Zusammenfassung Stochastik I

Der abstrakte Maß- und Wkts-Begriff

Def. 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 alle $A, B, C \in \mathfrak{A}$ gilt:

- $A \wedge A = A$
- $A \wedge U = U$
- $A \vee A = A$
- $A \vee U = A$
- $A \wedge (B \vee C) = (A \wedge B) \vee (A \wedge C)$
- $A \wedge B = B \wedge A$
- $A \wedge \overline{A} = U$
- $A \vee B = B \vee A$
- $A \vee \overline{A} = S$
- $A \wedge S = A$
- $A \wedge (B \wedge C) = (A \wedge B) \wedge C$
- $A \vee S = S$
- $A \vee (B \vee C) = (A \vee B) \vee C$

Def. Sei $\mathfrak{A}$ eine Ereignisalgebra und A, B Ereignisse.

- Durch $A \leq B \;:\Longleftrightarrow\; A \wedge B = B$ (gesprochen A **impliziert** B) ist auf $\mathfrak{A}$ eine Partialordnung definiert.
- A und B heißen **äquivalent**, falls $A \leq B$ und $B \leq A$.
- A und B heißen **unvereinbar**, falls $A \wedge B = U$.

Kor. In einer Ereignisalgebra $\mathfrak{A}$ gilt mit $A, B \in \mathfrak{A}$:

- $\overline{\overline{A}} = A$
- $\overline{A \vee B} = \overline{A} \wedge \overline{B}$
- $A \leq B \iff \overline{B} \leq \overline{A}$ (Kontraposition)
- $\overline{A \wedge B} = \overline{A} \vee \overline{B}$ (De Morgansche Regeln)

Kor. Durch Induktion folgt aus den De Morganschen Regeln:

$$\overline{\left(\bigvee_{i=1}^{n} A_i\right)} = \bigwedge_{i=1}^{n} \overline{A_i} \quad \text{und} \quad \overline{\left(\bigwedge_{i=1}^{n} A_i\right)} = \bigvee_{i=1}^{n} \overline{A_i} \quad \text{für } A_1, ..., A_n \in \mathfrak{A}.$$

Def. Eine **Algebra** (Mengenalgebra) über Ω ist ein System von Teilmengen $\mathfrak{A} \subset \mathcal{P}(\Omega)$ mit $\Omega \in \mathfrak{A}$, das unter endl. Vereinigungen und Komplementen stabil ist, d. h. für alle $A, B \in \mathfrak{A}$ gilt:

- $\Omega \in \mathfrak{A}$
- $A \cup B \in \mathfrak{A}$
- $A^c \coloneqq \Omega \setminus A \in \mathfrak{A}$

Bem. Aus den De Morganschen Regeln folgt, dass Mengenalgebren auch unter endlichen Schnitten stabil sind.

Satz (Isomorphiesatz von Stone)**.** Zu jeder Ereignisalgebra $\mathfrak{A}$ gibt es eine Menge Ω, sodass $\mathfrak{A}$ isomorph zu einer Mengenalgebra über Ω ist.

Notation. $A \triangle B \coloneqq (A \setminus B) \cup (B \setminus A)$ heißt **symm. Differenz**.

Def. Eine **σ-Algebra** über Ω ist eine Algebra $\mathfrak{A} \subset \mathcal{P}(\Omega)$ über Ω, die auch unter abzählbaren Vereinigungen stabil ist, d. h.

$$\bigcup_{n=0}^{\infty} A_n \in \mathfrak{A} \quad \text{für alle Folgen } (A_n)_{n \in \mathbb{N}} \text{ in } \mathfrak{A}.$$

Bem. Jede σ-Algebra ist auch unter abzählb. Schnitten stabil.

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

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

Bem. 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.

Def. Eine Folge $(A_n)_{n \in \mathbb{N}}$ in einer σ-Algebra $\mathfrak{A}$ **konvergiert** gegen $A \in \mathfrak{A}$, notiert $\lim_{n \to \infty} A_n = A$, falls $A = \liminf_{n \to \infty} A_n = \limsup_{n \to \infty} A_n$.

Satz. Für isotone / antitone Folgen $(A_n)_{n \in \mathbb{N}}$ in $\mathfrak{A}$ gilt:

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

Def. Ein **Ring** (Mengenring) über Ω ist ein System von Teilmengen $\mathfrak{R} \subset \mathcal{P}(\Omega)$ mit $\emptyset \in \mathfrak{R}$, das unter endlichen Vereinigungen und Differenzen stabil ist, d. h. für alle $A, B \in \mathfrak{R}$ gilt:

- $\emptyset \in \mathfrak{R}$
- $A \cup B \in \mathfrak{R}$
- $A \setminus B \in \mathfrak{R}$

Bem. Ein Ring ist auch unter Schnitten stabil, da

$$A \cap B = (A \cup B) \setminus (A \triangle B).$$

Def. Ein **σ-Ring** über Ω ist ein Ring $\mathfrak{R} \subset \mathcal{P}(\Omega)$ über Ω, der auch unter abzählbaren Vereinigungen stabil ist, d. h.

$$\bigcup_{n=0}^{\infty} A_n \in \mathfrak{A} \quad \text{für alle Folgen } (A_n)_{n \in \mathbb{N}} \text{ in } \mathfrak{A}.$$

Bem. Jeder σ-Ring ist auch unter abzählb. Schnitten stabil.

Satz. $\mathfrak{A}$ ist $(\sigma$-$)$ Algebra $\iff$ $\mathfrak{A}$ ist $(\sigma$-$)$ Ring mit $\Omega \in \mathfrak{A}$.

Satz. Sei $(\mathfrak{A}_i)_{i \in I}$ Familie von $(\sigma$-$)$ Ringen / $(\sigma$-$)$ Algebren über Ω. Dann ist der Schnitt $\bigcap_{i \in I} \mathfrak{A}_i$ ein $(\sigma$-$)$ Ring / eine $(\sigma$-$)$ Algebra über Ω.

Def. Sei $E \subset \mathcal{P}(\Omega)$. Setze

$$\mathcal{R}(E) \coloneqq \{\mathfrak{R} \subset \mathcal{P}(\Omega) \mid E \subset \mathfrak{R}, \mathfrak{R}\ \sigma\text{-Ring}\} \text{ und}$$
$$\mathcal{A}(E) \coloneqq \{\mathfrak{A} \subset \mathcal{P}(\Omega) \mid E \subset \mathfrak{A}, \mathfrak{A}\ \sigma\text{-Algebra}\}.$$

Dann heißen $\rho(E) \coloneqq \bigcap_{\mathfrak{R} \in \mathcal{R}(E)} \mathfrak{R}$, $\quad \sigma(E) \coloneqq \bigcap_{\mathfrak{A} \in \mathcal{A}(E)} \mathfrak{A}$

von E **erzeugter Ring** bzw. von E **erzeugte σ-Algebra**.

Def. Die **Borel-Mengen** in $\mathbb{R}^1$ sind $\mathfrak{B}(\mathbb{R}^1) \coloneqq \sigma(\mathcal{E})$, wobei wir $\mathcal{E}$ aus folgenden äquivalenten Optionen wählen dürfen:

- $\{(a, b] \mid a \leq b\}$
- $\{[a, b] \mid a \leq b\}$
- $\{(a, b) \mid a \leq b\}$
- $\{[a, b) \mid a \leq b\}$
- $\{G \subset \mathbb{R}^1 \mid G \text{ abgeschl.}\}$
- $\{F \subset \mathbb{R}^1 \mid F \text{ offen}\}$

Notation. $\overline{\mathbb{R}^1} \coloneqq \mathbb{R} \cup \{-\infty, +\infty\} = [-\infty, \infty]$

Def. Funktionen mit Wertebereich $\overline{\mathbb{R}^1}$ heißen **numerisch**.

Def. Sei $\mathfrak{R}$ ein Ring über Ω. Eine Fkt. $\mu : \mathfrak{R} \to [0, \infty]$ heißt

- **Inhalt** auf $\mathfrak{R}$, falls $\mu(\emptyset) = 0$ und $\mu(A \cup B) = \mu(A) + \mu(B)$ für alle $A, B \in \mathfrak{R}$ mit $A \cap B = \emptyset$ gilt.
- **Prämaß** auf $\mathfrak{R}$, wenn μ ein Inhalt ist und für alle Folgen $(A_n)_{n \in \mathbb{N}}$ in $\mathfrak{R}$ mit $A_i \cap A_j = \emptyset$ für $i \neq j$ und $\bigcup_{n=1}^{\infty} A_n \in \mathfrak{R}$ gilt:
$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n) \qquad (\sigma\text{-Additivität})$$
- **Maß**, wenn μ Prämaß und $\mathfrak{R}$ in Wahrheit sogar eine σ-Algebra ist. Dann ist die letzte Vorraussetzung in Punkt 2 immer erfüllt.

Def. Ein Inhalt/Maß μ auf einem Ring / einer σ-Algebra $\mathfrak{A}$

- heißt **endlich**, falls $\mu(\Omega) < \infty$,
- heißt **σ-endlich**, falls eine Folge A_n in $\mathfrak{A}$ existiert, sodass
$$\Omega = \bigcup_{n \in \mathbb{N}} A_n \quad \text{und} \quad \forall\, i \in \mathbb{N} : \mu(A_i) < \infty.$$

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

$$\chi_1 = \mathbb{1}_A : \Omega \to \mathbb{R}, \quad \omega \mapsto |\{\star \mid \omega \in A\}| = \begin{cases} 1, & \text{falls } \omega \in A \\ 0, & \text{falls } \omega \notin A \end{cases}$$

Indikatorfunktion von A.

