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: $A \in \mathfrak{A} \implies A^c \coloneqq \Omega \backslash A \in \mathfrak{A}$

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

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

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 $\mathfrak{L}(\mathbb{R}^1)$. Dann definiert $x \mapsto F_\mu(x) \coloneqq \mu(]-\infty, x])$ eine VF. Für eine VF $F : \mathbb{R} \to [0,1]$ definiert umgekehrt $\mu_F(]a,b]) \coloneqq F(b) - F(a)$ ein W-Maß auf $\mathfrak{L}(\mathbb{R}^1)$. 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 $\lambda > 0$:

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

Elementare Wahrscheinlichkeitsrechnung

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

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

auf $(\Omega, \mathfrak{L}(\Omega))$ **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_m^n, dass sich unter n gezogenen Kugeln genau $m \leq \min(n, M)$ schwarze Kugeln befinden,

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

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 $\mathbb{P}(B) > 0$

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

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

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

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 $A_k^{(r)}$, dass beim Versuch A_{k+r} der r-te Erfolg eintritt, gleich

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

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

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

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

Definition. Sei $(\Omega, \mathfrak{A})$ ein messbarer Raum mit zwei W-Maßen $\mathbb{P}_1$ und $\mathbb{P}_2$. Dann heißt

$$d_\infty(\mathbb{P}_1, \mathbb{P}_2) := \sup_{A \in \mathfrak{A}} |\mathbb{P}_1(A) - \mathbb{P}_2(A)|$$

Totalvariation des signierten Maßes $\mathbb{P}_1 - \mathbb{P}_2$.

Satz. Seien $\mathbb{P}_1$ und $\mathbb{P}_2$ zwei W-Maße auf $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, $\mathbb{P}_1(\{i\}) = p_i$, $\mathbb{P}_2(\{i\}) = q_i$ für alle $i \in \mathbb{N}$. Dann gilt

$$d_\infty(\mathbb{P}_1, \mathbb{P}_2) = \tfrac{1}{2} \sum_{i=0}^{\infty} |p_i - q_i|.$$

Lemma. Für $n, k \in \mathbb{N}$, $p \in [0,1]$ und $\mathbb{P}_1$ und $\mathbb{P}_2$ wie eben definiert durch $p_i := \binom{n}{k} p^k (1-p)^{n-k}$, $q_i := \frac{(np)^k}{k!} \exp(-np)$ gilt

$$d_\infty(\mathbb{P}_1, \mathbb{P}_2) \leq 2np^2.$$

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

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

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

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

also zusammengefasst $\mathbb{P}(A) \in \{0, 1\}$.

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

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

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.

Integrationstheorie

Definition. Seien $(\Omega_1, \mathfrak{A}_1)$ und $(\Omega_2, \mathfrak{A}_2)$ messbare Räume. Dann heißt $f : \Omega_1 \to \Omega_2$ **$(\mathfrak{A}_1, \mathfrak{A}_2)$-messbar**, falls

$$f^{-1}(A_2) \in \mathfrak{A}_1 \quad \text{für alle } A_2 \in \mathfrak{A}_2.$$

Notation. Für solches f schreiben wir $f : (\Omega_1, \mathfrak{A}_1) \to (\Omega_2, \mathfrak{A}_2)$.

Notation. Sei Ω eine Menge, $A \subset \Omega$. Dann heißt

$$\mathbb{1}_A : \Omega \to \mathbb{R}, \quad \omega \mapsto |\{\star \mid \omega \in A\}| = \begin{cases} 1, & \text{falls } \omega \in A \\ 0, & \text{falls } \omega \notin A \end{cases}$$

Indikatorfunktion von A.

Beobachtung. Sei $(\Omega, \mathfrak{A})$ messbarer Raum, $A \subset \Omega$, dann gilt

$$\mathbb{1}_A \ (\mathfrak{A}, \mathfrak{L}(\mathbb{R}^1))\text{-messbar} \iff A \in \mathfrak{A}.$$

Lemma. Die Verkettung messbarer Abbildungen ist messbar, d. h. für $f : (\Omega_1, \mathfrak{A}_1) \to (\Omega_2, \mathfrak{A}_2)$ und $g : (\Omega_2, \mathfrak{A}_2) \to (\Omega_3, \mathfrak{A}_3)$ gilt $g \circ f : (\Omega_1, \mathfrak{A}_1) \to (\Omega_3, \mathfrak{A}_3)$.

