Wahrscheinlichkeitstheorie
Vorlesungsnotizen

Stilianos Louca

Februar 2010

Inhaltsverzeichnis

1 Zufallsvariablen

1.1 Wahrscheinlichkeitsräume

1.1.1 Definition: Wahrscheinlichkeitsraum

Gegeben sei der Grundraum $\Omega \neq \emptyset$, dazu σ-Algebra $\mathfrak{S} \subseteq \mathscr{P}(\Omega)$ und Wahrscheinlichkeitsmaß $\mathcal{P}$ auf $(\Omega, \mathfrak{S})$, das heißt

(i) $\mathcal{P} : \mathfrak{S} \to [0, 1]$

(ii) $\mathcal{P}(\emptyset) = 0$ & $\mathcal{P}(\Omega) = 1$.

(iii) $\mathcal{P}$ ist σ-additiv: Für disjunkte $A_j \in \mathfrak{S},\ j \in \mathbb{N}$ gilt

$$\mathcal{P}\left[\bigcup_{j=1}^{\infty} A_j\right] = \sum_{j=1}^{\infty} \mathcal{P}(A_j)$$

Dann heißt das Tripel $(\Omega, \mathfrak{S}, \mathcal{P})$ *Wahrscheinlichkeitsraum*. Im Falle eines topologischen Raumes Ω ist typischerweise die σ-Algebra der Borelmengen $\mathfrak{S} = \mathscr{B}(T)$ impliziert.

Interpretation: In der Wahrscheinlichkeitstheorie stellt Ω den Raum aller Elementarereignisse dar, $\mathcal{P}(A)$ die Wahrscheinlichkeit ein Ereignis aus $A \in \mathfrak{S}$ zu beobachten.

1.1.2 Definition: Träger und Stetigkeitsbereich eines Maßes

Sei $(T, \mathscr{O})$ ein topologischer Raum, darauf das Borelmaß μ. Dann heißt

$$\operatorname{supp}(\mu) := \{x \in T : \forall\ x \in G \in \mathscr{O} : \mu(G) > 0\}$$

Träger von μ. Ist $\{x\} \in \mathscr{B}(T)\ \forall\ x \in T$, so heißt

$$\mathcal{C}(\mu) := \{x \in T : \mu(\{x\}) = 0\}$$

Stetigkeitsbereich von μ.

1.1.3 Eigenschaften von Wahrscheinlichkeitsräumen

Für Wahrscheinlichkeitsraum $(\Omega, \mathfrak{S}, \mathcal{P})$ gilt stets:

1. $0 \leq \mathcal{P}(A) \leq 1 \quad \forall\ A \in \mathfrak{S}$
2. $\mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(A \cap B)$ für $A, B \in \mathfrak{S}$
3. $\mathcal{P}(A^c) = 1 - \mathcal{P}(A) \quad \forall\ A \in \mathfrak{S}$
4. $A \subseteq B \quad \Rightarrow \quad \mathcal{P}(A) \leq \mathcal{P}(B)$ für $A, B \in \mathfrak{S}$
5. Für $A_n \nearrow A$, das heißt $A_1 \subseteq A_2 \subseteq \cdots \in \mathfrak{S}$ und $\bigcup\limits_{n \in \mathbb{N}} A_n = A$, gilt
$$\mathcal{P}(A_n) \overset{n\to\infty}{\longrightarrow} \mathcal{P}(A)$$
6. Für $A_n \searrow A$, das heißt $A_1 \supseteq A_2 \supseteq \cdots \in \mathfrak{S}$ und $\bigcap\limits_{n \in \mathbb{N}} A_n = A$, gilt
$$\mathcal{P}(A_n) \overset{n\to\infty}{\longrightarrow} \mathcal{P}(A)$$

Beispiele:

(1) Für $\Omega := \{\omega_1, \omega_2, \dots\}$ und $\mathfrak{S} := \mathscr{P}(\Omega)$ ist

$$\mathcal{P}(A) = \sum_{\{j : \omega_j \in A\}} \mathcal{P}(\{\omega_j\})$$

das heißt $\mathcal{P}$ ist eindeutig durch die Werte $p_j := \mathcal{P}(\{\omega_j\}),\ j \in \mathbb{N}$ bestimmt, sprich

$$\mathcal{P} = \sum_{j=1}^{\infty} p_j \cdot \delta_{\omega_j}$$

Somit kann der Raum $\mathfrak{M}_1$ aller Wahrscheinlichkeitsmaße auf Ω identifiziert werden mit

$$\left\{(p_j)_{j\in\mathbb{N}} \ :\ p_j \geq 0 \quad \wedge \quad \sum_{j=1}^{\infty} p_j = 1\right\}$$

Als Beispiel sei die so genannte *Poissonverteilung mit Parameter* $\lambda > 0$ erwähnt:

$$\mathcal{P}(\{k\}) = \frac{\lambda^k}{k!} e^{-\lambda} \quad , \quad k \in \mathbb{N}_0$$

(2) Für Borelmenge $\Omega \subseteq \mathbb{R}^n,\ 0 < \lambda^n(\Omega) < \infty$ definiert

$$\mathcal{P}(A) := \frac{\lambda^n(\Omega \cap A)}{\lambda^n(\Omega)} \quad , \quad A \in \mathscr{B}(\mathbb{R}^n)$$

die sog. *Gleichverteilung auf* Ω.

1.2 Verteilungsfunktion & absolute Stetigkeit

1.2.1 Definition: Verteilungsfunktion

Sei $\Omega := \mathbb{R}$ und $\mathfrak{S} := \mathscr{B}(\mathbb{R})$ die σ-Algebra der Borelmengen, dazu Wahrscheinlichkeitsmaß $\mathcal{P}$ auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$. Dann heißt die Abbildung $F : \mathbb{R} \to [0,1]$ definiert durch

$$F(t) := \mathcal{P}((-\infty, t]) \quad , \quad t \in \mathbb{R} \tag{1.2.1.1}$$

Verteilungsfunktion zu $\mathcal{P}$.

1.2.2 Satz: Charakterisierung von Verteilungsfunktionen

Die Verteilungsfunktion F eines W-Maßes $\mathcal{P}$ auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ hat folgende Eigenschaften:

1. $\lim\limits_{t\to-\infty} F(t) = 0 \ \ \& \ \ \lim\limits_{t\to\infty} F(t) = 1$

2. F ist nicht-fallend.

3. F ist rechtsseitig stetig.

Umgekehrt gilt: Erfüllt $F : \mathbb{R} \to [0,1]$ die Eigenschaften (1), (2) und (3), so existiert ein eindeutig bestimmtes W-Maß $\mathcal{P}$ mit Verteilungsfunktion F.

Bemerkung: Aufgrund dieser Identifizierung ist es in der Literatur auch oft üblich Integrale $\int f \ d\mathcal{P}$ in der Form $\int f \ dF$ zu schreiben.

Weitere Eigenschaften

(i) $\mathcal{P}((a,b]) = F(b) - F(a)$ für $a \leq b \in \mathbb{R}$

(ii) $F(t_0) - \lim\limits_{t \nearrow t_0} F(t) = \mathcal{P}(\{t_0\})$ für $t_0 \in \mathbb{R}$.
Insbesondere ist F stetig, genau dann wenn $\mathcal{P}(\{t\}) = 0 \;\; \forall\, t \in \mathbb{R}$.

(iii) F hat höchstens abzählbar viele Sprünge (Unstetigkeitsstellen).
Beispiel: Die Verteilungsfunktion des Wahrscheinlichkeitsmaßes

$$\mathcal{P}(A) = \frac{1}{2} \int\limits_A p(x)\ dx + \frac{1}{2}\delta_0(A) \quad , \quad A \in \mathscr{B}(\mathbb{R}) \tag{1.2.2.1}$$

(für irgendeine messbare $p : \mathbb{R} \to [0,\infty]$ mit $\int\limits_{\mathbb{R}} p\ d\lambda = \frac{1}{2}$), besitzt genau einen Sprung in der 0.

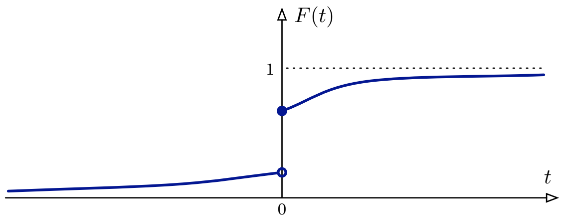

Abbildung 1: Zum Beispiel (1.2.2.1) einer Verteilungsfunktion.

1.2.3 Definition: Absolute Stetigkeit

Seien $\mathcal{P}, \mathcal{Q}$ W-Maße auf dem messbaren Raum $(\Omega, \mathfrak{S})$. Dann heißt $\mathcal{P}$ *absolut stetig* bzgl. $\mathcal{Q}$ ($\mathcal{P} \ll \mathcal{Q}$), falls

$$\forall\, A \in \mathfrak{S} : \mathcal{Q}(A) = 0 \quad \Rightarrow \quad \mathcal{P}(A) = 0$$

1.2.4 Satz über absolute Stetigkeit

Seien η, μ Maße auf dem messbaren Raum $(\Omega, \mathfrak{S})$ und η endlich. Dann ist $\eta \ll \mu \;\; \Leftrightarrow$

$$\forall\, \varepsilon > 0 : \exists\, \delta > 0 : \forall\, A \in \mathfrak{S} \;\; : \;\; \mu(A) < \delta \;\; \Rightarrow \;\; \eta(A) < \varepsilon$$

1.2.5 Satz über integrierbare Funktionen

Sei $(\Omega, \mathfrak{S}, \mu)$ ein Maßraum. Für jede integrierbare, messbare Abbildung $f : \Omega \to \mathbb{R}$ und Mengenfolge $(A_n)_n \subseteq \mathfrak{S}$ mit $\mu(A_n) \overset{n\to\infty}{\longrightarrow} 0$ gilt

$$\int\limits_{A_n} |f|\ d\mu \overset{n\to\infty}{\longrightarrow} 0$$

Erläuterung: $|f|$ induziert auf $(\Omega, \mathfrak{S})$ durch

$$\eta(A) := \int\limits_A |f|\ d\mu \quad , \quad A \in \mathfrak{S}$$

ein endliches Maß $\eta \ll \mu$. Nach Satz 1.2.4 folgt die Behauptung.
□

1.2.6 Satz von Radon-Nikodym

Seien $\mathcal{P}, \mathcal{Q}$ Maße auf $(\Omega, \mathfrak{S})$ und $\mathcal{Q}$ σ-endlich. Dann gilt $\mathcal{P} \ll \mathcal{Q}$ genau dann wenn eine messbare $p : \mathbb{R} \to [0, \infty)$ existiert mit

$$\mathcal{P}(A) = \int_A p \, d\mathcal{Q} \quad , \quad A \in \mathfrak{S}$$

Gegebenfalls ist p bis auf $\mathcal{Q}$-Nullmengen eindeutig bestimmt und heißt *Dichte* von $\mathcal{P}$ bzgl. $\mathcal{Q}$. Man schreibt auch

$$p = \frac{d\mathcal{P}}{d\mathcal{Q}}$$

1.2.7 Haupttypen von Wahrscheinlichkeitsmaßen auf $\mathbb{R}$

Man unterscheidet zwischen folgenden Hauptklassen von W-Maßen auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$:

1. *Diskrete* oder *atomare* W-Maße der Form

$$\mathcal{P}_{\mathrm{d}} = \sum_{k=1}^{\infty} p_k \cdot \delta_{x_k}$$

   für irgendwelche $x_k \in \mathbb{R}$, $p_k := \mathcal{P}_{\mathrm{d}}(\{x_k\})$.

2. Absolut-stetige W-Maße (bzgl. des Lebesgue-Maßes) $\mathcal{P}_{\mathrm{a}} \ll \lambda^1$.

3. Singulär-stetige W-Maße P_{s}, das heißt $\mathcal{P}_{\mathrm{s}}(\{x\}) = 0 \;\; \forall \, x \in \mathbb{R}$ und $\mathcal{P}_{\mathrm{s}} \perp \lambda^1$.[1] Als Beispiel sei das Haar-Maß auf der Cantorschen Menge erwähnt.

Dabei kann jedes beliebige W-Maß $\mathcal{P}$ auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ eindeutig in die oberen 3 Klassen zerlegt werden gemäß

$$\mathcal{P} = \varkappa_{\mathrm{d}} \cdot \mathcal{P}_{\mathrm{d}} + \varkappa_{\mathrm{a}} \cdot \mathcal{P}_{\mathrm{a}} + \varkappa_{\mathrm{s}} \cdot \mathcal{P}_{\mathrm{s}}$$

für irgendwelche $\varkappa_{\mathrm{d}}, \; \varkappa_{\mathrm{a}}, \; \varkappa_{\mathrm{s}} \geq 0$ mit $\varkappa_{\mathrm{d}} + \varkappa_{\mathrm{a}} + \varkappa_{\mathrm{s}} = 1$.

1.2.8 Wichtige Wahrscheinlichkeitsdichten auf $\mathbb{R}$

Folgende wichtigen Dichten auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ (bzgl. λ^1) seien erwähnt:

(i) Standardnormalverteilung $\mathcal{P} := \mathcal{N}_{0,1}$:

$$\frac{d\mathcal{P}}{d\lambda^1}(x) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right] \quad , \quad x \in \mathbb{R}$$

(ii) Normalverteilung um Mittelwert μ mit Varianz σ^2:

$$\frac{d\mathcal{P}}{d\lambda^1}(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \quad , \quad x \in \mathbb{R}$$

(iii) Exponentialverteilung $\mathcal{P} := \mathrm{Exp}_\lambda$ mit Parameter $\lambda > 0$:

$$\frac{d\mathcal{P}}{d\lambda^1}(x) = \begin{cases} \lambda e^{-\lambda x} & : x \geq 0 \\ 0 & : \text{sonst} \end{cases} \quad , \quad x \in \mathbb{R}$$

(iv) Cauchy-Verteilung:

$$\frac{d\mathcal{P}}{d\lambda^1}(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2} \quad , \quad x \in \mathbb{R} \tag{1.2.8.1}$$

[1] Zwei Maße μ, η auf $(\Omega, \mathfrak{S})$ heißen *singulär* falls

$$\exists \; B \in \mathfrak{S} \; : \; \mu(B) = \eta(B^c) = 0$$

Man schreibt $\mu \perp \eta$.

(v) Γ-Verteilung $\mathcal{P} := \Gamma_{a,b}$ mit Parametern $a, b > 0$:

$$\frac{d\Gamma_{a,b}}{d\lambda^1}(x) = \begin{cases} \frac{1}{a^b\Gamma(b)} \cdot x^{b-1}e^{-x/a} & : x > 0 \\ 0 & : \text{sonst} \end{cases} \quad , \quad x \in \mathbb{R} \tag{1.2.8.2}$$

1.3 Produktmaße

1.3.1 Definition: Produktraum

Zu messbaren Räumen $(\Omega_1, \mathfrak{S}_1), \ldots, (\Omega_n, \mathfrak{S}_n)$ und $\Omega := \Omega_1 \times \cdots \times \Omega_n$ definiert man die *Produkt-σ-Algebra*

$$\mathfrak{S} := \mathfrak{S}_1 \otimes \cdots \otimes \mathfrak{S}_n := \sigma \{A_1 \times \cdots \times A_n : A_j \in \mathfrak{S}_j\}$$

und nennt $(\Omega, \mathfrak{S})$ den entsprechenden *Produktraum*.

1.3.2 Charakterisierung von Produktmaßen

Seien $(\Omega_1, \mathfrak{S}_1, \mathcal{P}_1), \ldots, (\Omega_n, \mathfrak{S}_n, \mathcal{P}_n)$ W-Räume, dazu der Produktraum $(\Omega, \mathfrak{S})$. Dann existiert genau ein W-Maß $\mathcal{P} := \mathcal{P}_1 \otimes .. \otimes \mathcal{P}_n$ auf $(\Omega, \mathfrak{S})$ mit

$$(\mathcal{P}_1 \otimes \cdots \otimes \mathcal{P}_n)(A_1 \times \cdots \times A_n) = \mathcal{P}(A_1) \cdot .. \cdot \mathcal{P}(A_n) \quad \forall\ A_j \in \mathfrak{S}_j \tag{1.3.2.1}$$

Mann nennt $\mathcal{P}$ das *Produktmaß* von $\mathcal{P}_1, .., \mathcal{P}_n$.

Beispiele:

(i) Zu $\Omega_1 = \{x_1, x_2, \ldots\}$, $\Omega_2 = \{y_1, y_2, \ldots\}$, $\mathfrak{S}_i = \mathscr{P}(\Omega_i)$ ergibt sich die Darstellung für das Produktmaß

$$(\mathcal{P}_1 \otimes \mathcal{P}_2)(\{(x_i, y_j)\}) = (\mathcal{P}_1 \otimes \mathcal{P}_2)(\{x_i\} \times \{y_j\}) = \mathcal{P}_1(\{x_i\}) \cdot \mathcal{P}_2(\{y_j\})$$

(ii) Zu $\Omega_1 = \cdots = \Omega_n = \mathbb{R}$ und W-Maßen $\mathcal{P}_1, ..., \mathcal{P}_n$ auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ ist

$$\mathcal{P} := \mathcal{P}_1 \otimes \cdots \otimes \mathcal{P}_n$$

ein W-Maß auf $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$. Besitzen $\mathcal{P}_i$ jeweils die Dichten p_i bzgl. λ^1, so besitzt $\mathcal{P}$ die Dichte

$$p(x^1, .., x^n) = p_1(x^1) \cdot .. \cdot p_n(x^n) \quad , \quad \mathbf{x} \in \mathbb{R}^n$$

bzgl. λ^n. Sind etwa $\mathcal{P}_i = \mathcal{N}_{0,1}$, das heißt $p_i(t) = \frac{1}{\sqrt{2\pi}}e^{-t^2/2}$, so folgt

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\|\mathbf{x}\|^2/2}$$

Wahrscheinlichkeitstheoretische Interpretation: Der Produktraum $(\Omega, \mathfrak{S}, \mathcal{P})$ beschreibt n *unabhängige Versuche*, wobei $\mathcal{P}_1, \ldots, \mathcal{P}_n$ Wahrscheinlichkeitsverteilungen der einzelnen Versuchsereignisse sind. Die Wahrscheinlichkeit dass $A_j \in \mathfrak{S}_j$ im j-ten Versuch eintritt, ist gegeben durch

$$\mathcal{P}\big(\Omega_1 \times \cdots \times \underset{\substack{\uparrow \\ j}}{A_j} \times \cdots \times \Omega_n\big) = \underbrace{\mathcal{P}_1(\Omega_1)}_{1} \cdot .. \cdot \mathcal{P}_j(A_j) \cdot .. \cdot \underbrace{\mathcal{P}_n(\Omega_n)}_{1} = \mathcal{P}_j(A_j)$$

1.4 Zufallsvariablen

1.4.1 Definition: Zufallsvariable

Gegeben seien der Wahrscheinlichkeitsraum $(\Omega, \mathfrak{S}, \mathcal{P})$ und messbare Raum $(S, \mathcal{S})$, dazu messbare Abbildung $X : \Omega \to S$. Dann heißt X S-wertige *Zufallsvariable* (ZV) oder *Zufallsgröße* (Zufallsgröße, ZG).

Erinnerung: $X : \Omega \to S$ heißt *messbar*, falls

$$X^{-1}(B) \in \mathfrak{S} \ \ \forall\ B \in \mathcal{S}$$

Spezialfälle:

(i) Im Falle $(S, \mathcal{S}) = (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ heißt X *reelle Zufallsgröße.*

(ii) Im Falle $(S, \mathcal{S}) = (\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$ heißt X *(reeller), n-dimensionaler, zufälliger Vektor*.

1.4.2 Charakterisierung reeller Zufallsgrößen

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $X : \Omega \to \mathbb{R}^n$. Dann sind äquivalent:

1. $X : \Omega \to (\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$ ist zufälliger Vektor ($\Leftrightarrow$ messbar).
2. Die Komponenten $X^1, .., X^n : \Omega \to \mathbb{R}$ sind zufällige Größen ($\Leftrightarrow$ messbar).
3. $\forall\ t_1, .., t_n \in \mathbb{R} : \left\{X^1 \le t_1, \ldots, X^n \le t_n\right\} \in \mathfrak{S}$

1.4.3 Eigenschaften zufälliger Größen

Seien $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum, $(S, \mathcal{S})$, $(S', \mathcal{S}')$ messbare Räume. Es gilt:

1. Sind $X : \Omega \to S$, $f : S \to S'$ messbar, so ist auch $f \circ X : \Omega \to S'$ messbar.
2. Sind $X, Y : \Omega \to \mathbb{R}^n$ n-dimensionale, zufällige Vektoren und $\alpha, \beta \in \mathbb{R}$, so ist auch $\alpha X + \beta Y : \Omega \to \mathbb{R}^n$ ein n-dimensionaler, zufälliger Vektor.
3. Sind $X, Y : \Omega \to \mathbb{R}^n$ n-dimensionale Zufallsvektoren, so ist $\langle X, Y \rangle : \Omega \to \mathbb{R}$ eine reelle Zufallsgröße.
4. Sind $X, Y : \Omega \to \mathbb{R}$ reelle Zufallsgrößen mit $Y(\omega) \neq 0 \ \ \forall\ \omega \in \Omega$, so ist auch $X/Y : \Omega \to \mathbb{R}$ eine reelle Zufallsgröße.
5. Sind $X, Y : \Omega \to \mathbb{R}$ reelle Zufallsgrößen, so sind auch $X \vee Y := \max\{X, Y\} : \Omega \to \mathbb{R}$ und $X \wedge Y := \min\{X, Y\} : \Omega \to \mathbb{R}$ reelle Zufallsgrößen.

1.5 Verteilungsgesetz einer Zufallsvariablen

1.5.1 Definition: Verteilungsgesetz

Sei $X : (\Omega, \mathfrak{S}, \mathcal{P}) \to (S, \mathcal{S})$ eine S-wertige Zufallsvariable. Durch

$$\mathcal{P}_X(B) := \mathcal{P}\left(X^{-1}(B)\right) \quad , \quad B \in \mathcal{S}$$

wird auf $(S, \mathcal{S})$ ein Wahrscheinlichkeitsmaß, das sogenannte *Verteilungsgesetz* von X, definiert. In der Maßtheorie schreibt man auch $\mathcal{P}_X = \mathcal{P} \circ X^{-1}$.

Beispiele:

(i) Betrachtet sei $\Omega := \{0, 1\}^n$, $\mathfrak{S} := \mathscr{P}(\Omega)$, darauf die Gleichverteilung $\mathcal{P}$ und die Zufallsvariablen

$$X(\omega_1, .., \omega_n) = \omega_1 \quad , \quad Y(\omega_1, .., \omega_n) = \omega_n \quad , \quad \boldsymbol{\omega} \in \Omega$$

Offensichtlich sind sowohl $\mathcal{P}_X$ als auch $\mathcal{P}_Y$ gleichverteilt auf $\{0, 1\}$, sprich, X und Y besitzen das gleiche Verteilungsgesetz.

(ii) Die reellwertige, einfache Zufallsvariable

$$X = \sum_{j=1}^{n} \alpha_j \cdot 1_{A_j} \quad , \quad \alpha_j \in \mathbb{R} \ , \quad \{A_j\}_j : \text{disjunkte Zerlegung von } \Omega$$

nimmt die Werte $\{\alpha_1, .., \alpha_n\}$ an. Ihr Verteilungsgesetz $\mathcal{P}_X$ ist daher *konzentriert* auf $\{\alpha_1, .., \alpha_n\}$ und ergibt sich gemäß

$$\mathcal{P}_X(B) = \sum_{\alpha_j \in B} \mathcal{P}(A_j) \quad , \quad B \in \mathscr{B}(\mathbb{R})$$

Bemerkung: Ist X reell, so heißt X:

- *diskret* $:\Leftrightarrow$ $\mathcal{P}_X$ ist diskret.
- *absolut stetig* $:\Leftrightarrow$ $\mathcal{P}_X$ ist absolut stetig.
- *singulär stetig* $:\Leftrightarrow$ $\mathcal{P}_X$ ist singulär stetig.

(vgl. 1.2.7).

1.5.2 Definition: Identisch verteilte Zufallsgrößen

Es seien $X : (\Omega, \mathfrak{S}, \mathcal{P}) \to (S, \mathcal{S})$ und $Y : (\Omega', \mathfrak{S}', \mathcal{P}') \to (S, \mathcal{S})$ S-wertige Zufallsgrößen. Dann heißen X, Y *identisch verteilt* ($X \overset{\text{d.}}{=} Y$) falls $\mathcal{P}_X = \mathcal{P}'_Y$. Interpretationsgemäß, beschreiben X und Y den gleichen Vorgang.

1.6 Erwartungswert

1.6.1 Definition: Erwartungswert reeller Zufallsgrößen

Sei $X : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}$ eine reelle Zufallsgröße. Ist $X \geq 0$, so heißt

$$\mathbb{E}X := \int_\Omega X(\omega) \, d\mathcal{P}(\omega) \ \in \ [0, \infty]$$

Erwartungswert von X. Im allgemeinen Fall $X : \Omega \to \mathbb{R}$, besitzt X einen Erwartungswert wenn $\mathbb{E}\,|X| < \infty$, gegeben durch

$$\mathbb{E}X := \int_\Omega X(\omega) \, d\mathcal{P}(\omega)$$

1.6.2 Definition: Erwartungswert zufälliger Vektoren

Sei $X : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}^n$ ein n-dimensionaler zufälliger Vektor. Dann besitzt X den *Erwartungswert*

$$\mathbb{E}X := \left(\mathbb{E}X^1, \ldots, \mathbb{E}X^n\right)$$

insofern jede Komponente X^i einen Erwartungswert besitzt.

1.6.3 Transformationssatz für Erwartungswerte

Sei $X : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}$ eine reelle Zufallsgröße und $f : \mathbb{R} \to \mathbb{R}$ messbar. Dann gilt:

$$\mathbb{E}\,|f(X)| < \infty \quad \Leftrightarrow \quad \int_{-\infty}^{\infty} |f(t)| \ d\mathcal{P}_X(t) \ < \ \infty$$

Gegebenfalls gilt dann:

$$\mathbb{E}f(X) = \int_{-\infty}^{\infty} f(t) \ d\mathcal{P}_X(t)$$

(vgl. Transformationssatz der Maßtheorie).

Spezialfälle:

(i) Sei $\mathbb{E}|X| < \infty$. Dann ist

$$\mathbb{E}X = \int_{-\infty}^{\infty} t \, d\mathcal{P}_X(t)$$

(Spezialfall $f = \mathrm{Id}$). Insbesondere gilt für identisch verteilte X, Y: $\mathbb{E}X = \mathbb{E}Y$.

(ii) Ist $\mathcal{P}_X \ll \lambda^n$ mit Dichte p, so ist

$$\mathbb{E}X = \int_{\mathbb{R}^n} \mathbf{x} \, d\mathcal{P}_X(\mathbf{x}) = \int_{\mathbb{R}^n} \mathbf{x} \cdot p(\mathbf{x}) \, d\mathbf{x}$$

Allgemeiner:

$$\mathbb{E}f(X) = \int_{\mathbb{R}^n} f(\mathbf{x}) \cdot p(\mathbf{x}) \, d\mathbf{x}$$

(falls existent) für messbare $f : \mathbb{R}^n \to \mathbb{R}^m$.

(iii) Ist $Q \in \mathscr{B}(\mathbb{R}^n)$, $\lambda^n(Q) \neq 0$ eine beschränkte Borelmenge und

$$\mathcal{P}(X \in A) = \frac{\lambda^n(A \cap Q)}{\lambda^n(Q)} \quad , \quad A \in \mathscr{B}(\mathbb{R}^n)$$

so ist X gleichverteilt auf Q, mit Verteilungsdichte

$$p(\mathbf{y}) := \frac{d\mathcal{P}_X}{d\lambda^n}(\mathbf{y}) = \begin{cases} 1/\lambda^n(Q) & : \mathbf{y} \in Q \\ 0 & : \text{sonst} \end{cases}$$

und Erwartungswert

$$\mathbb{E}X = \int_{\mathbb{R}^n} \mathbf{y} \; p(\mathbf{y}) \, d\mathbf{y} = \frac{1}{\lambda^n(Q)} \int_Q \mathbf{y} \, d\mathbf{y}$$

(*Mittelpunkt von* Q).

(iv) X sei diskret, das heißt $X(\Omega) = \{x_1, x_2, ...\}$. Dann gilt

$$\mathbb{E}X = \sum_{i=1}^{\infty} x_i \cdot \mathcal{P}(X = x_i)$$

und allgemeiner

$$\mathbb{E}f(X) = \sum_{i=1}^{\infty} f(x_i) \cdot \mathcal{P}(X = x_i)$$

für messbare $f : \mathbb{R} \to \mathbb{R}$.

1.6.4 Alternative Berechnung von Erwartungswerten

Sei $X : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}$ eine reelle Zufallsvariable, $X \geq 0$ und es existiere $\mathbb{E}X$.

1. Es gilt:

$$\mathbb{E}X = \int_0^{\infty} \mathcal{P}(X > s) \, ds$$

2. Ist $g : (0, \infty) \to (0, \infty)$ streng monoton wachsend, differenzierbar, so ist

$$\mathbb{E}g(X) = \int_{g^{-1}(0)}^{g^{-1}(\infty)} \mathcal{P}(X > s) \cdot g'(s) \, ds$$

Beweis:

1.

$$\mathbb{E}X = \int_0^\infty t \; d\mathcal{P}_X(t) = \int_0^\infty \int_0^t ds \; d\mathcal{P}_X(t) = \int_0^\infty \int_s^\infty d\mathcal{P}_X(t) \; ds = \int_0^\infty \mathcal{P}(X \geq s) \; ds$$

$$\overset{\substack{\mathcal{P}(X \geq s) \text{ in} \\ \lambda^1\text{-fast allen } s \\ \text{stetig}}}{=} \int_0^\infty \mathcal{P}(X > s) \; ds$$

2.

$$\mathbb{E}g(X) = \int_0^\infty \mathcal{P}\left(g(X) > s\right) \; ds = \int_0^\infty \mathcal{P}\Big(X > \underbrace{g^{-1}(s)}_{=:t}\Big) \; ds = \int_{g^{-1}(0)}^{g^{-1}(\infty)} \mathcal{P}(X > t) \cdot g'(t) \; dt$$

□

Beispiel: Für $g(t) = t^q, \; q > 0$ ist

$$\mathbb{E}\left|X\right|^q = q \int_0^\infty t^{q-1} \mathcal{P}\left(\left|X\right| > t\right) \; dt$$

1.6.5 Korollar: Charakterisierung der Existenz des Erwartungswertes

Sei $X : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}$ ein reelle Zufallsvariable.

1. Es gelten die Ungleichungen

$$\sum_{n=1}^\infty \mathcal{P}\left(\left|X\right| \geq n\right) \leq \mathbb{E}\left|X\right| \leq 1 + \sum_{n=1}^\infty \mathcal{P}\left(\left|X\right| \geq n\right)$$

2. Nimmt X nur Werte in $\mathbb{N}$ an, so besteht sogar die Identität

$$\mathbb{E}X = \sum_{n=1}^\infty \mathcal{P}(X \geq n)$$

3. Es gilt genau dann $\mathbb{E}\left|X\right| < \infty$ wenn für ein $c > 0$ die Aussage

$$\sum_{n=1}^\infty \mathcal{P}\left(\left|X\right| \geq c \cdot n\right) < \infty \tag{1.6.5.1}$$

   richtig ist. In diesem Fall gilt dann (1.6.5.1) sogar für alle $c > 0$.

4. Für ein $p > 0$ gilt dann und nur dann $\mathbb{E}\left|X\right|^p < \infty$, wenn

$$\sum_{n=1}^\infty n^{p-1} \mathcal{P}\left(\left|X\right| \geq n\right) < \infty$$

Beweis:

1. Bekanntlich ist

$$\mathbb{E}\,|X| = \int_0^\infty \underbrace{\mathcal{P}(|X| \geq t)}_{=:G(t)}\ dt$$

Da die (messbare) Funktion $G : [0,\infty) \to [0,1]$ monoton fallend ist, gilt

$$G(\lceil t \rceil) \leq G(t) \leq G(\lfloor t \rfloor) \quad , \quad t \geq 0$$

so dass man abschätzen kann

$$\sum_{n=1}^\infty \underbrace{\int_{n-1}^{n} G(n)\ dt}_{\mathcal{P}(|X|\geq n)} = \int_0^\infty G(\lceil t \rceil)\ dt \leq \underbrace{\int_0^\infty G(t)\ dt}_{\mathbb{E}|X|} \leq \int_0^\infty G(\lfloor t \rfloor)\ dt = \underbrace{\int_0^1 G(0)\ dt}_{1} + \sum_{n=1}^\infty \underbrace{\int_n^{n+1} G(n)\ dt}_{\mathcal{P}(|X|\geq n)}$$

2. Nimmt X Werte in $\mathbb{N}$ an, das heißt $\mathcal{P}(|X| \geq t) = \mathcal{P}(|X| \geq \lceil t \rceil)$, so folgt

$$\mathbb{E}\,|X| = \int_{(0,\infty)} \mathcal{P}(|X| \geq t)\ dt = \int_{(0,\infty)} \mathcal{P}(|X| \geq \lceil t \rceil)\ dt = \sum_{n=1}^\infty \underbrace{\int_{(n-1,n]} \mathcal{P}(|X| \geq n)\ dt}_{\mathcal{P}(|X|\geq n)}$$

3. Zeigen zunächst: Gilt (1.6.5.1) für irgendein c_0, so gilt sie sogar für alle $c > 0$.

Tatsächlich, im Falle $c < c_0$ ist

$$\sum_{n=1}^\infty \mathcal{P}(|X| \geq c \cdot n) = \sum_{n=0}^\infty \underbrace{\mathcal{P}\Big(|X| \geq c_0 \cdot \frac{cn}{c_0}\Big)}_{\leq \mathcal{P}\left(|X| \geq c_0 \cdot \left\lfloor \frac{cn}{c_0} \right\rfloor\right)} \overset{(*)}{\leq} \left\lceil \frac{c_0}{c} \right\rceil \underbrace{\sum_{n=0}^\infty \mathcal{P}\left(|X| \geq c_0 n\right)}_{<\infty}$$

$$(*) \quad : \quad \#\left\{ m \in \mathbb{N} : \left\lfloor \frac{c(n+m)}{c_0} \right\rfloor = \left\lfloor \frac{cn}{c_0} \right\rfloor \right\} \leq \#\left\{ m \in \mathbb{N} : \frac{cm}{c_0} \leq 1 \right\} \leq \left\lceil \frac{c_0}{c} \right\rceil$$

und im Falle $c > c_0$ sowieso

$$\sum_{n=1}^\infty \underbrace{\mathcal{P}(|X| \geq c \cdot n)}_{\leq \mathcal{P}(|X|\geq c_0 \cdot n)} < \infty$$

Alternativer Beweis:

$$\infty > \frac{c_0}{c}\Big[\sum_{n=1}^\infty \underbrace{\mathcal{P}\left(|X| \geq c_0 n\right)}_{\mathcal{P}\left(\frac{|X|}{c_0} \geq n\right)} + 1\Big] \overset{(1)}{\geq} \frac{c_0}{c}\frac{\mathbb{E}\,|X|}{c_0} = \frac{\mathbb{E}\,|X|}{c} \overset{(1)}{\geq} \sum_{n=1}^\infty \mathcal{P}\left(\frac{|X|}{c} \geq n\right) = \sum_{n=1}^\infty \mathcal{P}\left(|X| \geq nc\right)$$

Zeigen nun: Gilt (1.6.5.1) für ein $c > 0$, so muss $\mathbb{E}\,|X| < \infty$. Tatsächlich gilt dann (1.6.5.1) insbesondere für $c = 1$, die Aussage folgt also aus (2).

Zu zeigen bleibt: Gilt $\mathbb{E}\,|X| < \infty$, so gilt (1.6.5.1) für irgendein $c > 0$. Doch dies folgt ebenfalls aus (2) (setze $c := 1$).

4. **Fall** $p \geq 1$**:** Analog zu vorhin gilt

$$p \cdot 0^{p-1} + \frac{p}{2^{p-1}} \sum_{n=2}^{\infty} \mathcal{P}(|X| \geq n) \cdot n^{p-1} \leq p \cdot 0^{p-1} + p \cdot \sum_{n=2}^{\infty} \mathcal{P}(|X| \geq n) \cdot \underbrace{(n-1)}_{\geq n/2}{}^{p-1}$$

$$= p \cdot \sum_{n=1}^{\infty} \mathcal{P}(|X| \geq n) \cdot (n-1)^{p-1} = p \cdot \int_0^{\infty} \mathcal{P}\left(|X| \geq \lceil s \rceil\right) \cdot \lceil s-1 \rceil^{p-1} \; ds$$

$$\leq \overbrace{\int_0^{\infty} \mathcal{P}(|X| \geq s) \cdot ps^{p-1} \; ds}^{\mathbb{E}|X|^p}$$

$$\leq p \cdot \int_0^{\infty} \mathcal{P}\left(|X| \geq \lfloor s \rfloor\right) \cdot \lfloor s+1 \rfloor^{p-1} \; ds = p \cdot \sum_{n=0}^{\infty} \mathcal{P}(|X| \geq n) \cdot (n+1)^{p-1} = p + \sum_{n=1}^{\infty} \mathcal{P}(|X| \geq n) \cdot \underbrace{(n+1)}_{\leq 2n}{}^{p-1}$$

$$\leq p + 2^{p-1} p \cdot \sum_{n=1}^{\infty} \mathcal{P}(|X| \geq n) \cdot n^{p-1}$$

Fall $0 < p < 1$**:**

$$p \cdot \sum_{n=1}^{\infty} \mathcal{P}(|X| \geq n) \cdot n^{p-1} = p \cdot \int_0^{\infty} \mathcal{P}\left(|X| \geq \lceil s \rceil\right) \cdot \lceil s \rceil^{p-1} \; ds$$

$$\leq \overbrace{\int_0^{\infty} \mathcal{P}(|X| \geq s) \cdot ps^{p-1} \; ds}^{\mathbb{E}|X|^p}$$

$$\leq \underbrace{p \int_0^1 s^{p-1} \; ds}_{\substack{<\infty \\ \text{da } p>0}} + p \cdot \int_1^{\infty} \mathcal{P}\left(|X| \geq \lfloor s \rfloor\right) \cdot \lfloor s \rfloor^{p-1} \; ds = \text{const} + p \cdot \sum_{n=1}^{\infty} \mathcal{P}(|X| \geq n) \cdot n^{p-1}$$

□

1.6.6 Eigenschaften des Erwartungswertes

Die Zufallsvariablen $X, Y : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}$ besitzen einen Erwartungswert. Dann gilt:

1. **Linearität:** Für $\alpha, \beta \in \mathbb{R}$ besitzt auch $\alpha X + \beta Y$ einen Erwartungswert:

$$\mathbb{E}(\alpha X + \beta Y) = \alpha \mathbb{E} X + \beta \mathbb{E} Y$$

2. **Monotonie:** Ist $X \leq Y$ fast sicher, so ist $\mathbb{E}X \leq \mathbb{E}Y$.

3. $|\mathbb{E}X| \leq \mathbb{E}\,|X|$

4. Für c : const ist $\mathbb{E}c = c$.

5. Für $A \in \mathfrak{S}$ ist $\mathbb{E}1_A = \mathcal{P}(A)$

6. $\forall\, \varepsilon > 0\ :\ \exists\, \delta > 0 : \forall\, \mathcal{P}(A) < \delta : |\mathbb{E}(X1_A)| < \varepsilon$
Erläuterung: Folgt aus 1.2.4 und der Tatsache dass $|X|$ als Dichte ein endliches, bzgl. $\mathcal{P}$ absolut stetiges, Maß erzeugt.

7. Gehen $X_n \underset{\text{punktweise}}{\overset{n\to\infty}{\longrightarrow}} X$ fast überall und $|X_n| \leq Y$ mit $\mathbb{E}Y < \infty$, so geht auch
$$\mathbb{E}X_n \overset{n\to\infty}{\longrightarrow} \mathbb{E}X$$
(Satz von Lebesgue).

8. Sind $0 \leq X_1 \leq X_2 \leq \ldots$ so ist
$$\lim_{n\to\infty} \mathbb{E}X_n = \mathbb{E} \lim_{n\to\infty} X_n$$
(vgl. Satz der monotonen Konvergenz).

Beispiele:

(i) Ist X Cauchy-verteilt (1.2.8.1), das heiß
$$\mathcal{P}(X \leq t) = \frac{1}{\pi} \int_{-\infty}^{t} \frac{dx}{1+x^2}$$
so ist
$$\mathbb{E}\,|X| = \frac{2}{\pi} \int_{0}^{\infty} \frac{x}{1+x^2} dx = \infty$$
das heißt X besitzt <u>keinen</u> Erwartungswert.

(ii) Ist X gleichverteilt auf $\{x_1, .., x_n\} \subseteq \mathbb{R}$ so gilt
$$\mathbb{E}X = \frac{1}{N} \sum_{i=1}^{n} x_i$$

(iii) Ist $X \sim \mathcal{N}_{\mu,\sigma^2}$ Normalverteilt, so ist
$$\mathbb{E}X = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \mu$$

1.7 Höhere Momente

1.7.1 Definition: n-tes Moment für reelle Größen

Sei $X : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}$ eine Zufallsvariable und $n \geq 1$ mit $\mathbb{E}\,|X|^n < \infty$. Dann existiert $\mathbb{E}X^n$ und heißt *n-tes Moment* von X.

Bemerkung: Besitzt X die Verteilungsdichte p bzgl. λ^1, so ist
$$\mathbb{E}X^n = \int_{-\infty}^{\infty} t^n \cdot p(t)\, dt$$

1.7.2 Definition: n-tes Moment für zufällige Vektoren

Der zufällige Vektor $X : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}^m$ *besitzt das n-te Moment*

$$\mathbb{E}(\underbrace{X \otimes \cdots \otimes X}_{\times n}) := (\mathbb{E}(X_{i_1} \cdot .. \cdot X_{i_n}))_{i_1,..,i_n=1}^{m}$$

insofern $\mathbb{E}\,|X_{i_1} \cdot .. \cdot X_{i_n}| < \infty \;\; \forall\; i_k = 1, .., m.$

1.7.3 Satz über n-te Momente

Sei X eine reelle Zufallsvariable und $n \in \mathbb{N}$.

1. Besitzt X ein n-tes Moment, so besitzt X für $1 \leq m \leq n$ auch ein m-tes Moment.
2. Besitzt X ein n-tes Moment, so besitzt auch die zentrierte Version $(X - \mathbb{E}X)$ ein n-tes Moment.
3. Besitzt X einen Erwartungswert und $(X - \mathbb{E}X)$ ein n-tes Moment, so besitzt auch X ein n-tes Moment.

Anmerkung: Eine lange Zeit unbeantwortete Frage war:

Sind X, Y reelle Zufallsvariablen mit $\mathbb{E}\,|X|^n < \infty,\; \mathbb{E}\,|Y|^n < \infty$ und $\mathbb{E}X^n = \mathbb{E}Y^n \;\; \forall\; n \in \mathbb{N}$, folgt dann auch $X \overset{\text{d.}}{=} Y$?

Es stellt sich heraus, dass die Antwort negativ ausfällt.

Beispiel: Die standardnormalverteilte Zufallsvariable $X \sim \mathcal{N}_{0,1}$ besitzt die Momente

$$\mathbb{E}X^n = \begin{cases} 0 & : n \text{ ungerade} \\ (n-1)! & : n \text{ gerade} \end{cases}$$

1.8 Varianz & Kovarianz

1.8.1 Definition: Kovarianz reeller Zufallsgrößen

Die reellen Zufallsvariablen $X, Y : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}$ besitzen jeweils 2. Momente. Dann heißt

$$\mathrm{Cov}(X, Y) := \mathbb{E}\left[(X - \mathbb{E}X)\,(Y - \mathbb{E}Y)\right]$$

Kovarianz von X und Y. Ferner heißt $\mathbb{V}X := \mathrm{Cov}(X, X)$ *Varianz* von X.