Bsp. Sei $\mathfrak{R}$ ein Ring über Ω und $\omega \in \Omega$. Die Abbildung

$$\delta_\omega : \mathfrak{R} \to [0, \infty], \quad A \mapsto \mathbb{1}_A(\omega)$$

ist dann ein Prämaß auf $\mathfrak{R}$, genannt **Dirac-(Prä)-Maß**.

Lem. Sei μ ein Inhalt auf einem Ring $\mathfrak{R}$. Seien $A, B \in \mathfrak{R}$ und $(A_n)_{n \in \mathbb{N}}$ Folge in $\mathfrak{R}$ mit $\mu(A) < \infty$ und $\forall\, n \in \mathbb{N} : \mu(A_n) < \infty$. Dann:

- $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$
- $A \subset B \implies \mu(A) \leq \mu(B)$ (Isotonie)
- $A \subset B \implies \mu(B \setminus A) = \mu(B) - \mu(A)$
- $\mu(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} \mu(A_i)$ (Subadditivität)

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

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

Satz. Sei μ ein Inhalt auf $\mathfrak{R} \subset \mathcal{P}(\Omega)$. Wir betrachten die Aussagen:

(i) μ ist ein Prämaß auf $\mathfrak{R}$.

(ii) Stetigkeit von unten: Für jede monoton wachsende Folge $(A_n)_{n \in \mathbb{N}}$ in $\mathfrak{R}$ mit $A \coloneqq \lim_{n \to \infty} A_n = \bigcup_{n=0}^{\infty} A_n \in \mathfrak{R}$ gilt $\lim_{n \to \infty} \mu(A_n) = \mu(A)$.

(iii) Stetigkeit von oben: Für jede monoton fallende Folge $(A_n)_{n \in \mathbb{N}}$ in $\mathfrak{R}$ mit $\mu(A_0) < \infty$ und $A \coloneqq \lim_{n \to \infty} A_n = \bigcap_{n=0}^{\infty} A_n \in \mathfrak{R}$ gilt $\lim_{n \to \infty} \mu(A_n) = \mu(A)$.

(iv) Für jede monoton fallende Folge $(A_n)_{n \in \mathbb{N}}$ in $\mathfrak{R}$ mit $\mu(A_0) < \infty$ und $\lim_{n \to \infty} A_n = \bigcap_{n=0}^{\infty} A_n = \emptyset$ gilt $\lim_{n \to \infty} \mu(A_n) = 0$.

Dann gilt (i) $\iff$ (ii) $\implies$ (iii) $\iff$ (iv).
Falls μ endlich ist, so gilt auch (iii) $\implies$ (ii).

Def. Seien $a = (a_1, ..., a_d), b = (b_1, ..., b_d) \in \mathbb{R}^d$. Wir schreiben $a \leq b$, falls $a_i \leq b_i$ für alle $i \in \{1, ..., d\}$ gilt. Dann heißt

$$(a, b] := \{(x_1, ..., x_d) \in \mathbb{R}^d \mid a_i < x_i \leq b_i \text{ für alle } i \in \{1, ..., d\}\}$$

von a und b aufgespannter **Elementarquader** in $\mathbb{R}^d$.

Def. Sei $f : \mathbb{R}^d \to \mathbb{R}^1$ eine Funktion, $x = (x_1, ..., x_d) \in \mathbb{R}^d$ und $h = (h_1, ..., h_d) \in \mathbb{R}^d_{\geq 0}$. Dann heißt

$$(\triangle f)((x, x + h]) := \sum_{\delta_1,...,\delta_d \in \{0,1\}} (-1)^{d-(\delta_1+...+\delta_k)} f(x_1 + \delta_1 h_1, ..., x_d + \delta_d h_d)$$

Zuwachs von f im Elementarquader $(x, x + h]$.

Def. $G : \mathbb{R}^d \to \mathbb{R}$ heißt **maßerzeugende Funktion**, falls

- G ist nicht-fallend in jedem Argument, d. h. für alle $k \in \{1, ..., d\}$ und $x_1, ..., x_d \in \mathbb{R}$ ist $f(x_1, ..., x_{k-1}, -, x_{k+1}, ..., x_d)$ nicht-fallend.
- G ist rechtsseitig stetig in jedem Argument, d. h. für alle $k \in \{1, ..., d\}$ und $x_1, ..., x_d \in \mathbb{R}$ gilt
$$\lim_{h \downarrow 0} f(x_1, ..., x_{k-1}, x_k + h, x_{k+1}, ..., x_d) = f(x_1, ..., x_d).$$
- Für alle $x \in \mathbb{R}^d$ und $h \in \mathbb{R}^d_{\geq 0}$ ist der Zuwachs $(\triangle G)((x, x + h]) \geq 0$.

Def. Eine maßerzeugende Funktion F heißt **Verteilungsfunktion** (VF) in $\mathbb{R}^d$, falls zusätzlich gilt:

$$\lim_{\substack{x_1 \to \infty \\ ... \\ x_d \to \infty}} F(x_1, ..., x_d) = 1 \quad \text{und} \quad \lim_{x_i \to -\infty} F(x_1, ..., x_d) = 0$$

für alle $i \in \{1, ..., d\}$ und $x_1, ..., \widehat{x_i}, ..., x_d \in \mathbb{R}$ fest.

Bem. Sei $G_i : \mathbb{R}^1 \to \mathbb{R}^1_{>0}$ für $i \in \{1, ..., d\}$ maßerzeugende Funktion im $\mathbb{R}^1$, dann ist

$$G : \mathbb{R}^d \to \mathbb{R}^1_{>0}, \quad (x_1, ..., x_d) \mapsto G_1(x_1) \cdot ... \cdot G_d(x_d)$$

eine maßerzeugende Funktion in $\mathbb{R}^d$ und es gilt für jeden Elementarquader $(a, b] \subset \mathbb{R}^d$ mit $a = (a_1, ..., a_d)$, $b = (b_1, ..., b_d)$:

$$(\triangle G)((a, b]) = \prod_{i=1}^d (G_i(b_i) - G_i(a_i)) .$$

Satz. Der Ring aller Elementarquader im $\mathbb{R}^d$ ist

$$\mathfrak{R} := \{ \bigcup_{i=1}^m (a_i, b_i] \mid m \in \mathbb{N} \text{ und } (a_1, b_1] , ..., (a_m, b_m] \text{ disjunkte Elementarquader im } \mathbb{R}^d\}$$

und für jede maßerzeugende Funktion $G : \mathbb{R}^d \to \mathbb{R}^1$ definiert

$$\mu_G : \mathfrak{R} \to [0, \infty) , \quad \bigcup_{i=1}^m (a_i, b_i] \mapsto \sum_{i=1}^m (\triangle G)((a_i, b_i])$$

einen Inhalt auf $\mathfrak{R}$, der sogar ein Prämaß ist.

Def. Eine numerische Funktion $\mu^* : \mathcal{P}(\Omega) \to \overline{\mathbb{R}}$ heißt **äußeres Maß** auf Ω, wenn gilt:

- $\mu^*(\emptyset) = 0$
- $A \subset B \implies \mu^*(A) \leq \mu^*(B)$ (Monotonie)
- Für eine Folge $(A_n)_{n \in \mathbb{N}}$ in $\mathcal{P}(\Omega)$ gilt $\mu^* \left(\bigcup_{n \in \mathbb{N}} A_n \right) \leq \sum_{n=0}^{\infty} \mu^*(A_n)$.

Bem. Wegen $\mu^*(\emptyset) = 0$ und der Monotonie nimmt ein äußeres Maß nur Werte in $[0, \infty]$ an.

Def. Eine Teilmenge $A \subset \Omega$ heißt **μ^*-messbar**, falls

$$\mu^*(Q) = \mu^*(Q \cap A) + \mu^*(Q \setminus A) \quad \text{für alle } Q \subset \Omega.$$

Satz (Carathéodory)**.** Für ein äußeres Maß $\mu^* : \mathcal{P}(\Omega) \to [0, \infty]$ ist

- $\mathfrak{A}^* := \{A \subset \Omega \mid A \text{ ist } \mu^*\text{-messbar } \}$ eine σ-Algebra und
- $\mu^*|_{\mathfrak{A}^*}$ ein Maß auf $\mathfrak{A}^*$.

Satz (**1. Fortsetzungssatz**)**.**
- Sei μ ein Prämaß auf einem Ring $\mathfrak{R}$ über Ω. Dann existiert eine Fortsetzung $\widetilde{\mu}$ von μ zu einem Maß auf der von $\mathfrak{R}$ erzeugten σ-Algebra $\mathfrak{A} := \sigma(\mathfrak{R})$, sodass $\widetilde{\mu}|_{\mathfrak{R}} = \mu$.
- Falls μ σ-endlich ist, so ist die Fortsetzung eindeutig.

Bem. Im Beweis wird ein äußeres Maß μ^* auf Ω so definiert:

$$\mathfrak{U}(Q) := \left\{ (A_n)_{n \in \mathbb{N}} \,\middle|\, Q \subset \bigcup_{n=0}^{\infty} A_n \text{ und } A_n \text{ Folge in } \mathfrak{R} \right\},$$

$$\mu^*(Q) := \inf \left(\left\{ \sum_{i=0}^{\infty} \mu(A_n) \,\middle|\, (A_n)_{n \in \mathbb{N}} \in \mathfrak{U}(Q) \right\} \cup \{\infty\} \right).$$

Das äußere Maß μ^* eingeschränkt auf $\mathfrak{A}^* \supset \mathfrak{A}(\mathfrak{R})$ ist ein Maß.