Lemma. Sei $f : \Omega \to \Omega'$ eine Abb. und $\mathcal{E}' \subset \mathcal{P}(\Omega)$, dann ist

$$\mathfrak{A}(f^{-1}(\mathcal{E}')) = f^{-1}(\mathfrak{A}(\mathcal{E}')).$$

Lemma. Sei $(\Omega, \mathfrak{A})$ ein messbarer Raum und $f : \Omega \to \Omega'$ eine Abbildung, sowie $\mathcal{E} \subset \mathbb{P}(\Omega')$. Dann gilt

$$f \text{ ist } (\mathfrak{A}, \sigma(\mathcal{E}))\text{-messbar} \iff f^{-1}(E) \in \mathfrak{A} \text{ für alle } E \in \mathcal{E}.$$

Notation. Seien $f, g : \Omega \to \overline{\mathbb{R}}$ zwei numerische Funktionen. Setze

$$\{f \leq g\} := \{\omega \in \Omega \mid f(\omega) \leq g(\omega)\} \subset \Omega$$

und definiere analog $\{f < g\}$, $\{f \geq g\}$, $\{f > g\}$, $\{f = g\}$, $\{f \neq g\}$.

Satz. Für eine numerische Fkt. $f : (\Omega, \mathfrak{A}) \to (\overline{\mathbb{R}}, \overline{\mathfrak{B}})$ sind äquivalent:
- f ist messbar
- $\forall\, a \in \mathbb{R} : \{f \geq a\} = f^{-1}([a, \infty]) \in \mathfrak{A}$
- $\forall\, a \in \mathbb{R} : \{f > a\} \in \mathfrak{A}$
- $\forall\, a \in \mathbb{R} : \{f \leq a\} \in \mathfrak{A}$
- $\forall\, a \in \mathbb{R} : \{f < a\} \in \mathfrak{A}$

Definition.
- Sei $(\Omega, \mathfrak{A})$ ein messbarer Raum, $f : \Omega' \to \Omega$ eine Abbildung, dann heißt

  $$\sigma(f) := f^{-1}(\mathfrak{A}) := \{f^{-1}(A) \mid A \in \mathfrak{A}\}$$

  die **von f erzeugte σ-Algebra**.
- Sei $(\Omega_i, \mathfrak{A}_i)_{i \in I}$ eine Familie von messbaren Räumen, $f_i : \Omega' \to \Omega_i$ für alle $i \in I$ eine Abbildung. Dann heißt

  $$\sigma((f_i)_{i \in I}) := \sigma(\bigcup_{i \in I} \sigma(f_i)) = \sigma(\bigcup_{i \in I} f_i^{-1}(\mathfrak{A}_i))$$

  die **von der Familie $(f_i)_{i \in I}$ erzeugte σ-Algebra**.

Definition. Sei $(\Omega', \mathfrak{A}', \mu')$ ein Maßraum, $(\Omega, \mathfrak{A})$ ein messbarer Raum, $f : (\Omega', \mathfrak{A}) \to (\Omega, \mathfrak{A})$. Dann ist durch

$$\mu'_f := \mu' \circ f^{-1} : \mathfrak{A} \to [0, \infty], \quad A \mapsto \mu'(f^{-1}(A))$$

ein Maß auf $(\Omega, \mathfrak{A})$, das sog. **Bildmaß** von μ' unter f, definiert.

Satz. Für zwei numerische Funktionen $f, g : (\Omega, \mathfrak{A}) \to (\overline{\mathbb{R}}, \overline{\mathfrak{B}})$ gilt:
- $\{f < g\} \in \mathfrak{A}$
- $\{f > g\} \in \mathfrak{A}$
- $\{f = g\} \in \mathfrak{A}$
- $\{f \leq g\} \in \mathfrak{A}$
- $\{f \geq g\} \in \mathfrak{A}$
- $\{f \neq g\} \in \mathfrak{A}$

Satz. Seien $f, g : (\Omega, \mathfrak{A}) \to (\overline{\mathbb{R}}, \overline{\mathfrak{B}})$ messbare numerische Funktionen und $\lambda, \mu \in \mathbb{R}$. Dann auch messbar (‡: falls $0 \notin \text{Bild}(f)$):
- $\lambda \cdot f$
- $f + \mu \cdot g$
- $f \cdot g$
- $\frac{1}{f}$ (‡)
- $\frac{g}{f}$ (‡)

Satz. Seien $f_n : (\Omega, \mathfrak{A}) \to (\overline{\mathbb{R}}, \overline{\mathfrak{B}}), n \in \mathbb{N}$ messbare numerische Funktionen, dann auch messbar:
- $\sup_{n \in \mathbb{N}} f_n$
- $\inf_{n \in \mathbb{N}} f_n$
- $\liminf_{n \in \mathbb{N}} f_n$
- $\limsup_{n \in \mathbb{N}} f_n$

Dabei werden Infimum, Supremum, usw. punktweise gebildet.