Bemerkungen:

(i) Die Kovarianz existiert als solche tatsächlich, denn

$$\mathbb{E}\,|X - \mathbb{E}X| \cdot |Y - \mathbb{E}Y| \leq \left(\mathbb{E}\,|X - \mathbb{E}X|^2\right)^{\frac{1}{2}} \left(\mathbb{E}\,|Y - \mathbb{E}Y|^2\right)^{\frac{1}{2}}$$

(ii) Es gilt die Darstellung

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

(iii) Mit $f : \mathbb{R}^2 \to \mathbb{R}^2,\; f(t, s) = (t - \mathbb{E}X)(s - \mathbb{E}Y)$ und

$$\mathrm{Cov}(X, Y) = \int_{\Omega} f(X, Y)\; d\mathcal{P} = \int_{\mathbb{R}^2} f\; d\mathcal{P}_{(X,Y)}$$

wird ersichtlich, dass $\mathrm{Cov}(X, Y)$ nur vom Verteilungsgesetz des Vektors (X, Y) abhängt.

(iv) Ist $p(x)$ die Verteilungsdichte von X bzgl. λ^1, so ist

$$\mathbb{V}X = \int_{-\infty}^{\infty} (t - \mathbb{E}X)^2 \cdot p(t)\; dt$$

1.8.2 Definition: Kovarianz zufälliger Vektoren

Für n- & m-dimensionale, reelle, zufällige Vektoren X, Y heißt

$$\operatorname{Cov}(X,Y) := \left(\operatorname{Cov}(X^i, Y^j)\right)_{i,j=1}^{n,m} \in \mathbb{R}^{n\times m}$$

Kovarianz (*Kovarianzmatrix*) von X und Y. Speziell heißt

$$\operatorname{Cov}(X) := \operatorname{Cov}(X,X) \in \mathbb{R}^{n\times n}$$

Kovarianzmatrix von X.

Bemerkungen:

(i) Für reelle, zufällige Vektoren X, Y gilt die Darstellung

$$\operatorname{Cov}(X,Y) = \mathbb{E}\left[(X-\mathbb{E}X)\otimes(Y-\mathbb{E}Y)\right] = \mathbb{E}(X\otimes Y) - (\mathbb{E}X)\otimes(\mathbb{E}Y)$$

(ii) Die Kovarianz des Vektors X ist das 2. Moment des zentrierten Vektors $(X-\mathbb{E}X)$.

(iii) Die Kovarianz des Vektors X existiert genau dann wenn X oder äquivalent dazu jede Komponente X^k zweites Moment besitzt.

(iv) Die Kovarianz $\operatorname{Cov}(\cdot,\cdot)$ ist linear in jeweils jeder Komponente.

(v) Sind X, Y jeweils n & m-dimensionale Zufallsvektoren und $N : \mathbb{R}^n \to \mathbb{R}^{n'}$, $M : \mathbb{R}^m \to \mathbb{R}^{m'}$ linear, so gilt

$$\operatorname{Cov}(NX, MY) = N\operatorname{Cov}(X,Y)M^T$$

Insbesondere für $\mathbf{k}\in\mathbb{R}^n$

$$\underbrace{\operatorname{Cov}(\langle\mathbf{k},X\rangle)}_{\mathbb{V}\langle\mathbf{k},X\rangle} = \langle\operatorname{Cov}(X)\mathbf{k},\mathbf{k}\rangle \tag{1.8.2.1}$$

(vi) Für reellen, n-dimensionalen zufälligen Vektor X ist $\operatorname{Cov}(X)$ symmetrisch und nicht-negativ definit, denn für $\mathbf{k}\in\mathbb{R}^n$ gilt

$$\langle\operatorname{Cov}(X)\mathbf{k},\mathbf{k}\rangle = \mathbb{V}\langle\mathbf{k},X\rangle \geq 0$$

1.8.3 Ungleichung von Markov

Ist $X : (\Omega,\mathfrak{S},\mathcal{P}) \to \mathbb{R}$ eine reelle Zufallsvariable und $c > 0$, so gilt

$$\mathcal{P}(|X|\geq c) \leq \frac{\mathbb{E}|X|}{c}$$

Beweis: Folgt direkt aus 1.6.4(1).

Spezialfälle:

(i) Ist $\varphi : \mathbb{R}\to[0,\infty)$ gerade und nicht-fallend auf $[0,\infty)$, so gilt

$$\mathcal{P}(|X|\geq c) \leq \frac{\mathbb{E}\varphi(X)}{\varphi(c)} \quad \forall\, c>0$$

(ii) Durch $\varphi(t) := t^2$ wird ersichtlich

$$\mathcal{P}(|X-\mathbb{E}X|\geq c) \leq \frac{\overbrace{\mathbb{E}|X-\mathbb{E}X|^2}^{\mathbb{V}X}}{c^2} \qquad \text{(Tschebyschow Ungleichung)}$$

Insbesondere ist $\mathbb{V}X = 0$ genau dann wenn X fast-sicher konstant ist.

2 Unabhängigkeit

2.1 Unendliche Produkträume

2.1.1 Definition: Zylindermenge & unendlicher Produktraum

Seien $(\Omega_i, \mathfrak{S}_i)_{i\in I}$ messbare Räume. Dann definiert man

$$\Omega := \bigtimes_{i\in I} \Omega_i := \{(\omega_i)_{i\in I} : \omega_i \in \Omega_i\}$$

Man nennt $Z \subseteq \Omega$ eine *Zylindermenge*, falls es paarweise verschiedene Indizes $i_1, .., i_k \in I$ und eine Menge $B \in \bigotimes_{k=1}^n \mathfrak{S}_{i_k}$ gibt, so dass

$$Z = \{(\omega_i)_{i\in I} \in \Omega : (\omega_{i_1}, .., \omega_{i_n}) \in B\}$$

Man schreibt $\mathcal{Z}(\Omega)$ für den Raum aller Zylindermengen auf Ω. Beachte dass $\mathcal{Z}(\Omega)$ eine Algebra ist. Schließlich setzt man

$$\bigotimes_{i\in I} \mathfrak{S}_i := \sigma(\mathcal{Z}(\Omega)) = \sigma\{Z \subseteq \Omega : Z \text{ Zylindermenge}\}$$

und nennt

$$\left(\bigtimes_{i\in I} \Omega_i, \bigotimes_{i\in I} \mathfrak{S}_i \right)$$

den (*unendlichen* falls $|I| = \infty$) *Produktraum* von $(\Omega_i, \mathfrak{S}_i)_{i\in I}$.

Bemerke: Ist die Indexmenge I endlich, so stimmen obige Definitionen mit den herkömmlichen überein.

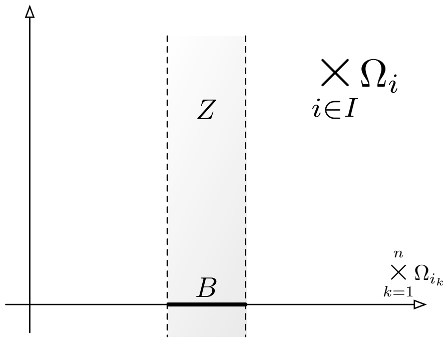

Abbildung 2: Zur Definition von Zylindermengen: Eine Zylindermenge ist stets durch Bedingungen an nur endlich vielen Komponenten ihrer Elemente definiert.

Beispiele:

(i) Ist I abzählbar, so enthält $\bigotimes_{k=1}^{\infty} \mathfrak{S}_k$ insbesondere alle Rechtecke

$$\bigtimes_{k=1}^{\infty} B_k \quad , \quad B_k \in \mathfrak{S}_k$$

denn

$$\bigtimes_{k=1}^{\infty} B_k = \bigcap_{n=1}^{\infty} \underbrace{(B_1 \times \cdots \times B_n \times \Omega_{n+1} \times \dots)}_{\text{Zylindermenge auf } \Omega} \in \mathfrak{S}$$

(ii) Sei $\Omega_1 = \Omega_2 = \cdots = \{0, 1\}$ und $\mathfrak{S}_1 = \mathfrak{S}_2 = \cdots = \mathscr{P}(\{0, 1\})$. Dann ist

$$\Omega := \underbrace{\bigtimes_{k=1}^{\infty} \Omega_i}_{\{0,1\}^{\mathbb{N}}} = \Big\{ \omega = (\omega_1, \omega_2, ..) : \omega_k \in \{0, 1\} \Big\}$$

Dabei gilt:

$$\bigotimes_{k=1}^{\infty} \mathfrak{S}_k \subsetneq \mathscr{P}(\Omega)$$

2.1.2 Satz: Erzeugung des unendlichen Produktraumes

Seien $(\Omega_i, \mathfrak{S}_i)_{i\in I}$ messbare Räume, dazu $\Omega := \bigtimes_{i\in I} \Omega_i$. Zu $B \in \bigotimes_{k=1}^{n} \mathfrak{S}_{i_k}$ sei

$$Z(B) := \{(\omega_i)_{i\in I} \in \Omega : (\omega_{i_1}, .., \omega_{i_n}) \in B\}$$

die entsprechende Zylindermenge. Dann gilt:

$$\sigma \underbrace{\left\{ Z(B) : B \in \bigotimes_{k=1}^{n} \mathfrak{S}_{i_k},\ \{i_1, .., i_n\} \subseteq I \right\}}_{\mathcal{Z}} = \sigma \underbrace{\left\{ Z\left(\bigtimes_{k=1}^{n} B_{i_k} \right) : B_{i_k} \in \mathfrak{S}_{i_k},\ \{i_1, .., i_n\} \subseteq I \right\}}_{\mathcal{Z}'}$$

das heißt $\bigotimes_{i\in I} \mathfrak{S}_i$ wird schon allein durch die *rechteckigen* Zylindermengen $Z(B_{i_1} \times .. \times B_{i_n})$ erzeugt.

Beweis: Seien zunächst $\{i_1, .., i_n\} \subseteq I$ fest. Seien $\mathcal{Z}_{i_1..i_n}$ und $\mathcal{Z}'_{i_1..i_n}$ jeweils die Menge aller Zylindermengen $Z(B),\ B \in \bigotimes_{k=1}^{n} \mathfrak{S}_{i_k}$ und die Menge aller rechteckigen Zylindermengen $Z(B_{i_1} \times .. \times B_{i_n}),\ B_{i_k} \in \mathfrak{S}_{i_k}$ auf Ω.
Behauptung: $\mathcal{Z}_{i_1..i_n} = \sigma(\mathcal{Z}'_{i_1..i_n})$.

Beweis: Offensichtlich ist $\mathcal{Z}_{i_1..i_n}$ eine σ-Algebra auf Ω mit $\mathcal{Z}'_{i_1..i_n} \subseteq \mathcal{Z}_{i_1..i_n}$. Zu zeigen wäre also: Ist $\mathcal{A} \subseteq \mathscr{P}(\Omega)$ eine weitere σ-Algebra mit $\mathcal{Z}'_{i_1..i_n} \subseteq \mathcal{A}$, so ist auch $\mathcal{Z}_{i_1..i_n} \subseteq \mathcal{A}$. Sei

$$\mathcal{B}_{\mathcal{A}} := \left\{ B \in \bigotimes_{k=1}^{n} \mathfrak{S}_{i_k} : Z(B) \in \mathcal{A} \right\}$$

Dann ist $\mathcal{B}_{\mathcal{A}}$ eine σ-Algebra auf Ω. Wegen $\mathcal{Z}'_{i_1..i_n} \subseteq \mathcal{A}$, ist $\bigtimes_{k=1}^{n} \mathfrak{S}_{i_k} \subseteq \mathcal{B}_{\mathcal{A}}$, daher

$$\bigotimes_{k=1}^{n} \mathfrak{S}_{i_k} = \sigma\left(\bigtimes_{k=1}^{n} \mathfrak{S}_{i_k} \right) \subseteq \mathcal{B}_{\mathcal{A}}$$

das heißt $\mathcal{Z}_{i_1..i_n} \subseteq \mathcal{A}$.

Wegen

$$\mathcal{Z} = \bigcup_{\substack{\{i_1,..,i_n\}\subseteq I \\ n\in\mathbb{N}}} \mathcal{Z}_{i_1..i_n} \quad , \quad \mathcal{Z}' = \bigcup_{\substack{\{i_1,..,i_n\}\subseteq I \\ n\in\mathbb{N}}} \mathcal{Z}'_{i_1..i_n}$$

gilt nun

$$\sigma(\mathcal{Z}') = \sigma\left[\bigcup_{\substack{\{i_1,..,i_n\}\subseteq I \\ n\in\mathbb{N}}} \underbrace{\sigma(\mathcal{Z}'_{i_1..i_n})}_{\mathcal{Z}_{i_1..i_n}} \right] = \sigma(\mathcal{Z})$$

□

2.1.3 Satz: Messbarkeit in unendlichen Produkträumen

Seien $(\Omega, \mathfrak{S})$ und $(S_i, \mathcal{S}_i)_{i\in I}$ messbare Räume und $X_i : (\Omega, \mathfrak{S}) \to (S_i, \mathcal{S}_i),\ i \in I$ messbare Abbildungen. Dann ist die Abbildung

$$(X_i)_{i\in I} : (\Omega, \mathfrak{S}) \to \left(\bigtimes_{i\in I} S_i, \bigotimes_{i\in I} \mathcal{S}_i \right) \quad , \quad (X_i)_{i\in I}(\omega) := (X_i(\omega))_{i\in I}$$

messbar.

Beweis: Sei $\mathcal{Z}$ die Menge aller Zylindermengen auf $\bigtimes_{i\in I} S_i$. Wegen $\bigotimes_{i\in I} \mathcal{S}_i = \sigma(\mathcal{Z})$ genügt es zu zeigen

$$X^{-1}(\mathcal{Z}) \subseteq \mathfrak{S}$$

Tatsächlich, ist

$$Z = \left\{ (s_i)_{i\in I} \in \bigtimes_{i\in I} S_i : (s_{i_1}, .., s_{i_n}) \in B \right\} \quad , \quad B \in \bigotimes_{k=1}^{n} \mathcal{S}_{i_k}$$

eine Zylindermenge aus $\mathcal{Z}$, so folgt

$$X^{-1}(Z) = \{\omega \in \Omega : (X_{i_1}(\omega), .., X_{i_n}(\omega)) \in B\} = \underbrace{(X_{i_1}, .., X_{i_n})^{-1}(B)}_{\substack{\in\mathfrak{S} \\ \text{da } (X_{i_1},..,X_{i_n}) \\ \text{messbar}}} \in \mathfrak{S}$$

□

2.1.4 Theorem: Existenz eines (unendlichen) Produktmaßes

Seien $(\Omega_i, \mathfrak{S}_i, \mathcal{P}_i)_{i\in I}$ Wahrscheinlichkeitsräume, dazu der Produktraum

$$(\Omega, \mathfrak{S}) := \left(\bigtimes_{i\in I} \Omega_i, \bigotimes_{i\in I} \mathfrak{S}_i \right)$$

Dann existiert ein eindeutig bestimmtes Wahrscheinlichkeitsmaß $\mathcal{P}$ auf $(\Omega, \mathfrak{S})$ mit

$$\mathcal{P}\left(\omega \in \Omega : \omega_{i_1} \in B_{i_1}, \dots, \omega_{i_n} \in B_{i_n}\right) = \prod_{k=1}^{n} \mathcal{P}_{i_k}(B_{i_k}) \quad \forall \underbrace{i_1, .., i_n}_{\substack{\text{paarweise} \\ \text{verschieden}}} \in I,\ B_{i_k} \in \mathfrak{S}_{i_k}$$

Dieses Wahrscheinlichkeitsmaß $\mathcal{P}$ heißt (*unendliches*) *Produktmaß*. Man schreibt auch

$$\mathcal{P} = \bigotimes_{i\in I} \mathcal{P}_i$$

Beweis: [2]

Bemerkung: Ist $I = \mathbb{N}$, so gilt sogar

$$\mathcal{P}\Big(\underbrace{\bigtimes_{k=1}^{\infty} B_k}_{\substack{\in\mathfrak{S} \\ \text{nach Bsp.} \\ 2.1.1(i)}} \Big) = \prod_{k=1}^{\infty} \mathcal{P}_k(B_k) \quad \forall\ B_k \in \Omega_k,\ k \in \mathbb{N}$$

denn

$$\mathcal{P}\left(\bigtimes_{k=1}^{\infty} B_k\right) = \lim_{n\to\infty} \mathcal{P}\big(\overbrace{B_1 \times \dots \times B_n \times \Omega_{n+1} \times \dots}^{\searrow_{n\to\infty} \bigtimes_{k=1}^{\infty} B_k}\big) = \lim_{n\to\infty} \prod_{k=1}^{n} \mathcal{P}_k(B_k) = \prod_{k=1}^{\infty} \mathcal{P}_k(B_k)$$

Beispiele:

(i) Betrachten

$$(\Omega, \mathfrak{S}, \mathcal{P}) := \left([0,1]^{\mathbb{N}},\ \bigotimes_{k=1}^{\infty} \mathscr{B}([0,1]),\ \bigotimes_{k=1}^{\infty} \lambda^1 \right) \quad ,$$

dazu das (unendlich-dimensionale) Rechteck

$$B := \bigtimes_{k=1}^{\infty} [0, 1 - \alpha_k] \quad , \quad \alpha_k \in (0,1)$$

Dann gilt

$$\mathcal{P}(B) = \prod_{k=1}^{\infty} \lambda^1([0, 1-\alpha_k]) = \prod_{k=1}^{\infty} (1-\alpha_k)$$

Insbesondere gilt

$$\mathcal{P}(B) > 0 \quad \Leftrightarrow \quad \sum_{k=1}^{\infty} \alpha_k < \infty$$

und speziell

$$\mathcal{P}\left(\bigtimes_{k=1}^{\infty} \left[0, 1 - \frac{1}{k}\right]\right) = 0$$

(ii) Betrachten ein unendlich oft mal wiederholtes Münzwurfexperiment, modelliert durch

$$(\Omega, \mathfrak{S}, \mathcal{P}) := \left(\{0,1\}^{\mathbb{N}}, \bigotimes_{k=1}^{\infty} \mathscr{P}(\{0,1\}), \bigotimes_{k=1}^{\infty} \frac{\delta_0 + \delta_1}{2}\right)$$

Dann ist z.B. die Wahrscheinlichkeit stets *Kopf* zu bekommen, gegeben durch

$$\mathcal{P}(\{1, 1, ...\}) = \prod_{k=1}^{\infty} \underbrace{\frac{1}{2}(\delta_0 + \delta_1)(\{1\})}_{\frac{1}{2}} = \lim_{n \to \infty} \frac{1}{2^n} = 0$$

Dagegen ist die Wahrscheinlichkeit unendlich oft Kopf zu bekommen gleich 1!

2.1.5 Alternativer Zugang zu (unendlichen) Produkträumen

Gegeben seien die messbaren Räume $(\Omega_i, \mathfrak{S}_i)_{i \in I}$, dazu

$$\Omega := \bigtimes_{i \in I} \Omega_i := \left\{ f : I \to \bigcup_{i \in I} \Omega_i,\ f(i) \in \Omega_i \right\} = \{(\omega_i)_{i \in I} : \omega_i \in \Omega_i\}$$

Für endliche $J = \{i_1, .., i_n\} \subseteq I$ sei definiert

$$\Pi_J : \Omega \to \bigtimes_{k=1}^{n} \Omega_{i_k} \quad , \quad \Pi_J(\omega) := (\omega_{i_k})_{k=1}^{n}$$

Dann ist

$$\mathfrak{S} := \bigotimes_{i \in I} \mathfrak{S}_i := \sigma\left\{ \Pi_J^{-1}(B) : B \in \bigtimes_{k=1}^{n} \mathfrak{S}_{i_k},\ J = \{i_1, .., i_k\} \subseteq I \right\} \overset{(2.1.2)}{=} \sigma\left\{ \Pi_{\{i\}}^{-1}(B) : B \in \mathfrak{S}_i,\ i \in I \right\} \qquad (2.1.5.1)$$

genau die Produkt-σ-Algebra der $\mathfrak{S}_i$. Dabei sind per Konstruktion alle

$$\Pi_J : (\Omega, \mathfrak{S}) \to \left(\bigtimes_{k=1}^{n} \Omega_{i_k}, \bigotimes_{k=1}^{n} \mathfrak{S}_{i_k} \right) \quad , \quad J = \{i_1, .., i_n\} \subseteq I$$

messbar.

Theorem (2.1.4) lautet nun wie folgt:
Es existiert ein eindeutig bestimmtes Wahrscheinlichkeitsmaß $\mathcal{P}$ auf $(\Omega, \mathfrak{S})$ mit

$$\mathcal{P} \circ \Pi_J^{-1} = \bigotimes_{k=1}^{n} \mathcal{P}_{i_k} \quad \forall\ J = \{i_1, .., i_n\} \subseteq I$$

das heißt

$$\mathcal{P}\left(\{\omega \in \Omega : \omega_{i_1} \in B_{i_1}, \ldots, \omega_{i_n} \in B_{i_n}\}\right) = \prod_{k=1}^{n} \mathcal{P}_{i_k}(B_{i_k}) \quad \forall\ B_{i_k} \in \mathfrak{S}_{i_k}$$

Beweisidee:

- Zu endlichem $J = \{i_1, .., i_n\} \subseteq I$ setze
$$\mathfrak{S}_J := \left\{ \Pi_J^{-1}(B) : B \in \bigotimes_{k=1}^{n} \mathfrak{S}_{i_k} \right\} = \mathcal{Z}_{i_1,..,i_n}$$
(vgl. Beweis von 2.1.2) und
$$\mathfrak{S}_0 := \bigcup_{\substack{J \subseteq I \\ \text{endlich}}} \mathfrak{S}_J$$
Dann ist $\mathfrak{S}_0$ eine Algebra und $\mathfrak{S} = \sigma(\mathfrak{S}_0)$. Definiere nun die Mengenfunktion $\mathcal{P}_0 : \mathfrak{S}_0 \to [0,1]$ gemäß
$$\mathcal{P}_0\left(\Pi_J^{-1}(B)\right) := \bigotimes_{k=1}^{n} \mathcal{P}_{i_k}(B) \quad , \quad J = {}_{i_1,..,i_n} \subseteq I, \; B \in \bigotimes_{k=1}^{n} \mathfrak{S}_{i_k}$$
Bemerke: Zwar sind J, B nicht eindeutig durch $\Pi_J^{-1}(B)$ bestimmt, doch ist der Wert $\mathcal{P}_0(\Pi_J^{-1}(B))$ tatsächlich unabhängig von deren speziellen Wahl, insbesondere wohldefiniert. Ferner kann gezeigt werden, dass $\mathcal{P}_0$ additiv ist auf $\mathfrak{S}_0$.
- Darauf folgend zeigt man, dass $\mathcal{P}_0$ sogar σ-Additiv ist, indem man für $Z_1 \supseteq Z_2 \supseteq \dots \in \mathfrak{S}_0$ mit $\inf_j \mathcal{P}_0(Z_j) > 0$ zeigt, dass
$$\bigcap_{j=1}^{\infty} Z_j \neq \emptyset$$
- Als σ-additive, endliche Mengenfunktion mit $\mathcal{P}_0(\emptyset) = 0$, kann $\mathcal{P}_0$ eindeutig zu einem Maß $\mathcal{P}$ auf $\mathfrak{S}$ fortgesetzt werden.

Bemerke: Die Einzelprojektionen $\Pi_{\{i\}}$, $i \in I$ sind unabhängige Zufallsvariablen mit Verteilungsgesetz $\mathcal{P}_i$ (siehe später 2.4.1).

2.2 Unabhängigkeit von Mengen & Mengensystemen

2.2.1 Definition: Unabhängige Mengen

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $(A_i)_{i\in I} \subseteq \mathfrak{S}$ eine Familie messbarer Mengen. Dann heißen $(A_i)_{i\in I}$ *unabhängig* :⇔
$$\forall\, I_0 \subseteq I,\; I_0 \text{ endlich} : \mathcal{P}\left(\bigcap_{i\in I_0} A_i\right) = \prod_{i \in I_0} \mathcal{P}(A_i)$$
Sie heißen *paarweise unabhängig* :⇔
$$\mathcal{P}(A_i \cap A_j) = \mathcal{P}(A_i) \cdot \mathcal{P}(A_j) \;\; \forall\, i \neq j \in I$$

Bemerke: Aus Unabhängigkeit folgt paarweise Unabhängigkeit. Jedoch gilt die Umkehrung im Allgemeinen nicht!

Beispiele:

(i) Zwei Mengen $A, B \in \mathfrak{S}$ sind unabhängig genau dann wenn $\mathcal{P}(A \cap B) = \mathcal{P}(A)\mathcal{P}(B)$.

(ii) Sind $(A_k)_{k\in\mathbb{N}} \subseteq \mathfrak{S}$ unabhängig (beachte: abzählbar viele), so gilt
$$\mathcal{P}\left(\bigcap_{k\in\mathbb{N}} A_k\right) = \prod_{k=1}^{\infty} \mathcal{P}(A_k)$$
Erläuterung: Die Folge $B_n := \bigcap_{k=1}^{n} A_k$ ist monoton fallend und $\bigcap_{n\in\mathbb{N}} B_n = \bigcap_{k\in\mathbb{N}} A_k$. Aufgrund der Stetigkeit des W-Maßes von oben, gilt
$$\mathcal{P}\left(\bigcap_{k\in\mathbb{N}} A_k\right) = \lim_{n\to\infty} \mathcal{P}(B_n) = \lim_{n\to\infty} \prod_{k=1}^{n} \mathcal{P}(A_k) = \prod_{k=1}^{\infty} \mathcal{P}(A_k)$$

2.2.2 Definition: Unabhängige Mengensysteme

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $(\mathcal{E}_j)_{j\in J}$ eine Familie von Mengensystemen $\mathcal{E}_j \subseteq \mathfrak{S}$. Dann heißen $(\mathcal{E}_j)_{j\in J}$ *unabhängig* $:\Leftrightarrow$

$$\forall\ E_j \in \mathcal{E}_j,\ j \in J\ :\ (E_j)_{j\in J} \text{ unabhängig} \tag{2.2.2.1}$$

Bemerkungen:

(i) Äquivalent zu (2.2.2.1) wäre

$$\forall\ J_0 \subseteq J,\ J_0 \text{ endlich} : \forall\ E_j \in \mathcal{E}_j,\ j \in J_0 \quad : \quad \mathcal{P}\Big(\bigcap_{j\in J_0} E_j\Big) = \prod_{j\in J_0} \mathcal{P}(E_j)$$

(ii) Ebenso äquivalent zu (2.2.2.1) wäre

$$\forall\ J_0 \subseteq J,\ J_0 \text{ abzählbar} : \forall\ E_j \in \mathcal{E}_j,\ j \in J_0 \quad : \quad \mathcal{P}\Big(\bigcap_{j\in J_0} E_j\Big) = \prod_{j\in J_0} \mathcal{P}(E_j)$$

(vgl. 2.2.1 Bsp. (ii))

(iii) Die Unabhängigkeit der $(\mathcal{E}_j)_{j\in J}$ macht keine Aussage über die Unabhängigkeit der Mengen in einem einzelnen $\mathcal{E}_j$.

Beispiel: $\Omega = [0,1]^2$, $\mathfrak{S} = \mathscr{B}[0,1]^2$ und $\mathcal{P}$ die Gleichverteilung auf Ω. Dann sind

$$\mathcal{E}_1 := \{A \times [0,1] : A \in \mathscr{B}[0,1]\} \quad , \quad \mathcal{E}_2 := \{[0,1] \times B : B \in \mathscr{B}[0,1]\}$$

unabhängig (siehe Abb. 3).

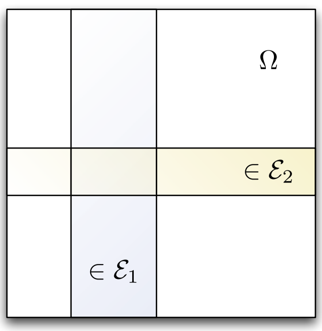

Abbildung 3: Zur Unabhängigkeit von Mengen(systemen).

2.2.3 Satz über unabhängige Mengensysteme

Seien $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $\mathcal{E}_1, .., \mathcal{E}_n \subseteq \mathfrak{S}$ eine endliche Familie von Mengensystemen mit einer von folgenden Eigenschaften:

1. $\forall\ j = 1, .., n\ :\ \Omega \in \mathcal{E}_j$

2. $\forall\ j = 1, .., n\ :\ \exists$ disjunkte $E_j^k \in \mathcal{E}_j,\ k \in \mathbb{N} : \Omega = \biguplus_{k=1}^{\infty} E_j^k$

3. $\forall\ j = 1, .., n\ :\ \exists\ E_j^1 \subseteq E_j^2 \subseteq \cdots : \bigcup_{k=1}^{\infty} E_j^k = \Omega$ das heißt $E_j^k \underset{k\to\infty}{\nearrow} \Omega$

Dann sind $(\mathcal{E}_j)_{j=1}^n$ unabhängig $\Leftrightarrow$

$$\mathcal{P}\Bigg(\bigcap_{j=1}^n E_j\Bigg) = \prod_{j=1}^n \mathcal{P}(E_j) \quad \forall\ E_j \in \mathcal{E}_j$$

Spezialfall: Sind $\mathfrak{S}_1, .., \mathfrak{S}_n$ σ-algebren, so sind diese unabhängig $\Leftrightarrow$

$$\mathcal{P}\Bigg(\bigcap_{j=1}^n A_j\Bigg) = \prod_{j=1}^n \mathcal{P}\left(A_j\right) \quad \forall\ A_j \in \mathfrak{S}_j$$

2.2.4 Eigenschaften unabhängiger Mengensysteme

Sei $(\mathcal{E}_j)_{j\in J}$ eine Familie unabhängiger Mengensysteme auf dem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{S}, \mathcal{P})$, $\mathcal{E}_j \subseteq \mathfrak{S}$.

1. Ist $J' \subseteq J$, so sind auch $(\mathcal{E}_j)_{j\in J'}$ unabhängig.
2. Sind $\mathcal{E}_j' \subseteq \mathcal{E}_j$, so sind auch $(\mathcal{E}_j')_{j\in J}$ unabhängig.

Anmerkung: Seien $A, B \in \mathfrak{S}$ unabhängig auf dem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{S}, \mathcal{P})$. Dann sind auch gleich die σ-Algebren

$$\sigma(\{A\}) = \{\emptyset, A, A^c, \Omega\} \quad , \quad \sigma(\{B\}) = \{\emptyset, B, B^c, \Omega\}$$

unabhängig.
Dies gibt Anlass zu folgender allgemeineren Frage: Folgt aus der Unabhängigkeit von Mengensystemen $(\mathcal{E}_j)_{j\in J}$ auch die Unabhängigkeit der σ-Algebren $(\sigma(\mathcal{E}_j))_{j\in J}$?

2.2.5 Definition: Dynkinsystem

Ein Mengensystem $\mathcal{D} \subseteq \mathscr{P}(\Omega)$ heißt *Dynkinsystem*, falls gilt:

1. $\Omega \in \mathcal{D}$
2. Für $D \in \mathcal{D}$ ist auch $D^c \in \mathcal{D}$.
3. Für disjunkte $D_1, D_2, ... \in \mathcal{D}$ ist auch $\biguplus_{n\in\mathbb{N}} D_n \in \mathcal{D}$.

Ist $\mathcal{E} \subseteq \mathscr{P}(\Omega)$, so sei $\delta(\mathcal{E})$ das kleinste, $\mathcal{E}$-umfassende Dynkinsystem. Allgemein gilt $\delta(\mathcal{E}) \subseteq \sigma(\mathcal{E})$.

2.2.6 Hilfslemma über Dynkinsysteme

Ist $\mathcal{E} \subseteq \mathscr{P}(\Omega)$ ein $\cap$-stabiles Mengensystem[2], so gilt

$$\delta(\mathcal{E}) = \sigma(\mathcal{E})$$

2.2.7 Theorem über unabhängige Mengensysteme

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $(\mathcal{E}_i)_{i\in I}$ eine Familie $\cap$-stabiler Mengensysteme $\mathcal{E}_i \subseteq \mathfrak{S}$. Sind $(\mathcal{E}_i)_{i\in I}$ unabhängig, so sind auch die jeweils erzeugten σ-Algebren $\mathfrak{S}_i := \sigma(\mathcal{E}_i)$ unabhängig.

[2] Ein Mengensystem $\mathcal{E} \subseteq \mathscr{P}(\Omega)$ heißt *durchschnittstabil* oder $\cap$-*stabil*, falls für zwei Mengen $A, B \in \mathcal{E}$ auch $A \cap B \in \mathcal{E}$ ist.

Beweis: Es genügt die Aussage für endlich viele $(\mathcal{E}_i)_{i=1}^n$ zu zeigen. Setze

$$\mathcal{D} := \{A \in \mathfrak{S} : \{A\}, \mathcal{E}_2, .., \mathcal{E}_n \text{ unabhängig}\}$$

das heißt $A \in \mathcal{D} \;\Leftrightarrow$

$$\forall\ J \subseteq \{2, .., n\},\ E_j \in \mathcal{E}_j\ :\ \mathcal{P}\Big(A \cap \bigcap_{j \in J} \mathcal{E}_j\Big) = \mathcal{P}(A) \cdot \prod_{j \in J} \mathcal{P}(E_j)$$

Behauptung: $\mathcal{D}$ ist Dynkin-System.

Beweis:

- Da $\mathcal{E}_2, .., \mathcal{E}_n$ unabhängig sind, ist $\Omega \in \mathcal{D}$.
- Ist $A \in \mathcal{D}$ so gilt für $J \subseteq \{2, .., n\},\ E_j \in \mathcal{E}_j$:

$$\mathcal{P}(A^c) \cdot \prod_{j \in J} \mathcal{P}(E_j) = [1 - \mathcal{P}(A)] \cdot \prod_{j \in J} \mathcal{P}(E_j) = \prod_{j \in J} \mathcal{P}(E_j) - \mathcal{P}(A) \cdot \prod_{j \in J} \mathcal{P}(E_j)$$

$$= \mathcal{P}\Big(\bigcap_{j \in J} E_j\Big) - \mathcal{P}\Big[A \cap \bigcap_{j \in J} E_j\Big] = \mathcal{P}\Big[\bigcap_{j \in J} E_j \setminus \Big(A \cap \bigcap_{j \in J} E_j\Big)\Big] = \mathcal{P}\Big(A^c \cap \bigcap_{j \in J} E_j\Big)$$

das heißt $A^c \in \mathcal{D}$.

- Sind $A_1, A_2, .. \in \mathcal{D}$ disjunkt und $J \subseteq \{2, .., n\},\ E_j \in \mathcal{E}_j$, so folgt

$$\mathcal{P}\Big(\biguplus_{i=1}^{\infty} A_i \cap \bigcap_{j \in J} E_j\Big) = \sum_{i=1}^{\infty} \mathcal{P}\Big[A_i \cap \bigcap_{j \in J} E_j\Big] = \sum_{i=1}^{\infty} \mathcal{P}(A_i) \cdot \mathcal{P}\Big(\bigcap_{j \in J} E_j\Big)$$

$$= \mathcal{P}\Big(\bigcap_{j \in J} E_j\Big) \cdot \sum_{i=1}^{\infty} \mathcal{P}(A_i) = \mathcal{P}\Big(\bigcap_{j \in J} E_j\Big) \cdot \mathcal{P}\Big(\biguplus_{i=1}^{\infty} A_i\Big)$$

sprich

$$\biguplus_{i=1}^{\infty} A_i \in \mathcal{D}$$

Offensichtlich ist auch $E_1 \in \mathcal{D}$ für jedes $E_1 \in \mathcal{E}_1$, sprich $\mathcal{E}_1 \subseteq \mathcal{D}$ und daher

$$\sigma(\mathcal{E}_1) \overset{2.2.6}{=} \delta(\mathcal{E}_1) \subseteq \mathcal{D}$$

das heißt $\sigma(\mathcal{E}_1), \mathcal{E}_2, .., \mathcal{E}_n$ sind unabhängig. Umgeschrieben: Die $\cap$-stabilen

$$\mathcal{E}_2, \sigma(\mathcal{E}_1), \mathcal{E}_3, .., \mathcal{E}_n$$

sind unabhängig, und obige Überlegungen liefern die Unabhängigkeit von

$$\sigma(\mathcal{E}_2), \sigma(\mathcal{E}_1), \mathcal{E}_3, .., \mathcal{E}_n$$

Induktive Fortsetzungen liefert schließlich die Behauptung.
$\square$

Beispiel: $\Omega := \mathbb{N} \times \mathbb{N},\ \mathfrak{S} = \mathscr{P}(\Omega)$, dazu

$$\mathcal{E}_1 := \{\emptyset\} \cup \{\{n\} \times \mathbb{N} : n \in \mathbb{N}\} \quad , \quad \mathcal{E}_2 := \{\emptyset\} \cup \{\mathbb{N} \times \{m\} : m \in \mathbb{N}\}$$

Die σ-Algebren

$$\mathfrak{S}_1 := \sigma(\mathcal{E}_1) = \{A \times \mathbb{N} : A \subseteq \mathbb{N}\} \quad , \quad \mathfrak{S}_2 := \sigma(\mathcal{E}_2) = \{\mathbb{N} \times B : B \subseteq \mathbb{N}\}$$

sind genau dann unabhängig, wenn

$$\mathcal{P}(A \times B) = \mathcal{P}(A \times \mathbb{N}) \cdot \mathcal{P}(\mathbb{N} \times B) \quad \forall\ A, B \subseteq \mathbb{N}$$

und nach Theorem 2.2.7 genau dann wenn

$$\mathcal{P}(\{(n, m)\}) = \mathcal{P}(\{n\} \times \mathbb{N}) \cdot \mathcal{P}(\mathbb{N} \times \{m\}) \quad \forall\ n, m \in \mathbb{N}$$

2.2.8 Satz: Blockbildung unabhängiger Mengensysteme

Gegeben sei die Familie $(\mathfrak{S}_j)_{j\in J}$ unabhängiger σ-Algebren $\mathfrak{S}_j \subseteq \mathfrak{S}$ auf dem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{S}, \mathscr{P})$. Sei $J = \biguplus_{i\in I} J_i$ eine disjunkte Zerlegung der Indexmenge J, dazu die Familie von σ-Algebren

$$\mathfrak{D}_i := \sigma\left(\bigcup_{j\in J_i} \mathfrak{S}_j\right) \quad , \quad i \in I$$

Sind die $(\mathfrak{S}_j)_{j\in J}$ unabhängig, so sind auch die $(\mathfrak{D}_i)_{i\in I}$ unabhängig.

Beweis: Definieren die Mengensysteme

$$\widetilde{\mathfrak{D}}_i := \left\{ \bigcap_{j\in L} A_j : A_j \in \mathfrak{S}_j, \quad L \subseteq J_i \text{ endlich} \right\} \quad , \quad i \in I$$

Dann gilt

$$\bigcup_{j\in J_i} \mathfrak{S}_j \subseteq \widetilde{\mathfrak{D}}_i \subseteq \mathfrak{D}_i$$

sprich

$$\sigma(\widetilde{\mathfrak{D}}_i) = \mathfrak{D}_i$$

Ferner sind die $\widetilde{\mathfrak{D}}_i,\ i \in I$ unabhängig und $\cap$-stabil, so dass nach Theorem 2.2.7 die Behauptung folgt.
□

2.3 Die von einer Zufallsvariablen erzeugte σ-Algebra

2.3.1 Einführendes Beispiel

Betrachtet sei die Grundmenge $\Omega := \{1, .., 6\}$ mit $\mathfrak{S} := \mathscr{P}(\Omega)$ und Wahrscheinlichkeitsmaß der Gleichverteilung, dazu die Zufallsvariable

$$X(\omega) := \begin{cases} 1 & : \omega \text{ gerade} \\ 0 & : \text{sonst} \end{cases}$$

Die Frage ist nun: Je nach Wert von X, über welche Teilmengen von Ω lassen sich Aussagen über deren Eintreffen machen? In diesem Fall sind es genau

$$\overbrace{\emptyset}^{X^{-1}(\emptyset)} \ , \ \overbrace{\Omega}^{X^{-1}(\{0,1\})} \ , \ \overbrace{\{2,4,6\}}^{X^{-1}(\{1\})}, \ \overbrace{\{1,3,5\}}^{X^{-1}(\{0\})}$$

2.3.2 Definition: $\sigma(X)$

Sei $X : (\Omega, \mathfrak{S}) \to (S, \mathcal{S})$ messbar. Dann heißt

$$\sigma(X) := X^{-1}(\mathcal{S}) := \left\{ X^{-1}(B) : B \in \mathcal{S} \right\}$$

die *von X erzeugte σ-Algebra.*

Bemerkungen:

(i) $\sigma(X)$ ist tatsächlich eine σ-Algebra.

(ii) $\sigma(X)$ ist die kleinste σ-Algebra bzgl. der X messbar ist.

Beispiele

(i) $X = 1_A$ für irgendein $A \in \mathfrak{S}$. Dann ist

$$\sigma(X) = \{\emptyset,\ A,\ A^c,\ \Omega\}$$

(ii) Für $\Omega = \{1, .., 6\}^n$, $\mathfrak{S} = \mathscr{P}(\Omega)$ und $X((\omega_1, .., \omega_n)) := (\omega_1, \omega_2)$ ist

$$\sigma(X) = \left\{A \times \{1, .., 6\}^{n-2} : A \subseteq \{1, .., n\}^2\right\}$$

(iii) $X : \mathbb{R}^n \to \mathbb{R}$ mit $X(\mathbf{x}) := \|\mathbf{x}\|$. Dann gilt $B \in \sigma(X)$ genau dann wenn

$$B = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| \in A\}$$

für irgendein $A \in \mathscr{B}(\mathbb{R})$.

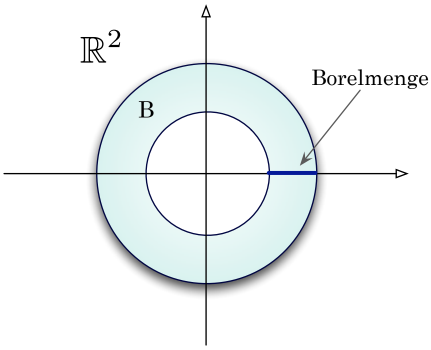

Abbildung 4: Zu Beispiel (iii): Typische Menge in $\sigma(X)$.

2.3.3 Satz über Erzeuger von $\sigma(X)$

Sei $X : \Omega \to S$ eine Abbildung und $\mathcal{E} \subseteq \mathscr{P}(S)$ ein beliebiges Mengensystem auf S. Dann gilt:

$$\sigma\left(X^{-1}(\mathcal{E})\right) = X^{-1}(\sigma(\mathcal{E}))$$

Erläuterung: Ist $X : (\Omega, \mathfrak{S}) \to (S, \mathcal{S})$ messbar und $\mathcal{E}$ ein Erzeuger von $\mathcal{S}$, so ist $X^{-1}(\mathcal{E})$ ein Erzeuger von $\sigma(X)$.