Def. Sei Ω eine Menge und $\mathfrak{A} \subset \mathcal{P}(\Omega)$ eine σ-Algebra auf Ω, sowie ggf. μ ein Maß auf $\mathfrak{A}$. Dann heißt

- das Tupel $(\Omega, \mathfrak{A})$ **messbarer Raum**,
- das Tripel $(\Omega, \mathfrak{A}, \mu)$ **Maßraum**.

Satz. Unter den Bedingungen des 1. Fortsetzungssatzes ist $\mathfrak{A}^*$ die größte σ-Algebra $\overline{\mathfrak{A}}$ mit $\mathfrak{A} \subset \overline{\mathfrak{A}}$, sodass $\mu^*|_{\overline{\mathfrak{A}}}$ ein Maß ist.

Def. Sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum. Eine Menge $N \subset \Omega$ heißt **Nullmenge**, falls es ein $A \in \mathfrak{A}$ gibt, sodass $N \subset A$ und $\mu(A) = 0$.

Def. Ein Maßraum $(\Omega, \mathfrak{A}, \mu)$ heißt **vollständig**, falls jede Nullmenge $N \subset \Omega$ ein Element von $\mathfrak{A}$ ist.

Satz. $(\Omega, \mathfrak{A}^*, \mu^*|_{\mathfrak{A}^*})$ ist vollständig für jedes bel. äußere Maß μ^*.

Satz. Jeder Maßraum $(\Omega, \mathfrak{A}, \mu)$ kann zu einem vollständigen Maßraum $(\Omega, \mathfrak{A}_c, \mu_c)$ erweitert werden mit

$$\mathfrak{A}_c := \{A \cup N \mid A \in \mathfrak{A},\ N\ \mu\text{-Nullmenge}\}, \quad \mu_c(A \cup N) := \mu(A).$$

Satz. Sei μ ein σ-endliches Prämaß auf dem Ring $\mathfrak{R}$ über Ω sowie $\widetilde{\mathfrak{A}} := \sigma(\mathfrak{R})$. Dann gilt $\mathfrak{A}^* = \mathfrak{A}_c$ und $\mu^*|_{\mathfrak{A}^*} = \widetilde{\mu}_c$, wobei $\widetilde{\mu}$ das eindeutig fortgesetzte Maß ist.

Sprechweise. Eine Eigenschaft oder Aussage gilt für **fast alle** $\omega \in \Omega$ oder **μ-fast-überall**, wenn es eine Nullmenge $N_0 \subset \Omega$ gibt, sodass die Aussage oder Eigenschaft für alle $\omega \in N_0^c$ gilt.

Def. Sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum.

- Dann heißt μ **diffus** (atomlos), falls $\mu(\{\omega\}) = 0$ für alle $\omega \in \Omega$.
- Sei η ein weiteres Maß auf $(\Omega, \mathfrak{A})$. Dann heißt μ **absolut stetig** bezüglich η (notiert $\mu \ll \eta$), falls
$$\eta(A) = 0 \implies \mu(A) = 0 \quad \text{für alle } A \in \mathfrak{A}.$$

Def. Die von den Elementarquadern im $\mathbb{R}^d$ erzeugte σ-Algebra heißt **Borel-σ-Algebra** $\mathfrak{B}(\mathbb{R}^d)$. Das von der maßerzeugenden Funktion

$$G : \mathbb{R}^d \to \mathbb{R}, \quad (x_1, ..., x_d) \mapsto x_1 \cdot ... \cdot x_d$$

erzeugte Prämaß auf dem von den Elementarquadern erzeugten Ring μ_G, das zu einem Maß $\widetilde{\mu}_G$ auf $\mathfrak{B}(\mathbb{R}^d)$ fortgesetzt wird, heißt **Lebesgue-Borel-Maß**. Die durch Hinzunahme aller Nullmengen vervollständigte σ-Algebra $\mathfrak{A}^* = \mathfrak{B}(\mathbb{R}^d)_c$ heißt **Lebesgue-σ-Algebra** und das fortgesetzte Maß $\lambda_d := \mu^*_G$ **Lebesgue-Maß**.

Satz. Das Lebesgue-Maß auf dem $\mathbb{R}^d$ ist bewegungsinvariant, d. h.

$$\forall\, A \in \mathfrak{B}(\mathbb{R}^d),\ O \in \mathrm{SO}_d,\ x \in \mathbb{R}^d \ : \ \lambda_d(O \cdot A + x) = \lambda_d(A).$$

Das Lebesgue-Maß ist bis auf einen multiplikativen Faktor das einzige verschiebungsinvariante Maß auf $(\mathbb{R}^d, \mathfrak{B}(\mathbb{R}^d))$.

Sprechweise. Sei $(\Omega, \mathfrak{A})$ ein messbarer Raum. Wir nennen Ω abstrakte Grundmenge und die Elemente von Ω Elementarereignisse. Die σ-Algebra $\mathfrak{A}$ enthält zufällige Ereignisse, unter anderem das sichere Ereignis Ω und das unmögliche Ereignis $\emptyset$.

Def. Sei $(\Omega, \mathfrak{A}, \mathbb{P})$ ein Maßraum mit $\mathbb{P}(\Omega) = 1$. Dann heißt $\mathbb{P}$ **Wahrscheinlichkeitsmaß** (W-Maß) und das Tripel $(\Omega, \mathfrak{A}, \mathbb{P})$ **Wahrscheinlichkeitsraum** (W-Raum).

Sprechweise. Sei $(\Omega, \mathfrak{A}, \mathbb{P})$ ein W-Raum und $A, B \in \mathfrak{A}$. Wir sagen:

- A ist **fast sicher**, wenn $\mathbb{P}(A) = 1$.
- A ist **fast unmöglich**, wenn $\mathbb{P}(A) = 0$.
- A und B sind **äquivalent**, wenn $\mathbb{P}(A \triangle B) = 0$.

Bem. Sei μ ein W-Maß auf $\mathfrak{L}(\mathbb{R}^1)$. 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 $\mathfrak{L}(\mathbb{R}^1)$. Analog funktioniert dies auf dem $\mathbb{R}^d$.

Def (Wichtige Verteilungsfunktionen)**.**

- **Exponentialverteilung** mit Parameter $\lambda > 0$:
$$F_\lambda(x) = \max(0, 1 - \exp(-\lambda x)) \qquad \mathrm{Exp}(\lambda)$$
- **Poisson-Verteilung** mit Parameter $\lambda > 0$:
$$F_\lambda(x) = \sum_{0 \leq n \leq x} \frac{\lambda^n}{n!} \exp(-\lambda) \qquad \mathrm{Poi}(\lambda)$$
- **Gleichverteilung** auf $(a, b]$: $F(x) = \min(1, \max(0, \frac{x-a}{b-a}))$
- **Normalverteilung** (Gaußverteilung) mit EW μ und Varianz σ^2:
$$F_{\mu\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} \exp\left(\frac{-(t-\mu)^2}{2\sigma^2} \right) \mathrm{d}t \qquad \mathrm{N}(\mu, \sigma^2)$$
besitzt die Dichte $F'_{\mu\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2} \right)$ und die Symmetrie $F_{\mu\sigma^2}(\mu - x) = 1 - F_{\mu\sigma^2}(\mu + x)$.

- d-dimensionale **Normalverteilung** mit Erwartungswertvektor $m \in \mathbb{R}^d$ positiv definiter Kovarianzenmatrix $C \in \mathbb{R}^{d\times d}$:

$$F(x) = \frac{1}{\sqrt{(2\pi)^d \det C}} \int_{(-\infty,x]} \exp(-\tfrac{1}{2}(y-m)^T C^{-1}(y-m))\,\mathrm{d}y$$

- **2-dimensionale Exponentialverteilung** mit $\lambda, \mu > 0$, $\nu \geq 0$:

$$F(x,y) = \begin{cases} 0, \text{ falls } x < 0 \text{ oder } y < 0, \text{ ansonsten:} \\ 1 - e^{-(\lambda+\nu)x} - e^{-(\mu+\nu)y} + e^{-(\lambda x + \mu y + \nu \max(x,y))} \end{cases}$$

Elementare Wahrscheinlichkeitsrechnung

Def. 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.

Bem. 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 $A \cap B = \emptyset$

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

Def. 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.

Bem. 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 $A_1 \cap A_2 = \emptyset$

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

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

Def. Sei Ω eine endliche Menge. Dann definiert

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

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

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

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

Lem. 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	$\frac{(n+r-1)!}{r!}$	$\binom{n}{r} := \frac{n!}{r!(n-r)!}$

Lem. 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,

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

Bem. 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**.

Def. 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.

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

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

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

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

Sprechweise. In der Bayesischen Statistik heißt

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

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

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

Bem.
- $A \in \mathfrak{A}$ 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 $\iff \mathbb{P}B \mid A = \mathbb{P}(B)$.

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

- **vollständig unabhängig**, falls

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

  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).

Def. 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**.

Lem. Voraussetzung wie 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 $m \in \mathbb{N}$, $\lambda \in \mathbb{R}_{>0}$ gilt

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

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

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

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

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

Lem. 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

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

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

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

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

Def. 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

Def. 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 $f : (\Omega_1, \mathfrak{A}_1) \to (\Omega_2, \mathfrak{A}_2)$.