Definition. Für $f : \Omega \to \overline{\mathbb{R}}$ heißen die Funktionen
- $|f| := \max(f, -f) : \Omega \to [0, \infty]$ **Betrag** von f
- $f^+ := \max(f, 0) : \Omega \to [0, \infty]$ **Positivteil** von f
- $f^- := -\min(f, 0) : \Omega \to [0, \infty]$ **Negativteil** von f

Satz. Falls $f : (\Omega, \mathfrak{A}) \to (\overline{\mathbb{R}}, \overline{\mathfrak{B}})$ messbar, dann auch $|f|$, f^+ und f^-.

Satz.
- Sei $(\Omega, \mathcal{O})$ ein topologischer Raum und $f : \Omega \to \mathbb{R}^n$ stetig. Dann ist f $(\sigma(\mathcal{O}), \mathfrak{B}(\mathbb{R}^n))$-messbar.
- $\sigma(\mathcal{O})$ ist die kleinste σ-Algebra, bezüglich der alle stetigen Funktionen $f : \Omega \to \mathbb{R}^n$ Borel-messbar sind.

Satz (von Lusin)**.** Sei $M \in \mathfrak{L}(\mathbb{R}^n)$ mit $\lambda_n(M) < \infty$ und $f : M \to \mathbb{R}$ beschränkt. Dann ist f genau dann Borel-messbar, wenn gilt:

$$\forall\, \epsilon > 0 : \exists\, K_\epsilon \subset M \text{ kompakt} : \lambda_n(M \setminus K_\epsilon) < \epsilon \text{ und } f|_{K_\epsilon} \text{ stetig.}$$

Definition. Eine Funktion $f : \mathbb{R} \to \mathbb{R}$ heißt **Càdlàg-Funktion** (continue à droite, limite à gauche), falls für alle $x \in \mathbb{R}$ gilt:

$$\lim_{y \uparrow x} f(y) \text{ existiert} \quad \text{und} \quad \lim_{y \downarrow x} f(y) = f(x).$$

Beobachtung. Jede kumulierte VF ist eine Càdlàg-Funktion.

Definition. Die **Variation** von $g : [a, b] \to \mathbb{R}$ bzgl. einer Zerlegung $Z = \{a = x_0 < \ldots < x_n = b\}$ von $[a, b]$ ist die nicht-negative Zahl

$$V(g, Z) := \sum_{j=1}^{n} |g(x_j) - g(x_{j-1})|.$$

Die **Totalvariation** von $g : [a, b] \to \mathbb{R}$ ist

$$V_a^b(g) := \sup\{V(g, Z) : Z \text{ Zerlegung von } [a, b]\} \in \mathbb{R}_{\geq 0} \cup \{\infty\}.$$

Falls $V_a^b(g) < \infty$, so heißt g **von beschränkter Variation**.

Satz. Es sind messbar:
- Monotone Funktionen
- Càdlàg-Funktionen
- Funktionen von beschränkter Variation

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

Bemerkung. Häufig fordert man zusätzlich $\mathbb{P}(\{X = \pm\infty\}) = 0$.

Definition. Das durch die ZG X induzierte Bildmaß

$$P_X : \mathfrak{L}(\mathbb{R}^1) \to [0,1], \quad B \mapsto \mathbb{P}(\{X \in B\}) = \mathbb{P}(X^{-1}(B))$$

heißt **Verteilungsgesetz** der ZG X und

$$F_X : \mathbb{R} \to \mathbb{R}, \quad x \mapsto P_X(]-\infty, x]) = \mathbb{P}(\{X \leq x\})$$

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

Satz. Sei F eine VF auf $\mathbb{R}^1$. Dann existiert ein W-Raum $(\Omega, \mathfrak{A}, \mathbb{P})$ und eine ZG X auf Ω derart, dass $F_X = F$.

Beweis. 1. Möglichkeit: Wähle $\Omega := \mathbb{R}^1$, $\mathfrak{A} := \mathfrak{L}(\mathbb{R}^1)$ und $\mathbb{P} := \mu_F$ als das von von F erzeugte Maß und setze $X := \mathrm{id}$.

