\to\infty} A_n := \bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_n \in \mathfrak{A}.$$

Bemerkung. In einer σ-Algebra, in der die Mengen mögliche Ereignisse beschreiben, ist der Limes Superior das Ereignis, das eintritt, wenn unendlich viele Ereignisse der Folge A_n eintreten. Der Limes Infinum tritt genau dann ein, wenn alle bis auf endlich viele Ereignisse der Folge A_n eintreten.

Definition. Ein **Ring** $\mathfrak{A} \subset \mathcal{P}(\Omega)$ ist ein System von Teilmengen einer Menge Ω mit $\emptyset \in \mathfrak{A}$, das unter folgenden Operation stabil ist:

- Vereinigung: $A, B \in \mathfrak{A} \implies A \cup B \in \mathfrak{A}$
- Differenz: $A, B \in \mathfrak{A} \implies B \backslash A = B \cap A^c \in \mathfrak{A}$

Ein Ring, der nicht nur unter endlicher, sondern sogar unter abzählbarer Vereinigung stabil ist, heißt **σ-Ring**.

Bemerkung. $\mathfrak{A}$ (σ-) Algebra $\iff$ $\mathfrak{A}$ (σ-) Ring und $\Omega \in \mathfrak{A}$.

Satz. Sei $(\mathfrak{A}_i)_{(i\in I)}$ eine Familie von (σ-) Ringen / (σ-) Algebren über einer Menge Ω. Dann ist auch $\cup_{i\in I}\mathfrak{A}_i$ ein (σ-) Ring / eine (σ-) Algebra über Ω.

Satz. Sei $\mathfrak{R}$ ein Ring und μ ein Inhalt. Es gelten für $n \in \mathbb{N}$ und $A_1, ..., A_n \in \mathfrak{R}$ die Ein- und Ausschlussformeln

$$\mu(A_1 \cup ... \cup A_n) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1\leq i_1<...<i_k\leq n} \mu(A_{i_1} \cap ... \cap A_{i_k}),$$

$$\mu(A_1 \cap ... \cap A_n) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1\leq i_1<...<i_k\leq n} \mu(A_{i_1} \cup ... \cup A_{i_k}).$$

Bemerkung. Sei μ ein W-Maß auf $\mathcal{L}(\mathbb{R}^1)$. Dann definiert $x \mapsto F_\mu(x) := \mu(]-\infty, x])$ eine VF. Für eine VF $F : \mathbb{R} \to [0, 1]$ definiert umgekehrt $\mu_F(]a, b]) := F(b) - F(a)$ ein W-Maß auf $\mathcal{L}(\mathbb{R}^1)$. Analog funktioniert dies auf dem $\mathbb{R}^d$.

Definition (Wichtige Verteilungsfunktionen)**.**

- **Normalverteilung** (Gaußverteilung) mit EW μ und Varianz σ^2:

$$F_{\mu\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} \exp\left(\frac{-(t-\mu)^2}{2\sigma^2}\right) \mathrm{d}t$$

  erfüllt $F'_{\mu\sigma^2}(x) = \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$, $F_{\mu\sigma^2}(\mu - x) = 1 - F_{\mu\sigma^2}(\mu + x)$
- **Exponentialverteilung** mit Parameter $\lambda > 0$:

$$F_\lambda(x) = \begin{cases} 0, & \text{für } x \leq 0 \\ 1 - \exp(-\lambda x), & \text{für } x > 0 \end{cases}$$

- **Poisson-Verteilung** mit Parameter $\lambda > 0$:

$$F_\lambda(x) = \sum_{0\leq n\leq x} \frac{\lambda^n}{n!} \exp(-\lambda)$$

Definition. Ein Ereignis $A \in \mathfrak{A}$ trete bei n Versuchen genau $h_n(A) \in \mathbb{N}$ mal auf. Dann heißt

- $h_n(A)$ **absolute Häufigkeit** von A,
- $H_n(A) := \frac{h_n(A)}{n}$ **relative Häufigkeit** von A.

Bemerkung. Unmittelbar klar:

- $H_n(A) \in [0, 1]$
- $H_n(A) \leq H_n(B)$ für $A \subset B$
- $H_n(A \sqcup B) = H_n(A) + H_n(B)$ für $A \cap B = \emptyset$

Bemerkung. Bei wachsendem n stabilisiert sich normalerweise der Wert $H_n(A)$. Dieser Grenzwert ist die Wahrscheinlichkeit von A.

Definition. Seien $A, B \in \mathfrak{A}$ Ereignisse, $n \in \mathbb{N}$ die Anzahl der Versuche. Dann heißt

$$H_n(A \mid B) := \frac{H_n(A\cap B)}{H_n(B)} = \frac{h_n(A\cap B)}{h_n(B)}$$

die **relative Wahrscheinlichkeit** von A unter der Bedingung B.

