Zusammenfassung Stochastik I

(cc) (by) Tim Baumann, http://timbaumann.info/uni-spicker

Der abstrakte Maß- und Wkts-Begriff

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

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

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

- Durch $A \leq B \;:\iff\; A \wedge B = B$ (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 $A \wedge B = U$.

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

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

Korollar. Durch Induktion folgt aus den De Morganschen Regeln:

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

Definition. 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:

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

Bemerkung. 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**.

Definition. 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.

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

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

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. 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:

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

Definition. 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:

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

Bemerkung. Ein Ring ist auch unter Schnitten stabil, da

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

Definition. 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.

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

Bemerkung. 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.

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

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

Dann heißen $\rho(E) \coloneqq \bigcap_{\mathfrak{R} \in \mathcal{R}(E)} \mathfrak{R}$, $\sigma(E) \coloneqq \bigcap_{\mathfrak{A} \in \mathcal{A}(E)} \mathfrak{A}$ von E **erzeugter Ring** bzw. von E **erzeugte σ-Algebra**.

Definition. 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:

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

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

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

Definition. 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.

Definition. 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

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

Indikatorfunktion von A.

Beispiel. 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**.

Lemma. 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).

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

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

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

Definition. Sei $f : \mathbb{R}^d \to \mathbb{R}^1$ eine Funktion, $x = (x_1, ..., x_d) \in \mathbb{R}^d$ und $h = (h_1, ..., h_d) \in \mathbb{R}^d_{\ge 0}$. 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 $]x, x + h]$.

Definition. $G : \mathbb{R}^d \to \mathbb{R}$ 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 $h \in \mathbb{R}^d_{\ge 0}$ ist der Zuwachs $(\triangle G)(]x, x + h]) \ge 0$.

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

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

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

Bemerkung. 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

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

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

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

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

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

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

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

Definition. 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

- $\mathfrak{A}^* := \{A \subset \Omega \,|\, A \text{ ist } \mu^*\text{-messbar }\}$ 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.

Bemerkung. 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.

Definition. 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.

Definition. 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 $\mu(A) = 0$.

Definition. 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

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

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.

Definition. 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
$$\eta(A) = 0 \implies \mu(A) = 0 \quad \text{für alle } A \in \mathfrak{A}.$$

Definition. 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 $\emptyset$.

Definition. 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 $\mathbb{P}(A \triangle B) = 0$.

Bemerkung. 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 $\mathbb{R}^d$.

Definition (Wichtige Verteilungsfunktionen)**.**

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

- 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

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

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

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

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

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 unabhängig**, falls

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

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

- **paarweise unabhängig**, falls

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

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 $0 \le m \le n$, $p \in [0,1]$ 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

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

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

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

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

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

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

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

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

und definiere analog $\{f < g\}$, $\{f \ge g\}$, $\{f > g\}$, $\{f = g\}$, $\{f \ne g\}$.

Satz. Für eine Funktion $f : (\Omega,\mathfrak{A}) \to (\overline{\mathbb{R}^1},\overline{\mathfrak{B}})$ sind äquivalent:

- f ist messbar
- $\forall\, a \in \mathbb{R} : \{f \ge a\} = f^{-1}([a,\infty]) \in \mathfrak{A}$
- $\forall\, a \in \mathbb{R} : \{f > a\} \in \mathfrak{A}$
- $\forall\, a \in \mathbb{R} : \{f \le a\} \in \mathfrak{A}$
- $\forall\, a \in \mathbb{R} : \{f < a\} \in \mathfrak{A}$

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

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

  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 $f,g : (\Omega,\mathfrak{A}) \to (\overline{\mathbb{R}^1},\overline{\mathfrak{B}})$ gilt:

- $\{f < g\} \in \mathfrak{A}$
- $\{f > g\} \in \mathfrak{A}$
- $\{f = g\} \in \mathfrak{A}$
- $\{f \le g\} \in \mathfrak{A}$
- $\{f \ge g\} \in \mathfrak{A}$
- $\{f \ne g\} \in \mathfrak{A}$

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:

- $\sup_{n\in\mathbb{N}} f_n$
- $\inf_{n\in\mathbb{N}} f_n$
- $\liminf_{n\in\mathbb{N}} f_n$
- $\limsup_{n\in\mathbb{N}} f_n$

Dabei werden Infimum, Supremum, usw. punktweise gebildet.