2. Möglichkeit: Wähle $\Omega := [0,1]$, $\mathfrak{A} := \mathcal{L}([0,1])$, $\mathbb{P} := \lambda_1$. Setze

$$X(w) := F^-(w) := \inf\{F \geq w\} \quad \text{für } w > 0, \quad X(0) := \lim_{w \downarrow 0} F^-(w)$$

Definition. Sei $X_1, ..., X_n$ eine endliche Familie von ZGen über $(\Omega, \mathfrak{A}, \mathbb{P})$. Diese Familie heißt **stochastisch unabhängig**, falls

$$\mathbb{P}(\bigcap_{i=1}^n \{X_i \in B_i\}) = \prod_{i=1}^n \mathbb{P}(\{X_i \in B_i\}) \quad \text{für alle } B_1, ..., B_n \in \mathfrak{L}(\overline{\mathbb{R}}^1).$$

Satz. Seien $X_1, ..., X_n$ unabhängige ZGen über $(\Omega, \mathfrak{A}, \mathbb{P})$ und $g_1, ..., g_n : \mathbb{R} \to \mathbb{R}$ Borel-messbar. Setze $Y_i := g_i(X_i) := g_i \circ X_i$ für $i = 1, ..., n$, dann sind auch $Y_1, ..., Y_n$ unabhängige ZGen.

Definition. Eine Funktion $f : (\Omega, \mathfrak{A}) \to (\mathbb{R}, \mathfrak{B})$ heißt **einfache Funktion** oder **Elementarfunktion** auf $(\Omega, \mathfrak{A})$, wenn gilt:

- f ist messbar
- $f(\Omega) \subset [0, \infty[$
- $f(\Omega)$ ist endlich

Die Menge aller elementaren Funktionen auf $(\Omega, \mathfrak{A})$ ist $\mathbb{E}(\Omega, \mathfrak{A})$.

Notation. $a \wedge b := \min\{a, b\}$ und $a \vee b := \max\{a, b\}$

Satz. Seien $f, g \in \mathbb{E}(\Omega, \mathfrak{A})$ und $a \geq 0$. Dann auch in $\mathbb{E}(\Omega, \mathfrak{A})$:

- $f + g$
- $f \cdot g$
- $f \vee g$
- $f \wedge g$
- $a \cdot f$

Definition. Sei $f \in \mathbb{E}(\Omega, \mathfrak{A})$ und $\Omega = A_1 \sqcup ... \sqcup A_k$ eine disjunkte Vereinigung von Mengen mit $A_j \in \mathfrak{A}$ für alle $j = 1, ..., k$, sodass $f(A_j) = \{y_j\}$, dann heißt die Darstellung

$$f = \sum_{j=1}^k y_j \cdot \mathbb{1}_{A_j} \quad \textbf{kanonische Darstellung}.$$

Definition. Sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum, $f : \Omega \to \mathbb{R}$ elementar. Dann heißt die (von der Darstellung $f = \sum_{j=1}^k y_j \cdot \mathbb{1}_{A_j}$ unabh.) Zahl

$$\int_\Omega f \, d\mu := \sum_{j=1}^k y_j \mu(A_j) \quad \mu\textbf{-Integral} \text{ von } f.$$

Satz. Es gilt für $f, g \in \mathbb{E}(\Omega, \mathfrak{A})$, $a, b \geq 0$:

- $\int_\Omega \mathbb{1}_A \, d\mu = \mu(A)$
- $f \leq g \implies \int_\Omega f \, d\mu \leq \int_\Omega g \, d\mu$
- $\int_\Omega a \cdot f + b \cdot g \, d\mu = a \cdot \int_\Omega f \, d\mu + b \cdot \int_\Omega g \, d\mu$

Satz. Sei $(f_n)_{n \in \mathbb{N}}$ eine isotone (= monotone) 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 $\int_\Omega f \, d\mu \leq \sup_{n \in \mathbb{N}} \int_\Omega f_n \, d\mu$.

Korollar. 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 $f : (\Omega, \mathfrak{A}, \mu) \to (\overline{\mathbb{R}}, \mathfrak{L}(\overline{\mathbb{R}}))$ nichtnegativ. Dann gibt es eine isotone Folge $(f_n)_{n \in \mathbb{N}}$ elementarer Funktionen mit $\sup_{n \in \mathbb{N}} f_n = f$.

Definition. Sei $f : (\Omega, \mathfrak{A}, \mu) \to (\overline{\mathbb{R}}, \mathfrak{L}(\overline{\mathbb{R}}))$ nichtnegativ und $(f_n)_{n \in \mathbb{N}}$ eine Folge elementarer Funktionen mit $\sup_{n \in \mathbb{N}} f_n = f$. Dann heißt

$$\int_\Omega f \, d\mu := \sup_{n \in \mathbb{N}} \int_\Omega f_n \, d\mu \quad \mu\textbf{-Integral} \text{ von } f.$$

Definition. Eine $\mathfrak{A}$-messbare, numerische Fkt. $f : (\Omega, \mathfrak{A}, \mu) \to (\overline{\mathbb{R}}, \mathfrak{L}(\overline{\mathbb{R}}))$ heißt **μ-integrierbar**, falls

$$\int_\Omega f^+ \, d\mu < \infty \quad \text{und} \quad \int_\Omega f^- \, d\mu < \infty.$$