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

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

Lem. 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 $g \circ f : (\Omega_1, \mathfrak{A}_1) \to (\Omega_3, \mathfrak{A}_3)$.

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

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

Lem. 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 $f, g : \Omega \to \overline{\mathbb{R}^1}$ zwei numerische Funktionen. Setze

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

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

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

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

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

  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

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

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

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

Satz. Seien $f_n : (\Omega, \mathfrak{A}) \to (\overline{\mathbb{R}^1}, \overline{\mathfrak{B}})$, $n \in \mathbb{N}$ messbare numerische Funktionen, dann auch messbar:
- $\sup\limits_{n\in\mathbb{N}} f_n$
- $\inf\limits_{n\in\mathbb{N}} f_n$
- $\liminf\limits_{n\in\mathbb{N}} f_n$
- $\limsup\limits_{n\in\mathbb{N}} f_n$

Dabei werden Infimum, Supremum, usw. punktweise gebildet.

Def. Für $f : \Omega \to \overline{\mathbb{R}^1}$ 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 $f : (\Omega, \mathfrak{A}) \to (\overline{\mathbb{R}^1}, \overline{\mathfrak{B}})$ 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:

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

Def. 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:

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

Beob. Jede kumulierte VF ist eine Càdlàg-Funktion.

Def. Die **Variation** von $g : [a,b] \to \mathbb{R}$ bzgl. einer Zerlegung $Z = \{a = x_0 < \dots < x_n = b\}$ 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

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

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

Notation. Für eine ZG X und eine Fkt. $g : \mathbb{R}^1 \to \overline{\mathbb{R}^1}$ schreiben wir

$$f(X) := f \circ X : \Omega \to \overline{\mathbb{R}^1}.$$

Def. 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 $X :=$ id.

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

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

Def. Sei $X_1, ..., X_n$ eine endliche Familie von ZGn 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 ZGn 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 ZGn.

Def. 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 $\mathbb{E}(\Omega, \mathfrak{A})$:

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

Def. 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

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

Def. 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

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

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 $\int_\Omega f \, d\mu \leq \sup_{n \in \mathbb{N}} \int_\Omega f_n \, d\mu$.

Kor. 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 $f : (\Omega, \mathfrak{A}, \mu) \to (\overline{\mathbb{R}^1}, \mathfrak{L}(\overline{\mathbb{R}^1}))$ nichtnegativ. Dann gibt es eine isotone Folge $(f_n)_{n \in \mathbb{N}}$ elementarer Funktionen mit $\sup_{n \in \mathbb{N}} f_n = f$.

Def. Sei $f : (\Omega, \mathfrak{A}, \mu) \to (\overline{\mathbb{R}^1}, \mathfrak{L}(\overline{\mathbb{R}^1}))$ nichtnegativ und $(f_n)_{n \in \mathbb{N}}$ eine Folge elementarer Funktionen mit $\sup_{n \in \mathbb{N}} f_n = f$. Dann heißt

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

Def. Eine $\mathfrak{A}$-messbare, numerische Fkt. $f : (\Omega, \mathfrak{A}, \mu) \to (\overline{\mathbb{R}^1}, \mathfrak{L}(\overline{\mathbb{R}^1}))$ 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!

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

$$L^p(\Omega, \mathfrak{A}, \mu) := \{f : \Omega \to \overline{\mathbb{R}^1} \mid f \text{ } p\text{-integrierbar, also } \int_\Omega |f|^p \, d\mu < \infty\},$$

$$L^\infty(\Omega, \mathfrak{A}, \mu) := \{f : \Omega \to \overline{\mathbb{R}^1} \mid \exists C > 0 : |f| \leq C \text{ fast überall}\}$$

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 $\|{-}\|_p$ ist auch konvergent.

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

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

Bem. 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

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

Satz (von der monotonen Konvergenz)**.** Sei für alle $n \in \mathbb{N}$ die Funktion $f_n : (\Omega, \mathfrak{A}, \mu) \to (\mathbb{R}^1, \mathfrak{L}(\mathbb{R}^1))$ nicht negativ und $\mathfrak{A}$-messbar, sodass $(f_n)_{n \in \mathbb{N}}$ eine isotone Folge ist. Dann gilt

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

Kor (Beppo Levi)**.** Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge nicht negativer, $\mathfrak{A}$-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. Sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum und $f : (\Omega, \mathfrak{A}) \to (\overline{\mathbb{R}^1}, \mathcal{L}(\overline{\mathbb{R}^1}))$ nichtnegativ und μ-integrierbar. Dann definiert

$$\nu : \mathfrak{A} \to [0, \infty], \quad A \mapsto \int_A f \, d\mu = \int_\Omega f \cdot \chi_A \, d\mu$$

ein zu μ absolut stetiges, endliches Maß auf $(\Omega, \mathfrak{A})$.

Lem (Fatou)**.** Sei für $n \in \mathbb{N}$ die Fkt. $f_n : (\Omega, \mathfrak{A}, \mu) \to (\mathbb{R}^1, \mathfrak{L}(\mathbb{R}^1))$ nicht negativ und $\mathfrak{A}$-messbar. Dann gilt

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

Falls $\int_\Omega \sup_{n \in \mathbb{N}} f_n \, d\mu < \infty$, gilt zusätzlich

$$\int_\Omega \limsup_{n \to \infty} f_n \, d\mu \geq \limsup_{n \to \infty} \int_\Omega f_n \, d\mu.$$

Def. Eine Folge $(f_n)_{n \in \mathbb{N}}$ $\mathfrak{A}$-messbarer, numerischer Fktn. über $(\Omega, \mathfrak{A}, \mu)$ **konvergiert μ-fast-überall** gegen $f : \Omega \to \mathfrak{A}$, falls

$$\lim_{n \to \infty} f_n(\omega) = f(\omega) \quad \text{für } \mu\text{-fast-alle } \omega \in \Omega \text{ gilt.}$$

Satz (Riesz)**.** Sei $(f_n)_{n \in \mathbb{N}}$ Folge in $L^p(\Omega, \mathfrak{A}, \mu)$, $f_n \xrightarrow[n \to \infty]{\text{f.ü.}} f$ mit $f \in L^p(\mu)$. Dann gilt

$$f_n \xrightarrow[n \to \infty]{L^p(\mu)} f \iff \int_\Omega |f_n|^p \, d\mu \xrightarrow{n \to \infty} \int_\Omega |f|^p \, d\mu.$$

Satz (von der majorisierten Konvergenz)**.** Sei $(f_n)_{n \in \mathbb{N}}$ Folge $\mathfrak{A}$-messbarer numerischer Funktionen auf $(\Omega, \mathfrak{A}, \mu)$ und $g \in L^1(\mu)$ nicht negativ, sodass $|f_n| \leq g$ für alle $n \in \mathbb{N}$. Sei desweiteren $f : \Omega \to \overline{\mathbb{R}^1}$ messbar mit $f_n \xrightarrow[n \to \infty]{\mu\text{-f.ü.}} f$. Dann ist

$$f \in L^1(\mu) \quad \text{mit} \quad \int_\Omega f \, d\mu = \lim_{n \to \infty} \int_\Omega f_n \, d\mu.$$

Satz. Sei $f : (\Omega, \mathfrak{A}, \mu) \to (\Omega', \mathfrak{A}')$ und $\mu' := \mu \circ f^{-1}$ das Bildmaß von μ unter f. Sei $g : (\Omega', \mathfrak{A}') \to (\mathbb{R}^1, \mathfrak{B}(\mathbb{R}^1))$ nicht negativ. Dann gilt

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

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 $f : \widetilde{U} \to \overline{\mathbb{R}^1}$ genau dann auf $\widetilde{U}$ Lebesgue-Borel-integrierbar, wenn $(f \circ \phi) \cdot |\det(D\phi)| : U \to \overline{\mathbb{R}^1}$ auf U Lebesgue-Borel-interierbar ist. In diesem Fall gilt

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

Obige Gleichung ist auch erfüllt, wenn lediglich $f \geq 0$ gilt.

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

$$\mathbb{E}X := \int_\Omega X \,\mathrm{d}\mathbb{P} \qquad \textbf{Erwartungswert} \text{ (EW) von } X.$$

Bem. Für eine konstante ZGe $X : \Omega \to \mathbb{R}$, also $\forall\, \omega \in \Omega : X(\omega) = x$, gilt $\mathbb{E}X = x$.

Satz. Der Erwartungswert ist linear, d. h. für ZGn X und Y und $\lambda \in \mathbb{R}$ gilt $\mathbb{E}(\lambda X + Y) = \lambda \mathbb{E}X + \mathbb{E}Y$.

Satz. $\mathbb{E}X = \int_{\mathbb{R}^1} \mathrm{id}\,\mathrm{d}P_X$, wobei $P_X := \mathbb{P} \circ X^{-1}$.

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

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

wobei rechts ein uneigentliches Riemann-Stieltjes-Integral steht.

Def. Für Zufallsvektoren $X = (X_1, ..., X_k)$ mit Werten in $\mathbb{R}^k$ definieren wir $\mathbb{E}X := (\mathbb{E}X_1, ..., \mathbb{E}X_k)$.

Bem. Sei $X = (X_1, ..., X_k)$ ein Zufallsvektor und $g : \mathbb{R}^k \to \mathbb{R}$ Borel-messbar und P_X-integrierbar. Dann ist

$$\mathbb{E}g(X_1, ..., X_k) = \int_{\mathbb{R}^k} g \,\mathrm{d}P_X.$$

Satz. Sei F_X VF einer ZG $X : (\Omega, \mathfrak{A}, \mathbb{P}) \to (\mathbb{R}^1, \mathfrak{L}(\mathbb{R}^1))$. Dann existiert für Lebesgue-fast-alle $x \in \mathbb{R}^1$ die Ableitung $F'(x)$.