Bemerkung. Offenbar gilt:

- $H_n(A \mid B) \in [0, 1]$
- $H_n(A_1 \mid B) \leq H_n(A_2 \mid B)$ für $A_1 \subset A_2$
- $H_n(A_1 \sqcup A_2 \mid B) = H_n(A_1 \mid B) + H_n(A_2 \mid B)$ für $A_1 \cap A_2 = \emptyset$

Definition. Sei $\Omega \in \mathcal{L}(\mathbb{R}^d)$ mit $\lambda_d(\Omega) > 0$. Dann heißt das W-Maß

$$\mathbb{P} : \mathcal{L}(\Omega) \to [0, 1], \quad A \mapsto \frac{\lambda_d(A)}{\lambda_d(\Omega)}$$

auf $(\Omega, \mathcal{L}(\Omega))$ **Gleichverteilung**.

Definition. Sei Ω eine endliche Menge. Dann definiert

$$\mathbb{P} : \mathcal{P} \to [0, 1], \quad A \mapsto \frac{|A|}{|\Omega|} = \frac{\#\text{ günstige Fälle}}{\#\text{ mögliche Fälle}}$$

ein W-Maß auf $(\Omega, \mathcal{P}(\Omega))$, genannt **Laplace'sche Wkt**.

Bemerkung. Damit sind Berechnungen von Wkten mit kombinatorischen Überlegungen möglich.

Lemma (Fundamentalprinzip des Zählens)**.** Seien $A_1, ..., A_n$ endliche Mengen. Dann gilt $|A_1 \times ... \times A_n| = |A_1| \cdots |A_n|$.

Lemma. Sei A eine endliche Menge, $r \leq n := |A| < \infty$. Dann ist die Anzahl der r-Tupel mit Elementen aus A gleich

	Mit Wdh.	Ohne Wdh.
Mit Ordnung	n^r	$\frac{n!}{(n-r)!}$
Ohne Ordnung	$\frac{(n+r-1)!}{r!}$	$\binom{n}{r} := \frac{n!}{r!(n-r)!}$

Lemma. Sei A eine endliche Menge, $n := |A| < \infty$. Dann ist die Anzahl der möglichen Zerlegungen von A in disjunkte Mengen $B_1, ..., B_k$ mit $|B_i| = n_i$ und $n_1 + ... + n_k = n$ gleich

$$\binom{n}{n_1, ..., n_k} := \frac{n!}{n_1! \cdots n_k!}. \quad \textbf{(Multinomialkoeffizient)}$$

Modell. Eine Urne enthalte N Kugeln, darunter $M \leq N$ schwarze. Dann ist ist die Wkt für das Ereignis A^n_m, dass sich unter n gezogenen Kugeln genau $m \leq \min(n, M)$ schwarze Kugeln befinden,

$$\mathbb{P}(A^n_m) = \frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}}. \quad \textbf{(hypergeometrische Verteilung)}$$

Bemerkung. Für Maximum-Likelihood-Schätzungen:

- Der Ausdruck $\binom{N-M}{n-m}/\binom{N}{n}$ wird maximal bei $N := \lfloor\frac{n-M}{m}\rfloor$.
- Der Ausdruck $\binom{M}{m} \cdot \binom{N-M}{n-m}$ wird maximal bei $M := \lfloor\frac{m(N-1)}{n}\rfloor$.

Modell. Eine Urne enthalte N Kugeln in $k \leq N$ verschiedenen Farben, darunter N_1 in der ersten Farbe, ..., N_k in der k-ten Farbe, $N_1 + ... + N_k = N$. Dann ist ist die Wkt für das Ereignis $A^n_{n_1,...,n_k}$, dass sich unter n gezogenen Kugeln genau $n_1 \leq N_1$ Kugeln der ersten Farbe, ..., und $n_k \leq N_k$ Kugeln der k-ten Farbe befinden, $n_1 + ... + n_k = n$, gleich

$$\mathbb{P}(A^n_{n_1,...,n_k}) = \frac{\binom{N_1}{n_1} \cdots \binom{N_k}{n_k}}{\binom{N}{n}}.$$

Diese W-Verteilung heißt **polyhypergeometrische Verteilung**.

Definition. Sei $(\Omega, \mathfrak{A}, \mathbb{P})$ ein W-Raum und $A, B \in \mathfrak{A}$. Dann heißt

$$\mathbb{P}(A \mid B) := \begin{cases} \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, & \text{falls } \mathbb{P}(B) > 0 \\ 0, & \text{falls } \mathbb{P}(B) = 0 \end{cases}$$

Wahrscheinlichkeit von A unter der Bedingung B.

Bemerkung. Falls $\mathbb{P}(B) > 0$ gilt, so ist $\mathbb{P}(- \mid B)$ ein W-Maß über B auf der Spur-σ-Algebra $\mathfrak{A}|_B$.