In diesem Fall definieren wir das **Lebesgue-Integral** von f als

$$\int_\Omega f \, d\mu := \int_\Omega f^+ \, d\mu - \int_\Omega f^- \, d\mu.$$

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

- f ist μ-integrierbar
- f^+ und f^- sind μ-integrierbar
- $|f|$ ist μ-integrierbar
- $\exists$ μ-integrierbare Funktion g mit $|f| \leq g$

Satz. Seien $f, g : (\Omega, \mathfrak{A}, \mu) \to (\overline{\mathbb{R}}^1, \mathfrak{L}(\overline{\mathbb{R}}^1))$ μ-integrierbar und $\alpha, \beta \in \mathbb{R}$. Dann sind auch μ-integrierbar:

- $f \pm g$
- $f \vee g$
- $f \wedge g$
- $\alpha \cdot f$

Es gilt:

- $\int_\Omega (\alpha \cdot f + \beta \cdot g) \, d\mu = \alpha \int_\Omega f \, d\mu + \beta \int_\Omega g \, d\mu$ (Linearität)
- $|\int_\Omega f \, d\mu| \leq \int_\Omega |f| \, d\mu$
- $f \leq g \implies \int_\Omega f \, d\mu \leq \int_\Omega g \, d\mu$ (Monotonie)

Achtung. Das Produkt $(f \cdot g)$ ist i. A. nicht μ-integrierbar!

Definition. Sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum und $f : (\Omega, \mathfrak{A}) \to (\overline{\mathbb{R}}, \mathcal{L}(\overline{\mathbb{R}}))$. Für $p \in [1, \infty[$ heißt f **p-integrierbar**, falls $|f|^p$ μ-integrierbar ist.

$$L^p(\Omega, \mathfrak{A}, \mu) := \{f : \Omega \to \overline{\mathbb{R}} \mid f \text{ } p\text{-integrierbar, also } \int_\Omega |f|^p \, d\mu < \infty\},$$

$$L^\infty(\Omega, \mathfrak{A}, \mu) := \{f : \Omega \to \overline{\mathbb{R}} \mid \exists C > 0 : |f| \leq C \text{ fast überall}\}$$

ist dann ein VR, genannt **Lebesgue-Raum** (L^p-Raum), mit Norm

$$\|f\|_p := (\int_\Omega |f|^p \, d\mu)^{\frac{1}{p}}$$

$$\|f\|_\infty := \operatorname*{ess\,sup}_{\omega \in \Omega} |f(\omega)| := \inf\{C \in \mathbb{R} \mid |f| \leq C \text{ fast-überall}\}$$

Wir betrachten in L^p zwei Funktionen als gleich, wenn sie bis auf einer Nullmenge übereinstimmen. Die $\triangle$-Ungleichung in L^p wird auch **Minkowski-Ungleichung** genannt.

Satz. Der $L^p(\mu)$ ist ein vollständiger normierter Raum, d. h. jede Cauchy-Folge bzgl. der Norm $\|\cdot\|_p$ ist auch konvergent.

Satz. Sei $f \in L^p(\Omega, \mathfrak{A}, \mu)$, $g \in Ł^q(\Omega, \mathfrak{A}, \mu)$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann ist $fg \in Ł^1(\Omega, \mathfrak{A}, \mu)$ und es gilt

$$\|fg\|_1 \leq \|f\|_p \cdot \|g\|_q \quad \textbf{(Hölder-Ungleichung).}$$

Bemerkung. Für $p = 2$ ist $L^p(\Omega, \mathfrak{A}, \mu)$ der Hilbertraum der quadratisch integrierbaren Funktionen mit dem Skalarprodukt

$$\langle f, g \rangle := \int_\Omega (f \cdot g) \, d\mu.$$

Mit $q = 2$ folgt aus der Hölder-Ungleichung

$$|\langle f, g \rangle| = \|fg\|_1 \leq \|f\|_2 \cdot \|g\|_2 \quad \textbf{(Cauchy-Schwarz-Ungl.)}$$

Satz. Sei $f : (\Omega, \mathfrak{A}, \mu) \to (\overline{\mathbb{R}}^1, \mathcal{L}(\overline{\mathbb{R}}^1))$ nichtnegativ. Dann gilt

$$\int_\Omega f \, d\mu = 0 \quad \iff \quad f \overset{\text{f.ü.}}{=} 0.$$

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

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

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:

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

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

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

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:

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

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 $g : \Omega' \to \mathbb{R}$:

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

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

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

Obige Gleichung ist auch erfüllt, wenn lediglich $f \geq 0$ 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 $X : (\Omega, \mathfrak{A}, \mathbb{P}) \to (\overline{\mathbb{R}}^1, \mathfrak{L}(\overline{\mathbb{R}}^1))$ 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 $X : (\Omega, \mathfrak{A}, \mathbb{P}) \to (\mathbb{R}^1, \mathfrak{L}(\mathbb{R}^1), P_X)$

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 $\sum_k (F_X(b_k) - F_X(a_k)) \leq \epsilon$ 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 $\alpha_d, \alpha_s, \alpha_a \geq 0$ und $\alpha_d + \alpha_s + \alpha_a = 1$.

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

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

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 $\sum_{j \geq 1} \lambda_k(I_j) \leq \delta$ 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 $f_{(X_1,...,Xk)}(x_1, ..., x_k) \geq 0$ mit $\int_{\mathbb{R}^1} f_{(X_1,...,X_k)} \, \mathrm{d}\lambda_k = 1$

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

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

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 $\sum_{i \geq 1} p_i = 1$

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

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

$g : \mathbb{R}^1 \to \mathbb{R}^1$ 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 $\max_{1 \leq k \leq k_n} (x_k^{(n)} - x_{k-1}^{(n)}) \xrightarrow{n \to \infty} 0$

$(x_k^{(n)})$ 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

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

$\int_a^b x \, \mathrm{d}F_X(x) = \lim_{a \to -\infty, b \to \infty} [x \cdot F_X(-x)]_0^{-a} - \int_0^{-a} F_X(-x) \, \mathrm{d}x + [x(F_X(x) - 1)]_0^b - \int_0^b ($

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

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

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

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

$\mathbb{E}X^2 = 2 \int_0^\infty x(1 - F_X(x) + F_X(-x)) \, \mathrm{d}x = \mathbb{P}(|X| > x)$ (falls F_X stetig)

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

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.

X sei ZG über $(\Omega, \mathfrak{A}, \mathbb{P})$ mit $\mathbb{E}|X|^n < \infty$ für alle $n \in \mathbb{N}$.

$$\mathbb{E}X^n = \int_{\mathbb{R}^n} x^n \mathbb{P}_X \, \mathrm{d}x = \int_{-\infty}^{\infty} x^n \, \mathrm{d}F_X(x)$$

heißt **n-tes Moment**
$\mathbb{E}X$ Erwartungswert (Schwerpunkt der W-Massenverteilung) „Lageparameter“
$\mathbb{D}^2 X = \mathrm{Var}(X) = \mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{E}(X^2) - (\mathbb{E}X)^2 \geq 0$
Eigenschaft: $\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$ für $a, b \in \mathbb{R}^1$
$\mathrm{Var}(X) \leq \mathbb{E}(X - c)^2$ für alle $c \in \mathbb{R}^1$
$\mathrm{Var}(X) = 0 \iff X - \mathbb{E}X = 0 \iff X = \mathrm{const}$ $\mathbb{P}$-fast-sicher

Satz. X sei eine ZG und $g : [0, \infty[\to [0, \infty]$ nichtfallend. Dann gilt $\mathbb{P}(|X| \geq \epsilon) \leq \dfrac{\mathbb{E}g(|X|)}{g(\epsilon)}$ für beliebiges $\epsilon > 0$.

Spezialfälle:

- $g(x) = x$, dann $\mathbb{P}(|X| \geq \epsilon) \leq \dfrac{\mathbb{E}X}{\epsilon}$ **Markow-Ungleichung**
- $g(x) = x^2$, dann $\mathbb{P}(|X - \mathbb{E}X| \geq \epsilon) \leq \dfrac{\mathbb{E}(X - \mathbb{E}X)^2}{\epsilon^2} = \dfrac{\mathrm{Var}(X)}{\epsilon^2}$ **Tschebyschew-Ungleichung**
- $g(x) = \exp(ax)$ für $a > 0$, dann $\mathbb{P}(|X| \geq \epsilon) \leq \exp(-a\epsilon) \mathbb{E} \exp(a|X|)$

$0 > B = \mathbb{E}\exp(a|X|) \geq \frac{a^n \mathbb{E}|X|^n}{n!}$

$\implies \mathbb{E}|X|^n \leq \frac{B}{a^n} n!$ für alle $n \in \mathbb{N}$

$\implies |\mathbb{E}X^n| \leq \frac{B}{a^n} n!$

$\implies \mathbb{E}\exp(zX)$ ist analytisch für $|z| < a$

Definition. $\mathbb{E}\exp(zX)$ heißt momenterzeugende Funktion der ZG X oder VF F_X.