Def. Sei F_X VF einer ZG $X : (\Omega, \mathfrak{A}, \mathbb{P}) \to (\mathbb{R}^1, \mathfrak{L}(\mathbb{R}^1))$.

- F_X heißt **diskret**, falls F_X höchstens abzählbar viele Sprungstellen $x_1, x_2, ... \in \mathbb{R}$ besitzt mit

$$\forall\, j \in J \subset \mathbb{N} : p_j := F_X(x_j) - \lim_{x \uparrow x_j} F_X(x) > 0, \qquad \sum_{j=1}^{\infty} p_j = 1.$$

  Dann ist F_X zwischen den Sprüngen konstant.
- F_X heißt **stetig** (diffus, atomlos), wenn F_X in jedem Punkt stetig ist. Dann gilt $P_X(\{X = x\}) = 0$ für alle $x \in \mathbb{R}$.

Für stetige Verteilungen ergibt sich eine weitere Unterteilung:

- F_X heißt **absolutstetig** (totalstetig), wenn es für alle $\epsilon > 0$ ein $\delta > 0$ gibt, sodass für höchstens abzählbar viele, disjunkte Intervalle $I_k = (a_k, b_k]$ mit $k \in J \subset \mathbb{N}$ gilt:

$$\sum_{k \in J} (b_k - a_k) < \delta \implies \sum_{k \in J} (F_X(b_k) - F_X(a_k)) < \epsilon.$$

- F_X heißt **singulärstetig** (stetig, aber nicht absolutstetig), wenn die Wachstumspunkte von F_X eine Nullmenge bilden, also

$$\lambda_1(\{x \in \mathbb{R}^1 \,|\, \forall\, \epsilon > 0 : F(x + \epsilon) - F(x - \epsilon) > 0\}) = 0$$

  oder äquivalent dazu die Ableitung fast-überall verschwindet, also

$$\lambda_1(\{x \in \mathbb{R}^1 \,|\, F_X'(x) = 0\}) = 1.$$

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

$$F = \alpha_d F_d + \alpha_s F_s + \alpha_a F_a \quad \text{mit} \quad \alpha_d, \alpha_s, \alpha_a \geq 0,\ \alpha_d + \alpha_s + \alpha_a = 1.$$

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

$$f = f_X = F_X' : \mathbb{R} \to \mathbb{R}_{\geq 0}, \quad x \mapsto \begin{cases} F_X'(x), & \text{falls Ableitung ex.} \\ 0, & \text{sonst} \end{cases}$$

(Wahrscheinlichkeits-)Dichte (WD) von F_X bzw. von X.

Bem. Dann gilt für alle $y \in \mathbb{R}$:

$$\int_{-\infty}^{y} f_X(x) \,\mathrm{d}x = F_X(y), \quad \text{also insbesondere} \quad \int_{-\infty}^{\infty} f_X(x) \,\mathrm{d}x = 1.$$

Bem. F_X ist als VF genau dann absolut stetig, wenn das Maß P_X bezüglich λ_1 absolut stetig ist (also $P_Y \ll \lambda_1$ gilt).

Satz.
$$\mathbb{E}X = \begin{cases} \int_{\mathbb{R}^1} x \cdot f_X(x) \,\mathrm{d}x, & \text{falls } F_X \text{ absolutstetig mit WD } f_X \\ \sum_{j \in J} x_j \cdot p_j, & \text{falls } F_X \text{ diskret mit Sprüngen } p_j \text{ bei } x_j,\ j \in J \subset \mathbb{N} \end{cases}$$

Satz (Erwartungswerte bekannter Zufallsverteilungen)**.**

- Für $X \sim \mathrm{Poi}(\lambda)$: $\mathbb{E}X = \lambda$
- Für $N \sim \mathrm{N}(\mu, \sigma^2)$: $\mathbb{E}X = \mu$
- Für $X \sim \mathrm{Exp}(\lambda)$: $\mathbb{E}X = \frac{1}{\lambda}$
- Für $N \sim \mathrm{N}(0, 1)$: $\mathbb{E}|X| = \sqrt{\frac{2}{\pi}}$

Bem. Die Cauchy-Verteilung hat die VF bzw. die W-Dichte

$$F(x) = \tfrac{1}{\pi} \arctan(x) + \tfrac{1}{2}, \qquad f(x) = \tfrac{1}{\pi(1+x^2)}.$$

Eine Cauchy-verteilte ZG X hat keinen EW, da $\int_{\mathbb{R}^1} |x| \cdot f(x) \,\mathrm{d}x = \infty$.

Def. Seien $X_1, ..., X_d : (\Omega, \mathfrak{A}, \mathbb{P}) \to (\mathbb{R}^1, \mathfrak{B}(\mathbb{R}^1))$ ZGn,

$$F = F_{(X_1, ..., X_d)} : \mathbb{R}^d \to [0, 1], \quad (x_1, ..., x_d) \mapsto \mathbb{P}(X_1 \leq x_1, ..., X_k \leq x_k)$$

die dazugehörige VF und $P = P_{(X_1, ..., X_d)}$ das von der VF induzierte Maß auf $\mathfrak{B}(\mathbb{R}^d)$.

- F heißt **diskret**, falls es eine höchstens abzählbare Menge $\{y_i \in \mathbb{R}^d \,|\, i \in I\}$ mit $I \subset \mathbb{N}$ gibt, sodass

$$\forall\, i \in I : P(\{y_i\}) > 0 \quad \text{und} \quad \sum_{i \in I} P(\{y_i\}) = 0.$$

- F heißt **stetig**, wenn $P(\{y\}) = 0$ für alle $y \in \mathbb{R}^1$.
- F heißt **absolut stetig**, falls das Maß P absolut stetig bzgl. dem Lebesgue-Maß ist, also $P \ll \lambda_d$ gilt. Dazu äquivalent: Für alle $\epsilon > 0$ gibt es ein $\delta > 0$, sodass für höchstens abzählbar viele, disjunkte Elementarquader $Q_j = (a_j, b_j]$ mit $j \in J \subset \mathbb{N}$ gilt:

$$\sum_{j \in J} \lambda_d(Q_j) \leq \delta \implies \sum_{j \in J} P(Q_j) = \sum_{j \geq J} (\triangle F) Q_j \leq \epsilon.$$

- F heißt **singulär stetig**, wenn F stetig ist und eine Lebesgue-Menge S mit $\lambda_d(S) = 0$ und $P(S) = 1$ existiert.

Bem. Falls F absolut stetig, ex. die W-Dichte, die f. ü. durch

$$f = f_{(X_1, ..., X_d)} : \mathbb{R}^d \to \mathbb{R}_{\geq 0}, \quad (x_1, ..., x_d) \mapsto \frac{\partial^d}{\partial x_1 \cdots \partial x_d} F(x_1, ..., x_d)$$

gegeben ist und $\int_{(-\infty, x]} f(y) \,\mathrm{d}y = F(x)$ für alle $x \in \mathbb{R}^d$ erfüllt.

Satz. Sei X eine ZG und $g : \mathbb{R}^1 \to \mathbb{R}^1$ messbar. Dann gilt

$$\mathbb{E}g(X) = \begin{cases} \int_{\mathbb{R}^1} g(x) \cdot f_X(x) \,\mathrm{d}x, & \text{falls } F_X \text{ absolutstetig mit WD } f_X \\ \sum_{j \in J} g(x_j) \cdot p_j, & \text{falls } F_X \text{ diskret mit Sprüngen } j_k \text{ bei } x_j,\ j \in J \subset \mathbb{N} \text{ (und wohldefiniert)} \end{cases}$$

Def. Eine **Zerlegung** eines Intervalls $[a, b]$ ist eine geordnete endliche Menge $Z = \{a = x_0 < x_1 < ... < x_k = b\} \subset [a, b]$. Eine weitere Zerlegung $\widetilde{Z}$ desselben Intervalls heißt **Verfeinerung** von Z, falls $\widetilde{Z} \supset Z$.

Notation. Die Menge aller Zerlegungen von $[a, b]$ ist $\mathcal{Z}([a, b])$.

Def. Eine Menge von Stützstellen bzgl. einer Zerlegung $\{x_0 < ... < x_k\}$ von $[a, b]$ ist eine Menge $\{\xi_1, ..., \xi_k\}$ mit $\xi_i \in (x_{i-1}, x_i)$ für $i \in \{1, ..., k\}$.

Def. Für zwei Funktionen $f, g : [a, b] \to \mathbb{R}$, eine Zerlegung $Z = \{a = x_0 < ... < x_n = b\}$ von $[a, b]$ und Stützstellen $\xi_1, ..., \xi_n$ bzgl. Z heißt die Summe

$$S(f, \mathrm{d}g, Z, \xi_1, ..., \xi_n) := \sum_{j=1}^{n} f(\xi_j)(g(x_j) - g(x_{j-1}))$$

Riemann-Stieltjes-Summe von f bzgl. g und der Zerlegung Z mit Stützstellen $\xi_1, ..., \xi_n$.