Lemma. Seien $A_1, ..., A_k \in \mathfrak{A}$, dann gilt die Pfadregel:

$$\mathbb{P}(A_1 \cap ... \cap A_k) = \mathbb{P}(A_1) \cdot \prod_{i=2}^{k} \mathbb{P}(A_i \mid A_1 \cap ... \cap A_{i-1}).$$

Satz. Sei $(\Omega, \mathfrak{A}, \mathbb{P})$ ein W-Raum und $A_1, ... \in \mathfrak{A}$ ein vollständiges Ereignissystem, d. h. paarweise disjunkt mit

$$\Omega = \coprod_{i=1}^{\infty} A_i.$$

Dann gilt für jedes $B \in \mathfrak{A}$ mit $\mathbb{P}(B) > 0$

$$\mathbb{P}(B) = \sum_{i=1}^{\infty} \mathbb{P}(B \mid A_i) \cdot \mathbb{P}(A_i) \qquad \text{(Formel der totalen Wkt)}$$

$$\mathbb{P}(A_n \mid B) = \frac{\mathbb{P}(B \mid A_n) \cdot \mathbb{P}(A_n)}{\sum_{i=1}^{\infty} \mathbb{P}(B \mid A_i) \cdot \mathbb{P}(A_i)} \qquad (\textbf{Bayessche Formel})$$

Sprechweise. In der Bayesischen Statistik heißt

- $\mathbb{P}(A_i)$ **A-priori-Wahrscheinlichkeit**,
- $\mathbb{P}(A_i \mid B)$ **A-posteriori-Wahrscheinlichkeit**.

Definition. Zwei Ereignisse $A, B \in \mathfrak{A}$ heißen **($\mathbb{P}$-)unabhängig**, falls

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

Bemerkung. • $A \in \mathfrak{A}$ mit $\mathbb{P}(A) = 0$ ist unabhängig zu jedem $B \in \mathfrak{A}$.
• Wenn $A, B \in \mathfrak{A}$ unabhängig, dann sind auch unabhängig:

$$(A^c, B), \quad (A, B^c), \quad (A^c, B^c)$$

Satz. $A, B \in \mathfrak{A}$ unabhängig $\iff$ $\mathbb{P}B \mid A = \mathbb{P}(B)$.

Definition. Sei $(A_i)_{i \in I}$ (I bel.) eine Familie von Ereignissen in $\mathfrak{A}$.

- **vollständig unabhängig**, falls

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_m}) = \mathbb{P}(A_{i_1}) \cdot \mathbb{P}(A_{i_2}) \cdots \mathbb{P}(A_{i_n})$$

für alle $i_1, ..., i_n \in I$ mit $2 \leq n < \infty$ und

- **paarweise unabhängig**, falls

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i) \cdot \mathbb{P}(A_j) \qquad \text{für alle } i, j \in I, i \neq j.$$

Achtung. Aus paarweiser Unabhängigkeit folgt nicht vollständige Unabhängigkeit (Gegenbeispiel: Bernsteins Tetraeder).

Definition. Sei $(\Omega, \mathfrak{A}, \mathbb{P})$ ein W-Raum und $\mathfrak{A}_1, \mathfrak{A}_2 \subset \mathfrak{A}$ Ereignissysteme. Dann heißen $\mathfrak{A}_1$ und $\mathfrak{A}_2$ **unabhängig**, falls

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \quad \text{für alle } A_1 \in \mathfrak{A}_1, A_2 \in \mathfrak{A}_2.$$

Satz. Seien $\mathfrak{A}_1, \mathfrak{A}_2 \subset \mathfrak{A}$ unabhängige Ereignissysteme, die Algebren sind. Dann sind auch die σ-Algebren $\sigma(\mathfrak{A}_1)$ und $\sigma(\mathfrak{A}_2)$ unabhängig.

Satz. Sei $(\Omega, \mathfrak{A}, \mathbb{P})$ ein W-Raum, $(A_i)_{i \in \mathbb{N}}$ Folge von unabhängigen Ereignissen mit gleicher Erfolgswkt $\mathbb{P}(A_i) = p$ für alle $i \in \mathbb{N}$. Für $k \leq n$, $k, n \in \mathbb{N}$ ist dann die Wahrscheinlichkeit, dass genau k Stück der Ereignisse $A_1, ..., A_n$ eintreten, genau

$$B(k, n, p) := \binom{n}{k} p^k \, (1-p)^{n-k}$$

Die zugehörige VF $x \mapsto \sum_{0 \leq k \leq x} B(k, n, p)$ heißt **Binomialverteilung**.

Lemma. Voraussetzung wir im vorherigen Satz. Sei $r, k \in \mathbb{N}$, $1 \leq r$, dann ist die Wkt für das Ereignis $A^{(r)}_k$, dass beim Versuch A_{k+r} der r-te Erfolg eintritt, gleich

$$\mathbb{P}(A^{(r)}_k) = \binom{k+r-1}{r-1} \, p^r \, (1-p)^k.$$

Im Spezialfall $r = 1$ ist $\mathbb{P}(A^{(1)}_k) = p \, (1-p)^k$.