2.3.4 Lemma über von diskreten Zufallsgrößen erzeugte σ-Algebren

Sei $\Omega \neq \emptyset$, $(S, \mathcal{S})$ ein messbarer Raum und $X : \Omega \to S$. Dann gilt:

1. $X^{-1}(\mathcal{S}) = X^{-1}(X(\Omega) \cap \mathcal{S})$ mit der Schnitt-σ-Algebra $X(\Omega) \cap \mathcal{S} := \{X(\Omega) \cap A : A \in \mathcal{S}\}$.
2. Ist $X(\Omega)$ abzählbar und $\{X(\omega)\} \in \mathcal{S}\ \ \forall\, \omega \in \Omega$, so gilt

$$X^{-1}(\mathcal{S}) = \sigma\left\{X^{-1}(\{s\}) : s \in X(\Omega)\right\}$$

Beweis:

1. Jede beliebige Menge $B \in \mathcal{S}$ kann zerlegt werden in $B = B_i \dot\cup B_o$, mit $B_i := B \cap X(\Omega)$, $B_o := B \setminus X(\Omega)$. Wegen

$$X^{-1}(B) = X^{-1}(B_i) \cup \underbrace{X^{-1}(B_o)}_{\emptyset} = X^{-1}(B \cap X(\Omega))$$

folgt die Behauptung.

2. Da $X(\Omega)$ abzählbar ist, und jede 1-Punkt-Menge $\{s\} \subseteq X(\Omega)$ in $X(\Omega) \cap \mathcal{S}$ enthalten ist, gilt

$$X(\Omega) \cap \mathcal{S} = \sigma\{\{s\} : s \in X(\Omega)\}$$

Nach obigem Satz (2.3.3) und Teil (1) folgt dann

$$X^{-1}(\mathcal{S}) = X^{-1}[X(\Omega) \cap \mathcal{S}] = \sigma\left\{X^{-1}(\{s\}) : s \in X(\Omega)\right\}$$

□

2.3.5 Definition: Vor mehreren Zufallsvariablen erzeugte σ-Algebra

Sei $\Omega \neq \emptyset$, $(S_i, \mathcal{S}_i),\ i \in I$ messbare Räume und $X_i : \Omega \to S_i,\ i \in I$. Dann heißt

$$\sigma(X_i : i \in I) := \sigma\left\{X_i^{-1}(B_i) : B_i \in \mathcal{S}_i,\ i \in I\right\}$$

die von $(X_i)_{i \in I}$ erzeugte σ-Algebra.

Bemerkung: $\sigma(X_i : i \in I)$ ist tatsächlich die kleinste σ-Algebra, bzgl. der alle X_i messbar sind.

Beispiel: Hinsichtlich eines n-fachen Würfelexperiments, sei $\Omega := \{1, .., 6\}^n$ und $X_i(\omega_1, .., \omega_n) := \omega_i$ das Ergebnis des i-ten Wurfs. Dann ist

$$\sigma(X_i) = \left\{\{1, .., 6\}^{i-1} \times A \times \{1, .., 6\}^{n-i} : A \subseteq \{1, .., 6\}\right\}$$

und ferner

$$\sigma(X_i : 1 \leq i \leq k) = \left\{B \times \{1, .., 6\}^{n-k} : B \subseteq \{1, .., 6\}^k\right\} \quad , \quad k \leq n$$

2.3.6 Satz: Erzeugung von $\sigma(X_i : i \in I)$ durch eine einzelne Zufallsvariable

Seien $X_i : \Omega \to (S_i, \mathcal{S}_i),\ i \in I$ Abbildungen von Ω in die messbaren Räume $(S_i, \mathcal{S}_i)$. Seien ferner

$$S := \bigtimes_{i \in I} S_i \quad , \quad \mathcal{S} := \bigotimes_{i \in I} \mathcal{S}_i$$

und

$$X : \Omega \to S \quad , \quad X(\omega) := (X_i(\omega))_{i \in I}$$

Dann gilt:

$$\sigma(X) = \sigma(X_i : i \in I)$$

Beweis: Folgt aus (2.1.5.1) und Satz 2.3.3.

2.4 Unabhängigkeit von Zufallsvariablen

2.4.1 Definition: Unabhängige Zufallsvariablen

Seien $X_i : (\Omega, \mathfrak{S}, \mathcal{P}) \to (S_i, \mathcal{S}_i),\quad i \in I$ Zufallsvariablen in die messbare Räume $(S_i, \mathcal{S}_i)$.
Dann heißen $(X_i)_{i \in I}$ *unabhängig* :⇔

$$(\sigma(X_i))_{i \in I}$$

sind unabhängige σ-Algebren (vgl. 2.2.2).
Allgemeiner heißen Familien $(X_i)_{i \in I_j},\ j \in J$ von Zufallsvariablen *unabhängig* :⇔

$$\sigma(X_i : i \in I_j),\ j \in J$$

sind unabhängige σ-Algebren.

Erläuterung: $(X_i)_{i\in I}$ sind genau dann unabhängig wenn

$$\forall\ I_0 \subseteq I \text{ endlich} : \forall\ B_i \in \mathcal{S}_i,\ i \in I_0 \quad : \quad \mathcal{P}\left(X_i \in B_i : i \in I_0\right) = \prod_{i\in I_0} \mathcal{P}(X_i \in B_i)$$

(vgl. Bemerkung 2.2.2 (i)) bzw. genau dann unabhängig wenn

$$\forall\ I_0 \subseteq I \text{ höchstens abzählbar} : \forall\ B_i \in \mathcal{S}_i,\ i \in I_0 \quad : \quad \mathcal{P}\left(X_i \in B_i : i \in I_0\right) = \prod_{i\in I_0} \mathcal{P}(X_i \in B_i)$$

(vgl. Bemerkung 2.2.2 (ii)).

Spezialfälle:

(i) Für $I = \{1, .., n\}$ sind $X_1, .., X_n$ unabhängig $\Leftrightarrow$

$$\forall\ B_i \in S_i \quad : \quad \mathcal{P}\left(X_1 \in B_1, .., X_n \in B_n\right) = \prod_{i=1}^{n} \mathcal{P}(X_i \in B_i)$$

(vgl. 2.2.3).

(ii) $(X_n)_{n\in\mathbb{N}}$ sind unabhängig, genau dann wenn $X_1, .., X_k$ unabhängig sind $\forall\ k \leq n$.

Beispiele:

(i) Sind $(X_k)_{k\in\mathbb{N}}$ unabhängig (beachte: abzählbar viele), so gilt

$$\mathcal{P}\left(X_k \in B_k,\ k \in \mathbb{N}\right) = \prod_{k=1}^{\infty} \mathcal{P}(X_k \in B_k) \qquad \forall\ B_k \in \mathcal{S}_k$$

(vgl. obige Erläuterung).

(ii) Gegeben $\Omega := [0,1]$, $\mathfrak{S} := \mathscr{B}([0,1])$ und $\mathcal{P} := \lambda^1$. Zu jedem $t \in [0,1]$ sei $t = 0.t_1t_2\ldots$ die Binärdarstellung[3] von t, sprich $t_i \in \{0,1\}$ und

$$t = \sum_{n=1}^{\infty} \frac{t_n}{2^n}$$

Dann sind die Zufallsvariablen $X_n : [0,1] \to \{0,1\}$, definiert durch

$$X_n(t) := t_n \quad , \quad n \in \mathbb{N}$$

unabhängig.

2.4.2 Charakterisierung der Unabhängigkeit von Zufallsvariablen

Seien $X_i : (\Omega, \mathfrak{S}, \mathcal{P}) \to (S_i, \mathcal{S}_i),\ i \in I$ Zufallsvariablen und

$$X := (X_i)_{i\in I} : \Omega \to \Big(\bigtimes_{i\in I} S_i, \bigotimes_{i\in I} \mathcal{S}_i\Big) \quad , \quad X(\omega) = (X_i(\omega))_{i\in I}$$

Dann sind $(X_i)_{i\in I}$ unabhängig $\Leftrightarrow$

$$\mathcal{P}_X = \bigotimes_{i\in I} \mathcal{P}_{X_i}$$

[3] Ambiguitäten der Art $0.* .. *01111.. = 0.* .. *10$ seien durch Wahl der Darstellung $0.* .. *10$ aufgelöst.

Beweis: Richtung "⇒": Nach Theorem 2.1.4 ist $\bigotimes_{i\in I} \mathcal{P}_{X_i}$ eindeutig bestimmt durch die Eigenschaft

$$\bigotimes_{i\in I} \mathcal{P}_{X_i}(Z) = \prod_{k=1}^{n} \mathcal{P}_{X_{i_k}}(B_{i_k})$$

für jegliche rechteckige Zylindermenge

$$Z = \left\{(s_i)_{i\in I} \in \bigtimes_{i\in I} S_i : s_{i_1} \in B_{i_1}, .., s_{i_n} \in B_{i_n}\right\} \quad , \quad B_{i_k} \in \mathcal{S}_{i_k}$$

Doch tatsächlich ist

$$\mathcal{P}_X(Z) = \mathcal{P}\left(\{X_{i_1} \in B_{i_1}, .., X_{i_n} \in B_{i_n}\}\right) \overset{\substack{X_{i_1},..,X_{i_n} \\ \text{unabhängig}}}{=} \prod_{k=1}^{n} \mathcal{P}\left(\{X_{i_k} \in B_{i_k}\}\right) = \prod_{k=1}^{n} \mathcal{P}_{X_{i_k}}(B_{i_k})$$

Richtung "⇐": Analog. □

Beispiel: Sei $\Omega := \left\{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| \leq 1\right\}$ und $\mathcal{P}$ die Gleichverteilung auf $(\Omega, \mathscr{B}(\Omega))$:

$$\mathcal{P}(A) = \frac{\lambda^2(A)}{\pi} \quad , \quad A \in \mathscr{B}(\Omega) \quad ,$$

dazu die 2 Zufallsvariablen $X_i : \Omega \to \mathbb{R},\ X_i(\mathbf{x}) := x_i$. Dann ist für $B_i \in \mathscr{B}(\mathbb{R})$ typischerweise

$$\mathcal{P}\left(X_1^{-1}(B_1)\right) \cdot \mathcal{P}\left(X_2^{-1}(B_2)\right) \neq \mathcal{P}\left(X_1^{-1}(B_1) \cap X_2^{-1}(B_2)\right)$$

sprich, X_1, X_2 sind abhängig (vgl. Abb. (5))!

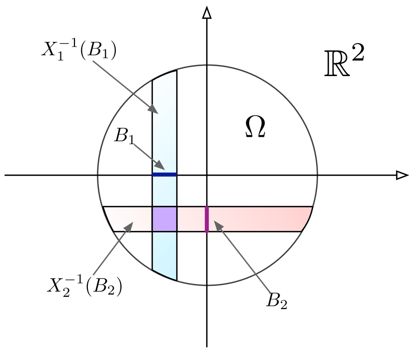

Abbildung 5: Zur Abhängigkeit von $X_1, X_2 : \Omega \to \mathbb{R}$.

Aus

$$\mathcal{P}_{X_i}((-1, t]) = \frac{1}{\pi} \int_{-1}^{t} 2\sqrt{1 - s^2}\, ds$$

wird ersichtlich, dass $\mathcal{P}_{X_1}$ und $\mathcal{P}_{X_2}$ beide die Dichte

$$\rho_i(s) = \frac{2}{\pi}\sqrt{1 - s^2} \quad , \quad -1 \leq s \leq 1$$

besitzen. Daher muss $\mathcal{P}_{X_1} \otimes \mathcal{P}_{X_2}$ die Dichte

$$\rho(\mathbf{x}) = \rho_1(x_1) \cdot \rho_2(x_2) = \frac{4}{\pi^2}\sqrt{1 - x_1^2} \cdot \sqrt{1 - x_2^2} \quad , \quad \mathbf{x} \in [-1, 1]^2$$

besitzen, ist also konzentriert auf dem Quadrat $[-1, 1]^2$. Anderseits ist $\mathcal{P}_{(X_1, X_2)} = \mathcal{P}$ konzentriert auf dem Ball Ω und insbesondere

$$\mathcal{P}_{(X_1, X_2)} \neq \mathcal{P}_{X_1} \otimes \mathcal{P}_{X_2}$$

2.4.3 Satz: Charakterisierung der Unabhängigkeit von Zufallsvariablen

Seien $X_i : (\Omega, \mathfrak{S}, \mathcal{P}) \to (S_i, \mathcal{S}_i)$ Zufallsvariablen in die messbaren Räume $(S_i, \mathcal{S}_i)_{i\in I}$. Ferner seien $\mathcal{E}_i \subseteq \mathcal{S}_i$ $\cap$-stabile Erzeuger von $\mathcal{S}_i$. Dann sind $(X_i)_{i\in I}$ unabhängig, genau dann wenn

$$X_i^{-1}(\mathcal{E}_i) := \left\{X_i^{-1}(E) : E \in \mathcal{E}_i\right\} \quad , \quad i \in I$$

unabhängige Mengensysteme sind.

Beweis: Richtung "⇒": klar.
Richtung "⇐": Vergleiche mit 2.3.3, bemerke dass auch die $X^{-1}(\mathcal{E}_i)$ $\cap$-stabil sind und verwende 2.2.7.

2.4.4 Satz über Unabhängigkeit reeller Zufallsvariablen

Reelle Zufallsvariablen $X_1, .., X_n : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}$ sind genau dann unabhängig, wenn

$$\mathcal{P}\left(X_1 \le t_1, \dots, X_n \le t_n\right) = \mathcal{P}\left(X_1 \le t_1\right) \cdot .. \cdot \mathcal{P}\left(X_n \le t_n\right) \quad \forall\, t_1, .., t_n \in \mathbb{R}$$

Beweis: Betrachte den $\cap$-stabilen Erzeuger

$$\mathcal{E} := \{(-\infty, t] : t \in \mathbb{R}\}$$

von $\mathscr{B}(\mathbb{R})$, dazu $\widetilde{\mathcal{E}}_j := X_j^{-1}(\mathcal{E})$. Nach 2.4.3 sind $X_1, .., X_n$ genau dann unabhängig, wenn $(\widetilde{\mathcal{E}}_j)_{j=1}^n$ unabhängig sind. Wegen

$$\Omega = \bigcup_{k=1}^{\infty} \underbrace{X_j^{-1}\left((-\infty, k]\right)}_{\substack{\in \widetilde{\mathcal{E}}_j \\ \text{monoton wachsend} \\ \text{in } k}} \quad \forall\, j = 1, .., n$$

sind diese nach 2.2.3 (3) genau dann unabhängig, wenn

$$\mathcal{P}\left(E_1 \cap \dots \cap E_n\right) = \mathcal{P}(E_1) \cdot .. \cdot \mathcal{P}(E_n) \;\; \forall\, E_j \in \widetilde{\mathcal{E}}_j$$

□

2.4.5 Satz über Unabhängigkeit diskreter Zufallsvariablen

Seien $(S_i, \mathcal{S}_i)_{i=1}^n$ messbare Räume und $X_i : (\Omega, \mathfrak{S}, \mathcal{P}) \to (S_i, \mathcal{S}_i)$ Zufallsvariablen mit abzählbaren Bildern $X_i(\Omega)$. Ferner gelte

$$\{s\} \in \mathcal{S}_i \quad \forall\, s \in S_i,\;\; i = 1, .., n$$

Dann sind $X_1, .., X_n$ genau dann unabhängig, wenn

$$\forall\, s_1 \in S_1, .., s_n \in S_n \;\;:\;\; \mathcal{P}\left(X_1 = s_1, \dots, X_n = s_n\right) = \mathcal{P}(X_1 = s_1) \cdot .. \cdot \mathcal{P}(X_n = s_n)$$

Beweis: Richtung "⇒" ist klar.
Richtung "⇐": Nach Lemma 2.3.4(2) gilt

$$\sigma(X_i) = \sigma \underbrace{\left[\left\{X_i^{-1}(\{s\}) : s \in S_i\right\} \cup \{\emptyset\}\right]}_{\widetilde{\mathcal{E}}_i}$$

wobei die $\widetilde{\mathcal{E}}_i$ $\cap$-stabil sind. Nach Theorem 2.2.7 genügt es also die Unabhängigkeit der $(\widetilde{\mathcal{E}}_i)_{i\in I}$ zu zeigen. Wegen

$$\Omega = \biguplus_{s \in X_i(\Omega)} \underbrace{X_i^{-1}\left(\{s\}\right)}_{\in \widetilde{\mathcal{E}}_i} \quad \forall\, i = 1, .., n$$

folgt nach Satz 2.2.3 (2) die Behauptung. □

Spezialfälle: Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum.

(i) Zwei Zufallsgrößen $X : \Omega \to S_1 := \{x_1, x_2, ..\}$ und $Y : \Omega \to S_2 := \{y_1, y_2, ..\}$, mit

$$p_i := \mathcal{P}(X = x_i) \quad , \quad q_i := \mathcal{P}(Y = y_i) \quad , \quad p_{ij} := \mathcal{P}(X = x_i, Y = y_j)$$

sind genau dann unabhängig, wenn

$$p_{ij} = p_i \cdot q_j \quad \forall\, i, j \in \mathbb{N}$$

(ii) Zwei zufällige Vektoren $X : \Omega \to \mathbb{R}^n$ und $Y : \Omega \to \mathbb{R}^m$ sind genau dann unabhängig, wenn

$$\mathcal{P}\left(X_1 \leq t_1, .., X_n \leq t_n,\ Y_1 \leq s_1, .., Y_m \leq s_m\right) = \mathcal{P}\left(X_1 \leq t_1, .., X_n \leq t_n\right) \cdot \mathcal{P}\left(Y_1 \leq s_1, .., Y_n \leq s_m\right) \ \ \forall\, t_i, s_j \in \mathbb{R}$$

Erläuterung: Betrachte die $\cap$-stabilen Erzeuger

$$\mathcal{E}_1 := \{(-\infty, t_1] \times .. \times (-\infty, t_n] : t_i \in \mathbb{R}\}$$
$$\mathcal{E}_2 := \{(-\infty, s_1] \times .. \times (-\infty, s_m] : s_j \in \mathbb{R}\}$$

von $\mathbb{R}^n$ bzw. $\mathbb{R}^m$ und verwende Satz 2.4.3.

2.4.6 Permanenzeigenschaften der Unabhängigkeit von Zufallsvariablen

1. Sind die Zufallsvariablen $(X_i)_{i \in I}$ unabhängig und $J \subseteq I$, so sind auch $(X_j)_{j \in J}$ unabhängig.

2. Sind $X_i : (\Omega, \mathfrak{S}, \mathcal{P}) \to (S_i, \mathcal{S}_i),\ i \in I$ unabhängige Zufallsvariablen und $f_i : (S_i, \mathcal{S}_i) \to (\widetilde{S}_i, \widetilde{\mathcal{S}}_i)$ messbar, so sind auch

$$(f_i(X_i))_{i \in I}$$

unabhängig.

3. Sind $X_i : (\Omega, \mathfrak{S}, \mathcal{P}) \to S_i,\ i \in I$ unabhängige Zufallsvariablen, $I = \biguplus_{j \in J} I_j$ eine disjunkte Zerlegung von I und

$$Y_j : \Omega \to \bigtimes_{i \in I_j} S_i \quad , \quad Y_j(\omega) := (X_i(\omega))_{i \in I_j} \quad , \quad j \in J$$

so sind auch $(Y_j)_{j \in J}$ unabhängig.
Erläuterung: Nach 2.3.6 gilt

$$\sigma(Y_j) \stackrel{(2.3.6)}{=} \sigma(X_i : i \in I_j) = \sigma\Big(\bigcup_{i \in I_j} \sigma(X_i) \Big)$$

Der Rest folgt aus dem für Mengensysteme entsprechenden Satz 2.2.8.

Beispiele:

(i) Sind die reellen Zufallsvariablen $X_1, .., X_n$ unabhängig, so sind auch

$$X_1 - \mathbb{E}X_1, \ldots, X_n - \mathbb{E}X_n$$

unabhängig.

(ii) Sind die reellen Zufallsvariablen $X_1, .., X_n$ unabhängig und $B_1, .., B_n \in \mathscr{B}(\mathbb{R})$, so sind auch

$$X_i \cdot \underbrace{1_{\{X_i \in B_i\}}}_{1_{B_i}(X_i)} \quad , \quad i = 1, .., n$$

unabhängig.

(iii) Sind $X_1, .., X_4$ unabhängige reelle Zufallsvariablen, so sind (X_1, X_2) und (X_3, X_4) nach 2.4.6 (3) unabhängige, reelle Zufallsvektoren. Ferner sind

$$X_1^2 + X_2^2 \ , \ X_3^2 + X_4^2$$

nach 2.4.6 (2) unabhängige, reelle Zufallsvariablen.

2.4.7 Lemma: Unabhängigkeit der Komponenten von Zufallsvektoren

Seien $(S_i, \mathcal{S}_i)_{i\in I}$ messbare Räume, dazu die identisch verteilten Zufallsvektoren

$$(X_i)_{i\in I}, (Y_i)_{i\in I} : (\Omega, \mathfrak{S}, \mathcal{P}) \to \Big(\bigtimes_{i\in I} S_i,\ \bigotimes_{i\in I} \mathcal{S}_i \Big)$$

Dann sind die Komponenten $X_i,\ i \in I$ unabhängig genau dann wenn die Komponenten $Y_i,\ i \in I$ unabhängig sind.

2.5 Rechenregeln für unabhängige Zufallsvariablen

2.5.1 Theorem über Erwartungswerte unabhängiger Zufallsvariablen

Seien $X_1, .., X_n$ unabhängige, reelle Zufallsvariablen. Dann gilt:

$$\mathbb{E}\,|X_1 \cdot .. \cdot X_n| < \infty \quad \Leftrightarrow \quad \mathbb{E}\,|X_i| < \infty \quad \forall\ i = 1, .., n$$

Gegebenfalls ist dann

$$\mathbb{E}\,(X_1 \cdot .. \cdot X_n) = (\mathbb{E}X_1) \cdot .. \cdot (\mathbb{E}X_n)$$

2.5.2 Formel von Bienaymé

Seien $X_1, .., X_n$ unabhängige, reelle Zufallsvariablen mit jeweils. 2. Moment. Dann gilt:

$$\mathbb{V}\left[\sum_{j=1}^{n} X_j \right] = \sum_{j=1}^{n} \mathbb{V} X_j$$

Beweis: Folgt direkt aus Theorem 2.5.1 und der Definition von $\mathbb{V}$.

2.6 Faltung von Maßen

2.6.1 Definition: Messbare Gruppe

Sei $(G, \circ)$ eine Gruppe, ausgestattet mit der σ-Algebra $\mathcal{G}$, bzgl. dieser die Abbildung

$$g : (G \times G, \mathcal{G} \otimes \mathcal{G}) \to (G, \mathcal{G}) \quad , \quad g(x, y) := x \circ y$$

messbar ist. Dann heißt $(G, \circ, \mathcal{G})$ *messbare Gruppe*.

Beispiel: So genannte *topologische* Gruppen, wie z.B. der $(\mathbb{R}^n, +)$.

2.6.2 Definition: Faltung von Wahrscheinlichkeitsmaßen

Sei $(G, \circ, \mathcal{G})$ eine *messbare* Gruppe mit $g(x, y) := x \circ y$. Für zwei Wahrscheinlichkeitsmaße $\mathcal{P}, \mathcal{Q}$ auf $(G, \mathcal{G})$, heißt das Wahrscheinlichkeitsmaß

$$P * Q := (P \otimes Q) \circ g^{-1}$$

auf $(G, \mathcal{G})$, *Faltung* von $\mathcal{P}$ und $\mathcal{Q}$.

Erläuterung: Für $B \in \mathcal{G}$ ist

$$(\mathcal{P} * \mathcal{Q})(B) = (\mathcal{P} \otimes \mathcal{Q})\big\{(x,y) \in G^2 : x \circ y \in B\big\} = \int_{G^2} 1_B(x \circ y)\ d(\mathcal{P} \otimes \mathcal{Q})(x,y)$$

$$\overset{\text{Fubini}}{=} \int_G \underbrace{\int_G 1_B(x \circ y)\ d\mathcal{P}(x)}_{\mathcal{P}(B \circ y^{-1})}\ d\mathcal{Q}(y) = \int_G \mathcal{P}(B \circ y^{-1})\ d\mathcal{Q}(y)$$

$$\overset{\text{analog}}{=} \int_G \mathcal{Q}(x^{-1} \circ B)\ d\mathcal{P}(x)$$

Beachte dass nach Satz A.1.1 $B \circ y^{-1}$ bzw. $x^{-1} \circ B$ tatsächlich messbar sind.

2.6.3 Eigenschaften der Faltung von Wahrscheinlichkeitsmaßen

Seien $\mathcal{P}, \mathcal{Q}$ Wahrscheinlichkeitsmaße auf der messbaren Gruppe $(G, \circ, \mathcal{G})$. Dann gilt:

1. $(\mathcal{P} * \mathcal{Q})(B) = \int_G \mathcal{P}(B \circ y^{-1})\ d\mathcal{Q}(y) = \int_G \mathcal{Q}(x^{-1} \circ B)\ d\mathcal{P}(x)$ für $B \in \mathcal{G}$.
2. $\mathcal{P} * \mathcal{Q} = \mathcal{Q} * \mathcal{P}$, falls $(G, \circ)$ kommutativ.
3. Sei $f : G \to \mathbb{R}$ messbar. Dann gilt

$$\int_G |f|\ d(\mathcal{P} * \mathcal{Q}) < \infty \quad \Leftrightarrow \quad \int_G \int_G |f(x \circ y)|\ d\mathcal{P}(x)\ d\mathcal{P}(y) < \infty$$

   und gegebenfalls

$$\int_G f\ d(\mathcal{P} * \mathcal{Q}) = \int_G \int_G f(x \circ y)\ d\mathcal{P}(x)\ d\mathcal{Q}(y)$$

4. Sei δ_0 konzentriert auf dem neutralen Element $0 \in G$. Dann gilt

$$\mathcal{P} * \delta_0 = \mathcal{P}$$

2.6.4 Spezialfall: Faltung absolut stetiger Wahrscheinlichkeitsmaße

Seien $\mathcal{P}, \mathcal{Q} \ll \lambda^n$ Wahrscheinlichkeitsmaße auf $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$ mit Dichten p, q bzgl. λ^n. Dann besitzt $(\mathcal{P} * \mathcal{Q})$ bzgl. λ die Dichte:

$$\frac{d(\mathcal{P} * \mathcal{Q})}{d\lambda} = p * q$$

wobei

$$p * q : \mathbb{R}^n \to \mathbb{R} \quad , \quad (p * q)(x) := \int_{\mathbb{R}^n} p(x-y)q(y)\ dy$$

die Faltung von p und q sei.

Beweis: Für Menge $B \in \mathscr{B}(\mathbb{R}^n)$ ist

$$(\mathcal{P} * \mathcal{Q})(B) = \int_{\mathbb{R}^n} \mathcal{P}(B-y)\ d\mathcal{Q}(y) = \int_{\mathbb{R}^n} \int_{B-y} p(x)\ dx\ q(y)\ dy$$

$$\overset{\substack{\lambda \\ \text{Translations-} \\ \text{invariant}}}{=} \int_{\mathbb{R}^n} \int_B p(x-y)\ dx\ q(y)\ dy \overset{\text{Fubini}}{=} \int_B \underbrace{\int_{\mathbb{R}^n} p(x-y)q(y)\ dy}_{(p*q)(x)}\ dx$$

□

2.6.5 Spezialfall: Faltung diskreter Wahrscheinlichkeitsmaße

Seien $\mathcal{P}, \mathcal{Q}$ Wahrscheinlichkeitsmaße auf $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$, konzentriert auf $\{x_1, x_2, ...\}$. Dann ist

$$(\mathcal{P} * \mathcal{Q})(\{x_n\}) = \sum_{k \in \mathbb{N}} \mathcal{P}(\{x_n - x_k\})\mathcal{Q}(\{x_k\})$$

Beispiele:

(i) $\delta_x * \delta_y = \delta_{x+y}$

(ii) Binomialverteilungen: $\mathcal{B}_{n,p} * \mathcal{B}_{m,p} = \mathcal{B}_{n+m,p}$

(iii) Poissonverteilungen: $\pi_\lambda * \pi_\varkappa = \pi_{\lambda+\varkappa}$

(iv) Normalverteilungen: $\mathcal{N}_{\mu_1,\sigma_1^2} * \mathcal{N}_{\mu_2,\sigma_2^2} = \mathcal{N}_{\mu_1+\mu_2,\sigma_1^2+\sigma_2^2}$

(v) Γ-Verteilungen: $\Gamma_{a,b_1} * \Gamma_{a,b_2} = \Gamma_{a,b_1+b_2}$

Man sagt: Die Familien

$$(\mathcal{B}_{n,p})_{n\in\mathbb{N}}\,,\ (\pi_\lambda)_{\lambda>0}\,,\ (\Gamma_{a,b})_{b>0}$$

bilden jeweils eine *Faltungshalbgruppe*.

2.6.6 Definition: n-fache Faltung von Wahrscheinlichkeitsmaßen

Sei $(G, \circ, \mathcal{G})$ eine messbare Gruppe ausgestattet mit den Wahrscheinlichkeitsmaßen $\mathcal{P}_1, .., \mathcal{P}_n$ und

$$g : \left(\bigtimes_{i=1}^n G, \bigotimes_{i=1}^n \mathcal{G} \right) \to (G, \mathcal{G}) \quad , \quad g(x_1, .., x_n) := x_1 \circ \cdots \circ x_n$$

(vgl. Satz A.1.2). Dann heißt das Wahrscheinlichkeitsmaß

$$\mathcal{P}_1 * \cdots * \mathcal{P}_n := (\mathcal{P}_1 \otimes \cdots \otimes \mathcal{P}_n) \circ g^{-1}$$

Faltung von $\mathcal{P}_1, .., \mathcal{P}_n$.

Eigenschaften:

(i) Es gilt $\mathcal{P}_1 * \cdots * \mathcal{P}_n = (\mathcal{P}_1 * \cdots * \mathcal{P}_{n-1}) * \mathcal{P}_n$.

(ii) Für $B \in \mathcal{G}$ gilt

$$(\mathcal{P}_1 * \cdots * \mathcal{P}_n)(B) = \int_G \ldots \int_G \mathcal{P}_1(B \circ x_n^{-1} \circ \cdots \circ x_2^{-1})\ d\mathcal{P}_2(x_2)\ \ldots\ d\mathcal{P}_n(x_n)$$

(iii) Ist $(G, \circ)$ kommutativ, so gilt

$$\mathcal{P}_1 * \cdots * \mathcal{P}_n = \mathcal{P}_{\pi(1)} * \cdots * \mathcal{P}_{\pi(n)} \quad \forall\ \pi \in \mathrm{Sym}(n)$$

(iv) Ist $(G, \circ, \mathcal{G}) = (\mathbb{R}^m, +, \mathscr{B}(\mathbb{R}^m))$, $\mathcal{P}_i \ll \lambda^m$, so besitzt $(\mathcal{P}_1 * \cdots * \mathcal{P}_n)$ bzgl. λ^m die Dichte

$$\frac{d}{d\lambda^m}(\mathcal{P}_1 * \cdots * \mathcal{P}_n) = \frac{d\mathcal{P}_1}{d\lambda^m} * \frac{d\mathcal{P}_2}{d\lambda^m} * \cdots * \frac{d\mathcal{P}_n}{d\lambda^m}$$

Beispiel: Betrachtet sei die Exponentialverteilung Exp_λ mit Dichte

$$p(x) = \begin{cases} \lambda e^{-\lambda x} & : x \geq 0 \\ \\ 0 & : x < 0 \end{cases}$$

bzgl. λ^1. Dann besitzt $\overbrace{\mathrm{Exp}_\lambda * \cdots * \mathrm{Exp}_\lambda}^{n \text{ mal}}$ die Dichte

$$p_n(x) = \begin{cases} \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} & : x \geq 0 \\ \\ 0 & : x < 0 \end{cases} \tag{2.6.6.1}$$

2.6.7 Theorem: Addition von Zufallsgrößen

Sei $(G, \circ, \mathcal{G})$ eine messbare Gruppe, dazu die unabhängigen, G-wertigen Zufallsvariablen $X_1, .., X_n$. Dann gilt:

$$\mathcal{P}_{X_1 \circ .. \circ X_n} = \mathcal{P}_{X_1} * \cdots * \mathcal{P}_{X_n}$$

Beweis: Sei

$$g : \left(\bigtimes_{i=1}^n G, \bigotimes_{i=1}^n \mathcal{G} \right) \to (G, \mathcal{G}) \quad , \quad g(x_1, .., x_n) := x_1 \circ \cdots \circ x_n$$

(vgl. Satz A.1.2). Dann gilt

$$\mathcal{P}_{X_1 \circ .. \circ X_n} = \mathcal{P}_{g(X_1, .., X_n)} = \mathcal{P}_{(X_1, .., X_n)} \circ g^{-1} \overset{\text{unabh.}}{=} (\mathcal{P}_{X_1} \otimes \cdots \otimes \mathcal{P}_{X_n}) \circ g^{-1} \overset{\text{def}}{=} \mathcal{P}_{X_1} * \cdots * \mathcal{P}_{X_n}$$

□

Beispiele:

(i) Seien X_1, X_2 unabhängig, gleichverteilt auf $\{1, .., 6\}$ (2 mal Würfeln). Dann besitzt $X_1 + X_2$ das Verteilungsgesetz $\mathcal{P}_{X_1} * \mathcal{P}_{X_2}$. So ist etwa

$$\mathcal{P}_{X_1 + X_2}(\{2\}) = \frac{1}{36} \quad , \quad \mathcal{P}_{X_1 + X_2}(\{3\}) = \frac{2}{36} \quad , \quad \ldots$$

(ii) Sind $X_1, .., X_n$ unabhängig, $\mathcal{P}_\lambda$ -verteilt, dann ist die Summe $X_1 + \cdots + X_n$ $\mathcal{P}_{n\lambda}$ -verteilt.

(iii) Sind $X_1, .., X_n$ unabhängig, standardnormalverteilt, so ist die Summe $X_1 + \cdots + X_n$ $\mathcal{N}_{0,n}$ -verteilt (vgl. Folgerung 4.3.6(iii)).

(iv) Sind $X_1, .., X_n$ unabhängig, $\mathcal{B}_{m,p}$ -verteilt, so ist deren Summe $\mathcal{B}_{nm,p}$ -verteilt.

2.6.8 Beispiel: Versagsrate von Glühbirnen

Betrachtet sei eine Art von Glühbirnen, deren Lebenszeit $X_i : (\Omega, \mathfrak{S}, \mathcal{P}) \to [0, \infty)$, $i \in \mathbb{N}$ (unabhängig von einander) exponentialverteilt $\sim \mathrm{Exp}_\lambda$ sei. Nach dem *Ausfall* der k-ten Glühbirne, sei diese ersetzt durch eine *neue*, $(k+1)$-te Glühbirne. So ist

$$S_n := \sum_{k=1}^n X_k$$

der Zeitpunkt, an dem die n-te Glühbirne ausfällt.

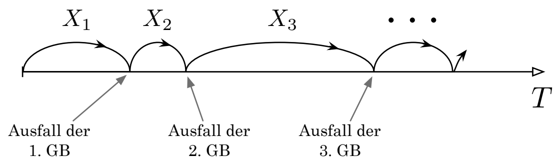

Abbildung 6: Typischer Verlauf des Glühbirnen-Austauschzyklus.

Gesucht ist nun die Wahrscheinlichkeit p_k, dass bis zum Zeitpunkt $T > 0$ genau k Glühbirnen ausfallen. Mit

$$A_k := \{S_k \leq T\} \quad , \quad k \in \mathbb{N}_0$$

gilt

$$p_k = \mathcal{P}\big(\underbrace{\{S_k \leq T,\ S_{k+1} > T\}}_{A_k \cap A_{k+1}^c}\big) = \mathcal{P}(A_k) - \mathcal{P}(A_k \cap A_{k+1}) \overset{A_{k+1} \subseteq A_k}{=} \mathcal{P}(A_k) - \mathcal{P}(A_{k+1})$$

$$\overset{(2.6.6.1)}{=} \int_0^T \frac{\lambda^k}{(k-1)!} x^{k-1} e^{-\lambda x}\ dx - \int_0^T \frac{\lambda^{k+1}}{k!} x^k e^{-\lambda x}\ dx = \frac{(\lambda T)^k}{k!} e^{-\lambda T} = \underbrace{\mathcal{P}_{\lambda T}}_{\substack{\text{Poisson-}\\\text{verteilung}}} (\{k\})$$

das heißt die Ausfälle bis zur Zeit T sind Poissonverteilt mit Parameter λT. Mit $\mathbb{E}\mathcal{P}_{\lambda T} = \lambda T$ und $\mathbb{E}\,\mathrm{Exp}_\lambda = 1/\lambda$, entspricht das Ergebnis der Intuition

$$\text{Mittlere Glühbirnenausfälle} = \frac{\text{Betriebszeit}}{\text{mittlere Lebensdauer}}$$

2.6.9 Beispiel: Umsatzverteilung eines Kaufhauses

Betrachtet sei ein Kaufhaus, dessen Kunden-Tagesanzahl $N \sim \mathcal{P}_\lambda$ Poissonverteilt sei. Jeder (i-te) Kunde kaufe unabhängig von den anderen Kunden, Waren eines zufälligen Wertes X_i ein, dessen Betrag das Verteilungsgesetz $\mathcal{Q}$ auf $\mathbb{R}$ besitze.

Gesucht sei nun das Verteilungsgesetz des täglichen Kaufhaus-Umsatzes G: Es ergibt sich

$$\mathcal{P}(G \in B) = \sum_{k=0}^{\infty} \underbrace{\mathcal{P}\left(G \in B \mid N = k\right)}_{\mathcal{P}(X_1 + \cdots + X_k \in B)} \cdot \underbrace{\mathcal{P}(N = k)}_{\frac{\lambda^k}{k!} e^{-\lambda}} = \sum_{k=0}^{\infty} \mathcal{Q}^{*k}(B) \cdot \frac{\lambda^k}{k!} e^{-\lambda}$$

sprich

$$\mathcal{P}_G = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \cdot \mathcal{Q}^{*k} \tag{2.6.9.1}$$

wobei

$$\mathcal{Q}^{*k} := \begin{cases} \overbrace{\mathcal{Q} * \cdots * \mathcal{Q}}^{k \text{ mal}} & : k \in \mathbb{N} \\ \delta_0 & : k = 0 \end{cases}$$

Das Verteilungsgesetz von G heißt *zusammengesetzte Poissonverteilung.*

2.7 Konstruktion unabhängiger Zufallsvariablen

2.7.1 Einführung

Eine zentrale Frage der Wahrscheinlichkeitstheorie ist, ob es zu vorgegebenen Wahrscheinlichkeitsräumen $(S_i, \mathcal{S}_i, \mathcal{P}_i)_{i\in I}$ stets einen Grundraum $(\Omega, \mathfrak{S}, \mathcal{P})$ und unabhängige Zufallsvariablen $X_i : \Omega \to S_i$ mit $\mathcal{P}_{X_i} = \mathcal{P}_i$ gibt.

Gibt es zum Beispiel überhaupt ein *Modell*, das einen unendlich wiederholten, fairen Würfelwurf beschreibt? Das heißt, gibt es einen Wahrscheinlichkeitsraum $(\Omega, \mathfrak{S}, \mathcal{P})$ mit unabhängigen, auf $\{1, .., 6\}$ gleichverteilten Zufallsvariablen $X_1, X_2, ... : \Omega \to \{1, .., 6\}$?

2.7.2 Satz: Existenz unabhängiger Zufallsvariablen

Seien $(S_i, \mathcal{S}_i, \mathcal{P}_i)_{i\in I}$ vorgegebene Wahrscheinlichkeitsräume. Dann existieren ein Wahrscheinlichkeitsraum $(\Omega, \mathfrak{S}, \mathcal{P})$ und unabhängige Zufallsvariablen

$$X_i : \Omega \to S_i \quad , \quad i \in I$$

mit $\mathcal{P}_{X_i} = \mathcal{P}_i,\ i \in I$.

Beweis: Man wähle

$$\Omega := \bigtimes_{i\in I} S_i \quad , \quad \mathfrak{S} := \bigotimes_{i\in I} \mathcal{S}_i \quad , \quad \mathcal{P} := \bigotimes_{i\in I} \mathcal{P}_i$$

und definiere

$$X_i : \Omega \to S_i \quad , \quad X_i\left((s_i)_{i\in I}\right) := s_i$$

Dann ist per Konstruktion $\mathcal{P}_{X_i} = \mathcal{P}_i$ und $(X_i)_{i\in I} = \mathrm{Id}_\Omega$ also $\mathcal{P}_{(X_i)_{i\in I}} = \mathcal{P}$. Nach Satz 2.4.2 folgt dann die Unabhängigkeit der X_i.
□

Beispiele:

(i) Der einzelne Münzwurf sei durch

$$(\Omega_i, \mathfrak{S}_i, \mathcal{P}_i) = \left(\{0, 1\}\,,\ \mathscr{P}(\{0, 1\}),\ \frac{1}{2}(\delta_0 + \delta_1)\right)$$

modelliert. Dann modelliert

$$(\Omega, \mathfrak{S}, \mathcal{P}) := \left(\{0, 1\}^n\,,\ \mathscr{P}\left(\{0, 1\}^n\right),\ \left(\frac{\delta_0 + \delta_1}{2}\right)^n \right) \quad , \quad X_i(\underbrace{(\omega_1, .., \omega_n)}_{\in\Omega}) := \omega_i\ ,\ i = 1, .., n$$

den n-maligen Münzwurf mit den Ergebnissen $X_1, .., X_n$.

(ii) Es sei die Länge einer am Fließband produzierten Werkzeugart n-mal gemessen. Die Ergebnisse seien dabei unabhängig, $\mathcal{N}_{\mu,\sigma^2}$ -verteilt. Ein mögliches Modell für dieses Vorgehen bietet

$$\Omega := \mathbb{R}^n\ ,\ \mathfrak{S} := \mathscr{B}(R^n)\ ,\ \mathcal{P} := \bigotimes_{k=1}^{n} \mathcal{N}_{\mu,\sigma^2}\ ,\ X_k(\underbrace{x_1, .., x_n}_{\in\mathbb{R}^n}) := \omega_k$$

Dabei besitzt $\mathcal{P}$ bzgl. λ^n die Dichte

$$\frac{d\mathcal{P}}{d\lambda^n}(\mathbf{x}) = \prod_{k=1}^{n} \frac{d\mathcal{N}_{\mu,\sigma^2}}{d\lambda^1}(x_k) = \frac{1}{(2\pi)^{\frac{n}{2}}} \frac{1}{\sigma^n} \cdot \exp\left[-\frac{1}{2\sigma^2} \sum_{k=1}^{n} (x_k - \mu)^2 \right]$$

Letztendlich ist die konkrete Gestalt des Grundraumes $(\Omega, \mathfrak{S}, \mathcal{P})$ und der Zufallsvariablen X_i belanglos.