$X = N(\mu, \sigma^2) \implies \mathbb{E}\exp(zX) = \exp\left(z\mu + \frac{\sigma^2}{2} z^2\right)$ für $z \in \mathbb{C}$

Höldersche Ungleichung:

$|\mathbb{E}XY| \leq \mathbb{E}|XY| \leq (\mathbb{E}|X|^p)^{\frac{1}{p}} \cdot (\mathbb{E}|X|^q)^{\frac{1}{q}}$ für $p, q \geq 1$

$\implies$ Cauchy-Schwarz-Bunjakowski-Ungleichung für $p = q = 2$:

$|\mathbb{E}XY| \leq \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}$

Verallgemeinerung:

$\mathbb{E}(X_1^{n_1} \cdots X_k^{n_k}) \leq (\mathbb{E}|X_1|^n)^{\frac{n_1}{n}} \cdots (\mathbb{E}|X_k|^n)^{\frac{n_k}{n}}$, $n = n_1 + ... + n_k$

Jensensche Ungleichung

Sei $g : \mathbb{R}^1 \to \mathbb{R}^1$ konvex auf einem Intervall J, d. h.

$g(\alpha x + (1-\alpha)y) \leq \alpha g(x) + (1-\alpha) g(y)$ für alle $x, y \in I$ und $\alpha \in [0,1]$

Per Induktion folgt: $g(\sum_{i=1}^n \alpha_i x_i) \leq \sum_{i=1}^n \alpha_i g(x_i)$ für $x_1, ..., x_n \in J$, $\alpha_i \geq 0$, $\alpha_1 + ... + \alpha_n = 1$

Satz. $g(\mathbb{E}X) \leq \mathbb{E}g(X)$, falls $\mathbb{P}(X \in J) = 1$ und $\mathbb{E}|X| < \infty$

$g(x) = |x|^{\frac{n}{m}}$ für $0 < m \leq n$, $\implies$ **Ljapunow-Ungleichung**

Problem (Momentenproblem)**.** Unter welchen Bedingungen ist eine Zahlenfolge $c_0 = 1, c_1, c_2, ...$ eine Momentenfolge einer ZG X, d. h. $c_n = \mathbb{E}X^n$.

Antwort.
$0 \leq \mathbb{E}(z_0 + z_1 X + ... + z_n X^n)^2 = \mathbb{E}(\sum_{i,j=1}^n z_i z_j X^{i+j}) = \sum_{i,j=1}^n z_i z_j c_{i+j}$

genau dann, wenn

$$\det \begin{pmatrix} c_0 & c_1 & \cdots & c_n \\ c_1 & c_2 & \cdots & c_{n+1} \\ \vdots & \vdots & & \vdots \\ c_n & c_{n+1} & \cdots & c_{2n} \end{pmatrix} \geq 0.$$

Problem. Wann ist die zugehörige VF F_X eindeutig festgelegt?
$c_n = \int_0^\infty x^n \,\mathrm{d}F_X(x)$ (Stieltjes-MP), $c_n = \int_{-\infty}^\infty x^n \,\mathrm{d}F_X(x)$ (Hamburger MP)

Hinreichende Bedingung für Bestimmtheit:

Stieltjes-MP: $\sum_{n=1}^\infty \frac{1}{c_{2n}^{\frac{1}{n}}} = \infty$

Hamburger MP: $\sum_{n=1}^\infty \frac{1}{c_{2n}^{\frac{1}{2n}}} = \infty$

(Carleman-Kriterien)

Definition. Sei $X = (X_1, ..., X_k)$ eine k-dimensionale ZV über $(\Omega, \mathfrak{A}, \mathbb{P})$. $X_1, ..., X_k$ heißen **stochastisch unabhängig**, falls

$$\mathbb{P}(\bigcap_{i=1}^k \{X_i \in B_i\}) = \prod_{i=1}^k \mathbb{P}(\{X_i \in B_i\})$$

für alle $B_1, ..., B_k \in \mathcal{L}(\mathbb{R}^1)$. Dies ist genau dann der Fall, wenn

$$\mathbb{P}(X_1 \leq x_1, ..., X_k \leq x_k) = \mathbb{P}(X_1 \leq x_1) \cdots \mathbb{P}(X_k \leq X_k)$$

für alle $x_1, ..., x_k \in \mathbb{R}$. Falls die W-Dichte $f_X(x_1, ..., x_k) = \frac{\partial}{\partial x_1 \cdots \partial x_k} F(x_1, ..., x_k)$ existiert (also F_X absolut stetig), ist dies äquivalent zu

$$f_X(x_1, ..., x_k) = f_{X_1}(x_1) \cdots f_{X_k}(x_k).$$

Für diskrete Verteilungen ist dies äquivalent zu

$$\mathbb{P}(X_1 = x_1, ..., X_k = x_k) = \mathbb{P}(X_1 = x_1) \cdots \mathbb{P}(X_k = x_k)$$

für alle $x_1, ..., x_k \in \mathbb{R}$.