Definition. Für $f : \Omega \to \overline{\mathbb{R}^1}$ 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:

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

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

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

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

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

Satz. Es gilt für $f, g \in \mathbb{E}(\Omega, \mathfrak{A})$, $a, b \geq 0$:
- $\int_\Omega \mathbb{1}_A \,\mathrm{d}\mu = \mu(A)$
- $f \leq g \implies \int_\Omega f \,\mathrm{d}\mu \leq \int_\Omega g \,\mathrm{d}\mu$
- $\int_\Omega a \cdot f + b \cdot g \,\mathrm{d}\mu = a \cdot \int_\Omega f \,\mathrm{d}\mu + b \cdot \int_\Omega g \,\mathrm{d}\mu$

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

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

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 $\sup_{n\in\mathbb{N}} f_n = f$.

Definition. 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

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

Definition. Eine $\mathfrak{A}$-messbare, numerische Fkt. $f : (\Omega, \mathfrak{A}, \mu) \to (\overline{\mathbb{R}^1}, \mathfrak{L}(\overline{\mathbb{R}^1}))$ heißt $\boldsymbol{\mu}$**-integrierbar**, falls

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Korollar (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

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

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

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

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

Lemma (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

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

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

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

Definition. Eine Folge $(f_n)_{n\in\mathbb{N}}$ $\mathfrak{A}$-messbarer, numerischer Fktn. über $(\Omega, \mathfrak{A}, \mu)$ **konvergiert** $\boldsymbol{\mu}$**-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

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

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

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

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

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

Satz (Transformationssatz)**.** Sei $U, \widetilde{U} \Subset \mathbb{R}^d$ und sei $\phi : U \to \widetilde{U}$ ein $\mathcal{C}^1$-Diffeomorphismus. Dann ist eine Funktion $f : \widetilde{U} \to \overline{\mathbb{R}^1}$ 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.

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

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

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

Korollar. 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.

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

Bemerkung. 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 $F'(x)$.

Definition. 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

  $$\forall\, k \in J \subset \mathbb{N} \,:\, p_k := F_X(x_k) - \lim_{x \uparrow x_k} F_X(x) > 0, \quad \sum_{k=1}^{\infty} p_k = 1.$$

  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:

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

- 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

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

Definition. 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.

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

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

Bemerkung. 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.

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

...

$F_(X_1, ..., X_k)$ 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

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

$F_{X_1,...,X_k}$ 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

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

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

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

$\mathbb{E}X = \int_0^\infty 1 - F_X(x) - F_X(-x)\,\mathrm{d}x$, 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. $\mathrm{Var}(X) := D^2X = \mathbb{E}(X-\mathbb{E}X)^2 = \mathbb{E}X^2$ 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"
$\mathbb{D}^2X = \mathrm{Var}(X) = \mathbb{E}(X-\mathbb{E}X)^2 = \mathbb{E}(X^2) - (\mathbb{E}X)^2 \geq 0$
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 $\mathbb{P}(|X|\geq\epsilon) \leq \frac{\mathbb{E}g(|X|)}{g(\epsilon)}$ für beliebiges $\epsilon > 0$.