Def. Seien $f, G : [a, b] \to \mathbb{R}$. Die Funktion f heißt **Riemann-Stieltjes-integrierbar** (RS-integrierbar) bzgl. der **Gewichtsfunktion** G, wenn gilt: Es gibt ein $\iota \in \mathbb{R}$, sodass für alle $\epsilon > 0$ eine Zerlegung Z_ϵ von $[a, b]$ existiert, sodass für alle Verfeinerungen $Z \supset Z_\epsilon$ und Wahlen von Stützstellen $\xi_1, ..., \xi_n$ gilt:

$$|\iota - S(f, \mathrm{d}G, Z, \xi_1, ..., \xi_n)| \leq \epsilon.$$

Dieses (eindeutig bestimmte) ι heißt **Riemann-Stieltjes-Integral** (RS-Integral) von f bzgl. G, geschrieben

$$\int_a^b f(x) \,\mathrm{d}G(x) := \iota.$$

Bem. Mit $G :=$ id erhalten wir aus dem RS-Integral das gewöhnliche Riemann-Integral.

Satz. Das RS-Integral ist sowohl in der integrierten Funktion als auch der Gewichtsfunktion linear.

Satz. Für f bzgl. G auf $[a,b]$ RS-int'bar und G stetig diff'bar gilt

$$\int_a^b f(x)\,\mathrm{d}G(x) = \int_a^b f(x)\cdot G'(x)\,\mathrm{d}x.$$

Satz (Partielle Integration)**.** Sei $f:[a,b]\to\mathbb{R}$ bzgl. $G:[a,b]\to\mathbb{R}$ RS-integrierbar. Dann ist auch G bzgl. f RS-integrierbar und es gilt

$$\int_a^b f(x)\,\mathrm{d}G(x) = [G(x)\cdot f(x)]_a^b - \int_a^b G(x)\,\mathrm{d}f(x).$$

Bem. Wir können uneigentliche RS-Integrale analog zu uneigentlichen Riemann-Integralen definieren.

Satz. Sei $\mathbb{E}|X| = \int_0^\infty 1 - F_X(x) + F_X(-x)\,\mathrm{d}x < \infty$ und $\lim_{x\to\infty} x\cdot F_X(-x) = \lim_{x\to\infty} x\cdot(1-F_X(x)) = 0$. Dann gilt

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

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

Def. Sei X eine ZG über $(\Omega,\mathfrak{A},\mathbb{P})$ und $k\in\mathbb{N}$. Dann heißt

- $\mathbb{E}X^k$ **k-tes Moment**,
- $\mathbb{E}|X|^k$ **k-tes absolutes Moment**,
- $\mathbb{E}(X-\mathbb{E}X)^k$ **k-tes zentriertes Moment**,
- $\mathbb{E}|X-\mathbb{E}X|^k$ **k-tes zentriertes absolutes Moment**,
- $\operatorname{Var}(X) := \mathbb{D}^2X := \mathbb{E}(X-\mathbb{E}X)^2$ **Varianz** (Dispersion, Streuung),
- $\sqrt{\operatorname{Var}(X)}$ **Standardabweichung** von X.

Lem. Es gilt $\operatorname{Var}(X) = \mathbb{E}(X-\mathbb{E}X)^2 = \mathbb{E}X^2 - (\mathbb{E}X)^2 \geq 0$.

Kor. $(\mathbb{E}X)^2 \leq \mathbb{E}(X^2)$

Lem. $\mathbb{E}X^k = \int_\Omega X^n\,\mathrm{d}\mathbb{P} = \int_{\mathbb{R}^1} \mathrm{id}^n\,\mathrm{d}P_X = \int_{-\infty}^\infty x^n\,\mathrm{d}F_X =$

$$= \begin{cases} \int_{\mathbb{R}^1} x^k\cdot f_X(x)\,\mathrm{d}x, & \text{falls } F_X \text{ absolutstetig mit WD } f_X \\ \sum_{j\in J} x_j^k\cdot p_j, & \text{falls } F_X \text{ diskret mit Sprüngen } p_j \\ & \text{bei } x_j,\ j\in J\subset\mathbb{N} \end{cases}$$

Bem. Falls $\mathbb{E}|X|^k<\infty$, dann existiert auch $\mathbb{E}X^k$.

Lem. Es gilt für eine ZG X und $a,b,c\in\mathbb{R}^1$:
- $\operatorname{Var}(aX+b) = a^2\operatorname{Var}(X)$ • $\operatorname{Var}(X)\leq\mathbb{E}(X-c)^2$
- $\operatorname{Var}(X)=0 \iff \mathbb{E}(X-\mathbb{E}X)=0 \iff X\equiv\text{const }\mathbb{P}\text{-fast-sicher}$

Satz (Verallgemeinerte Tschebyschow-Ungleichung)**.** Sei X sei eine ZG und $g:[0,\infty)\to[0,\infty)$ nicht fallend. Dann gilt für alle $\epsilon>0$:

$$\mathbb{P}(|X|\geq\epsilon)\leq\frac{\mathbb{E}g(|X|)}{g(\epsilon)}$$

Kor.
- **Markow-Ungleichung**: $\mathbb{P}(|X|\geq\epsilon)\leq\frac{\mathbb{E}|X|}{\epsilon}$.
- **Tschebyschow-Ungleichung**: $\mathbb{P}(|X-\mathbb{E}X|\geq\epsilon)\leq\frac{\operatorname{Var}(X)}{\epsilon^2}$.
- Für alle $a>0$ gilt $\mathbb{P}(|X|\geq\epsilon)\leq\frac{\mathbb{E}\exp(a|X|)}{\exp(a\epsilon)}$.

Def. Die Abbildung $t\mapsto\mathbb{E}\exp(tX) = \sum_{n=0}^\infty\frac{t^n}{n!}\mathbb{E}X^n$ heißt momenterzeugende Funktion der ZG X oder VF F_X.

Bsp. Für $X\sim N(\mu,\sigma^2)$ gilt $\mathbb{E}\exp(zX) = \exp\left(z\mu+\frac{\sigma^2}{2}z^2\right)$.

Satz. Für $p,q\in(1,\infty)$ mit $\frac{1}{p}+\frac{1}{q}=1$ und ZGen X,Y gilt

$$|\mathbb{E}XY|\leq\mathbb{E}|XY|\leq(\mathbb{E}|X|^p)^{\frac{1}{p}}\cdot(\mathbb{E}|X|^q)^{\frac{1}{q}}\quad\textbf{(Hölder-Ungl).}$$

Kor. **Cauchy-Schwarz-Ungl**: $|\mathbb{E}XY|\leq\sqrt{\mathbb{E}(X^2)\cdot\mathbb{E}(Y^2)}$

Def. Eine Funktion $g:J\to\mathbb{R}$ heißt **konvex**, falls gilt:

$$\forall\,x,y\in J\ :\ \forall\,t\in[0,1]\ :\ g(ta+(1-t)b)\leq tg(a)+(1-t)g(b)$$

Satz. Sei $g:\mathbb{R}^1\to\mathbb{R}^1$ konvex auf einem Intervall J und X eine ZG mit $\mathbb{P}(X\in J)=1$ und $\mathbb{P}|X|<\infty$. Dann gilt für $x,y\in I$, $t\in[0,1]$:

$$g(\mathbb{E}X)\leq\mathbb{E}g(X)\quad(\textbf{Jensen-Ungleichung}).$$

Kor. **Ljapunow-Ungleichung**: $|\mathbb{E}X|^{\frac{n}{m}}\leq\mathbb{E}|X|^{\frac{n}{m}}$

Frage (Momentenproblem (MP))**.** Unter welchen Bedingungen ist eine Folge $(c_i)_{i\in\mathbb{N}}$ eine Momentenfolge einer ZG X d. h. $c_i=\mathbb{E}X^i$?

Antwort. Genau dann, wenn

$$\forall\,n\in\mathbb{N}\ :\ \det\begin{pmatrix} c_0 & c_1 & \cdots & c_n \\ c_1 & c_2 & \cdots & c_{n+1} \\ \vdots & \vdots & & \vdots \\ c_n & c_{n+1} & \cdots & c_{2n}\end{pmatrix}\geq 0.$$

Frage. Wann ist die zugehörige VF F_X eindeutig festgelegt?

Bem. Dabei unterscheiden wir folgende Momentenprobleme:
- Stieltjes: $c_n = \int_0^\infty x^n\,\mathrm{d}F_X(x)$ • Hamburger: $c_n = \int_{-\infty}^\infty x^n\,\mathrm{d}F_X(x)$

Antwort. Hinreichende Bedingung für Bestimmtheit (Carleman):
- Stieltjes: $\sum_{n=1}^\infty c_n^{-\frac{1}{2n}}=\infty$ • Hamburger: $\sum_{n=1}^\infty c_{2n}^{-\frac{1}{2n}}=\infty$

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

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

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

$$\mathbb{P}(X_1\leq x_1,\dots,X_k\leq x_k) = \mathbb{P}(X_1\leq x_1)\cdot\ldots\cdot\mathbb{P}(X_k\leq x_k)$$

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

$$\forall\,x_1,\dots,x_k{\in}\mathbb{R}\ :\ f_X(x_1,\dots,x_k) = f_{X_1}(x_1)\cdot\ldots\cdot f_{X_k}(x_k).$$

Für diskrete Verteilungen ist dies äquivalent zu

$$\forall\,x_1,\dots,x_k{\in}\mathbb{R}\ :\ \mathbb{P}(X_1{=}x_1,\dots,X_k{=}x_k) = \mathbb{P}(X_1{=}x_1)\cdot\ldots\cdot\mathbb{P}(X_k{=}x_k).$$

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

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

mit $l\in\{1,\dots,k{-}1\}$ und $1\leq i_1<\dots<i_l\leq k$ **l-dimensionale Rand-(Marginal-)verteilungsfunktion**.