Definition. Für eine k-dimensionale ZV $X = (X_1, ..., X_k)$ heißt

$$F_{(X_{i_1}, ..., X_{i_l})}(x_{i_1}, ..., x_{i_l}) = \lim_{x_j \to \infty\, j \in \{1,...,k\} \setminus \{i_1,...,i_l\}} F_(X_1, ..., X_k)(x_1, ..., x_k)$$

für $1 \leq i_1 < ... < i_l \leq k$, $l = 1, ..., k-1$ **l-dimensionale Rand-(Marginal-)verteilungsfunktion**.

Falls $f_X(x_1, ..., x_k)$ existiert, so existieren sämtliche Randdichten

$$f_{(X_{i_1}, ..., X_{i_k})}(x_{i_1}, ..., x_{i_k}) = \int_{\mathbb{R}^{k-l}} f_{(X_1, ..., X_k)}(x_1, ..., x_k) \,\mathrm{d}(x_1, \widehat{x_{i_1}}, ..., \widehat{x_{i_k}}, ..., x$$

Analog folgt für eine diskrete ZV die Diskretheit der Randverteilungen $k = 2$:

$$\mathbb{P}(X_1 = x_m^{(1)}) = \sum_{x_m^{(2)}} \mathbb{P}(X_1 = x_m^{(1)}, X_2 = x_m^{(2)})$$

$$\mathbb{P}(X_2 = x_m^{(2)}) = \sum_{x_m^{(1)}} \mathbb{P}(X_1 = x_m^{(1)}, X_2 = x_m^{(2)})$$

wobei $x_m = (x_m^{(1)}, ..., x_m^{(k)})$ die Massenschwerpunkte sind.
Wichtig: Im allgemeinen bestimmen die Randverteilungen nicht die gemeinsame Verteilung des Vektors.

Definition. (X, Y) sei eine zweidimensionale ZVüber $(\Omega, \mathfrak{A}, \mathbb{P})$ mit $\mathbb{E}X^2 < \infty$, $\mathbb{E}Y^2 < \infty$. Dann heißt

$$\mathrm{Var}(X, Y) := \mathbb{E}(XY) - \mathbb{E}X\mathbb{E}Y = \mathbb{E}((X - \mathbb{E}X) \cdot (Y - \mathbb{E}Y))$$

Kovarianz von X und Y und

$$\mathrm{Var}(X, Y) := \frac{\mathrm{Var}(X, Y)}{\sqrt{\mathrm{Var}(X) \cdot \mathrm{Var}(Y)}}$$

Korrelation von X und Y.

Satz. • Falls X, Y unabhängig, so gilt $\mathrm{Var}(X, Y) = \mathrm{Var}(X, Y) = 0$

- $|\mathrm{Var}(X, Y)| \leq 1$
- $\mathrm{Var}(X, Y) = 1 \iff \exists\, a, b \in \mathbb{R}^1 : \mathbb{P}(Y = aX + b) = 1$.

Definition. Falls $\mathrm{Var}(X, Y) = 0$, so heißen X, Y **unkorreliert**.

Achtung. Aus Unkorreliertheit folgt i. A. nicht Unabhängigkeit!

Beispiel. Sei X eine ZG mit der symmetrischen Dichte $f_X(x) = f_X(-x)$ und $\int_{-\infty}^\infty |x|^3 f_X(x) \,\mathrm{d}x < \infty$, dann ist $\mathrm{Var}(X, X^2) = 0$, aber X und X^2 nicht unabhängig.

Bemerkung. • $\mathrm{Var}(X, Y) = 1$: positive Korrelation

- $\mathrm{Var}(X, Y) = -1$: negative Korrelation
- $\mathrm{Var}(X, Y) = 0$: Unkorreliertheit

Wichtig: Falls (X, Y) eine zweidimensionale Normalverteilung besitzt, so folgt aus $\mathrm{Var}(X, Y) = 0$ die Unabhängigkeit von X und Y.

Satz. $X_1, ..., X_n$ seien paarweise unkorrelierte ZGen mit $\mathbb{E}X_i^2 < \infty$ für $i = 1, ..., n$. Dann gilt

$$\mathrm{Var}(X_1 + ... + X_n) = \mathrm{Var}(X_1) + ... + \mathrm{Var}(X_n)$$