Spezialfälle:

- $g(x) = x$, dann $\mathbb{P}(|X|\geq\epsilon) \leq \frac{\mathbb{E}X}{\epsilon}$ **Markow-Ungleichung**
- $g(x) = x^2$, dann $\mathbb{P}(|X-\mathbb{E}X|\geq\epsilon) \leq \frac{\mathbb{E}(X-\mathbb{E}X)^2}{\epsilon^2} = \frac{\mathrm{Var}(X)}{\epsilon^2}$ **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:
$\mathbb{E}(X_1^{n_1}\cdots X_k^{n_k}) \leq (\mathbb{E}|X_1|^n)^{\frac{n_1}{n}}\cdots(\mathbb{E}|X_k|^n)^{\frac{n_k}{n}}$, $n = n_1+\ldots+n_k$

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 $x_1,\ldots,x_n\in J$, $\alpha_i\geq 0$, $\alpha_1+\ldots+\alpha_n = 1$

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 $c_0 = 1, c_1, c_2, \ldots$ eine Momentenfolge einer ZG X, d. h. $c_n = \mathbb{E}X^n$.

Antwort.
$0 \leq \mathbb{E}(z_0 + z_1X + \ldots + z_nX^n)^2 = \mathbb{E}(\sum_{i,j=1}^n z_iz_jX^{i+j}) = \sum_{i,j=1}^n z_iz_jc_{i+j}$
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 $X = (X_1,\ldots,X_k)$ eine k-dimensionale ZV über $(\Omega,\mathfrak{A},\mathbb{P})$. $X_1,\ldots,X_k$ 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 $B_1,\ldots,B_k\in\mathcal{L}(\mathbb{R}^1)$. Dies ist genau dann der Fall, wenn

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

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

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

Für diskrete Verteilungen ist dies äquivalent zu

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

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

Definition. Für eine k-dimensionale ZV $X = (X_1,\ldots,X_k)$ heißt
$F_{(X_{i_1},\ldots,X_{i_l})}(x_{i_1},\ldots,x_{i_l}) = \lim_{x_j\to\infty\, j\in\{1,\ldots,k\}\setminus\{i_1,\ldots,i_l\}} F_(X_1,\ldots,X_k)(x_1,\ldots,x_k)$

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

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

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

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 $x_m = (x_m^{(1)},\ldots,x_m^{(k)})$ 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

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

Kovarianz von X und Y und

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

Korrelation von X und Y.

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

Definition. Falls $\mathrm{Cor}(X,Y) = 0$, 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 $\int\limits_{-\infty}^{\infty} |x|^3 f_X(x)\,\mathrm{d}x < \infty$, dann ist $\mathrm{Cov}(X, X^2) = 0$, aber X und X^2 nicht unabhängig.

Bemerkung.
- $\mathrm{Cor}(X,Y) = 1$: positive Korrelation
- $\mathrm{Cor}(X,Y) = -1$: negative Korrelation
- $\mathrm{Cor}(X,Y) = 0$: Unkorreliertheit

Wichtig: Falls (X,Y) eine zweidimensionale Normalverteilung besitzt, so folgt aus $\mathrm{Cor}(X,Y) = 0$ 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

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

Wir wollen die Kovarianz der ZG X und Y berechnen.
1. Fall: Es existiert eine gemeinsame WD von (X,Y), $f_{(X,Y)}(x,y)$ (Lebesgue-messbar, ≥ 0, $\int\limits_{\mathbb{R}^2} f_{(X,Y)}(x,y)\,\mathrm{d}(x,y) = 1$)
Beispiel: 2-dimensionale Normalverteilungsdichte

$$f_{(X,Y)}(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\right)\right),$$