Satz. Sei $(\Omega, \mathfrak{A}, \mathbb{P})$ ein W-Raum, $A_1, ..., A_r \in \mathfrak{A}$ mit $p_i := \mathbb{P}(A_i)$ für $i = 1, ..., k$ und $p_1 + ... + p_r = 1$. Dann ist die Wahrscheinlichkeit, dass bei $n \in \mathbb{N}$ Versuchen A_1 genau n_1-mal, A_2 genau n_2-mal, ..., A_r genau n_r-mal auftritt ($n_1 + ... + n_r = n$), genau

$$B(n_1, ..., n_r, n, p_1, ..., p_r) := \binom{n}{n_1, ..., n_r} p_1^{n_1} \cdots p_r^{n_r}.$$

Diese W-Verteilung heißt **Multinomialverteilung**.

Satz. Für $0 \leq m \leq n$, $p \in [0, 1]$ gilt

$$\frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}} \xrightarrow[M/N \to p]{M, N \to \infty} \binom{n}{m} p^m \, (1-p)^{n-m}.$$

Satz (GWS von Poisson)**.** Für $k \in \mathbb{N}$, $\lambda \in \mathbb{R}_{>0}$ gilt

$$\binom{n}{m} p_n^m \, (1-p_n)^{n-m} \xrightarrow[np_n \to \lambda]{n \to \infty} \frac{\lambda^k}{k!} \exp(-\lambda).$$

Satz (von Lusin)**.** $f : ([a, b], \mathfrak{L}([a, b])) \to (\mathbb{R}^1, \mathfrak{L}(\mathbb{R}^1))$ ist Borel-messbar $\iff$ $\forall \epsilon > 0 : \exists K\epsilon \subset [a, b]$ abgeschlossen mit $\lambda_1(\mathbb{R}^1 \setminus K_\epsilon)$ und $f|_{K_\epsilon}$ stetig.

Satz. Folgerung: Es sind messbar

- monotone Funktionen
- Funktionen mit endlicher Variation
- Càdlàg-Funktionen, das sind Funktionen $f : [a, b] \to \mathbb{R}$ mit $\lim_{\epsilon \downarrow 0} f(x + \epsilon) = f(x)$ für alle $x \in [a, b[$.

Lemma (Borel-Cantelli)**.** Sei $(A_n)_{n \in \mathbb{N}}$ eine Folge von Ereignissen über $(\Omega, \mathfrak{A}, \mathbb{P})$. Dann gilt für $A = \limsup_{n \to \infty}$

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty \implies \mathbb{P}(A) = 0.$$

Falls die Ereignisse $(A_n)_{n \in \mathbb{N}}$ unabhängig sind, so gilt

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty \implies \mathbb{P}(A) = 1.$$

Definition. Sei $(\mathfrak{A}_n)_{n \in \mathbb{N}}$ Folge von σ-Algebren über Ω. Dann ist

$$\mathcal{T}_\infty = \cap_{n=1}^{\infty} \mathcal{T}_n \quad \text{mit} \quad \mathcal{T}_n := \sigma\left(\cup_{k=n}^{\infty} \mathfrak{A}_k\right)$$

die **terminale σ-Algebra** von $(\mathfrak{A}_n)_{n \in \mathbb{N}}$.

Satz (Null-Eins-Gesetz von Kolmogorow)**.** Sei $(\mathfrak{A}_n)_{n \in \mathbb{N}}$ eine Folge von unabhängigen Unter-σ-Algebren in einem W-Raum $(\Omega, \mathfrak{A}, \mathbb{P})$. Dann gilt $\mathbb{P}(A) \in \{0, 1\}$ für alle Ereignisse $A \in \mathcal{T}_\infty$ der terminalen σ-Algebra.

Definition. Eine $\mathfrak{A}$-messbare numerische Funktion X über einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, \mathbb{P})$ heißt **Zufallsgröße** (ZG) oder **Zufallsvariable**.

Definition. Die durch die ZG X auf $(\mathbb{R}^1, \mathfrak{L}(\mathbb{R}^1))$ induzierte Bildmaß P_X

$$P_X(B) = \mathbb{P}(X^{-1}(B)) = \mathbb{P}(\{\omega \in \Omega \,|\, X(\omega) \in B\})$$

heißt **Verteilung** der ZG X.

$$F_X(x) = P_X(]-\infty, x]) = \mathbb{P}(\{\omega \in \Omega \,|\, X(\omega) \leq x\})$$

heißt die **Verteilungsfunktion** (VF) der ZG X.