Bem. Falls $f_X(x_1,\dots,x_k)$ ex., so existieren alle Randdichten

$$f_{(X_{i_1},\dots,X_{i_l})}(x_{i_1},\dots,x_{i_l}) = \int_{\mathbb{R}^{k-l}} f_X(x_1,\dots,x_k)\,\mathrm{d}x_{i_1}\cdots\mathrm{d}x_{i_l}.$$

Für $k{=}2$ und ZV diskret mit Masseschwerpunkten $x_m=(x_m^1,x_m^2)$:

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

Bem. Im Allgemeinen bestimmen die Randverteilungen nicht die gemeinsame Verteilung des Vektors.

Def. (X,Y) sei eine zweidimensionale ZV über $(\Omega,\mathfrak{A},\mathbb{P})$ mit $\mathbb{E}X^2<\infty$ und $\mathbb{E}Y^2<\infty$. Dann heißen

$$\operatorname{Cov}(X,Y) := \mathbb{E}(X\cdot Y)-\mathbb{E}X\cdot\mathbb{E}Y = \mathbb{E}((X-\mathbb{E}X)\cdot(Y-\mathbb{E}Y))$$

$$\operatorname{Cor}(X,Y) := \frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)\cdot\operatorname{Var}(Y)}}$$

Kovarianz bzw. **Korrelation** von X und Y.

Satz.
- Falls X,Y unabhängig, so gilt $\operatorname{Cov}(X,Y)=\operatorname{Cor}(X,Y)=0$
- $|\operatorname{Cor}(X,Y)|\leq 1$ • $\operatorname{Cor}(X,Y){=}1 \iff \exists\,a,b\in\mathbb{R}:\mathbb{P}(Y{=}aX{+}b){=}1$

Def. Falls $\operatorname{Cor}(X,Y)=0$, so heißen X,Y **unkorreliert**.

Sprechweise.
- $\operatorname{Cor}(X,Y)\approx 1$: positive Korrelation
- $\operatorname{Cor}(X,Y)\approx -1$: negative Korrelation

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

Bsp. Sei X eine ZG mit der symmetrischen Dichte

$$f_X(x)=f_X(-x)\quad\text{und}\quad\int_{-\infty}^\infty|x|^3f_X(x)\,\mathrm{d}x<\infty,$$

dann ist $\operatorname{Cov}(X,X^2)=0$, aber X und X^2 nicht unabhängig.

Bem. Falls (X,Y) eine zweidimensionale Normalverteilung besitzt, so folgt aus $\operatorname{Cor}(X,Y)=0$ die Unabhängigkeit von X und Y.

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

$$\operatorname{Var}(X_1+\dots+X_n) = \operatorname{Var}(X_1)+\dots+\operatorname{Var}(X_n)$$

Bem. Seien X und Y ZGen. Gesucht: $\mathbb{E}(X\cdot Y)$

- Angenommen, es existiert eine gemeinsame WD $f_{(X,Y)}$. Dann:
$$\mathbb{E}(XY) = \int_{\mathbb{R}^2} xy \cdot f_{(X,Y)}(x,y)\,\mathrm{d}(x,y)$$
$$\mathbb{E}X = \int_{-\infty}^{\infty} x \left(\int_{-\infty}^{\infty} f_{(X,Y)}(x,y)\,\mathrm{d}y \right) \mathrm{d}x$$
- Angenommen, (X,Y) hat Werte in einer höchstens abzählbaren Menge $(x_i, y_j)_{(i,j)\in IJ}$ mit $IJ \subset \mathbb{N}{\times}\mathbb{N}$, die mit Wahrscheinlichkeit $p_{ij} > 0$ angenommen werden. Dann gilt
$$\mathbb{E}(XY) = \sum_{(i,j)\in IJ} x_i y_j p_{ij}$$
$$\mathbb{E}X = \sum_{i\in\pi_1(IJ)} x_i \cdot \mathbb{P}(X{=}x_i) = \sum_{(i,j)\in IJ} x_i p_{ij}$$
$$\mathrm{Cov}(X,Y) = \sum_{(i,j)\in IJ} x_i y_j \left(p_{ij} - \mathbb{P}(X{=}x_i)\cdot\mathbb{P}(Y{=}y_j)\right)$$
- Angenommen, (X,Y) ist singulär-stetig verteilt oder besitzt eine singulärstetige Komponente. Dann gilt:
$$\mathbb{E}(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\cdot y\,\mathrm{d}F(x,y)$$
$$\mathbb{E}(XY) = \int_0^{\infty}\int_0^{\infty} \mathbb{P}(X > x, Y > y)\,\mathrm{d}x\,\mathrm{d}y \quad \text{für } X, Y \geq 0.$$

Satz. Sei $Y := g(X)$, wobei $g : \mathbb{R}^1 \to \mathbb{R}^1$ mb und X eine ZG. Dann:
$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(X \in g^{-1}((-\infty, y])).$$

Satz. X sei absolut stetig mit Dichte f_X und $\mathbb{P}(X \in D) = 1$ für $D \subset \mathbb{R}^1$ offen und $g : D \to \mathbb{R}^1$ eine $\mathcal{C}^1$-Funktion mit $g'(x) > 0$ für alle $x \in D$. Dann ist $Y := g(X)$ absolut-stetig mit der Dichte
$$f_Y(y) = \begin{cases} 0, & \text{für } y \in \mathbb{R}^1 \setminus g(D) \\ \frac{f_X(g^{-1}(y))}{|g'(g^{-1}(y))|} & \text{für } y \in g(D) \end{cases}$$

Bem. Falls die Ableitung $g' : D \to \mathbb{R}^1$ wechselndes Vorzeichen besitzt, so muss in Monotoniebereiche unterteilt werden.

Satz. Sei $X = (X_1, ..., X_k)$ ein k-dimensionaler Zufallsvektor mit WD f_X, $G, H \Subset \mathbb{R}^k$ offen und $g = (g_1, ..., g_k) : G \to H$ ein $\mathcal{C}^1$-Diffeo. Dann findet man Funktionen $h_i : H \to \mathbb{R}^1$, $i = 1, ..., k$ mit $h(y_1, ..., y_k) = (h_1(y_1, ..., y_k), ..., h_k(y_1, ..., y_k))$ für $(y_1, ..., y_k) \in H$, sodass für die Dichte von $Y = (Y_1, ..., Y_k)$ gilt:
$$f_Y(y_1, ..., y_k) = \begin{cases} 0, & \text{für } (y_1, ..., y_k) \notin H, \text{ ansonsten} \\ \frac{f_X(h_1(y_1,...,y_k),...,h_k(y_1,...,y_k))}{|\det Dg(h_1(y_1,...,y_k),...,h_k(y_1,...,y_k))|} \end{cases}$$

Satz. (X,Y) besitze eine gemeinsame Dichte $f_{(X,Y)}$. Dann gilt für die Dichten von $Z_1 := X+Y$, $Z_2 := X \cdot Y$, $Z_3 := \frac{X}{Y}$:
$$f_{Z_1}(z) = \int_{-\infty}^{\infty} f_{(X,Y)}(z-y, y)\,\mathrm{d}y \overset{\text{Unabh.}}{=} \int_{-\infty}^{\infty} f_X(z-y)\cdot f_Y(y)\,\mathrm{d}y$$
$$f_{Z_2}(z) = \int_{-\infty}^{\infty} f_{(X,Y)}(\tfrac{z}{y}, y)\cdot\tfrac{1}{|y|}\,\mathrm{d}y \overset{\text{Unabh.}}{=} \int_{-\infty}^{\infty} f_X(\tfrac{z}{y})\cdot f_Y(y)\cdot\tfrac{1}{|y|}\,\mathrm{d}y$$
$$f_{Z_3}(z) = \int_{-\infty}^{\infty} f_{(X,Y)}(zy, y)\cdot|y|\,\mathrm{d}y \overset{\text{Unabh.}}{=} \int_{-\infty}^{\infty} f_X(zy)\cdot f_Y(y)\cdot|y|\,\mathrm{d}y$$

Bsp. Seien X, Y unabhängig, $N(0,1)$-verteilt, $Z := \frac{X}{Y}$. Dann:
$$f_Z(z) = ... = \tfrac{1}{\pi(1+z^2)} \quad \text{(Cauchy-verteilt).}$$

Def. Seien X und Y unabhängige ZGen. Dann gilt für $z \in \mathbb{R}$:
$$F_{X+Y}(z) = \mathbb{P}(X{+}Y \leq z) = \int_{\mathbb{R}^1} F_X(z-y)\,\mathrm{d}F_Y(y) =: (F_X * F_Y)(z).$$
Dabei heißt $(F_X * F_Y)$ **Faltung** von F_X und F_Y.

Prop.
- $F_X * F_Y = F_Y * F_X$ (Kommutativität)
- $F_{X+Y} = F_X * F_Y$, wenn X und Y unabhängig
- Falls X oder Y eine Dichte besitzen, so auch $X{+}Y$ bei Unabh.:
$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(z-y)\,\mathrm{d}F_Y(y) \overset{\text{oder}}{=} \int_{-\infty}^{\infty} f_Y(z-x)\,\mathrm{d}F_X(x)$$
$$\overset{\text{oder}}{=} \int_{-\infty}^{\infty} f_Y(z-x)\,\mathrm{d}f_X(x)y$$

Satz.
- Seien X_1 und X_2 unabhängig und $(X_1{+}X_2)$ normalverteilt (also $X_1 + X_2 \sim N(\mu, \sigma^2)$). Dann sind X_1 und X_2 normalverteilt.
- Seien X_1 und X_2 unabhängig und $(X_1{+}X_2)$ Poisson-verteilt (also $X_1 + X_2 \sim Poi(\lambda)$). Dann sind auch X_1 und X_2 Poisson-verteilt.