wobei $\mu_1 = \mathbb{E}X$, $\mu_2 = \mathbb{E}Y$, $\sigma_1^2 = \mathrm{Var}(X)$, $\sigma_2^2 = \mathrm{Var}(Y)$
$\mathbb{E}(X \cdot Y) = \int\limits_{\mathbb{R}^2} x \cdot y \cdot f_{(X,Y)}(x,y)\,\mathrm{d}(x,y) = \mu_1 \cdot \mu_2 + \sigma_1 \cdot \sigma_2 \cdot \rho$
Also: $\mathrm{Cov}(X,Y) = \sigma_1 \cdot \sigma_2 \cdot \rho$ und $\mathrm{Cor}(X,Y) = \rho \in [-1,1]$.
Randverteilungen:
$f_X(x) = \int\limits_{-\infty}^{\infty} f_{(X,Y)}(x,y)\,\mathrm{d}y$ $f_Y(y) = \int\limits_{-\infty}^{\infty} f_{(X,Y)}(x,y)\,\mathrm{d}x$
$\mathbb{E}X = \int\limits_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} x f_{(X,Y)}(x,y)\,\mathrm{d}x\,\mathrm{d}y$
2. Fall: (X,Y) mit Werten in einer höchstens abzählbaren Menge an (x_i, y_i) wird mit Wahrscheinlichkeit $p_{ij} > 0$ angenommen, $i, j = 1, 2, ...$

$$\mathbb{P}((X,Y) = (x_i, y_i)) = \mathbb{P}(X = x_1, Y = y_j)$$

$$\mathbb{E}(X \cdot Y) = \sum_{i=1}\sum_{j=1} x_i \cdot y_j p_{ij}$$

$$\mathbb{E}X = \sum_{i=1} x_i \mathbb{P}(X = x_i) = \sum_{i=1}\sum_{j=1} x_i p_{ij}$$

$$\mathbb{E}Y = \sum_{j=1} y_j \mathbb{P}(Y = y_j) = \sum_{j=1}\sum_{i=1} y_j p_{ij}$$

$$\mathrm{Cov}(X,Y) = \mathbb{E}(X \cdot Y) - (\mathbb{E}X) \cdot (\mathbb{E}Y) = \sum_{i=1}\sum_{j=1} x_i y_j (p_{ij} - p_i \cdot q_j)$$

mit $p_i := \sum\limits_{j=1} p_{ij}$, $q_j := \sum\limits_{i=1} p_{ij}$
3. Fall: (X,Y) ist singulär-stetig verteilt oder besitzt eine singulär-stetige Komponente
Beispiel: Zweidimensionale Exponentialverteilung ...

$$\overline{F}(x,y) = \mathbb{P}(X > x, Y > y) = \exp\left(-(\lambda x + \mu y + \nu \max(x,y))\right)$$

$$F(x,y) = \mathbb{P}(A^c \cap B^c) = \mathbb{P}((A \cup B)^c) = 1 - \mathbb{P}(A \cup B) = 1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A \cap B) = 1 - \overline{F}(x,0) - \overline{F}(0,y) + \overline{F}(x,y)$$

$$\mathbb{E}XY = \int\limits_{\mathbb{R}^2} x \cdot y\,\mathrm{d}F(x,y) = ... = -\int\limits_0^{\infty} [y \cdot \mathbb{P}(X > x, Y > y)]_{y=0}^{y=\infty}\,\mathrm{d}x + \int\limits_{\mathbb{R}^2_{\geq 0}} \mathbb{P}(X > x, Y > y)\,\mathrm{d}(x,y)$$

Für $(X,Y) \in \mathbb{R}^2$ gilt

$$\mathbb{E}(XY) = \int\limits_0^{\infty}\int\limits_0^{\infty} \underbrace{\mathbb{P}(X > x, Y > y)}_{=\overline{F}(x,y)}\,\mathrm{d}x\,\mathrm{d}y$$

Falls $(X,Y) \in \mathbb{R}^2$, so gilt:

$$\mathbb{E}(X \cdot Y) = \int\limits_0^{\infty}\int\limits_0^{\infty} (\mathbb{P}(X > x, Y > y) + \mathbb{P}(X \leq -x, Y \leq -y) - \mathbb{P}(X \geq x, Y \leq -y) - \mathbb{P}(X \leq -x, Y > y))\,\mathrm{d}y\,\mathrm{d}x$$