Satz. F sei eine VF auf $\mathbb{R}^1$. Dann existiert ein Wahrscheinlichkeits-Raum $(\Omega, \mathfrak{A}, \mathbb{P}$ und eine ZG X derart, dass

$$F_X(x) = F(x) \text{ für } x \in \mathbb{R}^1$$

Notation. Sei X eine Zufallsgröße und $B \in \mathfrak{L}(\overline{\mathbb{R}}^1)$. Dann schreibe

$$\{X \in B\} = X^{-1}(B).$$

Definition. Eine endliche Familie von Zufallsgrößen $X_1, ..., X_n$ heißt **stochastisch unabhängig**, falls

$$\mathbb{P}\left(\bigcap_{i=1}^{n} \{X_i \in B_i\}\right) = \prod_{i=1}^{n} \mathbb{P}(\{X_i \in B_i\}) \text{ für alle } B_i \in \mathcal{L}(\overline{R}^1), i = 1, ..., n.$$

Satz. Seien $X_1, ..., X_n$ unabhängige Zufallsgrößen über $(\Omega, \mathfrak{A}, \mathbb{P})$ von $g_1, ..., g_n$ Borel-messbare Funktionen von $\mathbb{R}^1$ nach $\mathbb{R}^1$. Dann sind auch die Zufallsgrößen $Y_i := g_i \circ X_i$ unabhängig über $(\Omega, \mathfrak{A}, \mathbb{P})$.

Satz. Sei $0 \leq f_1 \leq f_2 \leq ...$ eine isotone Folge elementarer Funktionen über $(\Omega, \mathfrak{A})$. Dann gilt für jede elementare Funktion f mit $f \leq \sup_{n \in N} f_n$ die Ungleichung $\int_\Omega f \, d\mu \leq \sup_{n \in N} \int_\Omega f_n \, d\mu$.

Satz. Seien $(f_n)_{n \in \mathbb{N}}$ und $(g_n)_{n \in \mathbb{N}}$ isotone Folgen elementarer Funktionen mit $\sup_{n \in \mathbb{N}} f_n = \sup_{n \in \mathbb{N}} g_n$. Dann ist $\sup_{n \in \mathbb{N}} \int_\Omega f_n \, d\mu = \sup_{n \in \mathbb{N}} \int_\Omega g_n \, d\mu$.

Satz. Sei $f : (\Omega, \mathfrak{A}, \mu) \to (\overline{R}^1, \mathfrak{L}(\mathbb{R}^1))$ sein $\mathfrak{A}$-messbar, numerisch. Dann sind äquivalent:

- f ist μ-integrierbar
- f^+ und f^- sind μ-integrierbar mit $\int_\Omega f^\pm \,\mathrm{d}\mu < \infty$
- $\int_\Omega |f| \,\mathrm{d}\mu < \infty$
- $\int_\Omega g \,\mathrm{d}\mu < \infty$ für eine $\mathfrak{A}$-messbare, numerische Funktion mit $|f| \le g$

Satz. Seien $f, g : (\Omega, \mathfrak{A}, \mu) \to (\mathbb{R}^1, \mathfrak{L}(\mathbb{R}^1))$ μ-integrierbar. Dann sind $f \pm g$, $f \vee g$, $f \wedge g$ und $\alpha \cdot f$ für $\alpha \in \mathbb{R}^1$ μ-integrierbar und es gilt

$$\int_\Omega \alpha \cdot f + \beta \cdot g \,\mathrm{d}\mu = \alpha \int_\Omega f \,\mathrm{d}\mu + \beta \int_\Omega g \,\mathrm{d}\mu, \qquad \left| \int_\Omega f \,\mathrm{d}\mu \right| \le \int_\Omega |f| \,\mathrm{d}\mu,$$

$$f \le g \implies \int_\Omega f \,\mathrm{d}\mu \le \int_\Omega g \,\mathrm{d}\mu$$

Definition. Mit $L^p(\mu) = L^p(\Omega, \mathfrak{A}, \mu)$ bezeichnen wir den normierten Vektorraum der aus den Funktionen $f : (\Omega, \mathfrak{A}, \mu) \to (\mathbb{R}^1, \mathfrak{L}(\mathbb{R}^1))$ mit $\int_\Omega |f|^p \,\mathrm{d}\mu < \infty$ für $1 \le p \le \infty$ besteht. Die Norm in diesem Raum wird durch

$$\|f\|_p := \left(\int_\Omega |f|^p \,\mathrm{d}\mu \right)^{1/p}$$

definiert. Es kann gezeigt werden, dass die Normeigenschaften erfüllt sind.

Bemerkung. Der $L^p(\mu)$ ist ein vollständiger normierter Raum, d. h. jede Cauchy-Folge bzgl. der Norm $\|\cdot\|_p$ ist auch konvergent. Im Spezialfall $p = 2$ heißt $L^p(\mu)$ Hilbertraum der quadratisch integrierbaren Funktionen mit Skalarprodukt $\langle f, g \rangle = \int_\Omega f \cdot g \,\mathrm{d}\mu$. Es gilt in diesem Fall außerdem die Cauchy-Schwarz-Bunjakowski-Ungleichung:

$$\|f \cdot g\|_1 \le \|f\|_2 \cdot \|g\|_2$$

Höldersche Ungleichung:

$$\|f \cdot g\|_1 \le \|f\|_p \cdot \|g\|_q$$

wobei $p, q \in [1, \infty]$ mit $\frac{1}{p} + \frac{1}{q} = 1$.