Satz. Seien $X_1, ..., X_n$ unabhängig. Dann gilt:
- $\frac{1}{n}(X_1 + ... + X_n) \xrightarrow[\mathbb{P}\text{-f.ü.}]{n\to\infty} \mathbb{E}X_1$, falls $X_1, ..., X_n$ identisch verteilt.
- $\sqrt{n}(\frac{1}{n}(X_1 + ... + X_n) - \mathbb{E}X_1) \xrightarrow{n\to\infty} N(0, \mathrm{Var}(X_1))$, falls $\mathrm{Var}(X_1) < \infty$

Produkt-Maße und der Satz von Fubini-Tonelli

Voraussetzung. Seien $(\Omega_i, \mathfrak{A}_i)$ für $i \in \{1, ..., n\}$ messbare Räume und $\Omega := \Omega_1 \times ... \times \Omega_n$ das kartesische Produkt der Mengen.

Notation. $\pi_i : \Omega \to \Omega_i$, $(\omega_1, ..., \omega_1) \mapsto \omega_i$ (Projektionsabbildung)

Def. Die **Produkt-σ-Algebra** von $\mathfrak{A}_1, ..., \mathfrak{A}_n$ ist definiert durch
$$\mathfrak{A} := \mathfrak{A}_1 \otimes ... \otimes \mathfrak{A}_n := \sigma\left(\pi_1^{-1}(\mathfrak{A}_1) \cup ... \cup \pi_n^{-1}(\mathfrak{A}_n)\right).$$

Satz. $\mathfrak{A} = \sigma(\{A_1 \times ... \times A_n \mid A_1 \in \mathfrak{A}_1, ..., A_n \in \mathfrak{A}_n\})$

Lem. Sei $\mathcal{E}_i$ mit $\sigma(E_i) = \mathfrak{A}_i$ und es existieren Folgen $(E_i^k)_{k\in\mathbb{N}}$ in $\mathcal{E}_i$ mit $E_i^k \uparrow \Omega_i$. Dann gilt:
$$\mathfrak{A}_1 \otimes ... \otimes \mathfrak{A}_n = \sigma(\{E_1 \times ... \times E_n \mid E_1 \in \mathcal{E}_1, ..., E_n \in \mathcal{E}_n\})$$

Bem. Auf die Zusatzvoraussetzung der monoton aufsteigenden Mengenfolge können wir nicht verzichten.

Satz. $\mathcal{L}(\mathbb{R}^n) = \mathcal{L}(\mathbb{R}^1) \otimes ... \otimes \mathcal{L}(\mathbb{R}^1)$

Lem. $(\widetilde{\Omega}', \widetilde{\mathfrak{A}})$ sei ein messbarer Raum. Dann gilt: $f : \widetilde{\Omega} \to \Omega_1 \times ... \times \Omega_n$ genau dann $(\widetilde{\mathfrak{A}}, \mathfrak{A})$-messbar, wenn für alle $i \in \{1, ..., n\}$ die Abbildung $f_i := \pi_i \circ f : \widetilde{\Omega} \to \Omega_i$ $(\widetilde{\mathfrak{A}}, \mathfrak{A}_i)$-messbar ist.

Satz. Sei μ_1 auf $(\Omega_1, \mathfrak{A}_1)$ und μ_2 auf $(\Omega_2, \mathfrak{A}_2)$ zwei σ-endliche Maße. Dann ex. genau ein Maß $\mu := \mu_1 \times \mu_2$ auf $\mathfrak{A} = \mathfrak{A}_1 \otimes \mathfrak{A}_2$ mit
$$\forall A_1 \in \mathfrak{A}_1, A_2 \in \mathfrak{A}_2 : (\mu_1 \times \mu_2)(A_1 \times A_2) = \mu_1(A_1)\cdot\mu_2(A_2).$$
Das Maß μ ist dann auch σ-endlich auf $(\Omega, \mathfrak{A})$.

Kor. Das Produkt-W-Maß $\mathbb{P} = \mathbb{P}_1 \times \mathbb{P}_2$ ist das einzige W-Maß auf $\mathfrak{A}_1 \otimes \mathfrak{A}_2$ mit der Eigenschaft:
- $\mathbb{P}(A_1 \times \Omega_2) = \mathbb{P}_1(A_1)$ und $\mathbb{P}(\Omega_1 \times A_2) = \mathbb{P}_2(A_2)$
- $\mathbb{P}(B_1 \cap B_2) = \mathbb{P}(B_1) \times \mathbb{P}(B_2)$ für $B_1 = A_1 \times \Omega_2$, $B_2 = \Omega_1 \times A_2$

Kor. Für $k, l \in \mathbb{N}_{>0}$ mit $k + l = n$ gilt $\lambda_n = \lambda_k \times \lambda_l$

Def. Sei $A \subset \Omega_1 \times \Omega_2$, $\omega_1 \in \Omega_1$, $\omega_2 \in \Omega_2$, $f : \Omega_1 \to \Omega_2 \to \mathbb{R}^1$. Dann heißen
$$A_{\omega_1} := \{\omega_2 \mid (\omega_1, \omega_2) \in A\} = \pi_2(A \cap \pi_1^{-1}(\omega_1)) \quad \omega_1\text{-}\textbf{Schnitt} \text{ von } A,$$
$$A_{\omega_2} := \{\omega_1 \mid (\omega_1, \omega_2) \in A\} = \pi_1(A \cap \pi_2^{-1}(\omega_2)) \quad \omega_2\text{-}\textbf{Schnitt} \text{ von } A,$$
$$f_{\omega_1} : \Omega_2 \to \mathbb{R}^1,\ \omega_2 \mapsto f(\omega_1, \omega_2) \quad \omega_1\text{-}\textbf{Schnitt von } f,$$
$$f_{\omega_2} : \Omega_1 \to \mathbb{R}^1,\ \omega_1 \mapsto f(\omega_1, \omega_2) \quad \omega_2\text{-}\textbf{Schnitt von } f.$$

Satz. Sei $A \in \mathfrak{A}_1 \otimes \mathfrak{A}_2$ und f sei $(\mathfrak{A}_1{\otimes}\mathfrak{A}_2)$-messbar. Dann gilt
$$\forall \omega_1 \in \Omega_1 : A_{\omega_1} \in \mathfrak{A}_2 \quad \text{und} \quad \forall \omega_2 \in \Omega_2 : A_{\omega_2} \in \mathfrak{A}_1$$
$$\forall \omega_1 \in \Omega_1 : f_{\omega_1} \text{ ist } \mathfrak{A}_2\text{-messbar} \quad \text{und} \quad \forall \omega_2 \in \Omega_2 : f_{\omega_2} \text{ ist } \mathfrak{A}_1\text{-messbar}$$

Lem. Sei μ_i ein σ-endliches Maß auf $(\Omega_i, \mathfrak{A}_i)$ für $i = 1, 2$. Dann sind für $A \in \mathfrak{A}_1 \otimes \mathfrak{A}_2$ die Funktionen $f_1(\omega_1) = \mu_2(A_{\omega_1})$ bzw. $f_2(\omega_2) = \mu_1(A_{\omega_2})$ nichtnegative $\mathfrak{A}_1$- bzw. $\mathfrak{A}_2$-messbare Funktionen auf Ω_1 bzw. Ω_2 und es gilt
$$\int_{\Omega_1} f_1\,\mathrm{d}\mu_1 = \mu(A) = \int_{\Omega_2} f_2\,\mathrm{d}\mu_2,$$
wobei $\mu = \mu_1 \times \mu_2$ das Produktmaß auf $\mathfrak{A}_1 \otimes \mathfrak{A}_2$ ist.

Satz (Fubini, Tonelli)**.** Seien $(\Omega_1, \mathfrak{A}_1, \mu_1)$ und $(\Omega_2, \mathfrak{A}_2, \mu_2)$ zwei σ-endliche Maßräume und sei
$$f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}^1}$$
eine nichtnegative, numerische $(\mathfrak{A}_1{\otimes}\mathfrak{A}_2)$-messbare Funktion. Dann sind
$$\omega_1 \mapsto \int_{\Omega_2} f_{\omega_1}\,\mathrm{d}\mu_2 \quad \text{und} \quad \omega_2 \mapsto \int_{\Omega_1} f_{\omega_2}\,\mathrm{d}\mu_1$$
$\mathfrak{A}_1$- bzw. $\mathfrak{A}_2$-messbare Funktionen und es gilt
$$\int_{\Omega_1\times\Omega_2} f\,\mathrm{d}(\mu_1{\times}\mu_2) = \int_{\Omega_1}\left(\int_{\Omega_2} f_{\omega_1}\,\mathrm{d}\mu_2\right)\mathrm{d}\mu_1 = \int_{\Omega_2}\left(\int_{\Omega_1} f_{\omega_2}\,\mathrm{d}\mu_1\right)\mathrm{d}\mu_2$$

Bem. Falls f $(\Omega_1{\otimes}\Omega_2)$-messbar mit Vorzeichenwechsel, so muss $\int_{\Omega_1\times\Omega_2} |f|\,\mathrm{d}(\mu_1{\times}\mu_2)$ vorausgesetzt werden.