2.7.3 Beispiel: Größtabschätzung normalverteilter Zufallsvariablen

Die Zufallsvariablen $X_k : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R},\ k \in \mathbb{N}$ seien unabhängig, $\sim \mathcal{N}_{0,1}$ standardnormalverteilt, dazu die Folge $(C_k)_{k\in\mathbb{N}} \subseteq (0, \infty)$. Gesucht sei die Wahrscheinlichkeit

$$\mathcal{P}\left(\sup_{k\in\mathbb{N}} \frac{|X_k|}{C_k} \leq 1\right)$$

dass alle X_k jeweils nicht größer als C_k sind. Es gilt

$$\mathcal{P}\left(\sup_{k\in\mathbb{N}} \frac{|X_k|}{C_k} \leq 1\right) = \mathcal{P}\left(|X_k| \leq C_k\ \forall\ k \in \mathbb{N}\right) \overset{\substack{\text{Erläuterung} \\ 2.4.1}}{=} \prod_{k=1}^{\infty} \mathcal{P}\left(|X_k| \leq C_k\right) = \prod_{k=1}^{\infty} \left[\frac{1}{\sqrt{2\pi}} \int\limits_{-C_k}^{C_k} e^{-t^2/2} dt\right]$$

$$= \prod_{k=1}^{\infty} \left[\sqrt{\frac{2}{\pi}} \int\limits_{0}^{C_k} e^{-t^2/2} dt\right] = \prod_{k=1}^{\infty} \left[1 - \sqrt{\frac{2}{\pi}} \int\limits_{C_k}^{\infty} e^{-t^2/2} dt\right] = (\clubsuit)$$

Insbesondere ist $(\clubsuit) > 0$ genau dann wenn

$$\sum_{k=1}^{\infty} \int\limits_{C_k}^{\infty} e^{-t^2/2} dt < \infty$$

Mit

$$F(x) := \int\limits_{x}^{\infty} e^{-t^2/2} dt \overset{x\to\infty}{\longrightarrow} 0 \quad , \quad G(x) := \frac{e^{-x^2/2}}{x} \overset{x\to\infty}{\longrightarrow} 0$$

gilt

$$\frac{F'(x)}{G'(x)} = \frac{1}{1 - \frac{1}{x^2}} \overset{x\to\infty}{\longrightarrow} 1$$

bzw.

$$\frac{F(x)}{G(x)} \overset{x\to\infty}{\longrightarrow} 1$$

Dementsprechend gilt

$$\sum_{k=1}^{\infty} F(C_k) < \infty \quad \Leftrightarrow \quad \sum_{k=2}^{\infty} G(C_k) < \infty \quad \Leftrightarrow \quad \sum_{k=2}^{\infty} \frac{e^{-C_k^2/2}}{C_k} < \infty$$

Speziell folgt für $C_k := \sqrt{2\alpha \ln k}$:

$$(\clubsuit) > 0 \quad \Leftrightarrow \quad \sum_{k=2}^{\infty} \frac{k^{-\alpha}}{\sqrt{2\alpha \ln k}} < \infty \quad \Leftrightarrow \quad \alpha > 1$$

sprich

$$\boxed{\mathcal{P}\left(\sup_{k\geq 2} \frac{|X_k|}{\sqrt{2\alpha \log k}} \leq 1\right) > 0 \quad \Leftrightarrow \quad \alpha > 1}$$

3 Gesetz der großen Zahlen

3.1 Oberer und unterer Limes von Mengen

3.1.1 Definition: Oberer & unterer Limes von Mengenfolgen

Gegeben sei eine Mengenfolge $(A_n)_{n\in\mathbb{N}}$. Dann heißt

$$\overline{\lim_{n\to\infty}}\, A_n := \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n$$

oberer Limes und

$$\underline{\lim_{n\to\infty}}\, A_n := \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n$$

unterer Limes von $(A_n)_n$.

Interpretation: Aus

$$\omega \in \overline{\lim_{n\to\infty}}\, A_n \Leftrightarrow \forall\, m \in \mathbb{N} : \exists\, n \geq m : \omega \in A_n$$

$$\Leftrightarrow \#\{n \in \mathbb{N} : \omega \in A_n\} = \infty$$

erkennt man, dass $\overline{\lim\limits_{n\to\infty}}\, A_n$ aus genau den Elementen besteht, die *unendlich oft* in den A_n vorkommen. Ähnlich gilt

$$\omega \in \underline{\lim_{n\to\infty}}\, A_n \Leftrightarrow \exists\, m \in \mathbb{N} : \forall\, n \geq m : \omega \in A_n$$

das heißt $\underline{\lim\limits_{n\to\infty}}\, A_n$ besteht aus den Elementen die *schließlich immer* in A_n sind.

3.1.2 Eigenschaften des unteren & oberen Mengenlimes

Sei $(A_n)_{n\in\mathbb{N}}$ eine Mengenfolge. Dann gilt stets:

1. $\underline{\lim\limits_{n\to\infty}}\, A_n \subseteq \overline{\lim\limits_{n\to\infty}}\, A_n$
2. $\left(\underline{\lim\limits_{n\to\infty}}\, A_n\right)^c = \overline{\lim\limits_{n\to\infty}}\, A_n^c$
3. $1_{\overline{\lim\limits_{n\to\infty}}\, A_n} = \overline{\lim\limits_{n\to\infty}}\, 1_{A_n}$
4. $1_{\underline{\lim\limits_{n\to\infty}}\, A_n} = \underline{\lim\limits_{n\to\infty}}\, 1_{A_n}$
5. Für jede Teilfolge $(A_{n_k})_{k=1}^{\infty}$ gilt $\overline{\lim\limits_{k\to\infty}}\, A_{n_k} \subseteq \overline{\lim\limits_{n\to\infty}}\, A_n$.
6. Für jede Teilfolge $(A_{n_k})_{k=1}^{\infty}$ gilt $\underline{\lim\limits_{k\to\infty}}\, A_{n_k} \supseteq \underline{\lim\limits_{n\to\infty}}\, A_n$.

3.1.3 Lemma über obere & untere Mengenlimite in Wahrscheinlichkeitsräumen

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $(A_n)_{n\in\mathbb{N}} \subseteq \mathfrak{S}$ eine Folge messbarer Mengen.

1. Gilt $\mathcal{P}(A_n) \geq \delta\ \forall\, n \in \mathbb{N}$ so gilt auch
$$\mathcal{P}\Big(\overline{\lim_{n\to\infty}}\, A_n\Big) \geq \delta$$
2. Gilt $\mathcal{P}(A_n) \leq \delta\ \forall\, n \in \mathbb{N}$ so gilt auch
$$\mathcal{P}\Big(\underline{\lim_{n\to\infty}}\, A_n\Big) \leq \delta$$

Beweis:

1.
$$\mathcal{P}\Big(\overline{\lim_{n\to\infty}}\ A_n\Big) = \mathcal{P}\Big[\bigcap_{m=1}^{\infty} \underbrace{\bigcup_{n=m}^{\infty} A_n}_{\text{fallend in } m}\Big] = \lim_{m\to\infty} \underbrace{\mathcal{P}\Big(\bigcup_{n=m}^{\infty} A_n\Big)}_{\geq\delta} \geq \delta$$

2.
$$\mathcal{P}\Big(\underline{\lim_{n\to\infty}}\ A_n\Big) = \mathcal{P}\Big[\bigcup_{m=1}^{\infty} \underbrace{\bigcap_{n=m}^{\infty} A_n}_{\text{wachsend in } m}\Big] = \lim_{m\to\infty} \underbrace{\mathcal{P}\Big(\bigcap_{n=m}^{\infty} A_n\Big)}_{\leq\delta} \leq \delta$$

□

3.2 Konvergenz zufälliger Größen

3.2.1 Definition: Konvergenz von Zufallsvariablen

Sei (E,d) ein separabler metrischer Raum, $(\Omega,\mathfrak{S},\mathcal{P})$ ein Wahrscheinlichkeitsraum und $(X_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen $X_n:\Omega\to E$.
Dann:

- (X_n) *konvergiert stochastisch* oder *in Wahrscheinlichkeit* gegen die Zufallsvariable $X:\Omega\to E$ falls

$$\forall\ \varepsilon>0\ :\ \lim_{n\to\infty} \underbrace{\mathcal{P}\left(d(X_n,X)>\varepsilon\right)}_{\mathcal{P}\{\omega\ :\ d(X_n(\omega),X(\omega))>\varepsilon\}} = 0$$

  Man schreibt $X_n \underset{p}{\overset{n\to\infty}{\longrightarrow}} X$.

- (X_n) konvergiert *fast sicher* gegen X, falls

$$\mathcal{P}\left(\lim_{n\to\infty} X_n = X\right) = \mathcal{P}\left(\left\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}\right) = 1$$

  das heißt X_n konvergiert $\mathcal{P}$-fast-überall punktweise (in E) gegen X. Man schreibt $X_n \underset{\text{f.s.}}{\overset{n\to\infty}{\longrightarrow}} X$.

- X_n konvergiert *sicher*, oder *punktweise* gegen X, falls

$$X_n(\omega) \overset{n\to\infty}{\longrightarrow} X(\omega)\quad \forall\ \omega\in\Omega$$

- (X_n) *konvergiert in Verteilung* oder *schwach* gegen X, wenn die Folge der induzierten Bildmaße $\mathcal{P}_{X_n}$ schwach gegen das Bildmaß $\mathcal{P}_X$ konvergiert (vgl. später: 4.4), das heißt für jede beschränkte, stetige $f:E\to\mathbb{R}$ geht

$$\lim_{n\to\infty} \underbrace{\int_E f\ d\mathcal{P}_{X_n}}_{\mathbb{E}f(X_n)} = \underbrace{\int_E f\ d\mathcal{P}_X}_{\mathbb{E}f(X)}$$

  Man schreibt $X_n \underset{d}{\overset{n\to\infty}{\longrightarrow}} X$. Ähnlich definiert man auch *Konvergenz in Verteilung* von (X_n) gegen irgendein Wahrscheinlichkeitsmaß $\mathcal{P}$.

- (X_n) konvergiert gegen X *im p-ten Mittel* $(0<p<\infty)$, falls

$$\lim_{n\to\infty} \mathbb{E}\left\{d\left(X_n,X\right)^p\right\} = \lim_{n\to\infty} \int_\Omega \left|d\left(X_n(\omega),X(\omega)\right)\right|^p\ d\mathcal{P}(\omega) = 0$$

  Im Spezialfall $p=1$ sagt man auch (X_n) konvergiert *im Mittel*, im Fall $p=2$ *im quadratischen Mittel* gegen X.

Bemerkung: Da (E, d) separabel ist, gilt $\mathscr{B}(E \times E) = \mathscr{B}(E) \otimes \mathscr{B}(E)$. Da $d : E \times E \to \mathbb{R}$ stetig ist, ist d bzgl. $\mathscr{B}(E) \times \mathscr{B}(E)$ messbar.

3.2.2 Satz über Konvergenz von Zufallsvariablen

Sei (E, d) ein separabler metrischer Raum und $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum. Dann gilt für Zufallsvariablen $X, X_n : \Omega \to (E, \mathscr{B}(E)),\ n \in \mathbb{N}$:

1. Sind $0 < p \leq q < \infty$, so impliziert Konvergenz im q-ten Mittel auch Konvergenz im p-ten Mittel:

$$\mathbb{E}\,|d(X_n, X)|^q \overset{n\to\infty}{\longrightarrow} 0 \quad \overset{p\leq q}{\Longrightarrow} \quad \mathbb{E}\,|d(X_n, X)|^p \overset{n\to\infty}{\longrightarrow} 0$$

2. Fast sichere Konvergenz impliziert Konvergenz in Wahrscheinlichkeit

$$X_n \underset{\text{f.s}}{\overset{n\to\infty}{\longrightarrow}} X \quad \Rightarrow \quad X_n \underset{\text{p}}{\overset{n\to\infty}{\longrightarrow}} X$$

Beweis: Angenommen $X_n \underset{\text{p}}{\overset{n\to\infty}{\not\longrightarrow}} X$, dann existieren $\varepsilon, \delta > 0$ und Teilfolge $(X_{n_k})_{k=1}^{\infty}$ mit

$$\mathcal{P}\left(d(X_{n_k}, X) \geq \varepsilon\right) \geq \delta$$

Nach 3.1.3(1) muss also

$$\mathcal{P}\Big(\overline{\lim_{k\to\infty}}\ \{d(X_{n_k}, X) \geq \varepsilon\}\Big) \geq \delta \quad \forall\, k \in \mathbb{N}$$

sein. Nach 3.1.2(5) also auch

$$\mathcal{P}\Big(\overline{\lim_{n\to\infty}}\ \{d(X_n, X) \geq \varepsilon\}\Big) \geq \delta$$

Doch wegen

$$\Big\{\lim_{n\to\infty} d(X_n, X) \neq 0\Big\} \supseteq \{d(X_n, X) \geq \varepsilon \text{ u.o.}\}$$

würde folgen $X_n \underset{\text{f.s.}}{\overset{n\to\infty}{\not\longrightarrow}} X$.

3. Konvergenz im p-ten Mittel impliziert Konvergenz in Wahrscheinlichkeit:

$$\mathbb{E}\,|d(X_n, X)^p| \overset{n\to\infty}{\longrightarrow} 0 \quad \Rightarrow \quad X_n \underset{\text{p}}{\overset{n\to\infty}{\longrightarrow}} X$$

Beweis:

$$\mathcal{P}\left(d(X_n, X) \geq \varepsilon\right) = \mathcal{P}\left(|d(X_n, X)|^p \geq \varepsilon^p\right) \overset{1.8.3}{\leq} \varepsilon^{-p} \cdot E\,|d(X_n, X)|^p$$

4. Konvergenz in Wahrscheinlichkeit impliziert stets Konvergenz in Verteilung:

$$X_n \underset{\text{p}}{\overset{n\to\infty}{\longrightarrow}} X \quad \Rightarrow \quad X_n \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} X$$

Die Umkehrung gilt im allgemeinen nicht!

5. Konvergiert X_n in Verteilung gegen eine Konstante C, so konvergiert X_n auch in Wahrscheinlichkeit gegen C:

$$X_n \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} C : \text{const} \quad \Rightarrow \quad X_n \underset{\text{p}}{\overset{n\to\infty}{\longrightarrow}} C$$

6. Seien $a_n \overset{n\to\infty}{\longrightarrow} a \in E$ und $X_n \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} X$. Dann gehen auch

$$a_n X_n \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} aX$$

7. Geht $X_n \underset{\text{p}}{\overset{n\to\infty}{\longrightarrow}} X$, so existieren $n_1 < n_2 < \ldots$ mit

$$\underbrace{X_{n_k}}_{\text{Teilfolge}} \underset{\text{f.s.}}{\overset{k\to\infty}{\longrightarrow}} X$$

8. Es geht $X_n \underset{\text{p}}{\overset{n\to\infty}{\longrightarrow}} X$ genau dann wenn:
$$\forall\, n_1 < n_2 < \cdots \in \mathbb{N} : \exists\, m_1 < m_2 < \cdots : X_{n_{m_k}} \underset{\text{f.s.}}{\overset{k\to\infty}{\longrightarrow}} X$$

9. Definiert man die Metrik
$$d_0(X,Y) := \mathbb{E} \min\{1, d(X,Y)\}$$
auf der Menge aller Zufallsvariablen $\Omega \to E$, so gilt:
$$X_n \underset{\text{p}}{\overset{n\to\infty}{\longrightarrow}} X \quad \Leftrightarrow \quad d_0(X_n, X) \overset{n\to\infty}{\longrightarrow} 0$$
das heißt, die Konvergenz in Wahrscheinlichkeit ist sogar eine metrische Konvergenz.

$$\begin{tikzcd}
 & \boxed{X_n \underset{\text{f.s}}{\overset{n\to\infty}{\longrightarrow}} X} \arrow[d, Rightarrow] & \boxed{\mathbb{E}\,|d(X_n,X)|^p \overset{n\to\infty}{\longrightarrow} 0} \arrow[dl, Rightarrow] \\
\boxed{\exists\,(n_k)_k : X_{n_k} \underset{\text{f.s.}}{\overset{k\to\infty}{\longrightarrow}} X} & \boxed{X_n \underset{\text{p}}{\overset{n\to\infty}{\longrightarrow}} X} \arrow[l, Rightarrow] \arrow[d, Rightarrow] & \\
 & \boxed{X_n \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} X} &
\end{tikzcd}$$

Abbildung 7: Zur Konvergenz von Zufallsgrößen in einem metrischen Raum.

3.2.3 Lemma über die Eindeutigkeit der Grenzwerte

Sei (E,d) ein separabler, metrischer Raum, $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $X, Y, (X_n)_{n\in\mathbb{N}}$ Zufallsvariablen $X, Y, X_n : \Omega \to E$. Dann gilt:

1. Konvergieren $X_n \underset{\text{p}}{\overset{n\to\infty}{\longrightarrow}} X$ und $X_n \underset{\text{p}}{\overset{n\to\infty}{\longrightarrow}} Y$ so muss $X = Y$ fast-sicher sein.
2. Konvergieren $X_n \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} X$ und $X_n \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} Y$ so muss $X \overset{\text{d.}}{=} Y$ sein.
3. Konvergieren $X_n \underset{\text{f.s.}}{\overset{n\to\infty}{\longrightarrow}} X$ und $X_n \underset{\text{f.s.}}{\overset{n\to\infty}{\longrightarrow}} Y$ so muss $X = Y$ fast-sicher sein.
4. Konvergieren $\mathbb{E}\,|d(X_n,X)|^p \to 0$ und $\mathbb{E}\,|d(X_n,Y)|^p \to 0$ für irgendein $0 < p < \infty$ so muss $X = Y$ fast-sicher sein.

Beweis:

1. Für jedes $\varepsilon > 0$ gilt
$$\mathcal{P}(d(X,Y) \geq \varepsilon) \leq \mathcal{P}\left[d(X,X_n) \geq \tfrac{\varepsilon}{2} \ \vee \ d(X_n,Y) \geq \tfrac{\varepsilon}{2}\right]$$
$$\leq \mathcal{P}\left[d(X,X_n) \geq \tfrac{\varepsilon}{2}\right] + \mathcal{P}\left[d(X_n,Y) \geq \tfrac{\varepsilon}{2}\right] \overset{n\to\infty}{\longrightarrow} 0$$
und daher
$$\mathcal{P}\left(d(X,Y) = 0\right) = 1 - \mathcal{P}\left[\bigcup_{n=1}^{\infty} \left\{d(X,Y) \geq \tfrac{1}{n}\right\}\right] = 1 - \lim_{n\to\infty} \underbrace{\mathcal{P}\left[d(X,Y) \geq \tfrac{1}{n}\right]}_{0} = 1$$

2. Siehe Satz 4.4.3.
3. Klar.
4. Im Falle $0 < p \leq 1$ gilt:
$$\mathbb{E}\,|d(X,Y)|^p = \|d(X,Y)\|_p^p \leq \|d(X,X_n) + d(X_n,Y)\|_p^p \leq \|d(X,X_n)\|_p^p + \|d(X_n,Y)\|_p^p \overset{n\to\infty}{\longrightarrow} 0$$
und analog auch für $1 \leq p < \infty$:
$$\mathbb{E}\,|d(X,Y)|^p = \|d(X,Y)\|_p^p \leq \|d(X,X_n) + d(X_n,Y)\|_p^p \leq \Big[\,\|d(X,X_n)\|_p + \|d(X_n,Y)\|_p\Big]^p \overset{n\to\infty}{\longrightarrow} 0$$
sprich $d(X,Y) = 0$ fast-überall.

$\square$

3.2.4 Lemma: Konvergenz zufälliger Vektoren

Seien $X^i, (X_n^i)_{n\in\mathbb{N}} : (\Omega, \mathfrak{S}, \mathcal{P}) \to (E_i, \mathscr{B}(E_i)),\ i = 1,..,m$ Zufallsvariablen in die separablen, metrischen Räume $(E_i, d_i),\ i = 1,..,m$. Dazu sei
$$E := \bigtimes_{i=1}^m E_i$$
ausgestattet mit der Metrik
$$d(\mathbf{x},\mathbf{y}) := \sqrt{\sum_{i=1}^m d_i(x^i,y^i)^2} \quad , \quad \mathbf{x} = (x^1,..,x^m), \mathbf{y} = (y^1,..,y^m) \in E$$
Beachte dass dann auch (E,d) separabel ist und [4]
$$\bigotimes_{i=1}^m \mathscr{B}(E_i,d_i) = \mathscr{B}(E,d)$$
Es gilt:

1. Es konvergieren
$$X_n^i \underset{\text{f.s.}}{\overset{n\to\infty}{\longrightarrow}} X^i \quad \forall\ i = 1,..,m$$
genau dann wenn
$$\mathbf{X}_n := (X_n^1,..,X_n^m) \underset{\text{f.s.}}{\overset{n\to\infty}{\longrightarrow}} (X^1,..,X^m) =: \mathbf{X}$$
in $(E, \mathscr{B}(E))$ konvergieren.
2. Es konvergieren
$$X_n^i \underset{\text{p}}{\overset{n\to\infty}{\longrightarrow}} X^i \quad \forall\ i = 1,..,m$$
genau dann wenn
$$\mathbf{X}_n := (X_n^1,..,X_n^m) \underset{\text{p}}{\overset{n\to\infty}{\longrightarrow}} (X^1,..,X^m) =: \mathbf{X}$$
in $(E, \mathscr{B}(E))$ konvergieren.
3. Sei $0 < p < \infty$ vorgegeben. Dann konvergieren
$$\mathbb{E}\left|d_i(X_n^i, X^i)\right|^p \overset{n\to\infty}{\longrightarrow} 0 \quad \forall\ i = 1,..,m$$
genau dann wenn
$$\mathbb{E}\,|d(\mathbf{X}_n,\mathbf{X})|^p \overset{n\to\infty}{\longrightarrow} 0$$
gehen.

[4]Notwendige Voraussetzung ist die separabilität der E_i.

Beweis:

1. Wegen
$$X_n^i(\omega) \overset{n\to\infty}{\longrightarrow} X^i(\omega) \;\; \forall\, i \quad \Leftrightarrow \quad \mathbf{X}_n(\omega) \overset{n\to\infty}{\longrightarrow} \mathbf{X}(\omega)$$
gilt
$$\left\{X_n^i \overset{n\to\infty}{\longrightarrow} X^i \;\; \forall\, i\right\} = \left\{\mathbf{X}_n \overset{n\to\infty}{\longrightarrow} \mathbf{X}\right\}$$
Anderseits ist
$$\left\{X_n^i \overset{n\to\infty}{\longrightarrow} X^i \;\; \forall\, i\right\}^c = \bigcup_{i=1}^m \left\{X_n^i \overset{n\to\infty}{\longrightarrow} X^i\right\}^c$$
genau dann in einer Nullmenge enthalten wenn alle $\left\{X_n^i \overset{n\to\infty}{\longrightarrow} X^i\right\}^c$ in jeweils einer Nullmenge enthalten sind, woraus die Behauptung folgt.

2. Wegen
$$\bigcup_{i=1}^m \left\{d_i(X_n^i, X^i) \geq \varepsilon\right\} \subseteq \left\{d(\mathbf{X}_n, \mathbf{X}) \geq \varepsilon\right\} \subseteq \bigcup_{i=1}^m \left\{d_i(X_n^i, X^i) \geq \frac{\varepsilon}{\sqrt{m}}\right\} \quad \forall\, \varepsilon > 0$$
gehen
$$\mathcal{P}\left(d(\mathbf{X}_n, \mathbf{X}) \geq \varepsilon\right) \overset{n\to\infty}{\longrightarrow} 0 \;\; \forall\, \varepsilon > 0$$
genau dann wenn alle
$$\mathcal{P}\left(d_i(X_n^i, X^i) \geq \varepsilon\right) \overset{n\to\infty}{\longrightarrow} 0 \;\; \forall\, \varepsilon > 0, \; i = 1, .., m$$
gehen.

3. Wegen
$$\mathbb{E}\left|d_i(X_n^i, X^i)\right|^p \leq \mathbb{E}\left|d(\mathbf{X}_n, \mathbf{X})\right|^p \quad \forall\, n, i$$
folgt Richtung "$\Leftarrow$" der Behauptung. Wegen
$$\mathbb{E}\left|d(\mathbf{X}_n, \mathbf{X})\right|^p = \left\|\sum_{i=1}^m d_i(X_n^i, X^i)^2\right\|_{L_{p/2}}^{p/2} \leq \sum_{i=1}^m \left\|d_i(X_n^i, X^i)^2\right\|_{L_{p/2}}^{p/2} = \sum_{i=1}^m \mathbb{E}\left|d_i(X_n^i, X^i)\right|^p$$
für $p < 2$ bzw.
$$\mathbb{E}\left|d(\mathbf{X}_n, \mathbf{X})\right|^p = \left\|\sum_{i=1}^m d_i(X_n^i, X^i)^2\right\|_{L_{p/2}}^{p/2} \leq \left[\sum_{i=1}^m \left\|d_i(X_n^i, X^i)^2\right\|_{L_{p/2}}\right]^{p/2} = \left[\sum_{i=1}^m \sqrt[p/2]{\mathbb{E}\left|d_i(X_n^i, X^i)\right|^p}\right]^{p/2}$$
für $p \geq 2$ gilt auch die Umkehrung.

$\square$

Beachte: Analoges gilt für die Konvergenz in Verteilung **nicht**! Sind z.B. $X : \Omega \to E_1, Y : \Omega \to E_2$ unabhängige, identisch verteilte Zufallsgrößen, so gehen zwar

$$X \underset{\mathrm{d}}{\longrightarrow} X, \;\; X \underset{\mathrm{d}}{\longrightarrow} Y$$

doch $(X, X) \underset{\mathrm{d}}{\not\longrightarrow} (X, Y)$.

3.2.5 Unterer & oberer Mengenlimes bei Zufallsvariablen

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum.

1. Sind $X_1, X_2, .. : \Omega \to \mathbb{R}$ Zufallsvariablen, so gilt für Konstante $C \in \mathbb{R}$:

$$\left\{\varlimsup_{n\to\infty} X_n = C\right\} = \bigcap_{k=1}^{\infty} \left[\left\{X_n \geq C - \frac{1}{k} \ \substack{\text{unendlich}\\ \text{oft}}\right\} \cap \left\{X_n \leq C + \frac{1}{k} \ \substack{\text{schließlich}\\ \text{immer}}\right\}\right]$$

$$= \bigcap_{k=1}^{\infty} \left[\varlimsup_{n\to\infty} X_n^{-1}\left([C - \tfrac{1}{k}, \infty)\right) \cap \varliminf_{n\to\infty} X_n^{-1}\left((-\infty, C + \tfrac{1}{k}]\right)\right]$$

$$\left\{\varlimsup_{n\to\infty} X_n \geq C\right\} = \bigcap_{k=1}^{\infty} \left\{X_n \geq C - \frac{1}{k} \ \substack{\text{unendlich}\\ \text{oft}}\right\} = \bigcap_{k=1}^{\infty} \varlimsup_{n\to\infty} X_n^{-1}\left((C - \tfrac{1}{k}, \infty]\right)$$

2. Ist $(E, \|\cdot\|)$ ein normierter Raum und $X_1, X_2, .. : \Omega \to (E, \mathscr{B}(E))$ Zufallsvariablen, so gilt

$$\left\{\lim_{n\to\infty} X_n = 0\right\} = \bigcap_{k=1}^{\infty} \left\{\|X_n\| \leq \tfrac{1}{k} \ \substack{\text{schließlich}\\ \text{immer}}\right\}$$

$$= \bigcap_{k=1}^{\infty} \left\{\omega \in \Omega : \exists\, n_{k,\omega} : \|X_n(\omega)\| \leq \tfrac{1}{k} \ \ \forall\, n \geq n_{k,\omega}\right\}$$

3. Ist (E, d) ein separabler, metrischer Raum und $X_1, X_2, .. : \Omega \to (E, \mathscr{B}(E))$ Zufallsvariablen, so gilt

$$\{\omega \in \Omega : (X_n(\omega))_{n\in\mathbb{N}} \text{ Cauchyfolge}\} = \bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=1}^{\infty} \left\{d(X_n, X_{n+m}) \leq \tfrac{1}{k}\right\}$$

4. Ist (E, d) ein separabler, metrischer Raum und $X, X_1, X_2, .. : \Omega \to (E, \mathscr{B}(E))$ Zufallsvariablen, so gilt

$$X_n \xrightarrow[\text{f.s.}]{n\to\infty} X \quad \Leftrightarrow \quad \forall\, \varepsilon > 0 \ : \ \mathcal{P}\left(\left\{d(X_n, X) \leq \varepsilon \ \substack{\text{schließlich}\\ \text{immer}}\right\}\right) = 1$$

Beweis:

$$X_n \xrightarrow[\text{f.s.}]{n\to\infty} X \Leftrightarrow \mathcal{P}\left(\bigcap_{k=1}^{\infty} \left\{d(X_n, X) \leq \tfrac{1}{k} \text{ s.i.}\right\}\right) = 1$$

$$\Leftrightarrow \lim_{k\to\infty} \underbrace{\mathcal{P}\left(d(X_n, X) \leq \tfrac{1}{k} \text{ s.i}\right)}_{\text{fallend in } k} = 1$$

$$\Leftrightarrow \forall\, k \in \mathbb{N} : \mathcal{P}\left(d(X_n, X) \leq \tfrac{1}{k} \text{ s.i}\right) = 1$$

□

3.3 Null-Eins-Gesetze

3.3.1 Definition: Terminale Ereignisse

Sei $(\Omega, \mathfrak{S})$ ein messbarer Raum und $\mathfrak{S}_1, \mathfrak{S}_2, .. \subseteq \mathfrak{S}$ Teil-σ-Algebren von $\mathfrak{S}$. Dann setzt man

$$\mathfrak{S}_{(m)} := \sigma\left(\bigcup_{n=m}^{\infty} \mathfrak{S}_n\right)$$

und nennt

$$\mathfrak{S}_\infty := \bigcap_{m=1}^{\infty} \mathfrak{S}_{(m)}$$

die σ-Algebra der *terminalen Ereignisse* (engl. *remote events* oder *tail events*).

Interpretation: Sind $X_n : \Omega \to (S_n, \mathcal{S}_n),\ n \in \mathbb{N}$ Zufallsvariablen und $\mathfrak{S}_n := \sigma(X_n)$, so liegen in

$$\mathfrak{S}_{(m)} = \sigma(X_n : n \geq m)$$

die Ereignisse, die sich aus Kenntnis des m-ten, $(m+1)$-ten usw. Zufallswertes entscheiden lassen. Ein Ereignis $A \in \mathfrak{S}$ liegt genau dann in

$$\mathfrak{S}_\infty = \bigcap_{m=1}^{\infty} \sigma(X_n : n \geq m)$$

wenn $A \in \mathfrak{S}_{(m)}\ \forall\ m \in \mathbb{N}$, es also gewissermaßen *stets von der Zukunft abhängt.*

Beispiele:

1. Betrachte ein unendlich mahl wiederholtes Würfelexperiment, modelliert durch den messbaren Raum $\Omega := \{1,..,6\}^{\mathbb{N}}$ und den Zufallsvariablen

$$X_n : \Omega \to (\{1,..,6\}, \mathscr{P}(\{1,..,6\})) \quad , \quad X_n(\omega) := \omega_n,\ \omega = (\omega_n)_{n\in\mathbb{N}} \in \Omega$$

Mit

$$\mathfrak{S}_n := \sigma(X_n) = \left\{ \{1,..,6\}^{n-1} \times B \times \{1,..,6\} \times \cdots : B \subseteq \{1,..,6\} \right\}$$

ist

$$\mathfrak{S}_{(m)} = \sigma\left(\bigcup_{n=m}^{\infty} \mathfrak{S}_n \right) = \sigma(X_n : n \geq m) = \left\{ \{1,..,6\}^{n-1} \times A : A \in \bigotimes_{m=n}^{\infty} \mathscr{P}\{1,..,6\} \right\}$$

So ist z.B.

$$A := \{\omega \in \Omega : \#\{n \in \mathbb{N} : X_n(\omega) = 6\} = \infty\} \qquad (\infty\text{-}mal\ die\ 6\ Würfeln)$$

ein terminales Ereignis, was gewissermaßen stets von zukünftigen Würfen abhängt. Tatsächlich, ist

$$A = \bigcap_{k=1}^{\infty} \underbrace{\bigcup_{n=k}^{\infty} \{\omega \in \Omega : X_n(\omega) = 6\}}_{\text{Fallend in } k}$$

$$= \bigcap_{k=m}^{\infty} \underbrace{\bigcup_{n=k}^{\infty} \{\omega : X_n(\omega) = 6\}}_{\in\sigma(X_n : n\geq m)} \in \sigma(X_n : n \geq m) \quad \forall\ m$$

$$\Rightarrow A \in \bigcap_{m=1}^{\infty} \sigma(X_n : n \geq m) =: \mathfrak{S}_\infty$$

2. $X_1, X_2, .. : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}$ seien reelle Zufallsgrößen. Dann ist

$$A := \left\{ \omega \in \Omega : \lim_{n\to\infty} X_n = 0 \right\}$$

ein terminales Ereignis, denn:

$$A = \bigcap_{k=1}^{\infty} \bigcup_{l=1}^{\infty} \underbrace{\bigcap_{n=l}^{\infty} \{\omega : |X_n(\omega)| \leq \tfrac{1}{k}\}}_{\text{wachsend in } l}$$

$$= \bigcap_{k=1}^{\infty} \underbrace{\bigcup_{l=m}^{\infty} \bigcap_{n=l}^{\infty} \{\omega : |X_n(\omega)| \leq \tfrac{1}{k}\}}_{\in\sigma(X_n : n \geq m)} \quad \in \sigma(X_n : n \geq m) \quad \forall\, m \in \mathbb{N}$$

$$\Rightarrow A \in \bigcap_{m=1}^{\infty} \sigma(X_n : n \geq m) =: \mathfrak{S}_\infty$$

3. Seien $X_1, X_2, .. : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}$ reelle Zufallsvariablen. Zu $a \in \mathbb{R}$ und reelle Folge $C_n \stackrel{n\to\infty}{\longrightarrow} 0$ ist

$$A := \left\{ \omega \in \Omega \ : \ C_n \cdot \sum_{k=1}^{n} X_k(\omega) \stackrel{n\to\infty}{\longrightarrow} a \right\}$$

ein terminales Ereignis, denn:

$$C_n \cdot \sum_{k=1}^{n} X_k(\omega) \stackrel{n\to\infty}{\longrightarrow} a \quad \stackrel{(\clubsuit)}{\Leftrightarrow} \quad C_n \sum_{k=m}^{n} X_k(\omega) \stackrel{n\to\infty}{\longrightarrow} a \ \ \forall\, m \qquad \Bigg| \qquad (\clubsuit) : C_n \cdot \sum_{k=1}^{m-1} X_k(\omega) \stackrel{n\to\infty}{\longrightarrow} 0 \ \ \forall\, m$$

Ähnlich sind auch die Ereignisse

$$\left\{ \omega : \overline{\lim_{n\to\infty}} \ X_n(\omega) = C : \text{const} \right\}$$

terminal.

3.3.2 Definition: Terminale Zufallsvariable

Seien $(\Omega, \mathfrak{S})$ und $(S, \mathcal{S})$ messbare Räume und $\mathfrak{S}_1, \mathfrak{S}_2, .. \subseteq \mathfrak{S}$ teil-σ-Algebren von $\mathfrak{S}$, dazu die σ-Algebra $\mathfrak{S}_\infty \subseteq \mathfrak{S}$ der terminalen Ereignisse. Eine Abbildung $f : \Omega \to S$ die messbar bzgl. $\mathfrak{S}_\infty$ ist, heißt *terminale* Abbildung, bzw. *terminale Zufallsvariable* falls Ω ein Wahrscheinlichkeitsraum ist.

3.3.3 Lemma über von Zufallsvariablen induzierte terminale Ereignisse

Seien $X_n : (\Omega, \mathfrak{S}, \mathcal{P}) \to (S_n, \mathcal{S}_n), \ n \in \mathbb{N}$ Zufallsvariablen und

$$f_n : \left(\bigtimes_{k=n}^{\infty} S_k, \bigotimes_{k=n}^{\infty} \mathcal{S}_k \right) \to \underbrace{(M_n, \mathcal{M}_n)}_{\substack{\text{messbarer} \\ \text{Raum}}} \quad , \quad n \in \mathbb{N}$$

messbare Funktionen, dazu die Zufallsvariablen $Y_n := f_n(X_n, X_{n+1}, ..)$. Dann sind alle bzgl. $(Y_n)_{\in\mathbb{N}}$ terminalen Ereignisse auch bzgl. $(X_n)_{n\in\mathbb{N}}$ terminal, das heißt

$$\bigcap_{m\in\mathbb{N}} \sigma(Y_n : n \geq m) \subseteq \bigcap_{m\in\mathbb{N}} \sigma(X_n : n \geq m)$$

Beweis: Wegen

$$\sigma(Y_n) = \sigma(f_n(X_n, X_{n+1}, ..)) \subseteq \sigma(X_k : k \geq n)$$

gilt

$$\sigma(Y_n : n \geq m) = \sigma\left[\bigcup_{n\geq m} \sigma(Y_n)\right] \subseteq \sigma\left[\bigcup_{n\geq m} \sigma(X_k : k \geq n)\right] = \sigma(X_n : n \geq m)$$

woraus die Behauptung folgt.
□

3.3.4 Terminalität bei Zufallsvariablen

Seien $X_1, X_2, .. : (\Omega, \mathfrak{S}) \to (S, \mathcal{S})$ messbar und $\mathfrak{S}_n := \sigma(X_n)$, dazu die σ-Algebra $\mathfrak{S}_\infty$ der terminalen Ereignisse. Dann gilt:

1. Für $B_n \in \mathcal{S},\ n \in \mathbb{N}$ sind die Ereignisse

$$\underbrace{\{X_n \in B_n \ {}^{\text{unendlich}}_{\text{oft}}\}}_{\overline{\lim}_{n\to\infty} \{X_n \in B_n\}} \in \mathfrak{S}_\infty$$

$$\underbrace{\{X_n \in B_n \ {}^{\text{schließlich}}_{\text{immer}}\}}_{\underline{\lim}_{n\to\infty} \{X_n \in B_n\}} \in \mathfrak{S}_\infty$$

   terminal.

2. Ist (S, d) ein separabler, metrischer Raum mit $\mathcal{S} = \mathscr{B}(S)$, so ist

$$\{(X_n) \text{ Cauchyfolge}\}$$

   ein terminales Ereignis.

3. Ist (S, d) ein vollständiger, separabler metrischer Raum und $\mathcal{S} = \mathscr{B}(S)$, so ist

$$\left\{\lim_{n\to\infty} X_n \text{ existiert in } S\right\}$$

   ein terminales Ereignis.

4. Ist S ein separabler Banachraum und $\mathcal{S} := \mathscr{B}(S)$, so ist

$$\left\{\sum_{k=1}^{\infty} X_n \text{ existiert in } S\right\}$$

   ein terminales Ereignis.

5. Ist S ein separabler, reeller Banachraum, $\mathcal{S} = \mathscr{B}(S)$ und $(C_n) \subseteq \mathbb{R},\ c_n \overset{n\to\infty}{\longrightarrow} 0$ so ist

$$A := \left\{\lim_{n\to\infty} C_n \cdot \sum_{k=1}^{n} X_k \text{ existiert in } S\right\}$$

   terminal.

6. Speziell für $X_n : \Omega \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ sind

$$\overline{\lim_{n\to\infty}}\ X_n : \Omega \to \mathbb{R}^* \quad , \quad \underline{\lim_{n\to\infty}}\ X_n : \Omega \to \mathbb{R}^*$$

   sogar $\mathfrak{S}_\infty$-messbar.

Beweis:

1. Folgt direkt aus Satz (3.3.5).
2. Da (S, d) separabel ist, ist $d : S \times S \to \mathbb{R}$ messbar bzgl. $\mathcal{S} \otimes \mathcal{S}$ und somit $\sigma(d(X_n, X_m) \subseteq \sigma(X_n, X_m)$. Demnach können wir schreiben
$$\{(X_n) \text{ Cauchyfolge}\} = \bigcap_{k\in\mathbb{N}} \bigcup_{m\in\mathbb{N}} \underbrace{\bigcap_{n=m}^{\infty} \left\{d(X_n, X_m) \leq \frac{1}{k}\right\}}_{\text{wachsend in } m}$$
$$\overset{\substack{l\in\mathbb{N} \\ \text{beliebig}}}{=} \bigcap_{k\in\mathbb{N}} \bigcup_{m=l} \underbrace{\bigcap_{n=m}^{\infty} \left\{d(X_n, X_m) \leq \frac{1}{k}\right\}}_{\in\sigma(X_n:n\geq l)} \in \sigma(X_n : n \geq l) \quad \forall\, l \in \mathbb{N}$$
woraus folgt
$$\{(X_n) \text{ Cauchyfolge}\} \in \bigcap_{l\in\mathbb{N}} \sigma(X_n : n \geq l) = \mathfrak{S}_\infty$$
3. Da S vollständig ist, sind die konvergenten Folgen genau die Cauchy-Folgen. Die Behauptung folgt nun aus Teil (2).
4. Setzt man
$$S_n := \sum_{k=1}^{n} X_k \quad , \quad n \in \mathbb{N}$$
so ist (S_n) genau dann konvergent wenn sie Cauchy ist, sprich
$$\{(S_n) \text{ konvergent in } S\} = \bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \underbrace{\bigcap_{n=m}^{\infty} \left\{\|S_n - S_m\| \leq \frac{1}{k}\right\}}_{\text{wachsend in } m}$$
$$\overset{\substack{l \\ \text{beliebig}}}{=} \bigcap_{k=1}^{\infty} \bigcup_{m=l}^{\infty} \underbrace{\bigcap_{n=m}^{\infty} \left\{\|X_{m+1} + \cdots + X_n\| \leq \frac{1}{k}\right\}}_{\in\sigma(X_n:n\geq l)} \in \sigma(X_n : n \geq l) \quad \forall\, l \in \mathbb{N}$$
woraus folgt
$$\{(S_n) \text{ konvergent in } S\} \in \bigcap_{l\in\mathbb{N}} \sigma(X_n : n \geq l) = \mathfrak{S}_\infty$$
5. Setzt man
$$S_n := \sum_{k=1}^{n} \alpha_k \cdot X_k$$
so ist $C_n S_n$ genau dann konvergent wenn $C_n \cdot (S_n - S_m)$ konvergent ist $\forall\, m \in \mathbb{N}$. Daher
$$\{C_n S_n \text{ konvergent}\} = \left\{\lim_{n\to\infty} C_n \cdot \sum_{k=m}^{n} X_k \text{ existiert in } S\right\} \in \sigma(X_n : n \geq m) \;\; \forall\, m \in \mathbb{N}$$
bzw.
$$\{C_n S_n \text{ konvergent}\} \in \bigcap_{m=1}^{\infty} \sigma(X_n : n \geq m) = \mathfrak{S}_\infty$$
6. Für jedes $c \in \mathbb{R}$ gilt
$$\left\{\overline{\lim_{n\to\infty}}\, X_n \geq c\right\} = \bigcap_{k=1}^{\infty} \underbrace{\left\{X_n \geq c - \tfrac{1}{k} \text{ unendlich oft}\right\}}_{\in\mathfrak{S}_\infty} \in \mathfrak{S}_\infty$$
Analog auch für $\underline{\lim}_{n\to\infty} X_n$.

□

3.3.5 Satz über obere und untere Mengengrenzwerte

Sei $(\Omega, \mathfrak{S})$ ein messbarer Raum, $\mathfrak{S}_1, \mathfrak{S}_2, ... \in \mathfrak{S}$ Teil-σ-Algebren und $A_n \in \mathfrak{S}_n,\ n \in \mathbb{N}$ beliebige Ereignisse. Dann gilt

$$\overline{\lim_{n\to\infty}}\ A_n \in \mathfrak{S}_\infty \quad , \quad \underline{\lim_{n\to\infty}}\ A_n \in \mathfrak{S}_\infty$$

wobei $\mathfrak{S}_\infty$ die σ-Algebra der terminalen Ereignisse sei (vgl. 3.3.1).