Satz. Sei $f_n : (\Omega, \mathfrak{A}, \mu) \to (\mathbb{R}^1, \mathfrak{L}(\mathbb{R}^1))$ $\mathfrak{A}$-messbar und $0 \le f_1 \le f_2 \le ....$ Dann gilt

$$\int_\Omega \sup_{n \in \mathbb{N}} f_n \,\mathrm{d}\mu = \sup_{n \in \mathbb{N}} \int_\Omega f_n \,\mathrm{d}\mu$$

Satz (von Beppo Levi)**.** Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge monotoner nichtnegativer, $\mathfrak{A}$-messbarer, numerischer Funktionen auf $(\Omega, \mathfrak{A}, \mu)$. Dann gilt:

$$\int_\Omega \sum_{n=1}^\infty f_n \,\mathrm{d}\mu = \sum_{n=1}^\infty \int_\Omega f_n \,\mathrm{d}\mu$$

Satz. f sei $\mathfrak{A}$-messbar, nichtnegativ und μ-integrierbar. Dann ist

$$\nu(A) := \int_A f \,\mathrm{d}\mu = \int_\Omega f \cdot \chi_A \,\mathrm{d}\mu$$

ein endliches Maß auf $(\Omega, \mathfrak{A})$.

Satz (Lemma von Fatou)**.** Sei $f_n : (\Omega, \mathfrak{A}, \mu) \to (\mathbb{R}^1, \mathfrak{L}(\mathbb{R}^1))$ eine Folge $\mathfrak{A}$-messbarer, nichtnegativer Funktionen. Dann gilt:

$$\int_\Omega \liminf_{n \to \infty} f_n \,\mathrm{d}\mu \le \liminf_{n \to \infty} \int_\Omega f_n \,\mathrm{d}\mu$$

Satz. Seien $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum, $(\Omega', \mathfrak{A}')$ ein messbarer Raum und $f : \Omega' \to \Omega$ messbar. Bezeichne mit $\mu' := \mu \circ f^{-1}$ das Bildmaß von μ unter f. Dann gilt für alle μ'-integrierbaren Funktionen $g : \Omega' \to \mathbb{R}$:

$$\int_{\Omega'} g \,\mathrm{d}\mu' = \int_\Omega g \circ f \,\mathrm{d}\mu$$

Satz (Transformationssatz)**.** Sei $U, \widetilde{U} \Subset \mathbb{R}^d$ und sei $\phi : U \to \widetilde{U}$ ein $\mathcal{C}^1$-Diffeomorphismus. Dann ist eine Funktion $f : \widetilde{U} \to \overline{\mathbb{R}}$ genau dann auf $\widetilde{U}$ Lebesgue-Borel-integrierbar, wenn $(f \circ \phi) \cdot |\det(D\phi)| : U \to \overline{\mathbb{R}}$ auf U Lebesgue-Borel-interierbar ist. In diesem Fall gilt

$$\int_U (f \circ \phi) \cdot |\det(D\phi)| \,\mathrm{d}\mu_{LB} = \int_{\phi(U)} f \,\mathrm{d}\mu_{LB} = \int_{\widetilde{U}} f \,\mathrm{d}\mu_{LB}.$$

Obige Gleichung ist auch erfüllt, wenn lediglich $f \ge 0$ gilt (also $f \in \overline{\mathbb{E}}(\widetilde{U}, \mathfrak{B}(\widetilde{U}))$; dann kann das Integral auch den Wert ∞ annehmen).

Definition. Für eine ZG $X : (\Omega, \mathfrak{A}, \mathbb{P}) \to (\overline{\mathbb{R}}^1, \mathcal{L}(\overline{\mathbb{R}}^1))$ heißt die Zahl

$$\mathbb{E}X := \int_\Omega X \,\mathrm{d}\mathbb{P} = \int_{\mathbb{R}^1} \mathrm{id} \,\mathrm{d}P_X$$

der **Erwartungswert** der ZG X, wobei $P_X = \mathbb{P} \circ X^{-1}$.

Korollar. Sei $g : \mathbb{R}^1 \to \mathbb{R}^1$ Borel-messbar und P_X-integrierbar. Dann gilt

$$\mathbb{E}g(X) = \int_{\mathbb{R}^1} g \,\mathrm{d}P_X = \int_{-\infty}^{\infty} g(x) \,\mathrm{d}F_X(x),$$

wobei das rechte Integral das uneigentliche Riemann-Stieltjes-Integral bzgl. F_X ist.