Im Beispiel:

$$\mathbb{E}(X \cdot Y) = \int\limits_0^{\infty}\int\limits_0^{\infty} \exp(-\lambda x - \mu y - \nu \max(x,y))\,\mathrm{d}x\,\mathrm{d}y = ... = \frac{1}{\lambda(\mu+\nu)} + \frac{1}{\mu(\lambda+\nu)} - \frac{1}{\lambda}\frac{1}{\mu}\frac{1}{\lambda+\mu+\nu}$$

$$\mathbb{E}X = \int\limits_0^{\infty} x\,\mathrm{d}F(x) = \int\limits_0^{\infty} \mathbb{P}(X > x)\,\mathrm{d}x = \frac{1}{\lambda+\nu}$$

$$\mathrm{Cor}(X,Y) = ... = \frac{\nu}{(\lambda+\nu)(\mu+\nu)(\lambda+\mu+\nu)}$$

Also: $\nu = 0$ genau dann, wenn X und Y unabhängig
$Y = g(X)$, $g : \mathbb{R}^1 \to \mathbb{R}^1$ Borel-messbar

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(Y in]-\infty, y[) = \mathbb{P}(X \in g^{-1}(]-\infty, y]))$$

Beispiel: $Y = X^2$, $y \geq 0$
$F_Y(y) = \mathbb{P}(X^2 \leq y) = \mathbb{P}(|X| \leq \sqrt{y}) = \mathbb{P}(-\sqrt{y} \leq X \leq \sqrt{y}) = \mathbb{P}(X = -\sqrt{y}) + F_X(\sqrt{y}) - F_X(-\sqrt{y}) = F_X(\sqrt{y}) - \lim\limits_{z \uparrow -\sqrt{y}} F_X(z)$

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

$$f_Y(y) = \begin{cases} 0, & \text{für } y \in \mathbb{R}^1 \setminus g(D) \\ \dfrac{f_X(g^{-1}(y))}{|g'(g^{-1}(y))|} & \text{für } y \in g(D) \end{cases}$$

Beispiel. $Y = e^{N(\mu,\sigma^2)} + c$
Dann heißt Y logarithmisch normalverteilt
$g(x) = e^x + c$
Ausgelassen: Rechnungen

Bemerkung. Falls die Ableitung $g'(x), x \in D$ wechselndes Vorzeichen besitzt, so muss in Monotoniebereiche unterteilt werden.

Beispiel. $Y = e^{N(\mu,\sigma^2)+c}$, speziell $\mu = 0$, $\sigma^2 = 1$, $c = 0$, dann

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{(\log y)^2}{2}).$$

Die Funktion

$$f_a(y) := f_Y(y) \cdot (1 + a \cdot \sin(2\pi \log y)), y \in \mathbb{R}^1_{>0}, a \in [0,1]$$

Für $a \in [0,1]$ ist dann $\int\limits_{-\infty}^{\infty} f_a(y)\,\mathrm{d}y = 1$

$$\mathbb{E}Y^n = \int\limits_0^{\infty} y^n f_Y(y)\,\mathrm{d}y = ... = \exp(-\frac{n^2}{2})$$

und

$$\int\limits_0^{\infty} y^n f_a(y)\,\mathrm{d}y = \exp(-\frac{n^2}{2})$$

Sei $X = (X_1, ..., X_k)$ ein k-dimensionaler Zufallsvektor, $g = (g_1, ..., g_k) : \mathbb{R}^k \to \mathbb{R}^k$ mit $g_i : \mathbb{R}^k \to \mathbb{R}^1$ für $i = 1, ..., k$ Borel-messbar. Dann ist $Y = g(X)$ ein k-dimensionaler ZV mit $P_Y = P_X \circ g^{-1}$. Falls $f_{(X_1,...,X_k)}$ eine W-Dichte von X ist und g ein $\mathcal{C}^1$-Diffeomorphismus ist, dann existiert die W-Dichte von $f_{(Y_1,...,Y_k)}$ von Y.