Beweis: Wegen

$$\underline{\lim_{n\to\infty}}\ A_n = \left(\overline{\lim_{n\to\infty}}\ A_n^c\right)^c$$

genügt es $\overline{\lim\limits_{n\to\infty}}\ A_n \in \mathfrak{S}_\infty$ zu zeigen. Tatsächlich gilt

$$\overline{\lim_{n\to\infty}}\ A_n = \bigcap_{k=1}^{\infty} \underbrace{\bigcup_{n=k}^{\infty} A_n}_{\text{fallend in } k} = \bigcap_{k=m}^{\infty} \underbrace{\bigcup_{n=k}^{\infty} A_n}_{\in\sigma\left(\bigcup_{n=m}^{\infty}\mathfrak{S}_n\right)} \in \sigma\left(\bigcup_{n=m}^{\infty} \mathfrak{S}_m\right) \quad \forall\ m$$

$$\Rightarrow \overline{\lim_{n\to\infty}}\ A_n \in \mathfrak{S}_\infty$$

□

3.3.6 Theorem: Null-Eins-Gesetz von Kolmogorov

Seien $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $\mathfrak{S}_1, \mathfrak{S}_2, .. \subseteq \mathfrak{S}$ unabhängige Teil-σ-Algebren. Dann gilt für jedes terminale Ereignis $A \in \mathfrak{S}_\infty$ entweder $\mathcal{P}(A) = 0$ oder $\mathcal{P}(A) = 1$.

Beweis:

- Nach Voraussetzung sind für jedes $m \in \mathbb{N}$ die σ-Algebren

  $$\sigma\left(\bigcup_{n=1}^{m} \mathfrak{S}_n\right) \quad , \quad \sigma\left(\bigcup_{n=m+1}^{\infty} \mathfrak{S}_n\right)$$

  unabhängig (vgl. Blockbildung 2.2.8). Demnach ist auch

  $$\mathfrak{S}_\infty \subseteq \sigma\left(\bigcup_{n=m+1}^{\infty} \mathfrak{S}_n\right)$$

  unabhängig von

  $$\sigma\left(\bigcup_{n=1}^{m} \mathfrak{S}_n\right)$$

  (für jedes $m \in \mathbb{N}$) bzw. von

  $$\mathcal{E}_0 := \bigcup_{m=1}^{\infty} \sigma\left(\bigcup_{n=1}^{m} \mathfrak{S}_n\right)$$

  Dabei ist $\mathcal{E}_0$ eine Algebra[5], da Vereinigung aufsteigender σ-Algebren, also insbesondere $\cap$-stabil, so dass nach Theorem 2.2.7 sogar $\sigma(\mathcal{E}_0)$ unabhängig von $\mathfrak{S}_\infty$ ist.

- Wegen $\mathfrak{S}_n \subseteq \mathcal{E}_0\ \forall\ n$ gilt

  $$\bigcup_{n=1}^{\infty} \mathfrak{S}_n \subseteq \mathcal{E}_0$$

[5] $\mathcal{E}_0$ lässt sich als die Menge aller Ereignisse, die mit endlich vielen Ergebnissen entscheidbar sind, interpretieren.

Daraus folgt

$$\sigma\left(\bigcup_{n=1}^{\infty} \mathfrak{S}_n\right) \subseteq \sigma(\mathcal{E}_0)$$

bzw.

$$\mathfrak{S}_\infty \subseteq \sigma\left(\bigcup_{n=1}^{\infty} \mathfrak{S}_n\right) \subseteq \sigma(\mathcal{E}_0)$$

das heißt $\mathfrak{S}_\infty$ ist unabhängig von $\mathfrak{S}_\infty$. Insbesondere ist jede Menge $A \in \mathfrak{S}_\infty$ unabhängig von sich selbst, sprich $\mathcal{P}(A) = \mathcal{P}(A)^2$ woraus die Behauptung folgt.

□

3.3.7 Null-Eins-Satz von Borel

Seien $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $A_1, A_2, ..\mathfrak{S}$ unabhängige Mengen. Dann gilt

$$\mathcal{P}\left(\overline{\lim_{n\to\infty}}\ A_n\right) \in \{0,1\}$$

$$\mathcal{P}\left(\underline{\lim_{n\to\infty}}\ A_n\right) \in \{0,1\}$$

Beweis: Nach Theorem 2.2.7 sind die σ-Algebren $\mathfrak{S}_n := \{\emptyset, A_n, A_n^c, \Omega\} = \sigma(\{A_n\}),\ n \in \mathbb{N}$ unabhängig. Nach 3.3.5 sind

$$\overline{\lim_{n\to\infty}}\ A_n \in \mathfrak{S}_\infty \quad , \quad \underline{\lim_{n\to\infty}}\ A_n \in \mathfrak{S}_\infty$$

Nach Kolmogorov 3.3.6 folgt schließlich die Behauptung.
□

Folgerung: Sind $X_n : (\Omega, \mathfrak{S}, \mathcal{P}) \to (S_n, \mathcal{S}_n),\ n \in \mathbb{N}$ unabhängig und $B_n \in \mathcal{S}_n$, so gilt

$$\mathcal{P}\left(X_n \in B_n \ ^{\text{unendlich}}_{\text{oft}}\right) \in \{0,1\}$$

$$\mathcal{P}\left(X_n \in B_n \ ^{\text{schließlich}}_{\text{immer}}\right) \in \{0,1\}$$

Erläuterung: Setze $A_n := X_n^{-1}(B_n)$ so dass

$$\{X_n \in B_n \ ^{\text{unendlich}}_{\text{oft}}\} = \overline{\lim_{n\to\infty}}\ A_n$$

$$\{X_n \in B_n \ ^{\text{schließlich}}_{\text{immer}}\} = \underline{\lim_{n\to\infty}}\ A_n$$

und verwende vorigen Satz 3.3.7.

Beispiel: Betrachtet sei ein unendlich oft wiederholtes Würfelexperiment, modelliert durch die unabhängigen Zufallsvariablen $X_n,\ n \in \mathbb{N}$:

$$\mathcal{P}(X_n = k) = \frac{1}{6} \quad , \quad k \in \{1, .., 6\}$$

Dann gilt

$$\mathcal{P}\left(X_n = 6 \ ^{\text{unendlich}}_{\text{oft}}\right) \in \{0,1\}$$

$$\mathcal{P}\left(X_n = 6 \ ^{\text{schließlich}}_{\text{immer}}\right) \in \{0,1\}$$

$$\mathcal{P}\left(\exists\ n \geq 1 : X_n = X_{n+1} = \cdots = X_{n+m}\right) \in \{0,1\} \qquad (m \in \mathbb{N} \text{ vorgegeben})$$

Es ist leicht zu sehen dass nur die 2. von drei Wahrscheinlichkeiten gleich 0 sein kann. Insbesondere gilt

$$\mathcal{P}\big(\underbrace{\{\forall\ m \in \mathbb{N} : \exists\ n \in \mathbb{N} : X_n = \cdots = X_{n+m}\}}_{\bigcap_{m\in\mathbb{N}}\{\exists\ n\in\mathbb{N}\ :\ X_n=..=X_n+m\}}\big) = 1$$

3.3.8 Satz über terminale Zufallsvariablen

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $\mathfrak{S}_1, \mathfrak{S}_2, .. \subseteq \mathfrak{S}$ unabhängige teil-σ-Algebren. Sei ferner (M, d) ein separabler, metrischer Raum und $f : (\Omega, \mathfrak{S}) \to (M, \mathscr{B}(M))$ messbar bzgl. der σ-Algebra der terminalen Ereignisse $\mathfrak{S}_\infty$. Dann ist f fast-sicher konstant.

Beweis: Sei $X \subseteq M$ abzählbar und dicht in M. Wegen

$$M = \bigcup_{x\in X} B^o_\varepsilon(x) \ \ \forall\, \varepsilon > 0$$

und

$$\mathcal{P}_f(B^o_\varepsilon(x)) = \mathcal{P}\big(\underbrace{f^{-1}(B^o_\varepsilon(x))}_{\in\mathfrak{S}_\infty}\big) \overset{(3.3.6)}{\in} \{0,1\} \quad \forall\, x \in M,\ \varepsilon > 0$$

existiert für jedes $n \in \mathbb{N}$ ein $x_n \in X$ mit $\mathcal{P}\left(f^{-1}(B^o_{1/n}(x_n))\right) = 1$. Offensichtlich ist

$$C_n := \bigcap_{k=1}^{n} B^o_{1/k}(x_k) \quad , \quad n \in \mathbb{N}$$

eine (in n) fallende Mengenfolge mit $\mathcal{P}_f(C_n) = 1$. Daher

$$\mathcal{P}_f\left(\bigcap_{n=1}^{\infty} C_n\right) = \lim_{n\to\infty} \mathcal{P}_f(C_n) = 1$$

Doch da die C_n jeweils einen Durchmesser $\leq \frac{2}{n}$ haben, kann $C := \bigcap_{n\in\mathbb{N}} C_n$ höchstens ein Element $x \in C$ enthalten (muss sogar, da $\mathcal{P}_f(C) = 1$), daher:

$$\mathcal{P}_f(\{x\}) = \mathcal{P}\left(\{f = x\}\right) = 1$$

sprich f ist fast-sicher konstant.
□

3.3.9 Korollar über Folgen unabhängiger Zufallsgrößen

Seien $X_1, X_2, .. : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}$ unabhängige, reelle Zufallsgrößen. Dann gilt:

1. Die Grenzwerte
$$\overline{\lim_{n\to\infty}}\ X_n : \Omega \to \mathbb{R}^* \quad , \quad \underline{\lim_{n\to\infty}}\ X_n : \Omega \to \mathbb{R}^*$$
sind fast-sicher konstant.
2. Der Grenzwert
$$\lim_{n\to\infty} X_n$$
existiert in $\mathbb{R}$ fast-immer oder fast-nie. Im ersteren Fall ist er fast-sicher konstant.
3. Seien $C_n \to 0$ und
$$S_n := \sum_{k=1}^{n} X_k$$
Dann existiert $\lim\limits_{n\to\infty} C_n S_n$ fast-immer oder fast-nie. Im ersteren Fall ist er sogar fast-sicher konstant.

Beweis: Seien $\mathfrak{S}_n := \sigma(X_n)$, dazu die σ-Algebra $\mathfrak{S}_\infty$ der terminalen Ereignisse.

1. Die σ-Algebren $\mathfrak{S}_n$ sind unabhängig und $\overline{\lim\limits_{n\to\infty}} X_n$ & $\underline{\lim\limits_{n\to\infty}} X_n$ nach Satz 3.3.4 (6) $\mathfrak{S}_\infty$-messbar. Nach vorigem Satz 3.3.8 sind sie fast-sicher konstant.
2. Nach Satz 3.3.4(3) ist
$$\left\{\exists \lim_{n\to\infty} X_n \in \mathbb{R}\right\}$$
terminal und besitzt nach Kolmogorov 3.3.6 das Maß 0 oder 1. Im zweiten Fall ist insbesondere fast sicher
$$\lim_{n\to\infty} X_n = \overline{\lim_{n\to\infty}} X_n$$
Da $\overline{\lim\limits_{n\to\infty}} X_n$ nach 3.3.4(6) messbar bzgl. $\mathfrak{S}_\infty$ ist, folgt aus vorigem Satz 3.3.8 dass $\lim\limits_{n\to\infty} X_n$ fast-sicher konstant ist.
3. Nach Satz 3.3.4 (5) ist
$$\left\{\lim_{n\to\infty} C_n \cdot S_n \text{ existiert in } \mathbb{R}\right\}$$
terminal und besitzt nach Kolmogorov 3.3.6 das Maß 0 oder 1. Im zweiten Fall ist
$$S := \begin{cases} \lim\limits_{n\to\infty} C_n S_n & : \text{ falls existent} \\ 0 & : \text{sonst} \end{cases}$$
messbar bzgl. $\mathfrak{S}_\infty$, denn für jedes $t \in \mathbb{R}$ ist $\lim\limits_{n\to\infty} C_n S_n \leq t$ genau dann wenn $\lim\limits_{n\to\infty} C_n(S_n - S_m) \leq t \ \ \forall\, m \in \mathbb{N}$, sprich
$$\left\{\exists \lim_{n\to\infty} C_n S_n \leq t\right\} = \left\{\lim_{n\to\infty} C_n(S_n - S_m) \leq t\right\} = \bigcap_{k\in\mathbb{N}} \bigcup_{n_0\in\mathbb{N}} \bigcap_{n=n_0}^{\infty} \underbrace{\left\{C_n(S_n - S_m) \leq t + \frac{1}{k}\right\}}_{\in\sigma(X_n : n\geq m)} \in \sigma(X_n : n \geq m)$$
$$\Rightarrow \left\{\lim_{n\to\infty} C_n S_n \leq t\right\} \in \bigcap_{m=1}^{\infty} \sigma(X_n : n \geq m) = \mathfrak{S}_\infty$$
Nach vorigem Satz 3.3.8 ist dann schließlich S fast-sicher konstant.

$\square$

Beispiel: Für die auf $\{\pm 1\}$ gleichverteilten, unabhängigen Zufallsvariablen $\varepsilon_n : (\Omega, \mathfrak{S}, \mathcal{P}) \to \{0,1\}$, $n \in \mathbb{N}$ und konstanten $\alpha_n \in \mathbb{R}$ existiert
$$\sum_{k=1}^{\infty} \alpha_k \varepsilon_k$$
fast-sicher oder divergiert fast-sicher (abhängig von den α_k!). Setzt man z.B. $\alpha_k := \frac{1}{k}$, so existieren zwar sowohl divergente Ereignisse, etwa
$$\varepsilon_k(\omega) = 1 \ \ \forall\, k \quad \Rightarrow \quad \sum_{k=1}^{\infty} \alpha_k \varepsilon_k(\omega) = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots = \infty$$
als auch konvergente Ereignisse wie etwa
$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$
doch ergibt sich tatsächlich
$$\mathcal{P}\left(\left\{\sum_{k=1}^{\infty} \frac{\varepsilon_k}{k} \text{ konvergent}\right\}\right) = 1$$
Es kann sogar gezeigt werden, dass
$$\mathcal{P}\left(\left\{\sum_{k=1}^{\infty} \frac{\varepsilon_k}{k^\alpha} \text{ konvergent}\right\}\right) = 1$$
für $\frac{1}{2} < \alpha < \infty$.

3.4 Lemma von Borel und Cantelli

Gegeben ein Wahrscheinlichkeitsraum $(\Omega, \mathfrak{S}, \mathcal{P})$ und unabhängige $A_1, A_2, .. \in \mathfrak{S}$ mit

$$\mathcal{P}\left(\overline{\lim_{n\to\infty}}\ A_n\right) \in \{0,1\}$$

(vgl. Null-Eins-Satz von Borel 3.3.7). Gesucht sind Kriterien dafür, zwischen den Fällen 0 und 1 zu unterscheiden.

3.4.1 Lemma von Borel und Cantelli

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $A_1, A_2, ... \in \mathfrak{S}$ beliebige Ereignisse. Dann gilt:

1. Ist
$$\sum_{n=1}^{\infty} \mathcal{P}(A_n) < \infty$$
so folgt
$$\mathcal{P}\left(\overline{\lim_{n\to\infty}}\ A_n\right) = 0$$

2. Sind die A_n paarweise unabhängig mit
$$\sum_{n=1}^{\infty} \mathcal{P}(A_n) = \infty$$
so folgt
$$\mathcal{P}\left(\overline{\lim_{n\to\infty}}\ A_n\right) = 1$$

Bemerkung: Die paarweise Unabhängigkeit in (2) ist tatsächlich notwendig! Als Beispiel sei betrachtet der Fall $A_1 = A_2 = \cdots = A$ mit $0 < \mathcal{P}(A) < 1$. Dann ist zwar

$$\sum_{n=1}^{\infty} \mathcal{P}(A_n) = \infty$$

doch trotzdem

$$\mathcal{P}\left(\overline{\lim_{n\to\infty}}\ A_n\right) = \mathcal{P}(A) < 1$$

Beweis des Lemmas:

1. Nach Voraussetzung gilt
$$\mathcal{P}\left(\overline{\lim_{n\to\infty}}\ A_n\right) = \mathcal{P}\left(\bigcap_{m=1}^{\infty} \underbrace{\bigcup_{n=m}^{\infty} A_n}_{\text{fallend in } m}\right) = \lim_{m\to\infty} \mathcal{P}\left(\bigcup_{n=m}^{\infty} A_n\right) \le \lim_{m\to\infty} \sum_{n=m}^{\infty} \mathcal{P}(A_n) = 0$$

2. • Es gelte
$$\sum_{n=1}^{\infty} \mathcal{P}(A_n) = \infty \qquad \text{und} \qquad \mathcal{P}(A_i \cap A_j) = \mathcal{P}(A_i) \cdot \mathcal{P}(A_j)\ \forall\ i \neq j$$
Zu zeigen wäre $\mathcal{P}(A) = 1$ für:
$$A := \overline{\lim_{n\to\infty}}\ A_n$$
Setze $X_n := 1_{A_n}$, dann gilt
$$A = \left\{\sum_{n=1}^{\infty} X_n = \infty\right\}$$
Setzt man
$$S_n := \sum_{k=1}^{n} X_k \quad , \quad S := \lim_{n\to\infty} S_n$$
so wäre zu zeigen: $\mathcal{P}(\{S = \infty\}) = 1$.

- Dabei sind die $X_1, X_2, ..$ nach Voraussetzung paarweise unabhängig, so dass gilt

$$\mathbb{V}S_n = \mathbb{E}\left[\sum_{k=1}^{n}(X_k - \mathbb{E}X_k)\right]^2 = \sum_{k,l=1}^{n} \underbrace{\mathbb{E}\left[(X_k - \mathbb{E}X_k)(X_l - \mathbb{E}X_l)\right]}_{0 \text{ falls } k \neq l} = \sum_{k=1}^{n}\mathbb{E}(X_k - \mathbb{E}X_k)^2 = \sum_{k=1}^{n}\mathbb{V}X_k$$

$$= \sum_{k=1}^{n}\left[\mathbb{E}X_k^2 - (\mathbb{E}X_k)^2\right] = \mathbb{E}S_n - \sum_{k=1}^{n}(\mathbb{E}X_k)^2 \leq \mathbb{E}S_n$$

Dabei gehen

$$\mathbb{E}S_n = \sum_{k=1}^{\infty}\mathbb{E}X_k = \sum_{k=1}^{n}\mathcal{P}(A_k) \overset{n\to\infty}{\longrightarrow} \infty \tag{$\spadesuit$}$$

- Für beliebiges $c > 0$ gilt nach Tschebyschow

$$\mathcal{P}\left(|S_n - \mathbb{E}S_n| \leq c\right) = 1 - \mathcal{P}\left(|S_n - \mathbb{E}S_n| > c\right) \overset{(ii)}{\geq} 1 - \frac{\mathbb{V}S_n}{c^2} \tag{$\clubsuit$}$$

Speziell folgt für $c := \frac{\mathbb{E}S_n}{2}$

$$\mathcal{P}\left(S_n \geq \frac{\mathbb{E}S_n}{2}\right) \geq \mathcal{P}\left(|S_n - \mathbb{E}S_n| \leq \frac{\mathbb{E}S_n}{2}\right) \overset{(\clubsuit)}{\geq} 1 - \frac{4\mathbb{V}S_n}{(\mathbb{E}S_n)^2} \geq 1 - \frac{4}{\mathbb{E}S_n} \underset{n\to\infty}{\overset{(\spadesuit)}{\longrightarrow}} 1$$

bzw.

$$\mathcal{P}(S = \infty) = \lim_{n\to\infty}\mathcal{P}\left(S \geq \frac{\mathbb{E}S_n}{2}\right) \geq \lim_{n\to\infty}\mathcal{P}\left(S_n \geq \frac{\mathbb{E}S_n}{2}\right) = 1$$

□

3.4.2 Korollar für unabhängige Mengen

Ist $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $A_1, A_2, .. \in \mathfrak{S}$ unabhängig, dann gilt:

$$\mathcal{P}\left(\overline{\lim_{n\to\infty}}\ A_n\right) = 0 \ \Leftrightarrow\ \sum_{n=1}^{\infty}\mathcal{P}(A_n) < \infty$$

$$\mathcal{P}\left(\overline{\lim_{n\to\infty}}\ A_n\right) = 1 \ \Leftrightarrow\ \sum_{n=1}^{\infty}\mathcal{P}(A_n) = \infty$$

Beweis: Folgt direkt nach Borel & Cantelli 3.4.1.

Beispiele:

1. Modelliert sei ein unendlich oft, wöchentlich durchgeführtes *Lotto-Spiel*, dazu das Ereignis $A_n,\ n \in \mathbb{N}$ in der n-ten Woche alle 6 Zahlen zu treffen. Mit $\mathcal{P}(A_n) > 0$ bzw.

$$\sum_{n=1}^{\infty}\mathcal{P}(A_n) = \infty$$

und der Annahme der Unabhängigkeit der $A_1, A_2, ..$ folgt

$$\mathcal{P}\left(A_n \text{ u.o.}\right) = 1$$

das heißt, ein sicherer Gewinn wird fast-sicher unendlich oft stattfinden!

2. Seien $U_1, U_2, \ldots$ unabhängig, gleichverteilt auf $[0,1]$ und $a_n \underset{n\to\infty}{\searrow} 0$ eine beliebige Nullfolge. Dann gilt die Äquivalenzkette:

$$\mathcal{P}\left(U_n \geq 1 - a_n \text{ u.o.}\right) = 1 \quad \Leftrightarrow \sum_{n=1}^{\infty} \mathcal{P}(U_n \geq 1 - a_n) = \infty$$

$$\Leftrightarrow \sum_{n=1}^{\infty} a_n = \infty$$

3. Betrachtet sei die Folge unabhängiger, auf $\{\pm 1\}$ gleichverteilter Zufallsvariablen ε_j (Bernoulli-Folge), dazu $N \in \mathbb{N}$ fest gewählt. Setzt man

$$\begin{aligned} Y_1 &:= \varepsilon_1 + \cdots + \varepsilon_N \\ Y_2 &:= \varepsilon_{N+1} + \cdots + \varepsilon_{2N} \\ &\vdots \end{aligned}$$

so sind $Y_1, Y_2, ..$ unabhängig. Wegen

$$\mathcal{P}(Y_k = N) = \frac{1}{2^N} > 0$$

bzw.

$$\sum_{k=1}^{\infty} \mathcal{P}(Y_k = N) = \infty$$

gilt

$$\mathcal{P}\Big(\underbrace{\{Y_k = N \text{ u.o.}\}}_{=:A_N}\Big) = 1$$

das heißt in der Bernoulli-Folge $\varepsilon_1, \varepsilon_2, ..$ tauchen fast-sicher unendlich viele '+1'-Blöcke der Länge N auf. Ferner gilt sogar

$$\mathcal{P}\Big(\bigcap_{N\in\mathbb{N}} A_N\Big) = 1$$

sprich, fast-sicher sind in $\varepsilon_1, \varepsilon_2, ..$ für jedes $N \in \mathbb{N}$ unendlich viele '+1'-Blöcke der Länge N zu finden!

3.4.3 Beispiel: Tendenzabschätzung normalverteilter Zufallsvariablen

Die Zufallsvariablen $X_k : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R},\ k \in \mathbb{R}$ seien unabhängig, $\sim \mathcal{N}_{0,1}$ standardnormalverteilt, sprich mit Verteilungsfunktion

$$\mathcal{P}(X_k \leq t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2} dx \quad , \quad t \in \mathbb{R}$$

Nach Korollar 3.4.2 gilt die Äquivalenzkette

$$\mathcal{P}\left(X_n \geq c \cdot \sqrt{\ln n} \text{ u.o.}\right) = 1 \Leftrightarrow \sum_{n=1}^{\infty} \mathcal{P}\left(X_n \geq c \cdot \sqrt{\ln n}\right) = \infty$$

$$\Leftrightarrow \sum_{n=1}^{\infty} \int_{c\sqrt{\ln n}}^{\infty} e^{-x^2/2} dx = \infty$$

$$\overset{\substack{\text{analog}\\ \text{zu (2.7.3)}}}{\Leftrightarrow} \sum_{n=2}^{\infty} \frac{e^{-(\sqrt{\ln n})^2 c^2/2}}{c\sqrt{\ln n}} = \infty$$

$$\Leftrightarrow \sum_{n=2}^{\infty} \frac{n^{-c^2/2}}{\sqrt{\ln n}} = \infty$$

$$\Leftrightarrow \quad c^2/2 \leq 1 \quad \Leftrightarrow \quad c \leq \sqrt{2}$$

sprich

$$\mathcal{P}\left(\frac{X_n}{\sqrt{\ln n}} \geq c \text{ u.o.}\right) = \begin{cases} 0 & : c > \sqrt{2} \\ \\ 1 & : c \leq \sqrt{2} \end{cases}$$

Insbesondere ist

$$\varlimsup_{n\to\infty} \frac{X_n}{\sqrt{\ln n}} = \sqrt{2}$$

fast-sicher. Aufgrund der Symmetrie der X_n ist auch

$$\varliminf_{n\to\infty} \frac{X_n}{\sqrt{\ln n}} = -\sqrt{2}$$

fast-sicher.

3.5 Summen unabhängiger Zufallsvariablen

3.5.1 Definition: Median

Ein Zahl $a \in \mathbb{R}$ heißt *Median* einer reellen Zufallsgröße $X : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}$, wenn gilt

$$\mathcal{P}(X \geq a) \geq \frac{1}{2} \quad \wedge \quad \mathcal{P}(X \leq a) \geq \frac{1}{2}$$

Bemerkungen:

(i) Schreiben $M(X)$ für den Median von X.

(ii) Der Median einer Zufallsgröße ist nicht unbedingt eindeutig bestimmt. In diesem Fall, repräsentiere $M(X)$ *irgendeinen* Median von X.

(iii) t_0 ist genau dann Median, falls für die Verteilungsfunktion F von X gilt:

$$F(t_0) \geq \frac{1}{2} \quad \wedge \quad F(t_0^-) \leq \frac{1}{2}$$

(iv) Jede reelle Zufallsgröße besitzt einen Median, z.B.

$$M(X) = \inf\left\{t \in \mathbb{R} : \mathcal{P}(X \leq t) \geq \tfrac{1}{2}\right\}$$

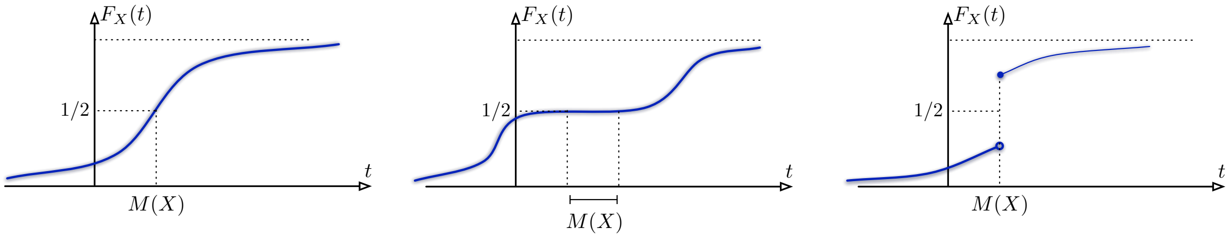

Abbildung 8: Zur Definition des Medians. Bestimmung des Medians anhand des Verteilungsfunktion F_X von X.

3.5.2 Eigenschaften des Medians

Seien X, Y reelle Zufallsvariablen.

1. Sind $X \stackrel{\text{d.}}{=} Y$, so ist jeder Median von X auch Median von Y.
2. Ist $M(X)$ Median von X und $c \in \mathbb{R}$, so ist $c \cdot M(X)$ Median von $c \cdot X$.
3. Ist $X \stackrel{\text{d.}}{=} -X$ symmetrisch, so ist 0 ein Median von X.
4. Gilt $\mathcal{P}(|X| \geq c) < \frac{1}{2}$ für irgendein $c > 0$, so muss $|M(X)| < c$ sein.

Beweis: (Aussage (4)): Aus

$$\frac{1}{2} > \mathcal{P}\left(|X| \geq c\right) = \mathcal{P}\left(X \geq c \ \vee \ X \leq -c\right) = \mathcal{P}(X \geq c) + \mathcal{P}(X \leq -c)$$

folgt

$$\mathcal{P}(X \geq c) < \frac{1}{2} \quad \Rightarrow \quad M(X) < c$$

und

$$\mathcal{P}(X \leq -c) < \frac{1}{2} \quad \Rightarrow \quad M(X) > -c$$

also insgesamt

$$|M(X)| \leq c$$

□

3.5.3 Levis Maximalungleichung

Seien $X_1, .., X_n$ unabhängige, reelle Zufallsgrößen und $\varepsilon > 0$, dazu

$$S_j := X_1 + \cdots + X_j \quad , \quad 1 \leq j \leq n$$

Dann gilt:

1. $\mathcal{P}\left(\max\limits_{1 \leq j \leq n} [S_j + M(S_n - S_j)] \geq \varepsilon\right) \leq 2 \cdot \mathcal{P}(S_n \geq \varepsilon)$
2. $\mathcal{P}\left(\max\limits_{1 \leq j \leq n} |S_j + M(S_n - S_j)| \geq \varepsilon\right) \leq 2 \cdot \mathcal{P}(|S_n| \geq \varepsilon)$

Interpretation: Sind die $X_1, .., X_n$ symmetrisch, so sind auch alle $(S_n - S_j)$ symmetrisch, sprich $M(S_n - S_j) = 0$. Nach (1) gilt dann

$$\mathcal{P}\left(\max_{1\leq j\leq n} S_j \geq \varepsilon\right) \leq 2\cdot\mathcal{P}\left(S_n \geq \varepsilon\right)$$

sprich, die Wahrscheinlichkeit, dass die Summen S_j irgendwann die Grenze ε überschreiten, verhält sich wie die Wahrscheinlichkeit (bis auf Vorfaktor 2) dass diese letztendlich durch S_n überschritten wird.

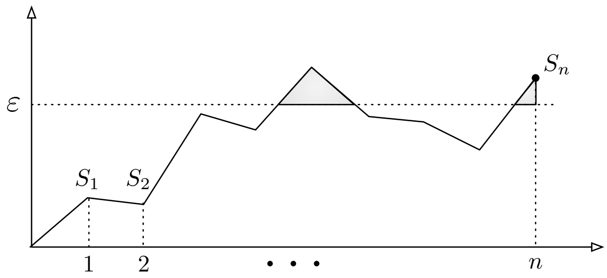

Abbildung 9: Zu Levis Maximalsatz: Typischer Verlauf der S_j für symmetrische X_j. Die Wahrscheinlichkeit, dass die S_j irgendwann ε überschreiten, verhält sich wie die Wahrscheinlichkeit dass ε letztendlich durch S_n überschritten wird.

Beweis:

1. - Setzen $M_j := M(S_j - S_n)$ und $B_j := \{S_j - S_n \leq M_j\}$. Definieren ferner die *Stoppzeit*

$$T(\omega) := \begin{cases} n+1 & : S_j(\omega) - M_j < \varepsilon \ \forall\ 1\leq j\leq n \\ \inf\{1\leq j\leq n : S_j(\omega) - M_j \geq \varepsilon\} & : \text{sonst} \end{cases}$$

   so dass gilt

$$\left\{\max_{1\leq j\leq n}\left[S_j - M_j\right] \geq \varepsilon\right\} = \{T \leq n\}$$

   Zu zeigen wäre also $\mathcal{P}(T\leq n) \leq 2\mathcal{P}(S_n \geq \varepsilon)$.

   - Für $\omega \in B_j \cap \{T = j\}$, sprich

$$S_j(\omega) - S_n(\omega) \leq M_j \quad \wedge \quad S_j(\omega) - M_j \geq \varepsilon$$

   gilt

$$S_n(\omega) \geq \varepsilon$$

   das heißt

$$B_j \cap \{T = j\} \subseteq \{S_n \geq \varepsilon\}$$

   bzw.

$$\bigcup_{j=1}^{n}\left(B_j \cap \{T = j\}\right) \subseteq \{S_n \geq \varepsilon\}$$

   Da die $\{T = j\}$, $1 \leq j \leq n$ ferner disjunkt sind, gilt

$$\sum_{j=1}^{n}\mathcal{P}\left(B_j \cap \{T = j\}\right) \leq \mathcal{P}\left(S_n \geq \varepsilon\right) \qquad (\clubsuit)$$

   - Wegen $\{T = j\} \in \sigma(X_1, .., X_j)$ und

$$B_j = \{\omega \in \Omega : -X_{j+1} - \cdots - X_n \leq M_j\} \in \sigma(X_{j+1}, .., X_n)$$

sind die B_j und $\{T = j\}$ unabhängig. Aus (♣) folgt daher

$$\sum_{j=1}^{n} \underbrace{\mathcal{P}(B_j)}_{\substack{\geq \frac{1}{2} \\ \text{nach} \\ \text{Voraussetzung}}} \cdot \mathcal{P}(T = j) \leq \mathcal{P}(S_n \geq \varepsilon)$$

bzw.

$$\mathcal{P}(T \leq n) = \sum_{j=1}^{n} \mathcal{P}(T = j) \leq 2 \cdot \mathcal{P}(S_n \geq \varepsilon)$$

wie gewünscht.

2. Durch Anwendung von (1) auf $-X_1, .., -X_n$ erhält man

$$\mathcal{P}\Big(\underbrace{\max_{1 \leq j \leq n} [-S_j - M(-S_j + S_n)]}_{-S_j + M(S_j - S_n)} \geq \varepsilon \Big) \leq 2 \cdot \mathcal{P}\left(-S_n \geq \varepsilon\right)$$

Sei nun

$$\max_{1 \leq j \leq n} |S_j(\omega) - M(S_j - S_n)| \geq \varepsilon$$

dann existiert ein j_1 mit $S_{j_1}(\omega) - M(S_{j_1} - S_n) \geq \varepsilon$ oder ein j_2 mit $S_{j_2}(\omega) - M(S_{j_2} - S_n) \leq -\varepsilon$, also

$$\mathcal{P}\left(\max_{1 \leq j \leq n} |S_j - M(S_j - S_n)| \geq \varepsilon \right) \leq \mathcal{P}\left(\max_{1 \leq j \leq n} [S_j - M(S_j - S_n)] \geq \varepsilon \right) + \mathcal{P}\left(\max_{1 \leq j \leq n} [-S_j + M(S_j - S_n)] \geq \varepsilon \right)$$

$$\stackrel{(1)}{\leq} 2\mathcal{P}(\underbrace{S_n \geq \varepsilon}_{\text{disjunkte..}}) + 2\mathcal{P}(\underbrace{-S_n \geq \varepsilon}_{\text{..Ereignisse}}) = 2\mathcal{P}\left(S_n \geq \varepsilon \ \vee \ -S_n \geq \varepsilon\right)$$

$$= 2 \cdot \mathcal{P}\left(|S_n| \geq \varepsilon\right)$$

□

Beispiel: Die unabhängigen, symmetrischen Zufallsvariablen $X_1, X_2, ..$ repräsentieren den Gewinn bzw. Verlust der 1., 2., ... Runde eines Gewinnspiels. Die Partialsummen

$$S_k := X_1 + \cdots + S_k \quad , \quad k \in \mathbb{N}$$

stellen den netto Gewinn bzw. Verlust nach k Spielrunden und

$$\max_{1 \leq k \leq n} |S_k|$$

den maximal erreichten Gewinn bzw. Verlust im Laufe der ersten n Spielrunden dar. Nach Levi gilt dann

$$\mathcal{P}\left(|S_n| \geq a\right) \leq \underbrace{\mathcal{P}\left(\max_{1 \leq k \leq n} |S_k| \geq a \right)}_{\substack{\text{maximal erreichter} \\ \text{Gewinn oder Verlust}}} \stackrel{(3.5.3)}{\leq} 2 \cdot \mathcal{P}\left(|S_n| \geq a\right)$$

für jegliches $a > 0$.

3.5.4 Folgerungen der Maximalungleichung

Seien $X_1, .., X_n$ unabhängige, reelle Zufallsgrößen, dazu die Partialsummen

$$S_j := X_1 + \cdots + X_j \quad , \quad 1 \leq j \leq n$$

1. Sind die X_j symmetrisch, so ist 0 ein Median von $M(S_j - S_n),\ 1 \leq j \leq n$ und nach Levi 3.5.3 gilt

$$\mathcal{P}\Big(\max_{1\leq j\leq n} |S_j| \geq \varepsilon\Big) \leq 2 \cdot \mathcal{P}\left(|S_n| \geq \varepsilon\right) \quad \forall\, \varepsilon > 0$$

   Ferner folgt

$$\mathbb{E}\Big(\max_{1\leq j\leq n} |S_j|\Big)^p \leq C_p \cdot \mathbb{E}\,|S_n|^p$$

   für geeignete, von n und der Verteilung der X_j unabhängige, Konstanten $C_p \geq 0$ (ohne Beweis).

2. Wendet man für $1 \leq m < n$ die Maximalgleichung auf die $X_{m+1}, .., X_n$ bzw. die entsprechenden Partialsummen

$$\widetilde{S}_j := X_{m+1} + \cdots + X_j \quad , \quad m < j \leq n$$

   an, erhält man

$$\mathcal{P}\Big(\max_{m<j\leq n} |(S_j - S_m) - M(S_j - S_n)| \geq \varepsilon\Big) \leq 2 \cdot \mathcal{P}\Big(|S_n - S_m| \geq \varepsilon\Big) \quad \forall\, \varepsilon > 0$$

3.5.5 Theorem von Paul Lévy zur Konvergenz von Summen

Seien $X_1, X_2, ...$ reelle, unabhängige Zufallsgrößen, dazu die Partialsummen

$$S_n := X_1 + \cdots + X_n$$

Dann gilt für jede reelle Zufallsvariable S:

$$S_n \overset{n\to\infty}{\underset{\text{p}}{\longrightarrow}} S \quad \Leftrightarrow \quad S_n \overset{n\to\infty}{\underset{\text{f.s.}}{\longrightarrow}} S$$

Beweis: Nach Satz 3.2.2 genügt es Richtung "⇒" zu zeigen. Sei also

$$S_n \overset{n\to\infty}{\underset{\text{p}}{\longrightarrow}} S$$

- Seien zunächst $0 < \varepsilon,\ 0 < \delta < \frac{1}{2}$ vorgegeben. Dann gilt

$$\mathcal{P}\left(|S_n - S_j| \geq \varepsilon\right) \leq \mathcal{P}\left(|S_n - S| \geq \tfrac{\varepsilon}{2} \ \vee\ |S_j - S| \geq \tfrac{\varepsilon}{2}\right)$$

$$\leq \mathcal{P}\left(|S_n - S| \geq \tfrac{\varepsilon}{2}\right) + \mathcal{P}\left(|S_j - S| \geq \tfrac{\varepsilon}{2}\right) \overset{n,j\to\infty}{\longrightarrow} 0$$

  sprich für geeignetes $n_{\varepsilon,\delta} \in \mathbb{N}$ gilt

$$\mathcal{P}\left(|S_n - S_j| \geq \varepsilon\right) \leq \delta \quad \forall\, n, j \geq n_{\varepsilon,\delta} \tag{3.5.5.1}$$

  Insbesondere

$$\mathcal{P}\left(|S_n - S_j| \geq \varepsilon\right) < \frac{1}{2}$$

  so dass nach 3.5.2(4) gilt

$$|M(S_j - S_n)| \leq \varepsilon \quad \forall\, n, j \geq n_{\varepsilon,\delta} \tag{3.5.5.2}$$

- Seien nun $n_{\varepsilon,\delta} \leq m < n$. Wegen

$$|M(S_j - S_n)| \leq \varepsilon \quad \forall\, n, j \geq n_{\varepsilon,\delta}$$

  würde

$$|S_j - S_m| \geq 2\varepsilon \tag{3.5.5.3}$$

  (für irgendein $m < j \leq n$) implizieren[6]

$$|(S_j - S_m) - M(S_j - S_n)| \geq \varepsilon$$

[6] Für $a, b \in \mathbb{R}$ mit $|a| \geq 2\varepsilon,\ |b| \leq \varepsilon$ kann man abschätzen $|a - b| \geq |a| - |b| \geq \varepsilon$.

bzw.

$$\max_{m<j\le n} |(S_j - S_m) - M(S_j - S_n)| \geq \varepsilon \tag{3.5.5.4}$$

Daher wäre (3.5.5.5) auch durch

$$\max_{m<j\le n} |S_j - S_m| \geq 2\varepsilon \tag{3.5.5.5}$$

impliziert. Dementsprechend kann man schreiben

$$\mathcal{P}\Big(\max_{m<j\le n} |S_j - S_m| \geq 2\varepsilon\Big) \leq \mathcal{P}\Big(\max_{m<j\le n} |(S_j - S_m) - M(S_j - S_n)| \geq \varepsilon\Big)$$

$$\overset{3.5.4(2)}{\leq} 2 \cdot \underbrace{\mathcal{P}\left(|S_n - S_m| \geq \varepsilon\right)}_{\leq \delta} \leq 2\delta$$

Setzt man nun (für festes m)

$$A_n := \Big\{\max_{m<j\le n} |S_j - S_m| \geq 2\varepsilon\Big\} \quad , \quad n > m$$

so sind die A_n aufsteigend in n und es gilt

$$\bigcup_{n=m+1}^{\infty} A_n \supseteq \Big\{\sup_{j>m} |S_j - S_m| \geq 3\varepsilon\Big\}$$

so dass

$$\mathcal{P}\Big(\sup_{j>m} |S_j - S_m| \geq 3\varepsilon\Big) \leq \lim_{n\to\infty} \underbrace{\mathcal{P}(A_n)}_{\leq 2\delta} \leq 2\delta \quad \forall\, m \geq n_{\varepsilon,\delta} \tag{3.5.5.6}$$

Aus (3.5.5.6) folgt nun schließlich

$$\lim_{m\to\infty} \mathcal{P}\Big(\sup_{j>m} |S_j - S_m| \geq \varepsilon\Big) = 0 \quad \forall\, \varepsilon > 0 \tag{3.5.5.7}$$

- Setzen nun

$$A := \{\omega \in \Omega : \{S_n(\omega)\}_{n=1}^{\infty} \text{ Cauchyfolge}\}$$

und wollen zeigen $\mathcal{P}(A) = 1$ (bzw. $\mathcal{P}(A^c) = 0$). Dabei ist

$$A^c = \bigcup_{k=1}^{\infty} \bigcap_{m=1}^{\infty} \bigcup_{j=m+1}^{\infty} \{|S_j - S_m| \geq \tfrac{1}{k}\}$$

das heißt

$$\mathcal{P}(A^c) = 0 \;\Leftrightarrow\; \mathcal{P}\Big(\bigcap_{m=1}^{\infty} \underbrace{\bigcup_{j=m+1}^{\infty} \{|S_j - S_m| \geq \tfrac{1}{k}\}}_{\text{fallend in } m}\Big) = 0 \quad \forall\, k \in \mathbb{N}$$

$$\Leftrightarrow\; \lim_{m\to\infty} \mathcal{P}\Big(\bigcup_{j=m+1}^{\infty} \{|S_j - S_m| \geq \tfrac{1}{k}\}\Big) = 0 \quad \forall\, k \in \mathbb{N}$$

Doch wegen

$$\bigcup_{j=m+1}^{\infty} \{|S_j - S_m| \geq \tfrac{1}{k}\} \subseteq \Big\{\sup_{j>m} |S_j - S_m| \geq \frac{1}{k}\Big\}$$

folgt aus (3.5.5.7):

$$\mathcal{P}\Big(\bigcup_{j=m+1}^{\infty}\{|S_j-S_m|\geq\tfrac{1}{k}\}\Big)\overset{m\to\infty}{\longrightarrow}0$$

bzw. $\mathcal{P}(A^c)=0$. Der Grenzwert

$$\lim_{n\to\infty}S_n(\omega)=:\widetilde{S}(\omega)$$

existiert also fast-überall.

- Nach 3.2.2(2) konvergiert auch $S_n\underset{\mathrm{p}}{\overset{n\to\infty}{\longrightarrow}}\widetilde{S}$. Nach Eindeutigkeitssatz (3.2.3) gilt schließlich $S=\widetilde{S}$.