Definition. Für Zufallsvektoren $X = (X_1, ..., X_k)$ mit Werten in $\mathbb{R}^k$ ist

$$\mathbb{E}X = (\mathbb{E}X_1, ..., \mathbb{E}X_k)$$

Sei $g : \mathbb{R}^k \to \mathbb{R}$ Borel-messbar und $P_{(X_1,...,X_k)}$-integrierbar. Dann ist

$$\mathbb{E}g(X_1, ..., X_k) = \int_{\mathbb{R}^k} g(x_1, ...) \,\mathrm{d}P_{(X_1,...,X_k)} = \int_{-\infty}^{\infty} g(x_1, ..., x_k) \,\mathrm{d}F_X(x_1, ..., x_k)$$

$F = F_X$ sei die VF einer ZG $X : (\Omega, \mathfrak{A}, \mathbb{P}) \to (\mathbb{R}^1, \mathcal{L}(\mathbb{R}^1), P_X)$

Definition. • F_X heißt diskret, falls F_X höchstens abzählbar viele Sprünge $x_1, x_2, ... \in \mathbb{R}$ mit $p_k := F(x_k) - \lim_{x \uparrow x_k} F(x) > 0$ mit $\sum_{k=1}^\infty p_k = 1$ besitzt (dann ist F_X) zwischen den Sprüngen konstant) item

- F_X heißt **stetig** (diffus, atomlos), wenn F_X in jedem Punkt stetig ist. Dann gilt $P_X(\{X = x\}) = 0$.
- F_X heißt **absolut stetig** (totalstetig), wenn es für alle $\epsilon > 0$ ein $\delta > 0$ gibt, sodass für abzählbare viele, disjunkte Intervalle $I_k =]a_k, b_k]$ mit $\sum_k (b_k - a_k) < \delta$ sich $\sum_k (F_X(b_k) - F_X(a_k)) \le \epsilon$ ergibt.
- **singulärstetig** (stetig, aber nicht absolutstetig), wenn die Wachstumspunkte VF F_X eine Lebesgue-Nullmenge bilden, also

$$\lambda_1(\{x \in \mathbb{R}^2 \,|\, \forall \epsilon > 0 : F(x + \epsilon) - F(x - \epsilon) > 0\}) = 0$$

Satz. $F_X'(x)$ existiert für Lebesgue-fast-alle $x \in \mathbb{R}^1$.

Satz. Jede VF F auf $\mathbb{R}^1$ besitzt eine eindeutige Zerlegung (Lebesgue-Zerlegung) als konvexe Linearkombination einer diskreten, singulär-stetigen und absolut-stetigen VF:

$$F = \alpha_d F_d + \alpha_s F_s + \alpha_a F_a$$

mit $\alpha_d, \alpha_s, \alpha_a \ge 0$ und $\alpha_d + \alpha_s + \alpha_a = 1$.

Definition. Falls F_X absolut-stetig, dann heißt die nicht-negative, Lebesgue-messbare Funktion

$$f f_X(x) := \begin{cases} F_X'(x) & \text{falls Ableitung ex.} \\ 0 & \text{sonst,} \end{cases}$$ welche $\int_{\mathbb{R}^1} f_X \,\mathrm{d}\lambda_1 = 1$ erfüllt,

die **W-Dichte** von F_X.

$$\mathbb{E}X = \int_{\mathbb{R}^1} x \,\mathrm{d}F_X = \int_{\mathbb{R}^1} \mathrm{id} \cdot f_X \,\mathrm{d}\lambda = \int_{-\infty}^{\infty} x \cdot f_X(x) \,\mathrm{d}x$$

Bemerkung. Für diskrete F_X, also

$$F(x) = c_i \quad \text{für alle} x \in]x_i, x_{i+1}[$$

für $x_1, ... \in \mathbb{R}$ und $c_1, ... \in \mathbb{R}$ gilt

$$\mathbb{E}X = \sum_{i=1} (x_i)$$

Deutung der $\mathbb{E}X$ als Massenschwerpunkte

...

$F_(X_1, ..., X_k)$ heißt **absolut stetig**, falls für alle $\epsilon > 0$ ein $\delta > 0$ existiert, sodass für $I_\alpha =]a_j, b_j]$, $j = 1, 2, ...$ mit $\sum\limits_{j\geq 1} \lambda_k(I_j) \leq \delta$ gilt:

$$\sum_{j\geq 1} \mathbb{P}_{(X_1,...,X_k)}(I) = \sum_{j\geq 1} (triangle F_{(X_1,...,X_k)}) I_j \leq \epsilon$$

Genau dann existiert eine (Lebesgue-) Borel-messbare Funktion $f_{(X_1,...,X_k)}(x_1, ..., x_k) \geq 0$ mit $\int\limits_{\mathbb{R}^1} f_{(X_1,...,X_k)} \,\mathrm{d}\lambda_k = 1$

Sei $g : \mathbb{R}^k \to \mathbb{R}^1$ Borel-messbar

$\mathbb{E}g(X_1, ..., X_k) = \int\limits_{\mathbb{R}^1} g \cdot f_{(X_1,...,X_k)} \,\mathrm{d}\lambda_k$

Falls $F_{(X_1,...,X_k)}$ "hinreichend glatt", so ergibt sich

$f_{X_1,...,X_k}(x_1, ..., x_k) = \dfrac{\partial^k}{\partial_{x_1} \cdot \partial_{x_k}} F_{(X_1,...,X_k)}(x_1, ..., x_k)$

$F_{X_1,...,X_k}$ heißt singulär-stetig, falls $P_{(X_1,...,X_k)}(\{x\}) = 0 \forall x \in \mathbb{R}^k$ und es existiert eine Lebesgue-messbare Menge S mit $\lambda_k(S) = 0$ und $P((X_1, ..., X_k))(S) = 1$.