Satz. Seien $G, H \Subset \mathbb{R}^k$ offen und $g = (g_1, ..., g_k) : G \to H$ ein $\mathcal{C}^1$-Diffeomorphismus. 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:

$$f_{(Y_1,...,Y_k)}(y_1, ..., y_k) = \begin{cases} \dfrac{f_{(X_1,...,X_k)}(h_1(y_1, ..., y_k), ..., h_k(y_1, ..., y_k))}{|\det Dg(h_1(y_1, ..., y_k), ..., h_k(y_1, ..., y_k))|} & \text{für } (y \\ 0, & \text{sonst} \end{cases}$$

Beachte:
$$\det D(h)(y_1, ..., y_k) = \frac{1}{\det D(g)(h_1(y_1, ..., y_k), ..., h_k(y_1, ..., y_k))}$$

Beispiel (Box-Muller-Transformation)**.**

$$Y_1 = g_1(X_1, X_2) = \sqrt{-2\log(X_1)}\sin(2\pi X_2)$$
$$Y_2 = g_2(X_1, X_2) = \sqrt{-2\log(X_1)}\cos(2\pi X_2)$$

mit X_1 und X_2 gleichverteilt auf $[0,1]$.

$$f_{(X_1)}(x_1, x_2) = \mathbb{1}_{]0,1[}$$

Behauptung: Dann ist $f_{(Y_1,Y_2)} = \phi(y_1) \cdot \phi(y_2)$, wobei $\phi(y) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2})$ oder $Y_1, Y_2 \sim N(0,1)$ und unabhängig.

Beispiel.

$$Y_1 = g_1(X_1, X_2) = \sqrt{X_1^2 + X_2^2}$$

$$Y_2 = g_2(X_1, X_2) = \begin{cases} \arctan(\frac{X_2}{X_1}) & \text{für } X_2 \geq 0,\ X_1 \in \mathbb{R}^1 \setminus \{0\} \\ \arctan(\frac{X_2}{X_1}) + \pi, & \text{für } X_2 < 0, X_1 \in \mathbb{R}^1 \setminus \{0\} \end{cases}$$

Wenn $X_1, X_2 \sim N(0, \sigma^2)$, unabhängig $\implies$
$f_{(X_1,X_2)}(x_1, x_2) = \phi(\frac{x_1}{\sigma})\phi(\frac{x_2}{\sigma})$

Ergebnis: $f_{(Y_1,Y_2)}(y_1, y_2) = \frac{1}{2\pi} \cdot \underbrace{\frac{y_1}{\sigma^2}\exp(-\frac{y_1^2}{\sigma^2})}_{\text{Dichte der Rayleigh-Verteilung}}$ für $(y_1, y_2) \in]0, \infty[\times]0, 2\pi]$

$$F_{Y_1}(y) = ... = 1 - \exp(-\frac{y^2}{2\sigma^2})$$

Satz. (X, Y) besitze eine gemeinsame Dichte $f_{(X,Y)}$ (also $\mathbb{P}(X = 0) = \mathbb{P}(Y = 0) = 0$). Dann gilt für die Dichten von $Z_1 := X + Y$, $Z_2 := X \cdot Y$, $Z_3 := \frac{X}{Y}$.

$$f_{Z_1}(z) = \int_{-\infty}^{\infty} f_{(X,Y)}(z - y, y)\,\mathrm{d}y \overbrace{=}^{\text{f. unabh.}}$$

$$f_{Z_2}(z) = \int_{-\infty}^{\infty} f_{(X,Y)}(z/y, y) \cdot \frac{1}{|y|}\,\mathrm{d}y = ...$$

$$f_{Z_3}(z) = \int_{-\infty}^{\infty} f_{(X,Y)}(x \cdot y, y) \cdot |y|\,\mathrm{d}y = ...$$