□

Interpretation des Beweises: Zentraler Punkt des Beweises war es, aus

$$\sup_{j>m}\mathcal{P}\left(|S_j-S_m|\geq\varepsilon\right)\overset{m\to\infty}{\longrightarrow}0\quad\forall\,\varepsilon>0$$

(Konvergenz in Wahrscheinlichkeit) mit Hilfe der Maximalungleichung

$$\mathcal{P}\left(\sup_{j>m}|S_j-S_m|\geq\varepsilon\right)\overset{m\to\infty}{\longrightarrow}0\quad\forall\,\varepsilon>0$$

(vgl. (3.5.5.7)) zu schlussfolgern.

3.5.6 Korollar über die Konvergenz von Partialsummen

Seien $X_1,X_2,..:(\Omega,\mathfrak{S},\mathcal{P})\to\mathbb{R}$ reelle, unabhängige Zufallsvariablen mit $\mathbb{E}X_j=0$.

1. Gilt
$$\sum_{j=1}^{\infty}\mathbb{V}X_j<\infty \tag{3.5.6.1}$$
so konvergieren die Partialsummen
$$S_n:=\sum_{k=1}^{n}X_k\quad,\quad n\in\mathbb{N}$$
fast-überall.
2. Sind $0<b_1\leq b_2\leq\dots\to\infty$ so das
$$\sum_{k=1}^{\infty}\frac{\mathbb{V}X_k}{b_k^2}<\infty$$
so geht
$$\frac{S_n}{b_n}:=\frac{1}{b_n}\sum_{k=1}^{n}X_k\overset{n\to\infty}{\longrightarrow}0\ ,\quad n\in\mathbb{N}$$
fast-überall.

Beweis:

1. Sei $\varepsilon>0$ beliebig, dann gilt nach Voraussetzung
$$\mathbb{E}\,|S_n-S_m|^2=\mathbb{E}\left|\sum_{j=m+1}^{n}X_j\right|^2=\sum_{i,j=m+1}^{n}\underbrace{\mathbb{E}(X_i\cdot X_j)}_{\substack{0\\ \text{falls }i\neq j}}=\sum_{j=m+1}^{n}\underbrace{\mathbb{E}\,|X_j|^2}_{\mathbb{V}X_j}<\varepsilon\quad\forall\,n>m\geq m_\varepsilon$$
für geeignetes $m_\varepsilon\in\mathbb{N}$, das heißt $(S_n)_{n=1}^{\infty}$ ist Cauchyfolge in $L_2(\Omega,\mathfrak{S},\mathcal{P})$. Aufgrund der Vollständigkeit von L_2 existiert ein $S\in L_2(\Omega,\mathfrak{S},\mathcal{P})$ mit
$$\mathbb{E}\,|S_n-S|^2\overset{n\to\infty}{\longrightarrow}0$$

Nach 3.2.2(3) geht dann auch $S_n \overset{n\to\infty}{\underset{\mathrm{p}}{\longrightarrow}} S$, nach Lévy 3.5.5 sogar

$$S_n \overset{n\to\infty}{\underset{\text{f.s.}}{\longrightarrow}} S$$

2. Nach (1) konvergiert die Reihe
$$\sum_{k=1}^{\infty} \frac{X_k}{b_k}$$
fast-überall. Nach Lemma von Kronecker A.5.2(1) geht dann auch
$$\frac{1}{b_n} \sum_{k=1}^{n} X_k \overset{n\to\infty}{\longrightarrow} 0$$
fast-überall.

□

Beispiel: Betrachtet seien die unabhängigen, auf $\{0,1\}$ gleichverteilten Zufallsvariablen $\varepsilon_1, \varepsilon_2, ..$ (Bernoulli-Folge), dazu die Folge $(c_k)_{k\in\mathbb{N}} \subseteq \mathbb{R}$ mit $\sum_{k=1}^{\infty} c_k^2 < \infty$. Dann gilt

$$\mathbb{E}\left(c_k \varepsilon_k\right) = 0 \quad , \quad \sum_{k=1}^{\infty} \mathbb{E}\left|c_k \varepsilon_k\right|^2 = \sum_{k=1}^{\infty} |c_k|^2 < \infty$$

so dass die Reihe
$$\sum_{k=1}^{\infty} c_k \varepsilon_k$$
fast-sicher konvergiert. Setzt man speziell
$$c_k := \frac{1}{\sqrt{k}\,[\ln k]^{\alpha}} \quad , \quad \alpha > \frac{1}{2}$$
so ist bekanntlich
$$\sum_{k=2}^{\infty} c_k^2 < \infty$$
also existiert
$$\sum_{k=2}^{\infty} \frac{\varepsilon_k}{\sqrt{k}\,[\ln k]^{\alpha}}$$
fast-immer.

3.6 Starkes Gesetz der großen Zahlen (SLLN)

3.6.1 Definition: SLLN

Seien $X_1, X_2, ..$ reelle Zufallsgrößen mit $\mathbb{E}X_k < \infty \;\forall\, k \in \mathbb{N}$. Dann erfüllt $(X_k)_{k=1}^{\infty}$ *das starke Gesetz der großen Zahlen (SLLN)*[7] wenn
$$\frac{1}{n} \sum_{k=1}^{n} (X_k - \mathbb{E}X_k) \overset{n\to\infty}{\underset{\text{f.s.}}{\longrightarrow}} 0$$
Sind $X_1, X_2, ..$ m-dimensionale, reelle zufällige Vektoren, so erfüllt $(X_k)_{k=1}^{\infty}$ das *SLLN* falls jede Komponentenfolge $(X_k^i)_{k=1}^{\infty}$, $i = 1, .., m$ das SLLN erfüllt (vgl. Lemma 3.2.4).

[7] **S**trong **L**aw of **L**arge **N**umbers

Spezialfall: Sind die $X_1, X_2, ..$ identisch verteilt, dann ist das SLLN genau dann erfüllt falls

$$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow[\text{f.s.}]{n\to\infty} \mathbb{E}X_1$$

3.6.2 SLLN Theorem von Khinchine

Seien $X_1, X_2, ..$ reelle, unabhängige Zufallsgrößen mit

$$\sum_{k=1}^{\infty} \frac{\mathbb{V}X_k}{k^2} < \infty$$

Dann erfüllt die Folge $(X_k)_{k=1}^{\infty}$ das SLLN, das heißt

$$\frac{1}{n}\sum_{k=1}^{n} (X_k - \mathbb{E}X_k) \xrightarrow[\text{f.s.}]{n\to\infty} 0$$

Beweis: Setze $Y_k := X_k - \mathbb{E}X_k$, dann ist $\mathbb{E}Y_k = 0,\ k \in \mathbb{N}$. Wegen $\mathbb{V}Y_k = \mathbb{V}X_k$ ist nach Voraussetzung

$$\sum_{k=1}^{\infty} \frac{\mathbb{V}Y_k}{k^2} < \infty$$

Nach Korollar 3.5.6(2) geht dann

$$\frac{1}{n}\sum_{k=1}^{n} Y_k \xrightarrow[\text{f.s.}]{n\to\infty} 0$$

(setze $b_k := k$).
□

Folgerung: Seien $X_1, X_2, ..$ reell, iid mit $\mathbb{V}X_1 < \infty$. Dann erfüllt $(X_k)_{k=1}^{\infty}$ das SLLN, das heißt

$$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow[\text{f.s.}]{n\to\infty} \mathbb{E}X_1$$

Problem: Die Forderung der Existenz des 2. Momentes der X_k erscheint angesichts der eigentlichen Aussage zunächst unnatürlich. Tatsächlich ist es möglich, das SLLN auch ohne dessen Existenz zu schlussfolgern!

3.6.3 Lemma über die Existenz des Erwartungswertes

Seien $X_1, X_2, \ldots$ reelle, iid Zufallsvariablen. Dann gilt:

$$\mathbb{E}\,|X_1| < \infty \quad\Leftrightarrow\quad \mathcal{P}\left(|X_n| \geq n \text{ u.o.}\right) = 0$$

Beweis: Es seien $X_k : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}$. Es gilt:

$$\mathcal{P}\left(|X_n| \geq n \text{ u.o.}\right) = 0 \overset{\substack{\text{Borel \&}\\ \text{Cantelli}}}{\Longleftrightarrow} \sum_{n=1}^{\infty} \mathcal{P}\left(|X_n| \geq n\right) < \infty$$

$$\Leftrightarrow \sum_{n=1}^{\infty} \mathcal{P}\left(|X_1| \geq n\right) < \infty$$

$$\overset{1.6.5(1)}{\Longleftrightarrow} \mathbb{E}\,|X_1| < \infty$$

□

3.6.4 SLLN Theorem von Kolmogorov

Die reellen Zufallsvariablen $X_1, X_2, \ldots$ seien unabhängig, identisch verteilt mit $\mathbb{E}|X_1| < \infty$. Dann erfüllen $(X_k)_{k=1}^{\infty}$ das SLLN, das heißt:

$$\frac{1}{n}\sum_{k=1}^{n} X_k \underset{\text{f.s}}{\overset{n\to\infty}{\longrightarrow}} \mathbb{E}X_1$$

Beweis: Seien $X_1, X_2, .. : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}$.

- Setze
$$Y_k := X_k \cdot 1_{\{|X_k| \leq k\}}$$
Dann sind $Y_1, Y_2, ..$ nach Bsp. 2.4.6(ii) unabhängig, jedoch nicht mehr identisch verteilt. Dabei gilt
$$|\mathbb{E}X_k - \mathbb{E}Y_k| = \left| \int\limits_{\{|X_k|>k\}} X_k \; d\mathcal{P} \right| \leq \int\limits_{\{|X_k|>k\}} |X_k| \; d\mathcal{P} = \int\limits_{|t|>k} |t| \; d\mathcal{P}_{X_1}(t)$$
$$= \int\limits_{A_k} |t| \; d\mathcal{P}_{X_1}(t) \underset{1.2.5}{\overset{k\to\infty}{\longrightarrow}} 0 \quad \Big| \quad A_k := \{|t| > k\} \underset{k\to\infty}{\searrow} 0$$

- Nach Lemma 3.6.3 ist
$$\mathcal{P}\left(|X_k| \geq k \text{ u.o.}\right) = 0$$
das heißt für fast-alle $\omega \in \Omega$ existiert ein $n_\omega \in \mathbb{N}$ mit
$$|X_k(\omega)| \leq k \quad \forall \; k \geq n_\omega$$
bzw.
$$Y_k(\omega) = X_k(\omega) \quad \forall \; k \geq n_\omega \tag{3.6.4.1}$$

- Die $(X_k)_{k=1}^{\infty}$ erfüllen genau dann das SLLN wenn dies auch die $(Y_k)_{k=1}^{\infty}$ tun.
Beweis: Wegen
$$\frac{1}{n}\sum_{k=1}^{n} X_k(\omega) \overset{(3.6.4.1)}{=} \frac{1}{n}\sum_{k=1}^{n_\omega} X_k(\omega) + \frac{1}{n}\sum_{k=n_\omega}^{n} Y_k(\omega)$$
kann man abschätzen
$$\left| \frac{1}{n}\sum_{k=1}^{n} (X_k(\omega) - \mathbb{E}X_k) - \frac{1}{n}\sum_{k=1}^{n} (Y_k(\omega) - \mathbb{E}Y_k) \right| \leq \underbrace{\left| \frac{1}{n}\sum_{k=1}^{n} \underbrace{(\mathbb{E}X_k - \mathbb{E}Y_k)}_{\overset{k\to\infty}{\longrightarrow} 0} \right|}_{\substack{\overset{n\to\infty}{\longrightarrow} 0 \\ \text{nach Kronecker} \\ A.5.2(2)}} + \underbrace{\left| \frac{1}{n}\sum_{k=1}^{n_\omega} (X_k(\omega) - Y_k(\omega)) \right|}_{\overset{n\to\infty}{\longrightarrow} 0} \overset{n\to\infty}{\longrightarrow} 0$$
woraus die Äquivalenz
$$\frac{1}{n}\sum_{k=1}^{n} (Y_k - \mathbb{E}Y_k) \underset{\text{f.s.}}{\overset{n\to\infty}{\longrightarrow}} 0 \quad \Leftrightarrow \quad \frac{1}{n}\sum_{k=1}^{n} (X_k - \mathbb{E}X_k) \underset{\text{f.s.}}{\overset{n\to\infty}{\longrightarrow}} 0 \tag{3.6.4.2}$$
folgt.

- Nach (3.6.4.2) und SLLN Theorem 3.6.2 genügt es zu zeigen:
$$\sum_{k=1}^{\infty} \frac{\mathbb{E}|Y_k|^2}{k^2} < \infty$$

Tatsächlich:

$$\sum_{k=1}^{\infty} \frac{\mathbb{E}\,|Y_k|^2}{k^2} = \sum_{k=1}^{\infty} \frac{1}{k^2} \int\limits_{|t|\le k} |t|^2 \; d\mathcal{P}_{X_1}(t) \le \sum_{k=1}^{\infty} \frac{1}{k^2} \sum_{j=1}^{k} \int\limits_{j-1\le|t|\le j} |t|^2 \; d\mathcal{P}_{X_1}(t)$$

$$= \sum_{j=1}^{\infty} \sum_{k=j}^{\infty} \frac{1}{k^2} \int\limits_{j-1\le|t|\le j} |t|^2 \; d\mathcal{P}_{X_1}(t) \le 2 \cdot \sum_{j=1}^{\infty} \frac{1}{j} \int\limits_{j-1\le|t|\le j} |t|^2 \; d\mathcal{P}_{X_1}(t)$$

$$\le 2 \cdot \sum_{j=1}^{\infty} \int\limits_{j-1\le|t|\le j} |t| \; d\mathcal{P}_{X_1}(t) \le 2 \int\limits_{\mathbb{R}} |t| \; d\mathcal{P}_{X_1}(t) = 2\mathbb{E}\,|X_1| < \infty$$

Beachte dass verwendet wurde

$$\sum_{k=j}^{\infty} \frac{1}{k^2} \le \frac{2}{j} \quad , \quad j \in \mathbb{N}$$

was durch

$$\sum_{k=j}^{\infty} \frac{1}{k^2} \le \int\limits_{j-1}^{\infty} \frac{dx}{x^2} = \frac{1}{j-1} \le \frac{2}{j}$$

für $j \ge 2$ und

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6} \le \frac{2}{1}$$

für $j = 1$ ersichtlich wird.

□

Frage: Angesichts der Tatsache dass für iid reelle $X_1, X_2, ..$ die Endlichkeit des Erwartungswertes $\mathbb{E}\,|X_1| < \infty$ das SLLN

$$\frac{1}{n} \sum_{k=1}^{n} X_k \xrightarrow[\text{f.s.}]{n\to\infty} \mathbb{E}X_1$$

impliziert, stellt sich die Frage ob umgekehrt im Falle der fast-sicheren Konvergenz[8] von

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k \overset{\text{f.s.}}{=} C : \text{const}$$

der entsprechende Grenzwert als *verallgemeinerte Grenzwert* aufgefasst werden kann. Wie sich heraus stellt, sind beide *Definitionen* äquivalent.

3.6.5 Theorem über die Existenz des Erwartungswertes aus dem SLLN

Seien $X_1, X_2, ..$ reelle iid Zufallsgrößen. Existiert der Grenzwert

$$\frac{1}{n} \sum_{k=1}^{n} X_k =: C$$

fast-sicher so ist auch $\mathbb{E}\,|X_1| < \infty$ und $\mathbb{E}X_1 = C$.

[8] Beachte dass nach Korollar 3.3.9 die Folge $\frac{1}{n} \sum_{k=1}^{n} X_k$ entweder fast-sicher gegen eine Konstante konvergiert oder fast-sicher divergiert.

Beweis: Setze

$$S_n := \sum_{k=1}^{n} X_k \quad , \quad n \in \mathbb{N}$$

dann gehen

$$\frac{S_n}{n} \xrightarrow[\text{f.s}]{n\to\infty} C$$

Mit $X_n = S_n - S_{n-1}$ folgt

$$\frac{X_n}{n} = \underbrace{\frac{S_n}{n}}_{\xrightarrow[\text{f.s.}]{n\to\infty} C} - \underbrace{\frac{n-1}{n}}_{\xrightarrow{n\to\infty} 1} \cdot \underbrace{\frac{S_{n-1}}{n-1}}_{\xrightarrow[\text{f.s.}]{n\to\infty} C} \xrightarrow[\text{f.s.}]{n\to\infty} 0$$

das heißt

$$\mathcal{P}\left(\tfrac{|X_n|}{n} \geq \varepsilon \text{ u.o.}\right) = 0 \quad \forall\, \varepsilon > 0$$

Speziell für $\varepsilon := 1$ folgt nach Lemma 3.6.3 die Existenz $\mathbb{E}\,|X_1| < \infty$. Nach dem SLLN Theorem von Kolmogorov 3.6.4 und der Eindeutigkeit des Grenzwertes muss $C = \mathbb{E}X_1$ sein.
□

Zusammenfassung: Aus obigen Überlegungen lässt sich für reelle, iid $X_1, X_2, ..$ zusammenfassen:

1. Die *gemittelten* Partialsummen

$$\frac{S_n}{n} := \frac{1}{n} \sum_{k=1}^{n} X_k$$

   konvergieren in $\mathbb{R}$ fast-sicher oder fast-nie.

2. Dabei konvergieren die S_n/n fast-sicher genau dann wenn $\mathbb{E}\,|X_1| < \infty$. In dem Fall geht

$$\frac{S_n}{n} \xrightarrow[\text{f.s.}]{n\to\infty} \mathbb{E}X_1$$

3. Gilt andernfalls $\mathbb{E}\,|X_1| = \infty$, so divergiert S_n/n fast-immer.

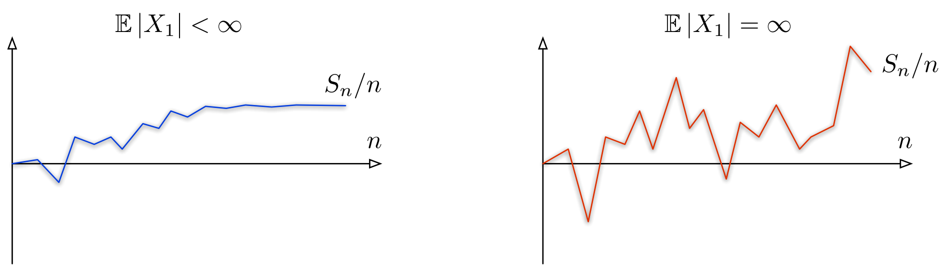

Abbildung 10: Zur Erfüllung des SLLN: Typischer Verlauf der S_n/n im Fall $\mathbb{E}\,|X_1| < \infty$ (blau) und $\mathbb{E}\,|X_1| = \infty$ (rot).

3.7 Folgerungen des SLLN

3.7.1 Gesetz der Häufigkeit

Unabhängig von obigem Formalismus, stellt sich in der Wahrscheinlichkeitstheorie die Frage nach dem Sinn der Wahrscheinlichkeit. Sprich, was bedeutet es bei vorgegebenem Wahrscheinlichkeitsraum $(S, \mathcal{S}, \mathcal{Q})$ für ein Ereignis $A \in \mathcal{S}$ das Wahrscheinlichkeitsmaß $\mathcal{Q}(A)$ zu besitzen?
Einsicht liefert folgender Satz.

3.7.2 Satz über die relative Häufigkeit von Ereignissen

Sei $(S, \mathcal{S}, \mathcal{Q})$ ein Wahrscheinlichkeitsraum, dazu die S-wertigen, unabhängigen, identisch verteilten Zufallsvariablen $X_1, X_2, .. : (\Omega, \mathfrak{S}, \mathcal{P}) \to (S, \mathcal{S})$ mit Verteilungsgesetz $\mathcal{Q}$ auf $(S, \mathcal{S})$. Für vorgegebene Menge $A \in \mathcal{S}$ bezeichne

$$r_n(\omega) := \frac{1}{n} \# \{k \leq n : X_k(\omega) \in A\} \quad , \quad \omega \in \Omega$$

die relative Häufigkeit des Eintretens von A nach n Versuchen. Dann gilt:

$$r_n \underset{\text{f.s.}}{\overset{n\to\infty}{\longrightarrow}} \mathcal{Q}(A)$$

das heißt die relative Eintreff-Häufigkeit des Ereignisses konvergiert fast-sicher gegen seine *Wahrscheinlichkeit.*[9]

Beweis: Setze

$$Y_k := 1_A \circ X_k \quad , \quad k \in \mathbb{N}$$

Dann sind die $Y_1, Y_2, ..$ unabhängig, identisch verteilt mit

$$\mathcal{P}(Y_k = 1) = \mathcal{P}(X_k \in A) = \mathcal{Q}(A)$$

Anderseits ist

$$\frac{1}{n} \sum_{k=1}^{n} Y_k = \frac{1}{n} \# \{k \leq n : X_k \in A\} = r_n$$

und nach Kolmogorov 3.6.4

$$r_n \underset{\text{f.s.}}{\overset{n\to\infty}{\longrightarrow}} \mathbb{E}Y_1 = \mathcal{Q}(A)$$

□

3.7.3 Deutung des Erwartungswertes

Entsprechend des SLLN für iid reelle Zufallsvariablen $X_1, X_2, ..$

$$\frac{1}{n} \sum_{k=1}^{n} X_k \underset{\text{f.s}}{\overset{n\to\infty}{\longrightarrow}} \mathbb{E}X_1$$

kann der Erwartungswert $\mathbb{E}X_1$ als asymptotisch durchschnittliches Ergebnis nach hinreichend vielen Versuchen gedeutet werden!

Beispiele:

(i) Ist X_i die, unabhängig von anderen Tagen, poissonverteilte $X_k \sim \pi_\lambda$ Anzahl der Kunden in einem Kaufhaus am i-ten Tag, sprich

$$\mathcal{P}(X_j = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad , \quad k \in \mathbb{N}_0$$

so entspricht

$$\frac{1}{n} \sum_{k=1}^{n} X_k$$

der durchschnittlichen Kundenanzahl der ersten n Tage. Nach Kolmogorov 3.6.4 geht diese

$$\frac{1}{n} \sum_{k=1}^{n} X_k \underset{\text{f.s}}{\overset{n\to\infty}{\longrightarrow}} \mathbb{E}X_1 = \lambda$$

[9] Beachte dass nicht desto trotz die Konvergenzgeschwindigkeit stets *zufällig* ist.

(ii) Die Lebensdauer L_1 von Glühbirnen einer bestimmten Art sei Exponentialverteilt $\sim \text{Exp}_\lambda$, sprich

$$\mathcal{P}(L_1 \geq t) = e^{-\lambda t} \quad , \quad t > 0$$

Entsprechen die $L_1, L_2, .., L_n$ den (unabhängigen) Lebensdauern von n Glühbirnen, so geht deren durchschnittliche Lebensdauer

$$\frac{1}{n} \sum_{k=1}^{n} L_k \overset{n\to\infty}{\underset{\text{f.s.}}{\longrightarrow}} \mathbb{E}L_1 = \frac{1}{\lambda}$$

Interpretationsgemäß ist $1/\lambda$ die asymptotisch durchschnittliche Lebenszeit von Glühbirnen dieser Art.

3.7.4 Monte Carlo Integrationsmethode

Für messbare $f : [0,1]^N \to \mathbb{R}$ sei (ggf. numerisch) zu berechnen

$$\int\limits_{[0,1]^N} f(x)\ dx \tag{3.7.4.1}$$

Die Tatsache dass diese Aufgabe für genügend großes N beliebig komplex werden kann (oft $N \sim 10^6$) gab in den 1940 ern Anlass zur Entwicklung der sogenannten *Monte Carlo Methode*. Das Prinzip sei im folgenden illustriert.

Gegeben sei die Folge unabhängiger, auf $[0,1]^N$ gleichverteilter Zufallsgrößen $\mathbf{U}_1, \mathbf{U}_2, ..$:

$$\mathbf{U}_n = \left(U_n^1, .., U_n^N\right) \sim \lambda^N \quad , \quad U_n^i \sim \lambda^1,\ i = 1, .., N$$

Dann sind auch $f(\mathbf{U}_1), f(\mathbf{U}_2), ..$ unabhängig, identisch verteilt. Nach Kolmogorov 3.6.4 geht dabei

$$\frac{1}{n} \sum_{k=1}^{n} f(\mathbf{U}_k) \overset{n\to\infty}{\underset{\text{f.s.}}{\longrightarrow}} \mathbb{E}f(\mathbf{U}_1) = \int\limits_{[0,1]^N} f(x)\ dx$$

Durch *zufälliges Abtasten* der Funktionswerte auf $[0,1]^N$ und entsprechend gewichteter Summierung kann der tatsächliche Integralwert (3.7.4.1) beliebig genau approximiert werden! Es stellt sich heraus, dass in vielen Fällen solch zufälliges Abtasten in eine bessere Konvergenz resultiert als etwa gleichmäßige Abtastmethoden!

3.7.5 Definition: Normale Zahlen

Sei $b \in \{2, 3, ..\}$ vorgegeben, dazu für $x \in [0,1]$ die b-adische Darstellung[10]

$$x = \sum_{j=1}^{\infty} \frac{x_j}{b^j} \quad , \quad x_j \in \{0, 1, .., b-1\}$$

Dann heißt $x \in [0,1]$ *b-normal*, falls

$$\frac{1}{n} \#\left\{j \leq n : x_j = a\right\} \overset{n\to\infty}{\longrightarrow} \frac{1}{b} \quad \forall\ a \in \{0, 1, .., b-1\}$$

Ferner heißt x *absolut normal*, falls x b-normal für jedes $b \in \{2, 3, ..\}$ ist.

3.7.6 Theorem über normale Zahlen (Émile Borel)

λ^1-fast jede Zahl $x \in [0,1]$ ist b-normal.

[10]Beachte dass solch eine Darstellung λ^1-fast sicher eindeutig ist.

Beweis: Setze $Z_j(x) := x_j$ für

$$x = \sum_{j=1}^{\infty} \frac{x_j}{b^j} \in [0,1] \quad , \quad x_j \in \{0,..,b-1\}$$

Dann gilt

$$\lambda^1(Z_j = a) = \frac{1}{b} \quad \forall\, a \in \{0,..,b-1\},\ j \in \mathbb{N}$$

wobei alle $Z_1, Z_2, ..$ unabhängig sind. Sei nun $a \in \{0,1,..,b-1\}$ vorgegeben, dazu $Y_j := 1_{\{a\}} \circ Z_j$. Dann sind auch $Y_1, Y_2, ..$ unabhängig und nach Kolmogorov 3.6.4 gilt

$$\frac{1}{n}\sum_{k=1}^{n} Y_k \xrightarrow[\text{f.s}]{n\to\infty} \mathbb{E}Y_1 = \frac{1}{b}$$

Wegen

$$\frac{1}{n}\sum_{k=1}^{n} Y_k = \frac{1}{n}\#\{j \le n : x_j = a\}$$

folgt die Behauptung.
□

Bemerkung: Es sind tatsächlich λ^1-fast alle $x \in [0,1]$ absolut normal, denn

$$\{x \in [0,1] : x \text{ absolut normal}\} = \underbrace{\bigcap_{b=2}^{\infty} \underbrace{\{x \in [0,1] : x\ b\text{-normal}\}}_{\text{fast-sichere Menge}}}_{\text{fast-sichere Menge}}$$

Offen bleibt jedoch die Frage nach der Normalität spezieller Zahlen wie z.B. $\pi - 3$ oder $\sqrt{2} - 1$.

3.7.7 Theorem über Bernstein Polynome (Bernstein)

Für stetige $f : [0,1] \to \mathbb{R}$ bezeichnet

$$B_n^f(x) := \sum_{k=0}^{n} \binom{n}{k} \cdot f\left(\tfrac{k}{n}\right) \cdot x^k (1-x)^{n-k} \tag{3.7.7.1}$$

das n*-te Bernstein Polynom* von f. Es gilt:

$$\lim_{n\to\infty} \sup_{0\le x\le 1} \left|f(x) - B_n^f(x)\right| = 0$$

das heißt, die Polynome B_n^f konvergieren gleichmäßig gegen f.

Beweis:

- Zeigen zunächst die punktweise Konvergenz. Für vorgegebenes $p \in [0,1]$ seien die reellen Zufallsgrößen $X_1, X_2, ..$ unabhängig, identisch verteilt gemäß

  $$\mathcal{P}(X_j = 1) = p\ ,\ \mathcal{P}(X_j = 0) = 1 - p$$

  (Bernoulli-Folge). Dann sind

  $$S_n := X_1 + \cdots + X_n \quad , \quad n \in \mathbb{N}$$

  binomialverteilt, sprich

  $$\underbrace{\mathcal{P}(S_n = k)}_{\mathcal{P}\left(\frac{S_n}{n} = \frac{k}{n}\right)} = \binom{n}{k} \cdot p^k (1-p)^{n-k}$$

Nach Kolmogorov 3.6.4 geht

$$\frac{S_n}{n} \xrightarrow[\text{f.s.}]{n\to\infty} \mathbb{E}X_1 = p$$

bzw.

$$f\left(\tfrac{S_n}{n}\right) \xrightarrow[\text{f.s.}]{n\to\infty} f(p)$$

Da f auf $[0,1]$ beschränkt ist, existiert ein $C \geq 0$ mit $\left|f\left(\frac{S_n}{n}\right)\right| \leq C$ (Majorante) und nach Lebesgue folgt

$$\mathbb{E}f\left(\tfrac{S_n}{n}\right) \xrightarrow{n\to\infty} \mathbb{E}f(p) = f(p)$$

Mit

$$\mathbb{E}f\left(\tfrac{S_n}{n}\right) = \sum_{k=0}^{n} \binom{n}{k} \cdot f\left(\tfrac{k}{n}\right) \cdot p^k(1-p)^{n-k} = B_n^f(p)$$

folgt schließlich die punktweise Konvergenz in p.

- Zu zeigen bleibt nun die gleichmäßige Konvergenz. Sei $\varepsilon > 0$ vorgegeben. Da $f : [0,1] \to \mathbb{R}$ gleichmäßig stetig ist, existiert ein $\delta_\varepsilon > 0$ so dass

$$\forall\, x, y \in [0,1] : |x-y| < \delta_\varepsilon \;\Rightarrow\; |f(x)-f(y)| \leq \varepsilon$$

Dementsprechend gilt

$$\mathbb{E}\left|f\left(\tfrac{S_n}{n}\right) - f(p)\right| \cdot 1_{\left\{\left|\frac{S_n}{n}-p\right|<\delta_\varepsilon\right\}} \leq \mathbb{E}\varepsilon = \varepsilon \tag{3.7.7.2}$$

Da f beschränkt ist, existiert ein $C \geq 0$ mit $|f| \leq C$ und daher

$$\mathbb{E}\underbrace{\left|f\left(\tfrac{S_n}{n}\right) - f(p)\right|}_{\leq 2C} \cdot 1_{\left\{\left|\frac{S_n}{n}-p\right|\geq\delta_\varepsilon\right\}} \leq 2C \cdot \mathbb{E}1_{\left\{\left|\frac{S_n}{n}-p\right|\geq\delta_\varepsilon\right\}} = 2C \cdot \mathcal{P}\Big(\big|\tfrac{S_n}{n} - \underbrace{p}_{\mathbb{E}\frac{S_n}{n}}\big| \geq \delta_\varepsilon\Big)$$

$$\overset{1.8.3(ii)}{\leq} 2C \cdot \frac{\mathbb{V}\left(\frac{S_n}{n}\right)}{\delta_\varepsilon^2} = \frac{2C \cdot \mathbb{V}S_n}{n^2\delta^2} = \frac{2C \cdot \overbrace{p(1-p)}^{\leq 1/4}}{n\delta_\varepsilon^2}$$

$$\leq \frac{C}{2n\delta_\varepsilon^2} \leq \varepsilon \quad \forall\, n \geq n_\varepsilon \quad \Big| \quad \text{für geeignetes } n_\varepsilon \in \mathbb{N}$$

Zusammen mit (3.7.7.2) folgt dann

$$\Big|\overbrace{\mathbb{E}f\left(\tfrac{S_n}{n}\right)}^{B_n^f(p)} - f(p)\Big| \leq \mathbb{E}\left|f\left(\tfrac{S_n}{n}\right) - f(p)\right|$$

$$= \underbrace{\mathbb{E}\left|f\left(\tfrac{S_n}{n}\right) - f(p)\right| \cdot 1_{\left\{\left|\frac{S_n}{n}-p\right|\geq\delta_\varepsilon\right\}}}_{\leq\varepsilon \quad \forall\, n\geq n_\varepsilon} + \underbrace{\mathbb{E}\left|f\left(\tfrac{S_n}{n}\right) - f(p)\right| \cdot 1_{\left\{\left|\frac{S_n}{n}-p\right|<\delta_\varepsilon\right\}}}_{\leq\varepsilon}$$

$$\leq 2\varepsilon \quad \forall\, n \geq n_\varepsilon\ \forall\, p \in [0,1]$$

sprich

$$\left\|B_n^f - f\right\|_\infty \leq 2\varepsilon \quad \forall\, n \geq n_\varepsilon$$

□

Weiteres Beispiel: Zu den unabhängigen, Poissonverteilten $X_1, X_2, .. \sim \pi_\lambda$, sprich

$$\mathcal{P}(X_j = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

setze

$$S_n := X_1 + \cdots + X_n \ \sim \ \pi_{n\lambda} \quad , \quad n \in \mathbb{N}$$

Dann ist für beschränkte, stetige $f : [0, \infty) \to \mathbb{R}$

$$\mathbb{E} f\left(\tfrac{S_n}{n}\right) = \sum_{k=0}^{\infty} f\left(\tfrac{k}{n}\right) \cdot \frac{(n\lambda)^k}{k!} e^{-\lambda n}$$

und nach Kolmogorov 3.6.4 bzw. Lebesgue gilt

$$\mathbb{E} f\left(\tfrac{S_n}{n}\right) \overset{n\to\infty}{\longrightarrow} f(\lambda)$$

sprich

$$e^{-\lambda n} \sum_{k=0}^{\infty} f\left(\tfrac{k}{n}\right) \cdot \frac{(n\lambda)^k}{k!} \overset{n\to\infty}{\longrightarrow} f(\lambda) \tag{3.7.7.3}$$

3.8 Das schwache Gesetz der großen Zahlen (WLLN)

3.8.1 Definition: Schwaches Gesetz der großen Zahlen (WLLN)

Die Folge reeller, zufälliger Größen $X_1, X_2, ..$ mit $\mathbb{E}|X_k| < \infty$ erfüllt das *schwache Gesetz der großen Zahlen (WLLN*[11]*)* falls

$$\frac{1}{n} \sum_{k=1}^{n} (X_k - \mathbb{E}X_k) \underset{\mathrm{P}}{\overset{n\to\infty}{\longrightarrow}} 0 \tag{3.8.1.1}$$

das heißt

$$\forall\, \varepsilon > 0 : \mathcal{P}\left(\frac{1}{n} \left| \sum_{k=1}^{n} (X_k - \mathbb{E}X_k) \right| \geq \varepsilon \right) \overset{n\to\infty}{\longrightarrow} 0$$

Bemerke dass das SLLN (z.B. im Falle $X_1, X_2, ..$ iid) das WLLN impliziert. Interessant ist das WLLN z.B. für nicht-identisch verteilte Zufallsgrößen.

3.8.2 WLLN-Satz von Khinchine

Die reellen Zufallsgrößen $X_1, X_2, ..$ seien unabhängig mit $\mathbb{E}|X_k|^2 < \infty$. Gilt

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{k=1}^{n} \mathbb{V} X_k = 0$$

so erfüllt $X_1, X_2, ..$ das WLLN.

Beweis: Sei $\varepsilon > 0$. Dann gilt

$$\mathcal{P}\left(\frac{1}{n} \left| \sum_{k=1}^{n} (X_k - \mathbb{E}X_k) \right| \geq \varepsilon \right) \overset{1.8.3(ii)}{\leq} \frac{1}{n^2 \varepsilon^2} \mathbb{V}\left(\sum_{k=1}^{n} X_k \right) = \frac{1}{n^2 \varepsilon^2} \sum_{k=1}^{n} \mathbb{V} X_k \overset{n\to\infty}{\longrightarrow} 0$$

□

[11] **W**eak **L**aw of **L**arge **N**umbers.

$$\begin{array}{ccc} \displaystyle\sum_{k=1}^{\infty} \frac{\mathbb{V}X_k}{k^2} < \infty & \overset{\text{Kronecker}}{\Longrightarrow} & \displaystyle\frac{1}{n^2}\sum_{k=1}^{n} \mathbb{V}X_k \overset{n\to\infty}{\longrightarrow} 0 \\ \Downarrow & & \Downarrow \\ \text{SLLN} & \Longrightarrow & \text{WLLN} \end{array}$$

Abbildung 11: Zusammenhang zwischen dem SLLN und dem WLLN für unabhängige, reelle Zufallsgrößen $X_1, X_2, \ldots$

Beispiel: Die unabhängigen, reellen Zufallsgrößen $X_1, X_2, ..,$ verteilt gemäß

$$\mathcal{P}(X_k = \pm k) = \frac{1}{2k\ln(k+1)} \quad , \quad \mathcal{P}(X_k = 0) = 1 - \frac{1}{k\ln(k+1)}$$

erfüllen das WLLN, aber nicht das SLLN.
Erläuterung: Wegen

$$\frac{1}{n^2}\sum_{k=1}^{n} \mathbb{V}X_k = \frac{1}{n^2}\sum_{k=1}^{n} \underbrace{\frac{k}{\log(k+1)}}_{\leq \frac{n}{\log(n+1)}} \leq \frac{1}{\log(n+1)} \overset{n\to\infty}{\longrightarrow} 0$$

erfüllen die $(X_n)_{n=1}^{\infty}$ das WLLN. Anderseits gilt

$$\sum_{k=1}^{\infty} \underbrace{\mathcal{P}(X_k \geq k)}_{\frac{1}{k\ln(k+1)}} = \infty$$

bzw. nach Borel & Cantelli 3.4.1

$$\mathcal{P}\left(\tfrac{|X_k|}{k} \geq 1 \text{ u.o}\right) = 1 \tag{3.8.2.1}$$

Würde nun $(X_k)_{k=1}^{\infty}$ das SLLN erfüllen

$$\frac{S_n}{n} := \frac{1}{n}\sum_{k=1}^{n} X_k \underset{\text{f.s.}}{\overset{n\to\infty}{\longrightarrow}} 0$$

so müsse auch

$$\frac{X_n}{n} = \frac{S_n - S_{n-1}}{n} = \underbrace{\frac{S_n}{n}}_{\underset{\text{f.s.}}{\overset{n\to\infty}{\longrightarrow}} 0} - \underbrace{\frac{n-1}{n}}_{\overset{n\to\infty}{\longrightarrow} 1} \underbrace{\frac{S_{n-1}}{n-1}}_{\underset{\text{f.s.}}{\overset{n\to\infty}{\longrightarrow}} 0} \underset{\text{f.s}}{\overset{n\to\infty}{\longrightarrow}} 0$$

ein Widerspruch zu (3.8.2.1).
□

3.9 Weitere Eigenschaften von Summen unabhängiger Zufallsgrößen

3.9.1 Vorbetrachtung

Zu reellen, unabhängigen, identisch verteilten $X_1, X_2, ..$ sei

$$S_n := X_1 + \cdots + X_n \quad , \quad n \in \mathbb{N}$$

und $\mathbb{E}\,|X_1| < \infty$. Nach Kolmogorov geht dann

$$\frac{S_n}{n} \underset{\text{f.s.}}{\overset{n\to\infty}{\longrightarrow}} \mathbb{E}X_1$$

Im Fall $\mathbb{E}X_1 > 0$ insbesondere

$$\mathcal{P}\left(S_n \geq n \cdot \tfrac{\mathbb{E}X_1}{2} \text{ s.i.}\right) = 1$$

sprich

$$S_n \underset{\text{f.s.}}{\overset{n\to\infty}{\longrightarrow}} \infty$$

Analog $S_n \underset{\text{f.s.}}{\overset{n\to\infty}{\longrightarrow}} -\infty$ im Fall $\mathbb{E}X_1 < 0$.

Nun stellt sich die Frage, was im Falle $\mathbb{E}X_1 = 0$ (o.B.d.A. $\mathcal{P}(X_1 = 0) < 1$) passiert. So gilt z.B. [1]

$$\mathcal{P}\Big(\varlimsup_{n\to\infty} S_n = \infty\Big) = 1 = \mathcal{P}\Big(\varliminf_{n\to\infty} S_n = -\infty\Big) \tag{3.9.1.1}$$

Allgemein ist die *minimale* Folge $0 < a_1 \leq a_2 \leq \dots \to \infty$ gesucht, für die

$$\varlimsup_{n\to\infty} \frac{S_n}{a_n} = 1 \text{ fast sicher}$$

So ist z.B. $a_n := n$ *zu klein*, $a_n := \sqrt{n}$ oft *zu groß* (vgl. unteres Lemma 3.9.2).

3.9.2 Lemma: Gewichtete Abschätzung von Summen unabhängiger Zufallsvariablen

Seien $X_1, X_2, ..$ reelle, iid Zufallsgrößen mit $\mathcal{P}(X_1 > t) > 0 \ \ \forall\, t \in \mathbb{R}$. Dann gilt

$$\sup_{n\in\mathbb{N}} \frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k = \infty \qquad \text{fast sicher}$$

Beispiel: Für standardnormalverteilte $\xi_1, \xi_2, .. \sim \mathcal{N}_{0,1}$ sind

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \xi_k \sim \mathcal{N}_{0,1}$$

auch standardnormalverteilt (jedoch nicht unbedingt unabhängig!), und bilden fast-sicher eine nach oben unbeschränkte Folge.

Beweis: Seien wie üblich

$$S_n := X_1 + \dots + X_n \quad , \quad n \in \mathbb{N}$$

und $C > 0$ beliebig. Die Mengen

$$A_m := \left\{\tfrac{S_n}{\sqrt{n}} \leq C \ \ \forall\, n \geq m\right\} \quad , \quad m \in \mathbb{N}$$

sind monoton wachsend in m und es gilt

$$\bigcup_{m\in\mathbb{N}} A_m = \left\{\tfrac{S_n}{\sqrt{n}} \leq C \text{ s.i.}\right\}$$

Ferner

$$\mathcal{P}(A_m) \leq \mathcal{P}\left(\tfrac{S_m}{\sqrt{m}} \leq C\right) =: \delta < 1 \qquad \text{nach Voraussetzung}$$

Daher

$$\mathcal{P}\left(\tfrac{S_n}{\sqrt{n}} \leq C \text{ s.i.}\right) = \lim_{n\to\infty} \mathcal{P}(A_m) \leq \delta < 1$$

Anderseits ist $\left\{\tfrac{S_n}{\sqrt{n}} \leq C \text{ s.i}\right\}$ terminal und kann nach Kolmogorov 3.3.6 nur die Wahrscheinlichkeit 0 oder 1 besitzen. Also

$$\mathcal{P}\left(\tfrac{S_n}{\sqrt{n}} \leq C \text{ s.i.}\right) = 0$$

bzw.

$$\mathcal{P}\left(\tfrac{S_n}{\sqrt{n}} > C \text{ u.o.}\right) = 1 \quad \forall\, C > 0$$

Aus

$$\mathcal{P}\left(\tfrac{S_n}{\sqrt{n}} > C \text{ u.o. } \forall\, C\right) = \mathcal{P}\Big(\bigcap_{k\in\mathbb{N}} \left\{\tfrac{S_n}{\sqrt{n}} > k \text{ u.o.}\right\}\Big) = \lim_{k\to\infty} \mathcal{P}\left(\tfrac{S_n}{\sqrt{n}} > k \text{ u.o.}\right) = 1$$

folgt schließlich die Behauptung.
□

3.9.3 Theorem vom iterierten Logarithmus (Hartman, Wintner 1991)

Seien $X_1, X_2, ..$ reelle, iid Zufallsgrößen mit $\mathbb{E}X_1 = 0$ und $\mathbb{V}X_1 < \infty$. Dann gilt:

$$\mathcal{P}\Big(\overline{\lim_{n\to\infty}} \frac{S_n}{\sqrt{2n \ln\ln n}} = \mathbb{V}X_1\Big) = 1$$

$$\mathcal{P}\Big(\underline{\lim_{n\to\infty}} \frac{S_n}{\sqrt{2n \ln\ln n}} = -\mathbb{V}X_1\Big) = 1$$

(ohne Beweis).