$F_{(X_1,...,X_k)}$ heißt diskret, falls eine höchstens abzählbare Punktmenge $S = \{x_1, ...\} \subset \mathbb{R}^k$ und $p_i = P_{(X_1,...,X_k)}(\{x_i\}) > 0$ mit $\sum\limits_{i\geq 1} p_i = 1$

Sei $x_i = (x_i^{(1)}, ..., x_i^{(k)}) \in \mathbb{R}^k$

$\mathbb{E}g(X_1, ..., X_k) = \sum\limits_{i\geq 1} g(x_i^{(1)}, ..., x_i^{(k)}) p_i$

$g : \mathbb{R}^1 \to \mathbb{R}^1$ sei zunächst beliebig (stetig oder hinreichend glatt)

Sei $a = \xi_0^{(n)} < \xi_1^{(n)} < ... < x_{k_n}^{(n)}$ und $x_k^{(n)} \in \left]\xi_{k-1}^{(n)}, \xi_k^{(n)}\right[$.

Definition. (ξ_n) sei eine Zerlegungsfolge mit $\max\limits_{1\leq k\leq k_n} (x_k^{(n)} - x_{k-1}^{(n)}) \xrightarrow{n\to\infty} 0$

$(x_k^{(n)})$ sei eine Zwischenwertfolge

$$\lim_{n\to\infty} \sum_{k=1}^{k_n} g(x_k^{(n)})(F(x_k^{(n)}) - F(x_{k-1}^{(n)})) = \int_a^b g(x) \,\mathrm{d}F(x) = \int_{[a,b]} g \,\mathrm{d}F\lambda_1$$

wobei $F : \mathbb{R}^1 \to \mathbb{R}^1$ zunächst beliebig (monoton oder von beschränkter Variation)

Sei g bzgl. F R-S-integrierbar, d. h. der Grenzwert oben existiert

Dann ist auch F bzgl. g R-S-integrierbar und es gilt

$$\int_a^b g(x) \,\mathrm{d}F(x) = [g(x) \cdot F(x)]_a^b - \int_a^b F(x) \,\mathrm{d}g(x)$$

Ausnutzen der partiellen Integration zur Berechnung von Erwartungswerten

$$\mathbb{E}X = \int_{-\infty}^{\infty} x \,\mathrm{d}F(x)$$

$\int\limits_a^b x \,\mathrm{d}F_X(x) = \lim\limits_{a\to-\infty, b\to\infty} [x \cdot F_X(-x)]_0^{-a} - \int\limits_0^{-a} F_X(-x) \,\mathrm{d}x + [x(F_X(x) - 1)]_0^b - \int\limits_0^b (F_X(x) - 1) \,\mathrm{d}x$

Falls $\lim x \to \infty x F_X(-x) = \lim x \to \infty x(1 - F_X(x)) = 0$, so gilt

$\mathbb{E}X = \int\limits_0^\infty 1 - F_X(x) - F_X(-x) \,\mathrm{d}x$, falls $\mathbb{E}|X| < \infty$

$\mathbb{E}|X| = \int\limits_0^\infty 1 - F_X(x) - F_X(-x) \,\mathrm{d}x$

Genauso werden Erwartungswerte von Funktionen von X berechnet, z.B. mit $x^2 F_X(-x) \xrightarrow{x\to\infty} 0$ und $x^2(1 - F_X(x)) \xrightarrow{x\to\infty} 0$

$\mathbb{E}X^2 = 2\int\limits_0^\infty x(1 - F_X(x) + F_X(-x)) \,\mathrm{d}x = \mathbb{P}(|X| > x)$ (falls F_X stetig)

$\mathbb{E}|X|^k = k\int\limits_0^\infty x^{k-1}(1 - F_X(x) + F_X(-x)) \,\mathrm{d}x$

Definition. $\mathbb{E}X^k$ ($\mathbb{E}|X|^k$) heißt k-tes (absolutes) Moment der ZG X. $\mathbb{E}(X - \mathbb{E}X)^k$ heißt k-tes zentriertes Moment der ZG X.

$\mathrm{Var}(X) := D^2 X = \mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{E}X^2$ heißt **Streuung** (Dispersion, Varianz) der ZG X.