Beispiel. Seien X, Y unabhängig, $N(0, 1)$-verteilt, $Z := \frac{X}{Y}$. Dann:

$$f_Z(z) = ... = \frac{1}{\pi(1+z^2)}$$

Beispiel. Sei nun $X \sim F_X$ unabhängig von $Y \sim F_Y$. Dann

$$\mathbb{E}g(X, Y) = \int_{\mathbb{R}^2} g(x, y)\,\mathrm{d}P_{(X,Y)} = ... = \int_{\mathbb{R}^1}\int_{\mathbb{R}^1} g(x, y)\,\mathrm{d}P_X(x)\,\mathrm{d}P_Y(y)$$

Mit Verteilungsfunktionen kann man schreiben:

$$\mathbb{E}g(X, Y) = \int_{\mathbb{R}^2} g(x, y)\,\mathrm{d}F_{X,Y}(x, y) = \int_{\mathbb{R}^1}\int_{\mathbb{R}^1} g(x, y)\,\mathrm{d}F_X(x)\,\mathrm{d}F_Y(y)$$

$$\mathbb{P}(X + Y \leq z) = ... = \int_{\mathbb{R}^1} F_X(z - y)\,\mathrm{d}F_Y(y) =: (F_X * F_Y)(z)$$

Faltung von F_X und F_Y

Bemerkung.
- $F_X * F_Y = F_Y * F_X$
- $F_{X+Y} = F_X * F_Y$
- Falls X oder Y eine Dichte besitzen, so auch $X + Y$:

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(z-y)\,\mathrm{d}F_Y(y) = \int_{-\infty}^{\infty} f_Y(z-x)\,\mathrm{d}F_X(x) = \int_{-\infty}^{\infty} f_Y(z-x)\,\mathrm{d}f_X(x)y$$

- Verallgemeinerung: $F_{X_1+...+X_n} = F_{X_1} * ... * F_{X_n}$, falls $X_1, ..., X_n$ unabhängig.

Gewisse Verteilungen sind „faltungsstabil", d. h. bei der Faltung bleibt der Verteilungstyp erhalten

Beispiel. X_i sei Poisson-verteilt mit Paramter $\lambda_i > 0$, $i = 1, 2$. Dann ist $X_1 + X_2$ Poisson-verteilt mit Parameter $\lambda_1 + \lambda_2$, falls X_1 und X_2 unabhängig sind.

Beispiel. Sei $X_i \overbrace{=}^{d} N(\mu_i, \sigma_i^2)$, $i = 1, 2$ unabhängig. Dann:
$X_1 + X_2 \overbrace{=}^{d} N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Satz. Falls X_1 und X_2 unabhängig und $X_1 + X_2$ normalverteilt (also $X_1 + X_2 \overbrace{=}^{d} N(\mu, \sigma^2)$), dann sind X_1 und X_2 normalverteilt. Falls X_1 und X_2 unabhängig und $X_1 + X_2$ Poisson-verteilt (also $X_1 + X_2 \overbrace{=}^{d} Poi(\lambda)$), dann sind X_1 und X_2 Poisson-verteilt.

Wichtige Ergebnisse über Summen unabhängiger ZGen:

- $\frac{1}{n}(X_1 + ... + X_n) \xrightarrow[\mathbb{P}\text{-fast-überall}]{n\to\infty} \mathbb{E}X_1$ falls $X_1, ..., X_n$ unabhängig, identisch verteilt und $\mathbb{E}|X_1| < \infty$ (GGZ)
- $\sqrt{n}(\frac{1}{n}(X_1 + ... + X_n) - \mathbb{E}X_1) \xrightarrow[d]{n\to\infty} N(0, \mathrm{Var}(X_1))$, falls $\mathrm{Var}(X_1) < \infty$