3.9.4 Theorem: Arcussinussatz

Seien $X_1, X_2, .. : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}$ reelle, iid Zufallsgrößen mit $\mathbb{E}X_1 = 0$ und $\mathbb{V}X_1 < \infty$. Zu $n \in \mathbb{N}$ bezeichne wie üblich

$$S_n := \sum_{k=1}^{n} X_k$$

und

$$r_n(\omega) := \frac{1}{n}\#\Big\{k \leq n : S_k(\omega) > 0\Big\} \quad , \quad \omega \in \Omega$$

die *relative Häufigkeit* zu der die ersten n Partialsummen S_k positiv sind. Dann gilt

$$\mathcal{P}\left(r_n \leq t\right) \overset{n\to\infty}{\longrightarrow} 1 - \frac{2}{\pi}\arcsin(\sqrt{t}) = \int_0^t \underbrace{\frac{dx}{\pi\sqrt{x(1-x)}}}_{=:\rho(x)}$$

Bemerkung: *Asymptotisch* gesehen, ist die relative Häufigkeit der Vorzeichen zwar symmetrisch um 50% verteilt, doch ist sie *Wahrscheinlicher* (höhere Dichte) an den Extremwerten. So gilt z.B.

$$\lim_{n\to\infty} \mathcal{P}\left(\tfrac{1}{8} \leq r_n \leq \tfrac{3}{8}\right) > \lim_{n\to\infty} \mathcal{P}\left(\tfrac{3}{8} \leq r_n \leq \tfrac{5}{8}\right)$$

(vgl. Abbildung (12)).

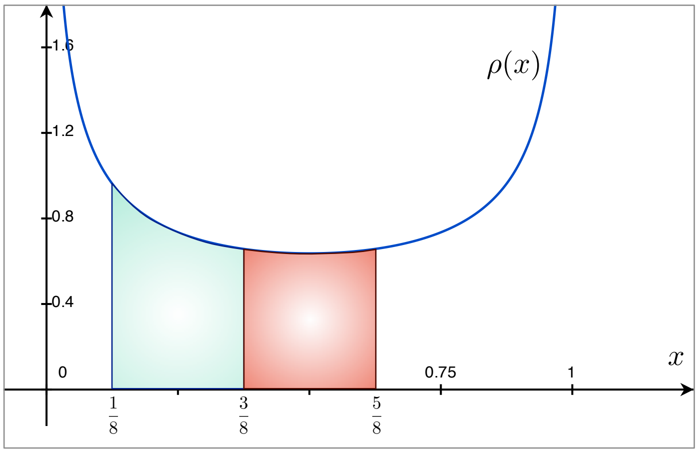

Abbildung 12: Zum Arcuscosinussatz: Die relative Häufigkeit der Partialsummen-Vorzeichen besitzt an den Extremwerten um 0 bzw. um 1 die höchste Dichte.

4 Charakteristische Funktion

4.1 $\mathbb{C}$-messbare Funktionen

4.1.1 Begriffe

Sei $(\Omega, \mathfrak{S}, \mu)$ ein Maßraum und $f : \Omega \to \mathbb{C}$ eine komplexwertige Funktion dargestellt durch

$$f(\omega) = \underbrace{f_{\text{re}}(\omega)}_{\text{reell}} + i \cdot \underbrace{f_{\text{im}}(\omega)}_{reell} \quad , \quad \omega \in \Omega$$

Dann ist $f : \mathfrak{S} \to \mathscr{B}(\mathbb{C})$-messbar genau dann wenn die reellwertigen $f_{\text{re}}, f_{\text{im}} : \mathfrak{S} \to \mathscr{B}(\mathbb{R})$ messbar sind.

Eine $\mathbb{C}$-wertige Funktion $f : \Omega \to \mathbb{C}$ heißt *Treppenfunktion*, falls sie die Gestalt

$$f = \sum_{j=1}^{n} z_j \cdot 1_{A_j} \quad , \quad z_1, .., z_n \in \mathbb{C},\ A_1, .., A_n \in \mathfrak{S} \tag{4.1.1.1}$$

hat.

Eine $\mathbb{C}$-wertige, messbare Funktion $f : \Omega \to \mathbb{C}$ heißt *μ-integrabel*, falls

$$\int\limits_{\Omega} |f| \ d\mu < \infty$$

bzw. äquivalent dazu

$$\int\limits_{\Omega} |f_{\text{re}}| \ d\mu < \infty \quad \wedge \quad \int\limits_{\Omega} |f_{\text{im}}| \ d\mu < \infty$$

Gegebenfalls setzt man dann

$$\int\limits_{\Omega} f \ d\mu := \int\limits_{\Omega} f_{\text{re}} \ d\mu + i \cdot \int\limits_{\Omega} f_{\text{im}} \ d\mu$$

Im Spezialfall einer Treppenfunktion (vgl. (4.1.1.1)) ist

$$\int\limits_{\Omega} f \ d\mu = \sum_{j=1}^{n} z_j \cdot \mu(A_j)$$

4.1.2 Eigenschaften des Integrals komplexer Funktionen

Seien $(\Omega, \mathfrak{S}, \mu)$ ein Maßraum und $f, g : \Omega \to \mathbb{C}$ messbar.

1. Für beliebige $c_1, c_2 \in \mathbb{C}$ gilt

$$\int\limits_{\Omega} (c_1 f + c_2 g) \ d\mu = c_1 \cdot \int\limits_{\Omega} f \ d\mu + c_2 \cdot \int\limits_{\Omega} g \ d\mu$$

2. Komplexe Konjugation & Integration sind vertauschbar:

$$\left[\int\limits_{\Omega} f \ d\mu \right]^{*} = \int\limits_{\Omega} f^{*} \ d\mu$$

3. Es lässt sich abschätzen

$$\left| \int\limits_{\Omega} f \ d\mu \right| \leq \int\limits_{\Omega} |f| \ d\mu$$

4. Der Raum $L_1(\Omega,\mu)$ aller integrablen Funktionen[12] $f:\Omega\to\mathbb{C}$ wird mit
$$\|f\|_{L_1} := \int_\Omega |f| \ d\mu$$
zu einem komplexen Banachraum.

5. Setzt man für $f,g\in L_2(\Omega,\mu)$
$$\langle f,g\rangle := \int_\Omega f^*\cdot g \ d\mu$$
so wird $(L_2(\Omega,\mu),\ \langle\cdot,\cdot\rangle)$ zu einem Hilbertraum.

4.2 Charakteristische Funktion

4.2.1 Definition: Charakteristische Funktion

Sei $(E,\|\cdot\|)$ ein reeller, normierter Raum, dazu der Dualraum E' aller stetigen Linearformen auf E und $U\in\mathscr{B}(E)$. Zu einem Wahrscheinlichkeitsmaß $\mathcal{P}$ auf $(U,\mathscr{B}(U))$ heißt die Abbildung $\mathcal{F}(\mathcal{P}):E'\to\mathbb{C}$, definiert durch
$$\mathcal{F}(\mathcal{P})(k) := \int_U e^{i\cdot k(x)} \ d\mathcal{P}(x) \quad , \quad k\in E'$$
charakteristische Funktion von $\mathcal{P}$.
Zu U-wertigen Zufallsvariablen $X:(\Omega,\mathfrak{S},\mathcal{Q})\to U\subseteq E$ heißt
$$\mathcal{F}(X) := \mathcal{F}(\mathcal{Q}_X) \quad , \quad \mathcal{F}(X)(k) = \int_U e^{i\cdot k(x)} \ d\mathcal{Q}_X(x) = \mathbb{E}e^{i\cdot k(X)} \quad , \quad k\in E'$$
charakteristische Funktion von X.

Bemerkungen:

(i) Wegen $\left|e^{i\cdot k(x)}\right|\leq 1$ ist $\mathcal{F}(\mathcal{P})$ tatsächlich wohldefiniert, komplex.

(ii) Man schreibt auch $\mathcal{F}(\mathcal{P})=\widehat{\mathcal{P}}$ und $\mathcal{F}(X)=\widehat{X}$.

(iii) Ist $\mathcal{P}_0$ Wahrscheinlichkeitsmaß auf $(U,\mathscr{B}(U))$, so ist $\mathcal{P}$, definiert durch
$$\mathcal{P}(A) := \mathcal{P}_0(A\cap U) \quad , \quad A\in\mathscr{B}(E)$$
Wahrscheinlichkeitsmaß auf $(E,\mathscr{B}(E))$ und es gilt $\mathcal{F}(\mathcal{P}_0)=\mathcal{F}(\mathcal{P})$. Insbesondere gilt für jede U-wertige Zufallsvariable X
$$\mathcal{F}(X) = \mathcal{F}(\underbrace{I_U\circ X}_{\substack{(E,\mathscr{B}(E))\\ \text{-wertig}}})$$
wobei
$$I_U:U\to E \quad , \quad I_U(x)=x,\ x\in U$$
die Inklusion von U in E sei. Werden also im folgenden o.B.d.A. $U=E$ annehmen.

(iv) Im reellen Fall $E=\mathbb{R}^n$ identifiziert man $E'=\mathbb{R}^n$ und
$$\mathbf{k}(\mathbf{x}) := \mathbf{k}\cdot\mathbf{x} := \langle\mathbf{k},\mathbf{x}\rangle := \sum_{j=1}^n k_j x^j$$
für $\mathbf{x},\mathbf{k}\in\mathbb{R}^n$.

(v) Der allgemeine Fall lässt sich durch
$$\mathcal{F}(X)(k) = \mathcal{F}(k(X))(1)$$
leicht auf den reellen Fall zurückführen.

[12] Genauer gesagt: Äquivalenzklassen, sich nur auf Nullmengen unterscheidenden Funktionen.

4.2.2 Permanenzeigenschaften der charakteristischen Funktion

Seien E, F reelle, normierte Räume, dazu die Dualräume E', F'.

1. Sind zwei E-wertige Zufallsvariablen X, Y identisch verteilt $X \stackrel{\text{d.}}{=} Y$, so sind auch $\mathcal{F}(X) = \mathcal{F}(Y)$.
2. Für diskretes Wahrscheinlichkeitsmaß
$$\mathcal{P} = \sum_{j=1}^{m} \lambda_j \cdot \delta_{x_j} \quad , \quad \lambda_j \geq 0, \ \sum_{j=1}^{m} \lambda_j = 1, \ x_j \in U$$
auf $(E, \mathscr{B}(E))$ ist
$$\mathcal{F}(\mathcal{P})(k) = \sum_{j=1}^{m} \lambda_j e^{i \cdot k(x_j)} \quad , \quad k \in E'$$
3. Besitzt $\mathcal{P}$ auf $(\mathbb{R}^n \mathscr{B}(\mathbb{R}^n), \lambda^n)$ die Dichte p, so ist
$$\mathcal{F}(\mathcal{P})(-\mathbf{k}) = \int\limits_{\mathbb{R}^n} e^{-i\mathbf{k}\cdot\mathbf{x}} p(\mathbf{x}) \ d\lambda^n(\mathbf{x}) =: \mathcal{F}(p)(\mathbf{k})$$
wobei $\mathcal{F}(p)$ genau die Fourier-Transformation von $p : \mathbb{R}^n \to \mathbb{R}$ ist.
4. Sind $\alpha \in \mathbb{R}, \ b \in E$ und X eine E-wertige Zufallsgröße, so gilt
$$\mathcal{F}(\alpha X + b)(k) = e^{i \cdot k(b)} \cdot \mathcal{F}(X)(\alpha k) \quad , \quad k \in E'$$
5. Allgemeiner: Ist $A : E \to F$ linear, stetig und $b \in F$, so gilt für U-wertige Zufallsgröße X
$$\mathcal{F}(\underbrace{AX + b}_{F-\text{wertig}})(k) = e^{i \cdot k(b)} \cdot \mathcal{F}(X)(A'k) \quad , \quad k \in F' \tag{4.2.2.1}$$
wobei $A' : F' \to E'$ der zu A duale Operator ist.
Bemerke: Im Fall $E = \mathbb{R}^n, F = \mathbb{R}^m$ und der Identifizierung $E' = \mathbb{R}^n, \ F' = \mathbb{R}^m$ nimm Gl. (4.2.2.1) die Form
$$\mathcal{F}(AX + \mathbf{b})(\mathbf{k}) = e^{i\mathbf{k}\cdot\mathbf{b}} \cdot \mathcal{F}(X)(A^{\dagger}\mathbf{k}) \quad , \quad \mathbf{k} \in \mathbb{R}^m$$
mit dem zu A adjungierten Operator $A^{\dagger} : \mathbb{R}^m \to \mathbb{R}^n$ an.
6. Sind X, Y unabhängig, $\mathbb{R}^n$-wertig, so gilt
$$\mathcal{F}(X + Y) = \mathcal{F}(X) \cdot \mathcal{F}(Y)$$
Analog gilt für Wahrscheinlichkeitsmaße $\mathcal{P}, \mathcal{Q}$ auf $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$:
$$\mathcal{F}(\mathcal{P} * \mathcal{Q}) = \mathcal{F}(\mathcal{P}) \cdot \mathcal{F}(\mathcal{Q})$$
7. Sind $X_1, X_2, .., X_n$ unabhängig, reellwertig und $\mathbf{X} := (X_1, .., X_n)$ so ist
$$\mathcal{F}(\mathbf{X})(\mathbf{k}) = \prod_{j=1}^{n} \mathcal{F}(X_j)(k_j) \quad , \quad \mathbf{k} \in \mathbb{R}^n$$

Beweis:

1. Klar.
2. Klar.
3. Klar.
4. Spezialfall von (5).

5.
$$\mathcal{F}(AX+b)(k) = \mathbb{E}e^{i\cdot k(AX+b)} = e^{i\cdot k(b)} \cdot \mathbb{E}\underbrace{e^{i\cdot k(AX)}}_{e^{i\cdot(A'k)X}} = e^{i\cdot k(b)} \cdot \mathcal{F}(X)(A'k) \quad , \quad k \in F'$$

6. Da X, Y unabhängig sind, sind auch $e^{i\mathbf{k}\cdot X}$ und $e^{i\mathbf{k}\cdot Y}$ unabhängig für jedes $\mathbf{k} \in \mathbb{R}^n$. Daher
$$\mathcal{F}(X+Y)(\mathbf{k}) = \mathbb{E}\underbrace{e^{i\mathbf{k}\cdot(X+Y)}}_{e^{i\mathbf{k}\cdot X}e^{i\mathbf{k}\cdot Y}} = \left(\mathbb{E}e^{i\mathbf{k}\cdot X}\right) \cdot \left(\mathbb{E}e^{i\mathbf{k}\cdot Y}\right) \quad , \quad \mathbf{k} \in \mathbb{R}^n$$

7.
$$\mathcal{F}(\mathbf{X})(\mathbf{k}) = \mathbb{E}\exp\left[i\sum_{j=1}^n k_j X_j\right] = \mathbb{E}\prod_{j=1}^n \underbrace{e^{ik_jX_j}}_{\text{unabhängig}} = \prod_{j=1}^n \mathbb{E}e^{ik_jX_j} = \prod_{j=1}^n \mathcal{F}(X_j)(k_j)$$

□

Beispiel: Sind alle $X_1, .., X_m$ iid, reell, so ist
$$\mathcal{F}\left[\sum_{j=1}^m X_j\right] = [\mathcal{F}(X_1)]^m$$

4.2.3 Satz: Eigenschaften der charakteristischen Funktion

Sei $\mathcal{P}$ ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$, dazu die charakteristische Funktion $\widehat{\mathcal{P}}$. Dann gilt:

1. $\widehat{\mathcal{P}}(0) = 1$
2. $\widehat{\mathcal{P}} : \mathbb{R}^n \to \mathbb{C}$ ist gleichmäßig stetig.
3. $\widehat{\mathcal{P}}(-\mathbf{k}) = \widehat{\mathcal{P}}^*(\mathbf{k}), \quad \mathbf{k} \in \mathbb{R}^n$.
4. $\widehat{\mathcal{P}}$ ist nicht-negativ definit, sprich die Matrix
$$\left(\widehat{\mathcal{P}}(\mathbf{k}_k - \mathbf{k}_l)\right)_{k,l=1}^m \in \mathbb{C}^{m\times m}$$
ist für beliebige $\mathbf{k}_1, .., \mathbf{k}_m \in \mathbb{R}^n$ nicht-negativ-definit[13].
5. $\left|\widehat{\mathcal{P}}\right| \leq 1$.
6. Ist $\mathcal{P}$ symmetrisch, sprich $\mathcal{P}(B) = \mathcal{P}(-B)$ für beliebige Borelmenge $B \in \mathscr{B}(\mathbb{R}^n)$, so ist $\widehat{\mathcal{P}} : \mathbb{R}^n \to \mathbb{R}$ reellwertig.

Beweis: Nach 2.7.2 können wir o.B.d.A. $\mathcal{P} = \mathcal{Q}_X$ für irgendeine Zufallsvariable $X : (\Omega, \mathfrak{S}, \mathcal{Q}) \to \mathbb{R}^n$ annehmen. Zu dieser sei $\widehat{X}$ die charakteristische Funktion.

1. Klar.
2. Zu zeigen wäre $\left\|\widehat{X}(\cdot+\mathbf{h}) - \widehat{X}(\cdot)\right\|_\infty \overset{\mathbf{h}\to 0}{\longrightarrow} 0$. Tatsächlich kann man wegen
$$\left|\widehat{X}(\mathbf{k}+\mathbf{h}) - \widehat{X}(\mathbf{k})\right| = \left|\mathbb{E}e^{i\mathbf{k}X}\left(e^{i\mathbf{h}X} - 1\right)\right| \leq \mathbb{E}\left|e^{i\mathbf{h}X} - 1\right| \quad \forall\, \mathbf{k} \in \mathbb{R}^n$$
abschätzen
$$\left\|\widehat{X}(\cdot+\mathbf{h}) - \widehat{X}(\cdot)\right\|_\infty \leq \mathbb{E}\left|e^{i\mathbf{h}X} - 1\right|$$
Wegen $\left|e^{i\mathbf{h}X} - 1\right| \leq 2$ (integrable Majorante) geht nach Lebesgue
$$\mathbb{E}\left|e^{i\mathbf{h}X} - 1\right| \overset{\mathbf{h}\to 0}{\longrightarrow} \mathbb{E}\underbrace{\lim_{\mathbf{h}\to 0}\left|e^{i\mathbf{h}X} - 1\right|}_{\substack{0 \\ \text{(punktweise)}}} = 0$$
woraus schließlich die Behauptung folgt.

[13] Beachte dass nach (3) diese stets hermitesch ist.

3.
$$\widehat{X}(-\mathbf{k}) = \mathbb{E}e^{-i\mathbf{k}X} = \left[Ee^{i\mathbf{k}X}\right]^* = \widehat{X}^*(\mathbf{k})$$

4. Zu $\mathbf{k}_1,..,\mathbf{k}_m \in \mathbb{R}^n$ und $c_1,..,c_m \in \mathbb{C}$ ist
$$\sum_{l,k=1}^{m} c_k c_l^* \cdot \widehat{X}(\mathbf{k}_k - \mathbf{k}_l) = \sum_{k,l=1}^{m} c_k c_l^* \mathbb{E}e^{i(\mathbf{k}_k-\mathbf{k}_l)X} = \mathbb{E}\sum_{k,l=1}^{m} c_k c_l^* \cdot e^{i(\mathbf{k}_k-\mathbf{k}_l)X}$$
$$= \mathbb{E}\left\{\left[\sum_{k=1}^{m} c_k e^{i\mathbf{k}_k X}\right] \cdot \left[\sum_{l=1}^{m} c_l e^{i\mathbf{k}_l X}\right]^*\right\} = \mathbb{E}\left|\sum_{k=1}^{m} c_k e^{i\mathbf{k}_k X}\right|^2 \geq 0$$

5.
$$\left|\widehat{X}(\mathbf{k})\right| = \left|\mathbb{E}e^{i\mathbf{k}X}\right| \leq \mathbb{E}\left|e^{i\mathbf{k}X}\right| \leq 1$$

6. Ist $X \overset{\text{d.}}{=} -X$ so ist
$$\widehat{X}(\mathbf{k}) = \mathbb{E}e^{i\mathbf{k}X} \overset{X\overset{\text{d.}}{=}-X}{=} \mathbb{E}e^{-i\mathbf{k}X} = \widehat{X}(-\mathbf{k}) = \widehat{X}^*(\mathbf{k})$$
sprich $\widehat{X}(\mathbf{k}) \in \mathbb{R}$.

□

4.2.4 Theorem über charakteristische Funktionen (Bochner)

Eine Abbildung $\varphi : \mathbb{R}^n \to \mathbb{C}$ ist charakteristische Funktion eines Wahrscheinlichkeitsmaßes auf $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$ genau dann wenn sie erfüllt:

1. $\varphi(0) = 1$
2. φ ist stetig
3. φ ist nicht-negativ-definit, sprich für beliebige $\mathbf{k}_1,..,\mathbf{k}_m \in \mathbb{R}^n$ ist die Matrix
$$\Big(\varphi(\mathbf{k}_k - \mathbf{k}_l)\Big)_{k,l=1}^{m} \in \mathbb{C}^{m\times m}$$
nicht-negativ-definit.

(Ohne Beweis)

4.2.5 Satz: Ableitungen der charakteristischen Funktion

Sei $X : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}$ eine reelle Zufallsvariable mit $\mathbb{E}|X|^n < \infty$. Dann ist deren charakteristische Funktion $\widehat{X}$ n-mal stetig differenzierbar und speziell
$$\widehat{X}^{(k)}(0) = i^k \cdot \mathbb{E}X^k \quad , \quad k = 0,..,n$$

Beweis: (nur für $n = 1$) Setze $f(t,\omega) := e^{itX(\omega)}$, dann ist $\partial_t f(t,\omega) = iX(\omega)e^{itX(\omega)}$ und
$$|\partial_t f(t,\omega)| \leq |X(\omega)| =: h(\omega)$$
Durch Anwendung des Differenzierungslemmas A.3.1 auf f erhält man
$$\frac{\partial}{\partial t}\widehat{X}(t) = \frac{\partial}{\partial t}\mathbb{E}f(t,\cdot) \overset{A.3.1}{=} \mathbb{E}\frac{\partial}{\partial t}f(t,\cdot) = i\cdot\mathbb{E}Xe^{itX} \tag{$\clubsuit$}$$
und insbesondere $\widehat{X}'(0) = \mathbb{E}X$. Ferner kann man abschätzen
$$\left|\widehat{X}'(t+h) - \widehat{X}'(t)\right| \overset{(\clubsuit)}{\leq} \mathbb{E}\left\{|X|\cdot\left|e^{i(t+h)X} - e^{itX}\right|\right\} = \mathbb{E}\left\{|X|\cdot\left|e^{ihX} - 1\right|\right\}$$

Aus $|X|\cdot\left|e^{ihX}-1\right| \leq 2\,|X|$ (integrable Majorante) und $|X|\cdot\left|e^{ihX}-1\right| \underset{\text{punktweise}}{\xrightarrow{h\to 0}} 0$ folgt schließlich nach Lebesgue

$$\lim_{h\to 0}\left|\widehat{X}(t+h)-\widehat{X}'(t)\right| = 0 \quad \forall\, t$$

sprich $\widehat{X}'$ ist stetig.
□

4.2.6 Folgerung: Entwicklung von charakteristischen Funktionen

Sei X eine reelle Zufallsvariablen mit $\mathbb{E}\,|X|^n < \infty$ und charakteristischer Funktion $\widehat{X}$. Dann gilt

$$\widehat{X}(t) = \sum_{k=0}^{n} \mathbb{E}X^k \cdot \frac{(it)^k}{k!} + \frac{t^n}{n!}\cdot\Theta(t)$$

für geeignetes Restglied $\Theta:\mathbb{R}\to\mathbb{C}$ mit

$$\lim_{t\to 0}\Theta(t) = 0$$

Beweis: Folgt aus vorigem Satz 4.2.5 und Taylor A.5.4.

Beispiele charakteristischer Funktionen

(i) Nimmt die $\mathbb{R}$-wertige Zufallsvariable X nur Werte in $\mathbb{Z}$ an, so ist $\mathcal{F}(X)$ 2π-periodisch, denn

$$\mathcal{F}(X)(t+2\pi) = \mathbb{E}e^{itX}\underbrace{e^{i2\pi X}}_{1} = \mathcal{F}(X)(t)$$

(ii) Für $\mathbf{c}\in\mathbb{R}^n$ ist $\mathcal{F}(\delta_{\mathbf{c}})(\mathbf{k}) = e^{i\mathbf{k}\cdot\mathbf{c}},\ \mathbf{k}\in\mathbb{R}^n$. Insbesondere $\mathcal{F}(\delta_0)\equiv 1$.

(iii) Für $\mathcal{P}(X=\pm 1) = 1/2$ ist

$$\mathcal{F}(X)(t) = \frac{1}{2}\left(e^{it}+e^{-it}\right) = \cos(t)$$

(iv) Die Binomialverteilung

$$\mathcal{B}_{n,p} = \sum_{k=1}^{n}\binom{n}{k}\cdot p^k(1-p)^{n-k}\delta_k$$

besitzt charakteristische Funktion

$$\mathcal{F}(\mathcal{B}_{n,p})(t) = \sum_{k=0}^{n}\binom{n}{k}\cdot p^k(1-p)^{n-k}e^{itk} \quad , \quad t\in\mathbb{R}$$

Alternativ kann man sich diesbezüglich überzeugen, wenn man beachtet dass

$$\mathcal{B}_{n,p} = \sum_{i=1}^{n} X_i \quad , \quad \mathcal{F}(X_1)(t) = (1-p)+pe^{it}$$

für n unabhängige, identisch verteilte $X_i\sim\mathcal{B}_{1,p}$ und daher

$$\mathcal{F}(\mathcal{B}_{n,p})(t) = \left[\mathcal{F}(X_1)(t)\right]^n = \left[(1-p)+pe^{it}\right]^n \quad , \quad t\in\mathbb{R}$$

(v) Für poissonverteilte $X\sim\pi_\lambda$ ist

$$\mathcal{F}(X)(t) = \sum_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}e^{itk} = e^{-\lambda}\exp\left[\lambda e^{it}\right] = \exp\left[-\lambda\left(1-e^{it}\right)\right] = \exp\left[-\lambda\left(1-\mathcal{F}(\delta_1)(t)\right)\right]$$

Bemerke den Zusammenhang zwischen π_λ und δ_1.

(vi) Ist X gleichverteilt auf $[a,b]$, so ist

$$\mathcal{F}(X)(t) = \frac{1}{b-a}\int_a^b e^{itx}\,dx = \frac{e^{itb}-e^{ita}}{it(b-a)} \quad , \quad t \in \mathbb{R}$$

Speziell für $a=-1,\ b=1$

$$\mathcal{F}(X)(t) = \frac{\sin t}{t} \quad , \quad t \in \mathbb{R} \tag{4.2.6.1}$$

(vii) Für standardnormalverteilte $X \sim \mathcal{N}_{0,1}$ ist

$$\mathcal{F}(X)(t) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{itx-x^2/2}dx \overset{\substack{\text{Cauchyscher}\\ \text{Integralsatz}}}{=} e^{-t^2/2} \tag{4.2.6.2}$$

Ist allgemeiner $X \sim \mathcal{N}_{\mu,\sigma^2}$, sprich $X = \sigma X_0 + a$ für irgendeine $X_0 \sim \mathcal{N}_{0,1}$, so folgt direkt

$$\mathcal{F}(X)(t) = e^{ita}\mathcal{F}(X_0)(\sigma t) = \exp\left[ita - \frac{\sigma^2 t^2}{2}\right] \tag{4.2.6.3}$$

(viii) Sei $X \sim \Gamma_b$, sprich mit Verteilungsdichte

$$\frac{d\mathcal{P}_X}{d\lambda^1}(x) = \frac{1}{\Gamma(b)}x^{b-1}e^{-x}$$

Für $|t|<1$ ist

$$\mathcal{F}(X)(t) = \frac{1}{\Gamma(b)}\int_0^\infty e^{itx}x^{b-1}e^{-x}dx = \frac{1}{\Gamma(b)}\int_0^\infty \sum_{k=0}^\infty \frac{(it)^k}{k!}x^{b+k-1}e^{-x}dx$$

$$\overset{(\clubsuit)}{=} \frac{1}{\Gamma(b)}\sum_{k=0}^\infty \frac{(it)^k}{k!}\underbrace{\int_0^\infty x^{b+k-1}e^{-x}dx}_{\substack{\Gamma(b+k)\\ =(b+k-1)\cdot\ldots\cdot b\cdot\Gamma(b)}} = \sum_{k=0}^\infty (it)^k(-1)^k \underbrace{\frac{(-b)(-b-1)\ldots(-b-k+1)}{k!}}_{\binom{-b}{k}}$$

$$= \sum_{k=0}^\infty \binom{-b}{k}(-it)^k = \frac{1}{(1-it)^b} \quad , \quad |t|<1$$

wobei verwendet wurde dass

$$(\clubsuit)\ :\ \frac{1}{\Gamma(b)}\sum_{k=0}^\infty\int_0^\infty \frac{|t|^k}{k!}x^{b+k-1}e^{-x}dx \overset{\text{analog}}{=} \sum_{k=0}^\infty\binom{-b}{k}(-|t|)^k = \frac{1}{(1-|t|)^b} < \infty$$

für $|t|<1$, nach Fubini also Integration & Summierung vertauscht werden können.

(ix) Ist allgemeiner $X \sim \Gamma_{a,b}$ (vgl. (1.2.8.2)), sprich $X = aX_0$ für irgendeine $X_0 \sim \Gamma_b$, so folgt

$$\mathcal{F}(X)(t) = \mathcal{F}(X_0)(at) = \frac{1}{(1-iat)^b} \quad , \quad |at|<1$$

(x) Die Exponentialverteilung $\text{Exp}_\lambda = \Gamma_{\frac{1}{\lambda},1}$ besitzt daher charakteristische Funktion

$$\mathcal{F}(\text{Exp}_\lambda)(t) = \frac{1}{1-\frac{it}{\lambda}} = \frac{\lambda}{\lambda - it}$$

für $|t/\lambda|<1$.

(xi) Die $\chi_n^2 = \Gamma_{2,n/2}$ Verteilung besitzt charakteristische Funktion

$$\mathcal{F}(\chi_n^2)(t) = \frac{1}{(1-2it)^{n/2}}$$

für $|2t|<1$.

Mehrdimensionale Beispiele:

(i) Sind $X_1, .., X_n \sim \mathcal{N}_{0,1}$ unabhängig, reellwertig, so heißt

$$\mathbf{X} := (X_1, .., X_n)$$

standardnormalverteilter n-dimensionaler Zufallsvektor. Nach 4.2.2(7) besitzt dieser die charakteristische Funktion

$$\mathcal{F}(\mathbf{X})(\mathbf{k}) \overset{4.2.2(7)}{=} \prod_{j=1}^{n} \mathcal{F}(X_j)(k_j) \overset{(4.2.6.2)}{=} \prod_{j=1}^{n} e^{-k_j^2/2} = \exp\left[-\|\mathbf{k}\|^2/2\right]$$

Ist nun $A : \mathbb{R}^n \to \mathbb{R}^m$ linear so ist

$$\mathcal{F}(AX)(\mathbf{k}) \overset{4.2.2(5)}{=} \mathcal{F}(X)(A^\dagger \mathbf{k}) = \exp\left[-\left\|A^\dagger \mathbf{k}\right\|^2/2\right] = \exp\left[-\langle R\mathbf{k}, \mathbf{k}\rangle/2\right]$$

wobei $R = AA^\dagger$. Anderseits gilt

$$\mathbb{E}\left[(AX)_k (AX)_l\right] = \mathbb{E}\left(\sum_{j=1}^{n} A_{kj} X_j\right)\left(\sum_{r=1}^{n} A_{lr} X_r\right) = \sum_{j,r=1}^{n} A_{kj} A_{lr} \underbrace{\mathbb{E} X_j X_r}_{\delta_{jr}} = \sum_{j=1}^{n} A_{kj} A_{lj} = (AA^\dagger)_{kl}$$

sprich $R = AA^\dagger$ ist genau die Kovarianzmatrix von AX.

Ferner kann für $n = m$ und positiv definite $R = AA^\dagger$ gezeigt werden, dass AX die Verteilungsdichte

$$\frac{d\mathcal{P}_{AX}}{d\lambda^n}(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{|\det R|}} e^{-\langle R^{-1}\mathbf{x}, \mathbf{x}\rangle/2} \quad , \quad \mathbf{x} \in \mathbb{R}^n$$

besitzt (vgl. eindimensionalen Fall (4.2.6.3) mit $R = \mathbb{V}\mathcal{N}_{0,\sigma^2} = \sigma^2$).

(ii) Ist $\mathbf{X}$ gleichverteilt auf $[-1,1]^n$, sprich $\mathbf{X} = (X_1, .., X_n)$ für auf $[-1,1]$ gleichverteilte, unabhängige $X_1, .., X_n$, so ist

$$\mathcal{F}(\mathbf{X})(\mathbf{k}) \overset{4.2.2(7)}{=} \prod_{j=1}^{n} \mathcal{F}(X_j)(k_j) \overset{(4.2.6.1)}{=} \prod_{j=1}^{n} \frac{\sin k_j}{k_j}$$

4.3 Umkehrformel & Eindeutigkeitssatz

4.3.1 Lemma über die sinc-Funktion

Sei $\alpha \in \mathbb{R} \setminus \{0\}$. Dann gilt:

1. Für $y > 0$ ist

$$0 \le \operatorname{sgn}(\alpha) \cdot \int_0^y \frac{\sin(\alpha x)}{x}\, dx \le \int_0^\pi \frac{\sin x}{x}\, dx$$

2.

$$\int_0^\infty \frac{\sin(\alpha x)}{x}\, dx = \operatorname{sgn}(\alpha) \cdot \frac{\pi}{2}$$

3.

$$\int_0^\infty \frac{1 - \cos(\alpha x)}{x^2}\, dx = |\alpha| \cdot \frac{\pi}{2}$$

Beweis: Durch die Substitution $u := \alpha x$ sind alle Aussagen auf den Fall $\alpha = 1$ zurückzuführen, nehmen also o.B.d.A. $\alpha = 1$ an.

1. Aus dem Verlauf der sinc-Funktion (vgl. Abb. (13)) wird ersichtlich

$$\int\limits_{(2k-1)\pi}^{(2k+1)\pi} \frac{\sin x}{x}\, dx \leq 0 \quad , \quad k \in \mathbb{N}$$

   Ist nun $(2k-1)\pi \leq y \leq 2k\pi$ für geeignetes $k \in \mathbb{N}$, so folgt

$$\int\limits_0^y \frac{\sin x}{x}\, dx \leq \int\limits_0^{(2k-1)\pi} \frac{\sin x}{x}\, dx = \int\limits_0^{\pi} \frac{\sin x}{x}\, dx + \underbrace{\int\limits_{\pi}^{3\pi} \frac{\sin x}{x}\, dx}_{\leq 0} + \dots + \underbrace{\int\limits_{(2k-3)\pi}^{(2k-1)\pi} \frac{\sin x}{x}\, dx}_{\leq 0} \leq \int\limits_0^{\pi} \frac{\sin x}{x}\, dx$$

   Ist anderseits $2k\pi \leq y \leq (2k+1)\pi$ so kann man ebenfalls abschätzen

$$\int\limits_0^y \frac{\sin x}{x}\, dx \leq \int\limits_0^{(2k+1)\pi} \frac{\sin x}{x}\, dx = \int\limits_0^{\pi} \frac{\sin x}{x}\, dx + \underbrace{\int\limits_{\pi}^{3\pi} \frac{\sin x}{x}\, dx}_{\leq 0} + \dots + \underbrace{\int\limits_{(2k-1)\pi}^{(2k+1)\pi} \frac{\sin x}{x}\, dx}_{\leq 0} \leq \int\limits_0^{\pi} \frac{\sin x}{x}\, dx$$

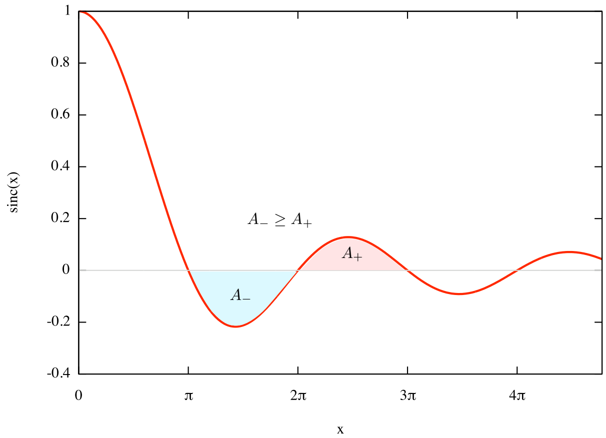

Abbildung 13: Verlauf der sinc-Funktion auf der positiven Achse.

   Analog lässt sich auch die andere Ungleichung zeigen.

2. Schreiben

$$\int\limits_0^{\infty} \frac{\sin x}{x}\, dx = \lim_{R\to\infty} \int\limits_0^R \underbrace{\frac{\sin x}{x}}_{\int_0^\infty e^{-tx} \sin x \; dt} dx \overset{\text{Fubini}}{=} \lim_{R\to\infty} \int\limits_0^{\infty} \int\limits_0^R e^{-tx} \sin x \; dx \; dt$$

$$= \lim_{R\to\infty} \int\limits_0^{\infty} \underbrace{\left[-\frac{(t \sin R + \cos R)}{1+t^2} e^{-tR} + \frac{1}{1+t^2} \right]}_{\xrightarrow[\text{punktweise}]{R\to\infty} \frac{1}{1+t^2}} dt \overset{\text{Lebesgue}}{=} \int\limits_0^{\infty} \frac{dt}{1+t^2} = \frac{\pi}{2}$$

wobei verwendet wurde dass

$$\left|-\frac{(t\sin R+\cos R)}{1+t^2}e^{-tR}+\frac{1}{1+t^2}\right| \leq \frac{(t+1)}{1+t^2}e^{-tR}+\frac{1}{1+t^2} \overset{(\clubsuit)}{\leq} \underbrace{\frac{3}{1+t^2}}_{\substack{\text{integrable}\\ \text{Majorante}}}$$

für genügend großes R ($\clubsuit$).

3.

$$\int_0^\infty \frac{1-\cos x}{x^2}\ dx = \underbrace{\left.\frac{1-\cos x}{-x}\right|_0^\infty}_{0} + \int_0^\infty \frac{\sin x}{x}\ dx \overset{(3)}{=} \frac{\pi}{2}$$

□

4.3.2 Definition: Cauchyscher Hauptwert

Sei $f : \mathbb{R} \to \mathbb{R}$ messbar und das Integral

$$\int_{-T}^{T} f(x)\ dx$$

existiere für jedes $T \geq 0$. Dann heißt

$$\lim_{T\to\infty} \int_{-T}^{T} f(x)\ dx$$

Cauchyscher Hauptwert des Integrals $\int_{-\infty}^{\infty} f(x)\ dx$.

Bemerke: Im Falle

$$\int_{-\infty}^{\infty} |f(x)|\ dx < \infty$$

gilt nach Lebesgue

$$\lim_{T\to\infty} \int_{-T}^{T} f(x)\ dx = \int_{-\infty}^{\infty} f(x)\ dx$$

Allgemein impliziert jedoch die Existenz des Cauchyschen Hauptwertes **nicht** die Integrabilität von f, wie etwa für $f(x) = 1/x$ der Fall ist.

4.3.3 Theorem: Umkehrformel für Maße auf $\mathbb{R}$

Sei $\mathcal{P}$ ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ mit charakteristischer Funktion $\widehat{\mathcal{P}}$. Dann gilt für beliebige $a < b \in \mathbb{R}$:

$$\mathcal{P}((a,b)) + \frac{1}{2}\big[\mathcal{P}(\{a\}) + \mathcal{P}(\{b\})\big] = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita}-e^{-itb}}{it}\ \widehat{\mathcal{P}}(t)\ dt$$

Bemerkungen:

(i) Der Integrant ist wohldefiniert. Insbesondere geht

$$\frac{e^{-ita}-e^{-itb}}{it} \overset{t\to 0}{\longrightarrow} b-a$$

(ii) Mit
$$\frac{e^{-ita}-e^{-itb}}{it} = \int\limits_a^b e^{-itx}\ dx$$
folgt die Darstellung
$$\lim_{T\to\infty}\frac{1}{2\pi}\int\limits_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it}\ \widehat{\mathcal{P}}(t)\ dt = \frac{1}{2\pi}\lim_{T\to\infty}\int\limits_{-T}^{T}\int\limits_a^b e^{-itx}\widehat{\mathcal{P}}(t)\ dx\ dt \overset{\text{Fubini}}{=} \frac{1}{2\pi}\lim_{T\to\infty}\int\limits_a^b\int\limits_{-T}^{T} e^{-itx}\widehat{\mathcal{P}}(t)\ dt\ dx$$
Beachte dass der Grenzwert $\lim\limits_{T\to\infty}$ im allgemeinen **nicht** mit dem Integral $\int\limits_a^b$ vertauschbar ist, z.B. im Fall $\mathcal{P}=\delta_0$.

Beweis: Wegen
$$\int\limits_{-\infty}^{\infty}\int\limits_{-T}^{T}\underbrace{\left|\frac{e^{it(x-a)}-e^{it(x-b)}}{it}\right|}_{\left|\int\limits_{x-b}^{x-a} e^{itu}\ du\right|\leq\int\limits_{x-b}^{x-a} du=b-a}\ dt\ d\mathcal{P}(x) \leq \int\limits_{-\infty}^{\infty} 2T(b-a)\ d\mathcal{P}(x) < \infty \tag{4.3.3.1}$$
und
$$\left|\int\limits_0^T \frac{\sin(t(x-a))-\sin(t(x-b))}{\pi t}\ dt\right| \overset{4.3.1(1)}{=} \frac{2}{\pi}\int\limits_0^{\pi}\frac{\sin s}{s}\ ds =: \underbrace{C:\mathrm{const}}_{\substack{\mathcal{P}\text{-integrable}\\ \text{Majorante}}} < \infty \tag{4.3.3.2}$$
gilt
$$\lim_{T\to\infty}\frac{1}{2\pi}\int\limits_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it}\ \widehat{\mathcal{P}}(t)\ dt = \lim_{T\to\infty}\int\limits_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{2\pi it}\int\limits_{-\infty}^{\infty} e^{itx}d\mathcal{P}(x)\ dt$$
$$\overset{\substack{\text{Fubini}\\(4.3.3.1)}}{=} \lim_{T\to\infty}\int\limits_{-\infty}^{\infty}\underbrace{\int\limits_{-T}^{T}\frac{e^{it(x-a)}-e^{it(x-b)}}{2\pi it}\ dt}_{\substack{\int\limits_{-T}^{T}\frac{\sin(t(x-a))-\sin(t(x-b))}{2\pi t}\\ \text{da }\cos(\text{const}\cdot t)/t\\ \text{ungerade in }t}}\ d\mathcal{P}(x)$$
$$= \lim_{T\to\infty}\int\limits_{-\infty}^{\infty}\underbrace{\int\limits_0^T \frac{\sin(t(x-a))-\sin(t(x-b))}{\pi t}\ dt}_{\substack{\xrightarrow[4.3.1(3)]{T\to\infty}\frac12\operatorname{sgn}(x-a)-\frac12\operatorname{sgn}(x-b)\\ =1_{(a,b)}(x)+\frac12\left[1_{\{a\}}(x)+1_{\{b\}}(x)\right]}}\ d\mathcal{P}(x)$$
$$\overset{\substack{\text{Lebesgue}\\(4.3.3.2)}}{=} \int\limits_{-\infty}^{\infty}\left[1_{(a,b)}+\frac12\left(1_{\{a\}}+1_{\{b\}}\right)\right]\ d\mathcal{P} = \mathcal{P}\left((a,b)\right)+\frac12\left[\mathcal{P}\left(\{a\}\right)+\mathcal{P}\left(\{b\}\right)\right]$$
□

4.3.4 Satz: Umkehrformel für Maße auf $\mathbb{R}^n$

Sei $\mathcal{P}$ ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$ mit charakteristischer Funktion und

$$Q := \bigtimes_{j=1}^{n} (a_j, b_j) \quad , \quad a_j < b_j \in \mathbb{R},\ j = 1, .., n$$

ein Quader in $\mathbb{R}^n$ mit $\mathcal{P}(\partial Q) = 0$. Dann gilt

$$\mathcal{P}(Q) = \lim_{T\to\infty} \frac{1}{(2\pi)^n} \int\limits_{[-T,T]^n} \prod_{j=1}^{n} \left[\frac{e^{-it_j a_j} - e^{-it_j b_j}}{it_j}\right] \cdot \widehat{\mathcal{P}}(\mathbf{t})\ d\mathbf{t}$$

Beweis: Fahren analog zum Beweis von 4.3.3 fort. Setze

$$I(T, x, a, b) := \int\limits_{-T}^{T} \frac{e^{it(x-a)} - e^{it(x-b)}}{2\pi i t}\ dt \quad , \quad T > 0,\ x \in \mathbb{R},\ a < b \in \mathbb{R}$$

Dann geht

$$I(T, x, a, b) = \int\limits_{0}^{T} \frac{\sin(t(x-a)) - \sin(t(x-b))}{\pi t}\ dt \underset{4.3.1(3)}{\xrightarrow{T\to\infty}} \frac{1}{2}\left[\operatorname{sgn}(x-a) - \operatorname{sgn}(x-b)\right] = 1_{(a,b)}(x) + \frac{1}{2}\left[1_{\{a\}}(x) + 1_{\{b\}}(x)\right]$$

wobei die Abschätzung

$$|I(T, x, a, b)| = \left|\int\limits_{0}^{T} \frac{\sin(t(x-a)) - \sin(t(x-b))}{\pi t}\ dt\right| \overset{4.3.1(1)}{\leq} \frac{2}{\pi}\int\limits_{0}^{T} \frac{\sin u}{u} = \text{const} < \infty$$

gilt. Damit geht auch

$$\prod_{j=1}^{n} I(T, x_j, a_j, b_j) \underset{\text{punktweise}}{\xrightarrow{T\to\infty}} \prod_{j=1}^{n} \left[1_{(a_j,b_j)}(x_j) + \frac{1}{2}\left(1_{\{a_j\}}(x_j) + 1_{\{b_j\}}(x_j)\right)\right] \overset{\mathcal{P}(\partial Q)=0}{=} 1_Q(\mathbf{x}) \qquad \mathcal{P}\text{-fast-sicher} \tag{4.3.4.1}$$

wobei auch

$$\left|\prod_{j=1}^{n} I(T, x_j, a_j, b_j)\right| \leq \text{const}^n < \infty$$

abgeschätzt werden kann. Können also schreiben

$$\frac{1}{(2\pi)^n} \int\limits_{[-T,T]^n} \prod_{j=1}^{n} \left[\frac{e^{-it_j a_j} - e^{-it_j b_j}}{2\pi i t_j}\right] \cdot \widehat{\mathcal{P}}(\mathbf{t})\ d\mathbf{t} = \int\limits_{[-T,T]^n} \prod_{j=1}^{n} \left[\frac{e^{-it_j a_j} - e^{-it_j b_j}}{2\pi i t_j}\right] \cdot \int\limits_{\mathbb{R}^n} e^{i\langle \mathbf{t}, \mathbf{x}\rangle} d\mathcal{P}(\mathbf{x})\ d\mathbf{t}$$

$$\underset{4.3.3}{\overset{\text{Fubini analog zu}}{=}} \int\limits_{\mathbb{R}^n} \underbrace{\prod_{j=1}^{n} \int\limits_{-T}^{T} \left[\frac{e^{it_j(x_j-a_j)} - e^{it_j(x_j-b_j)}}{2\pi i t_j}\right] dt_j}_{\substack{\xrightarrow[\mathcal{P}\text{-fast-sicher}]{T\to\infty} 1_Q(\mathbf{x}) \\ \text{nach } (4.3.4.1)}}\ d\mathcal{P}(\mathbf{x})$$

$$\underset{\text{Lebesgue}}{\xrightarrow{T\to\infty}} \int\limits_{\mathbb{R}^n} 1_Q\ d\mathcal{P} = \mathcal{P}(Q)$$

□

4.3.5 Korollar für Verteilungsfunktionen

Sei $\mathcal{P}$ ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ mit Verteilungsfunktion F und charakteristischer Funktion $\widehat{\mathcal{P}}$. Ist F stetig in $a < b \in \mathbb{R}$ so gilt

$$F(b) - F(a) = \frac{1}{2\pi} \lim_{T\to\infty} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \, \widehat{\mathcal{P}}(t) \; dt$$

Beweis: Direkte Folgerung aus Theorem 4.3.3.

4.3.6 Theorem: Charakterisierung von Maßen auf $\mathbb{R}^n$ durch charakteristische Funktionen

Seien $\mathcal{P}, \mathcal{Q}$ Wahrscheinlichkeitsmaße auf $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$ mit charakteristischen Funktionen $\widehat{\mathcal{P}}, \widehat{\mathcal{Q}}$. Dann gilt:

$$\mathcal{P} = \mathcal{Q} \quad \Leftrightarrow \quad \widehat{\mathcal{P}} = \widehat{\mathcal{Q}}$$

Beweis: Richtung "$\Rightarrow$" ist trivial. Für Richtung "$\Leftarrow$" genügt es zu zeigen

$$\mathcal{P}(Q) = \mathcal{Q}(Q)$$

für alle n-dimensionale Quader der Form

$$Q := \bigtimes_{j=1}^{n} (a_j, b_j) \quad , \quad a_j < b_j \in \mathbb{R}^n$$

Tatsächlich, ist $\mathcal{P}(\partial Q) = 0 = \mathcal{Q}(\partial Q)$, so folgt nach Theorem 4.3.4

$$\mathcal{P}(Q) = \lim_{T\to\infty} \frac{1}{(2\pi)^n} \int\limits_{[-T,T]^n} \prod_{j=1}^{n} \left[\frac{e^{-it_ja_j} - e^{-it_jb_j}}{it_j} \right] \cdot \underbrace{\widehat{\mathcal{P}}}_{\widehat{\mathcal{Q}}}(\mathbf{t}) \; d\mathbf{t} = \mathcal{Q}(Q)$$

Für beliebigen anderen Quader Q existiert nach A.3.4 eine Quaderfolge $Q^k \subseteq \mathbb{R}^n$ mit

$$Q^k \underset{k\to\infty}{\nearrow} Q \quad , \quad \mathcal{P}(\partial Q^k) = 0 = \mathcal{Q}(\partial Q^k)$$

so dass folgt

$$\mathcal{P}(Q) = \lim_{k\to\infty} \mathcal{P}(Q^k) = \lim_{k\to\infty} \mathcal{Q}(Q^k) = \mathcal{Q}(Q)$$

□

Folgerungen: Seien X, Y reelle Zufallsgrößen mit charakteristischen Funktionen $\widehat{X}, \widehat{Y}$. Dann gilt:

(i) $X \overset{\text{d.}}{=} Y \quad \Leftrightarrow \quad \widehat{X} = \widehat{Y}$

(ii) $X \overset{\text{d.}}{=} -X \quad \Leftrightarrow \quad \widehat{X} : \mathbb{R} \to \mathbb{R}$

(iii) Sind $X \sim \mathcal{N}_{\mu,\sigma^2}$, $Y \sim \mathcal{N}_{\nu,s^2}$ normalverteilt, unabhängig, so sind

$$\widehat{X}(t) = e^{it\mu} e^{-\frac{t^2}{2}\sigma^2} \quad , \quad \widehat{Y}(t) = e^{it\nu} e^{-\frac{t^2}{2}s^2} \quad , \quad t \in \mathbb{R}$$

und damit

$$\mathcal{F}(X+Y)(t) = e^{it(\mu+\nu)} e^{-\frac{t^2}{2}(\sigma^2+s^2)}(t) = \mathcal{F}\left\{\mathcal{N}_{\mu+\nu,\sigma^2+s^2}\right\}(t)$$

Nach Charakterisierungstheorem 4.3.6 ist also auch die Summe

$$(X+Y) \sim \mathcal{N}_{\mu+\nu,\sigma^2+s^2}$$

normalverteilt.

4.3.7 Satz: Charakteristische Funktion des Dirac-Maßes

Sei $\widehat{\mathcal{P}}$ charakteristische Funktion eines Wahrscheinlichkeitsmaßes $\mathcal{P}$ auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$. Ist

$$\widehat{\mathcal{P}}\big|_{B_\varepsilon^o(0)} = 1$$

für irgendein $\varepsilon > 0$, so folgt $\mathcal{P} = \delta_0$.

Beweisidee: Zunächst zeigt man dass für $t \in \mathbb{R}$ allgemein gilt

$$1 - \Re\widehat{\mathcal{P}}(2t) \leq 4\left[1 - \Re\widehat{\mathcal{P}}(t)\right]$$

Mit Hilfe dieser Abschätzung folgt durch konsekutive Verdopplung des Radius ε die Behauptung.
□

4.3.8 Satz über integrierbare charakteristische Funktionen

Sei $\mathcal{P}$ Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ mit charakteristische Funktion $\widehat{\mathcal{P}}$, so dass

$$\int_{-\infty}^{\infty} \left|\widehat{\mathcal{P}}(t)\right| \, dt < \infty$$

Dann ist $\mathcal{P} \ll \lambda^1$ absolut stetig und besitzt die Dichte

$$\frac{d\mathcal{P}}{d\lambda^1}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx}\widehat{\mathcal{P}}(t) \, dt$$

Beweis: Es genügt zu zeigen, dass

$$\rho(x) := \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx}\widehat{\mathcal{P}}(t) \, dt \quad , \quad x \in \mathbb{R}$$

tatsächlich Dichte von $\mathcal{P}$ ist. Nach Voraussetzung ist ρ wohldefiniert, und wegen

$$\rho^*(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{itx} \underbrace{\widehat{\mathcal{P}}^*(t)}_{\widehat{\mathcal{P}}(-t)} \, dt \overset{\substack{\text{sub.}\\ u:=-t}}{=} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux}\widehat{\mathcal{P}}(u) \, du = \rho(x)$$

stets reell. Außerdem ist ρ stetig, denn

$$|\rho(x) - \rho(x+h)| \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} \underbrace{\underbrace{\left|1 - e^{ith}\right|}_{\xrightarrow{h\to 0} 0} \cdot \left|\widehat{\mathcal{P}}(t)\right|}_{\substack{\leq 2\left|\widehat{\mathcal{P}}(t)\right| \\ \text{integrable} \\ \text{Majorante}}} \, dt \overset{h\to 0}{\longrightarrow} 0$$

Bekanntlich ist die Verteilungsfunktion von $\mathcal{P}$ in höchstens abzählbar vielen Punkten unstetig, sprich

$$D := \{s \in \mathbb{R} : \mathcal{P}(\{s\}) = 0\}$$

ist dicht in $\mathbb{R}$. Daher gilt für $a < b \in D$:

$$\int_a^b \rho(x) \, dx \overset{\text{Fubini}}{=} \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_a^b e^{-itx} dx \, \widehat{\mathcal{P}}(t) \, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-ita} - e^{-itb}}{it} \widehat{\mathcal{P}}(t) \, dt \overset{4.3.3}{=} \mathcal{P}\left((a,b)\right)$$

Nach Lemma A.3.2 ist dann ρ Dichte von $\mathcal{P}$.
□

Beispiel: Für die Normalverteilung $\mathcal{N}_{0,\sigma^2}$ folgt der bekannte Zusammenhang

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx} \underbrace{e^{-\sigma^2\frac{t^2}{2}}}_{\mathcal{F}(\mathcal{N}_{0,\sigma^2})(t)} dt = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma^2}} = \frac{d\mathcal{N}_{0,\sigma^2}}{d\lambda^1}(x)$$

4.3.9 Theorem von Cramér-Wold

Zu $t \in \mathbb{R}$ und $\mathbf{k} \in \mathbb{R}^n$ definiere die Halbräume

$$H(\mathbf{k},t) := \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{x}, \mathbf{k}\rangle \leq t\}$$

Sind $\mathcal{P}_1, \mathcal{P}_2$ Wahrscheinlichkeitsmaße auf $\mathbb{R}^n$ mit

$$\mathcal{P}_1\left(H(\mathbf{k},t)\right) = \mathcal{P}_2\left(H(\mathbf{k},t)\right) \quad \forall\ \mathbf{k} \in \mathbb{R}^n,\ t \in \mathbb{R}$$

so gilt $\mathcal{P}_1 = \mathcal{P}_2$.

Umformulierung für Zufallsvariablen: Sind X, Y zwei $\mathbb{R}^n$-wertige Zufallsvariablen mit

$$\langle X, \mathbf{k}\rangle \overset{\text{d.}}{=} \langle Y, \mathbf{k}\rangle \quad \forall\ \mathbf{k} \in \mathbb{R}^n$$

so sind $X \overset{\text{d.}}{=} Y$.

Interpretation: Stimmen die Wahrscheinlichkeitsmaße $\mathcal{P}_1, \mathcal{P}_2$ auf <u>allen</u> Halbräumen $H(\mathbf{k},t)$ überein, so müssen $\mathcal{P}_1 = \mathcal{P}_2$ sein. Beachte dass es nicht reicht wenn $\mathcal{P}_1, \mathcal{P}_2$ auf den achsenorientierten Halbräumen $H(\mathbf{e}_j, t)$ übereinstimmen.

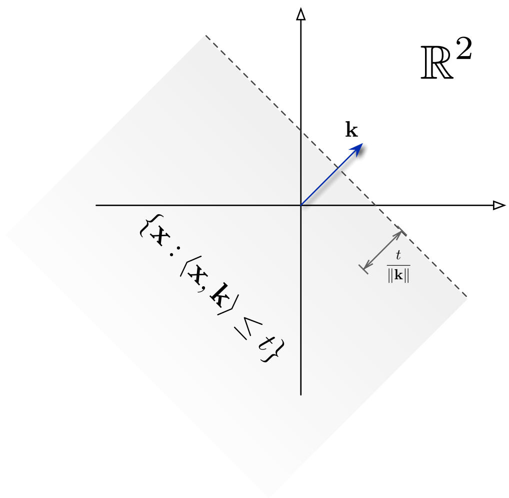

Abbildung 14: Zur Interpretation des Theorems von Cramér-Wold: Der Halbraum $H(\mathbf{k},t) \subseteq \mathbb{R}^2$.

Beweis: Nach Theorem 4.3.6 genügt es zu zeigen $\mathcal{F}(X) = \mathcal{F}(Y)$. Doch tatsächlich ist

$$\mathcal{F}(X)(\mathbf{k}) = \mathcal{F}(\underbrace{\langle X, \mathbf{k}\rangle}_{\overset{\text{d.}}{=}\langle Y,\mathbf{k}\rangle})(1) = \mathcal{F}(\langle Y, \mathbf{k}\rangle)(1) = \mathcal{F}(Y)(\mathbf{k}) \quad \forall\ \mathbf{k} \in \mathbb{R}^n$$

$\square$

4.4 Schwache Konvergenz

Angesichts der Charakterisierung von Verteilungen durch deren charakteristischen Funktionen, stellt sich die Frage inwiefern die Konvergenz der ersteren mit dem Konvergenzverhalten der letzteren in Beziehung steht.

4.4.1 Definition: Schwache Konvergenz von Wahrscheinlichkeitsmaßen

Sei (E, d) ein metrischer Raum, dazu der Raum $\mathcal{C}_b(X)$ aller beschränkter, stetiger, reeller Funktionen auf E. Seien $(\mathcal{P}_n)_{n\in\mathbb{N}}$, $\mathcal{P}$ Wahrscheinlichkeitsmaße auf $(E, \mathscr{B}(E))$. Dann *konvergieren* die $\mathcal{P}_n$ *schwach* (*weak*) gegen $\mathcal{P}$:$\Leftrightarrow$

$$\int\limits_E f\ d\mathcal{P}_n \overset{n\to\infty}{\longrightarrow} \int\limits_E f\ d\mathcal{P} \quad \forall\ f \in \mathcal{C}_b(E)$$

Man schreibt auch $\mathcal{P}_n \underset{\text{weak}}{\overset{n\to\infty}{\longrightarrow}} \mathcal{P}$.

Bemerkung: Nach A.2.3 ist der schwache Grenzwert einer Folge von Wahrscheinlichkeitsmaßen (falls existent) eindeutig.

Beispiele:

(i) Wegen

$$\int\limits_{\mathbb{R}} f\ d\delta_{\frac{1}{n}} = f\left(\tfrac{1}{n}\right) \overset{n\to\infty}{\longrightarrow} f(0) = \int\limits_{\mathbb{R}} f\ d\delta_0 \quad \forall\ f \in \mathcal{C}_b(\mathbb{R})$$

gehen $\delta_{\frac{1}{n}} \underset{\text{weak}}{\overset{n\to\infty}{\longrightarrow}} \delta_0$.

(ii) Für $\sigma_n \geq 0,\ \sigma_n \overset{n\to\infty}{\longrightarrow} 0$ gehen die normalverteiltenNormalverteilung $\mathcal{P}_n \sim \mathcal{N}_{0,\sigma_n^2}$ schwach gegen δ_0, denn

$$\frac{1}{\sqrt{2\pi}\sigma_n}\int\limits_{\mathbb{R}} f(x)\cdot e^{-\frac{x^2}{2\sigma^2}}dx = \frac{1}{\sqrt{2\pi}}\int\limits_{\mathbb{R}} \underbrace{f(\sigma_n y)}_{\underset{\text{punktweise}}{\overset{n\to\infty}{\longrightarrow}} f(0)}\ e^{-\frac{y^2}{2}}\ dy$$

$$\underset{\text{Lebesgue}}{\overset{n\to\infty}{\longrightarrow}} \frac{f(0)}{\sqrt{2\pi}}\int\limits_{\mathbb{R}} e^{-\frac{y^2}{2}}\ dy = f(0) = \int\limits_{\mathbb{R}} f\ d\delta_0 \quad \forall\ f \in \mathcal{C}_b(\mathbb{R})$$

(iii) Ist $\mathcal{P}_n$ gleichverteilt auf $\left\{\frac{1}{n}, \frac{2}{n}, .., 1\right\}$, sprich

$$\mathcal{P}\left(\left\{\tfrac{k}{n}\right\}\right) = \frac{1}{n} \quad , \quad k = 1, .., n$$

so gehen die $\mathcal{P}_n$ schwach gegen die Gleichverteilung auf $(0, 1)$, denn

$$\int\limits_{\mathbb{R}} f\ d\mathcal{P}_n = \frac{1}{n}\sum_{k=1}^{n} f\left(\tfrac{k}{n}\right) \underset{(\clubsuit)}{\overset{n\to\infty}{\longrightarrow}} \underbrace{\int\limits_0^1 f(x)\ dx}_{\substack{\text{im Riemann}\\ \text{-Sinn}}} \overset{\substack{f\\ \text{stetig}}}{=} \int\limits_0^1 f\ d\lambda^1(x) \quad \forall\ f \in \mathcal{C}_b(\mathbb{R})$$

wobei verwendet wurde dass f beschränkt & stetig ist ($\clubsuit$).

4.4.2 Bemerkungen zur Bezeichnung der *schwachen* Konvergenz

Ist (K, d) ein kompakter, metrischer Raum und $\mathcal{C}(K)$ der Raum aller stetigen, reellen Funktionen auf K, so gilt bekanntlich $\mathcal{C}(K) = \mathcal{C}_b(K)$ und die Norm

$$\|f\| := \sup_{x\in K}|f(x)| \quad , \quad f \in \mathcal{C}_b(K)$$

macht $\mathcal{C}_b(K)$ zu einem (separablen) Banachraum. Nach Frigyes Riesz ist dann der Dualraum $\mathcal{C}_b(K)'$ identifizierbar mit dem Raum

$$\Lambda(K) := \{\mu : \text{signiertes, endliches Maß auf } (K, \mathscr{B}(K))\}$$

durch

$$f \overset{\mu}{\mapsto} \mu(f) := \int_K f \, d\mu \quad , \quad f \in \mathcal{C}_b(K), \ \mu \in \Lambda(K)$$

In der Hinsicht sind die signierten, endlichen Maße auf K stetige Linearformen auf $\mathcal{C}_b(K)$. Sind nun $\mu_n, \mu \in \Lambda$ so dass

$$\mu_n(f) \overset{n\to\infty}{\longrightarrow} \mu(f) \quad \forall\, f \in \mathcal{C}_b(K)$$

so konvergieren die Linearformen $\mu_n \in \mathcal{C}_b(K)'$ definitionsgemäß *schwach** gegen μ (vgl. A.4.7). Deren schwache Konvergenz als Maße $\in \Lambda(K)$ ist also in Wirklichkeit eher eine schwache*-Konvergenz auf dem Dualraum $\mathcal{C}_b(K)'$.

Allerdings ist für allgemeinen metrischen Raum (E,d) lediglich $\mathcal{C}_b(E)' \supsetneq \Lambda(E)$. Als Beispiel diene der Raum $(\mathbb{N}, |\cdot|)$, für den gilt

$$\mathcal{C}_b(\mathbb{N}) = l_\infty = \{(a_n)_{n=1}^\infty \subseteq \mathbb{R} \text{ beschränkt}\}$$

Dann ist der Banachlimes $L \in l'_\infty$ nicht darstellbar durch irgendein Maß $\in \Lambda(\mathbb{N})$.

4.4.3 Satz: Eindeutigkeit des schwachen Grenzwertes

Sei (E,d) ein metrischer Raum und $\mathcal{P}_n, \mathcal{P}, \mathcal{Q}$ Wahrscheinlichkeitsmaße auf $(E, \mathscr{B}(E))$ mit $\mathcal{P}_n \underset{\text{weak}}{\overset{n\to\infty}{\longrightarrow}} \mathcal{P}$ und $\mathcal{P}_n \underset{\text{weak}}{\overset{n\to\infty}{\longrightarrow}} \mathcal{Q}$. Dann gilt $\mathcal{P} = \mathcal{Q}$.

Beweis: Nach Voraussetzung gilt

$$\int_E f \, d\mathcal{P} = \lim_{n\to\infty} \int_E f \, d\mathcal{P}_n = \int_E f \, d\mathcal{Q}$$

für jede $f \in \mathcal{C}_b(E)$. Nach Eindeutigkeitssatz A.2.3 folgt $\mathcal{P} = \mathcal{Q}$.
□

4.4.4 Theorem über die schwache Konvergenz

Sei (E,d) ein metrischer Raum, darauf die Wahrscheinlichkeitsmaße $\mathcal{P}_n, \mathcal{P}$. Dann sind folgende Aussagen äquivalent:

1. $\mathcal{P}_n \underset{\text{weak}}{\overset{n\to\infty}{\longrightarrow}} \mathcal{P}$

2. Für $A \subseteq E$ abgeschlossen gilt
$$\overline{\lim_{n\to\infty}}\ \mathcal{P}_n(A) \leq \mathcal{P}(A)$$

3. Für $G \subseteq T$ offen gilt
$$\underline{\lim_{n\to\infty}}\ \mathcal{P}_n(G) \geq \mathcal{P}(G)$$

4. Für $B \subseteq E$ mit $\mathcal{P}(\partial B) = 0$ gilt:
$$\lim_{n\to\infty} \mathcal{P}_n(B) = \mathcal{P}(B)$$

Beweis:

(1)⇒(2): Sei $A \subseteq E$ abgeschlossen, dazu Funktionen $f_k \in \mathcal{C}_b(E)$ mit $f_k \underset{k\to\infty}{\searrow} 1_A$ punktweise (vgl. A.2.2). Dann gilt

$$\overline{\lim_{n\to\infty}}\ \mathcal{P}_n(A) = \overline{\lim_{n\to\infty}} \int\limits_E 1_A\ d\mathcal{P}_n \overset{f_k \geq 1_A}{\leq} \overline{\lim_{n\to\infty}} \int\limits_E f_k\ d\mathcal{P}_n \overset{(1)}{=} \int\limits_E f_k\ d\mathcal{P}$$

für jedes $k \in \mathbb{N}$. Nach Lebesgue geht

$$\int\limits_E f_k\ d\mathcal{P} \overset{k\to\infty}{\longrightarrow} \int\limits_E 1_A\ d\mathcal{P} = \mathcal{P}(A)$$

woraus die Behauptung folgt.

(2)⇒(3): Sei $G \subseteq E$ offen, dann ist G^c abgeschlossen und es gilt

$$1 - \mathcal{P}(G) = \mathcal{P}(G^c) \overset{(2)}{\geq} \overline{\lim_{n\to\infty}}\ \underbrace{\mathcal{P}_n(G^c)}_{1-\mathcal{P}_n(G)} = 1 - \underline{\lim_{n\to\infty}}\ \mathcal{P}_n(G)$$

sprich

$$\underline{\lim_{n\to\infty}}\ \mathcal{P}_n(G) \geq \mathcal{P}(G)$$

(3)⇒(2): Ganz analog zu (2)⇒(3).

(2)∧(3)⇒(4): Sei $B \subseteq E$ mit $\mathcal{P}(\partial B) = 0$, sprich

$$\mathcal{P}(\mathrm{cl}(B)) = \mathcal{P}(B) = \mathcal{P}(\mathrm{int}(B)) \tag{4.4.4.1}$$

Dann gilt

$$\mathcal{P}(B) \overset{(4.4.4.1)}{=} \mathcal{P}(\mathrm{int}(B)) \overset{(3)}{\leq} \underline{\lim_{n\to\infty}}\ \mathcal{P}_n(\mathrm{int}(B)) \leq \underline{\lim_{n\to\infty}}\ \mathcal{P}_n(B)$$

$$\leq \overline{\lim_{n\to\infty}}\ \mathcal{P}_n(B) \leq \overline{\lim_{n\to\infty}}\ \mathcal{P}_n(\mathrm{cl}(B)) \overset{(2)}{\leq} \mathcal{P}(\mathrm{cl}(B)) \overset{(4.4.4.1)}{=} \mathcal{P}(B)$$

sprich

$$\lim_{n\to\infty} \mathcal{P}_n(B) = \mathcal{P}(B)$$

(4)⇒(2): Sei $A \subseteq E$ abgeschlossen. Zu $\delta > 0$ definiere

$$A_\delta := \{x \in E : d(x, A) \leq \delta\}$$

(vgl. (A.2.1.1)). Dabei sind alle A_δ abgeschlossen (da $d(\cdot, A)$ stetig) und monoton wachsend in δ. Wegen

$$\underbrace{\{x \in E : d(x, A) < \delta\}}_{\subseteq A_\delta,\ \text{offen}} \subseteq \mathrm{int}(A_\delta)$$

gilt außerdem[14]

$$\partial A_\delta = \underbrace{\mathrm{cl}(A_\delta)}_{A_\delta} \setminus \mathrm{cl}(A_\delta) \subseteq A_\delta \setminus \{x \in E : d(x, A) < \delta\} = \{x \in E : d(x, A) = \delta\} \tag{4.4.4.2}$$

und wegen

$$\{x \in E : d(x, A) = \delta_1\} \cap \{x \in E : d(x, A) = \delta_2\} = \emptyset \quad \forall\ \delta_1 \neq \delta_2 > 0$$

[14]Beachte: Im allgemeinen gilt $\partial A_\delta \subsetneq \{x \in E : d(x, A) = \delta\}$.

sind alle Ränder ∂A_δ, $\delta > 0$ disjunkt. Nach A.3.3 gibt es daher höchstens abzählbar viele $\delta > 0$ mit $\mathcal{P}(\partial A_\delta) > 0$. Insbesondere existiert Folge $(\delta_k)_{k=1}^\infty \subseteq (0,\infty)$ mit $\delta_k \underset{k\to\infty}{\searrow} 0$ und $\mathcal{P}(\partial A_{\delta_k}) = 0$. Dabei gilt

$$\bigcap_{k\in\mathbb{N}} A_{\delta_k} = \{x \in E : d(x,A) = 0\} \overset{(A.2.1.1)}{=} A \tag{4.4.4.3}$$

Aus

$$\overline{\lim_{n\to\infty}}\ \mathcal{P}_n(A) \leq \overline{\lim_{n\to\infty}}\ \mathcal{P}_n(A_{\delta_k}) \overset{(4)}{=} \mathcal{P}(A_{\delta_k}) \quad \forall\, k \in \mathbb{N}$$

und

$$\mathcal{P}(A_{\delta_k}) \underset{(4.4.4.3)}{\overset{k\to\infty}{\longrightarrow}} \mathcal{P}(A)$$

folgt schließlich die Behauptung.

(2)⇒(1): Geht für $f \in \mathcal{C}_b(E)$

$$\int_E f\ d\mathcal{P}_n \overset{n\to\infty}{\longrightarrow} \int_E f\ d\mathcal{P} \tag{4.4.4.4}$$

so geht auch

$$\int_E (\alpha f + \beta)\ d\mathcal{P}_n \overset{n\to\infty}{\longrightarrow} \int_E (\alpha f + \beta)\ d\mathcal{P}$$

für beliebige $\alpha, \beta \in \mathbb{R}$. Da beliebige $f \in \mathcal{C}_b(E)$, $-R < f < R$ durch

$$\widetilde{f} := \frac{f+R}{2R}$$

so skaliert werden kann dass $0 < \widetilde{f} < 1$, genügt es (4.4.4.4) für den Fall $0 < f < 1$ zu zeigen. Sei also o.B.d.A. $f \in \mathcal{C}_b(E)$ mit $0 < f < 1$. Dann sind für festes $k \in \mathbb{N}$ die Mengen

$$A_j := \left\{x \in E : f(x) \geq \tfrac{j}{k}\right\} \quad , \quad j = 0,..,k$$

abgeschlossen und monoton fallend

$$E = A_0 \supseteq A_1 \supseteq A_2 \supseteq \cdots \supseteq A_k = \emptyset \tag{4.4.4.5}$$

wobei

$$A_{j-1} \setminus A_j = \left\{\tfrac{j-1}{k} \leq f \leq \tfrac{j}{k}\right\} \tag{4.4.4.6}$$

Setzt man nun

$$g := \sum_{j=0}^{k} \frac{j-1}{k} \cdot 1_{A_{j-1}\setminus A_j}$$

$$h := \sum_{j=0}^{k} \frac{j}{k} \cdot 1_{A_{j-1}\setminus A_j}$$

so kann man Abschätzen

$$g \leq f \leq h \tag{4.4.4.7}$$

bzw.

$$\frac{1}{k}\sum_{j=1}^{k-1} \mathcal{P}(A_j) = \sum_{j=1}^{k} \frac{j-1}{k}\left[\mathcal{P}(A_{j-1}) - \mathcal{P}(A_j)\right] = \int_E g\ d\mathcal{P} \overset{(4.4.4.7)}{\leq} \int_E f\ d\mathcal{P} \overset{(4.4.4.7)}{\leq} \int_E h\ d\mathcal{P}$$

$$= \sum_{j=1}^{k} \frac{j}{k}\left[\mathcal{P}(A_{j-1}) - \mathcal{P}(A_j)\right] = \frac{1}{k} + \frac{1}{k}\sum_{j=1}^{k-1} \mathcal{P}(A_j)$$

$$\overline{\lim_{n\to\infty}} \int\limits_E f\, d\mathcal{P}_n \leq \frac{1}{k} + \frac{1}{k}\sum_{j=1}^{k-1} \overline{\lim_{n\to\infty}}\, \mathcal{P}_n(A_j) \leq \frac{1}{k} + \frac{1}{k}\sum_{j=1}^{k-1} \mathcal{P}(A_j) \leq \frac{1}{k} + \int\limits_E f\, d\mathcal{P}$$

Da $k \in \mathbb{N}$ beliebig war, folgt

$$\overline{\lim_{n\to\infty}} \int\limits_E f\, d\mathcal{P}_n \leq \int\limits_E f\, d\mathcal{P} \tag{4.4.4.8}$$

Offensichtlich gilt (4.4.4.8) auch für $(f-1)$, bzw. allgemein für $-1 < f < 0$, insbesondere auch für $(-f)$:

$$-\underline{\lim_{n\to\infty}} \int\limits_E f\, d\mathcal{P}_n = \overline{\lim_{n\to\infty}} \int\limits_E (-f)\, d\mathcal{P}_n \overset{(4.4.4.8)}{\leq} \int\limits_E (-f)\, d\mathcal{P} = -\int\limits_E f\, d\mathcal{P}$$

bzw.

$$\underline{\lim_{n\to\infty}} \int\limits_E f\, d\mathcal{P}_n \geq \int\limits_E f\, d\mathcal{P} \tag{4.4.4.9}$$

Zusammen mit (4.4.4.8) folgt dann die Behauptung.

□

Beispiel: Betrachtet sei die Folge $\delta_{\frac{1}{n}},\ n \in \mathbb{N}$ von Dirac-Maßen auf $\mathbb{R}$ für die bekanntlich

$$\delta_{\frac{1}{n}} \underset{\text{weak}}{\overset{n\to\infty}{\longrightarrow}} \delta_0$$

gilt. Ist z.B. $A := [-1, 0]$ so ist $\delta_{\frac{1}{n}}(A) = 0,\ \delta_0(A) = 1$ und tatsächlich

$$\overline{\lim_{n\to\infty}}\ \delta_{\frac{1}{n}}(A) < \delta_0(A)$$

Ähnlich gilt für die offene Menge $G := (0, 1)$

$$\overline{\lim_{n\to\infty}}\ \delta_{\frac{1}{n}}(G) > \delta_0(G)$$

Beachte dass

$$\delta_{\frac{1}{n}}(A) \overset{n\to\infty}{\not\longrightarrow} \delta_0(A) \quad , \quad \delta_{\frac{1}{n}}(G) \overset{n\to\infty}{\not\longrightarrow} \delta_0(G)$$

in Übereinstimmung ist mit der Tatsache dass $\delta_0(\partial A) \neq 0 \neq \delta_0(\partial G)$.

4.4.5 Satz: Hinreichende Bedingung zur schwachen Konvergenz

Sei (E, d) metrischer Raum, darauf die Wahrscheinlichkeitsmaße $\mathcal{P}_n, \mathcal{P},\ n \in \mathbb{N}$. Sei $\mathcal{A} \subseteq \mathscr{B}(E)$ ein Mengensystem mit:

1. $\mathcal{A}$ ist $\cap$-stabil.
2. $\mathcal{A}$ erzeugt alle offenen Mengen, das heißt

$$\forall\ \emptyset \neq G \subseteq E \text{ offen } :\ \exists\ A_1, A_2, .. \in \mathcal{A}\ :\ G = \bigcup_{k=1}^{\infty} A_k$$

3. Es gilt

$$\lim_{n\to\infty} \mathcal{P}_n(A) = \mathcal{P}(A) \quad \forall\ A \in \mathcal{A}$$

Dann gehen $\mathcal{P}_n \underset{\text{weak}}{\overset{n\to\infty}{\longrightarrow}} \mathcal{P}$.

Beweis: Sei $G \subseteq E$ offen (o.B.d.A. $G \neq \emptyset$) mit der Darstellung

$$G = \bigcup_{k=1}^{\infty} A_k \quad , \quad A_k \in \mathcal{A} \tag{4.4.5.1}$$

und $\varepsilon > 0$. Wegen

$$\mathcal{P}(G) = \lim_{n\to\infty} \mathcal{P}\Bigg(\underbrace{\bigcup_{k=1}^{n} A_k}_{\nearrow_{n\to\infty} G} \Bigg)$$

existiert ein $n_\varepsilon \in \mathbb{N}$ mit

$$\mathcal{P}\left(\bigcup_{k=1}^{n_\varepsilon} A_k \right) \geq \mathcal{P}(G) - \varepsilon \tag{4.4.5.2}$$

Wegen

$$\mathcal{P}\left(\bigcup_{k=1}^{n_\varepsilon} A_k \right) = \sum_{m=1}^{n_\varepsilon} \sum_{1\leq k_1 <..<k_m \leq n_\varepsilon} \mathcal{P}\Bigg(\underbrace{\bigcap_{i=1}^{m} A_{k_i}}_{\in\mathcal{A}} \Bigg) \tag{4.4.5.3}$$

(analog auch für $\mathcal{P}_n$) geht

$$\lim_{n\to\infty} \mathcal{P}_n \left(\bigcup_{k=1}^{n_\varepsilon} A_k \right) \overset{(4.4.5.3)}{=} \sum_{m=1}^{n_\varepsilon} \sum_{1\leq k_1 <..<k_m \leq n_\varepsilon} \underbrace{\lim_{n\to\infty} \mathcal{P}_n \Bigg(\underbrace{\bigcap_{i=1}^{m} A_{k_i}}_{\in\mathcal{A}} \Bigg)}_{\mathcal{P}(\bigcap_{i=1}^{m} A_{k_i})} \overset{(4.4.5.3)}{=} \mathcal{P}\left(\bigcup_{k=1}^{n_\varepsilon} A_k \right)$$

bzw.

$$\varliminf_{n\to\infty} \mathcal{P}_n(G) \overset{(4.4.5.1)}{\geq} \varliminf_{n\to\infty} \mathcal{P}_n \left(\bigcup_{k=1}^{n_\varepsilon} A_k \right) \overset{(4.4.5.3)}{=} \mathcal{P}\left(\bigcup_{k=1}^{n_\varepsilon} A_k \right) \overset{(4.4.5.2)}{\geq} \mathcal{P}(G) - \varepsilon \tag{4.4.5.4}$$

Da (4.4.5.4) für beliebiges $\varepsilon > 0$ gilt, muss

$$\varliminf_{n\to\infty} \mathcal{P}_n(G) \geq \mathcal{P}(G)$$

sein. Nach 4.4.4 folgt daraus die Behauptung.
□

4.4.6 Satz: Schwache Konvergenz & Einschränkung auf Unterräume

Sei (E, d) ein metrischer Raum, dazu $U \subseteq \mathscr{B}(E)$ ausgestattet mit der Schnitttopologie[15] aus E. Seien $X, X_1, .., X_n$ U-wertige Zufallsvariablen. Dann gehen $X_n \overset{n\to\infty}{\underset{\text{d}}{\longrightarrow}} X$ in $(U, \mathscr{B}(U))$ genau dann wenn $X_n \overset{n\to\infty}{\underset{\text{d}}{\longrightarrow}} X$ in $(E, \mathscr{B}(E))$.[16]

Beweis: **Richtung "⇒":** Es gehen $X_n \overset{n\to\infty}{\underset{\text{d}}{\longrightarrow}} X$ in $(U, \mathscr{B}(U))$. Sei $f \in \mathcal{C}_b(E)$ beliebig, dann ist auch die Einschränkung $f\big|_U \in \mathcal{C}_b(U)$ und es gilt

$$\mathbb{E}f(X_n) \overset{\substack{X_n \\ U\text{-wertig}}}{=} \mathbb{E}f\big|_U(X_n) \overset{n\to\infty}{\longrightarrow} \mathbb{E}f\big|_U(X) \overset{\substack{X \\ U\text{-wertig}}}{=} \mathbb{E}f(X)$$

[15] Gleichzeitig erzeugt durch die Einschränkung von d auf U.
[16] Beachte dass X, X_n auch als E-wertige Funktionen messbar sind da $\mathscr{B}(U) = U \cap \mathscr{B}(E)$.

Richtung "⇐": Es gehen $X_n \overset{n\to\infty}{\underset{d}{\longrightarrow}} X$ in $(E, \mathscr{B}(E))$. Sei $O \in \mathscr{B}(U)$ in U offen, das heißt $O = U \cap O'$ für irgendeine in E offene $O' \in \mathscr{B}(E)$. Dann gilt:

$$\varliminf_{n\to\infty} \mathcal{P}_{X_n}(O) = \varliminf_{n\to\infty} \mathcal{P}_{X_n}(U \cap O') \overset{\substack{X_n \\ U\text{-wertig}}}{=} \varliminf_{n\to\infty} \mathcal{P}_{X_n}(O') \geq \mathcal{P}_X(O') \overset{\substack{X \\ U\text{-wertig}}}{=} \mathcal{P}_X(U \cap O') = \mathcal{P}_X(O)$$

Nach Theorem 4.4.4 folgt die Behauptung.
□

4.4.7 Charakterisierung der schwachen Konvergenz auf $\mathbb{R}$

Seien $\mathcal{P}_n,\ n \in \mathbb{N}$ und $\mathcal{P}$ Wahrscheinlichkeitsmaße auf $\mathbb{R}$ mit jeweils Verteilungsfunktionen F_n und F. Dann gilt:

$$\mathcal{P}_n \overset{n\to\infty}{\underset{\text{weak}}{\longrightarrow}} \mathcal{P} \quad \Leftrightarrow \quad F_n(t) \overset{n\to\infty}{\longrightarrow} F(t) \quad \forall\, t \in \mathcal{C}(\mathcal{P})$$

Beweis: Richtung "⇒" ist nach Theorem 4.4.4 trivial, da $\mathcal{P}(\partial(-\infty, t]) = \mathcal{P}(\{t\}) = 0 \quad \forall\, t \in \mathcal{C}(\mathcal{P})$.
Richtung "⇐": Betrachtet sei das $\cap$-stabile Mengensystem

$$\mathcal{A} := \{(a, b) \subseteq \mathbb{R} \text{ Intervall} : a \leq b \in \mathcal{C}(\mathcal{P})\}$$

Da $\mathcal{C}(\mathcal{P})$ nach A.3.3 dicht in $\mathbb{R}$ liegt, erzeugt $\mathcal{A}$ nach Korollar A.4.3 jede offene Menge $G \subseteq \mathbb{R}$, sprich

$$G = \bigcup_{k=1}^{\infty} I_k$$

für irgendwelche $I_k \in \mathcal{A}$.
Behauptung: Für $t \in \mathcal{C}(\mathcal{P})$ gehen $\mathcal{P}_n((-\infty, t)) \overset{n\to\infty}{\longrightarrow} \mathcal{P}((-\infty, t))$.

> **Beweis:** Nach Voraussetzung
>
> $$\varlimsup_{n\to\infty} \mathcal{P}_n((-\infty, t)) \leq \varlimsup_{n\to\infty} \mathcal{P}_n((-\infty, t]) = \mathcal{P}((-\infty, t]) = \mathcal{P}((-\infty, t)) \tag{4.4.7.1}$$
>
> Anderseits existiert zu jedem $\varepsilon > 0$ ein $s \in \mathcal{C}(\mathcal{P}),\ s < t$ mit $\mathcal{P}((-\infty, s]) \geq \mathcal{P}((-\infty, t)) - \varepsilon$. Daher
>
> $$\varliminf_{n\to\infty} \mathcal{P}_n((-\infty, t)) \geq \varliminf_{n\to\infty} \mathcal{P}_n((-\infty, s]) = \mathcal{P}((-\infty, s]) \geq \mathcal{P}((-\infty, t)) - \varepsilon \tag{4.4.7.2}$$
>
> bzw. aufgrund der Beliebigkeit von ε
>
> $$\varliminf_{n\to\infty} \mathcal{P}_n((-\infty, t)) \geq \mathcal{P}((-\infty, t))$$
>
> Zusammen mit (4.4.7.1) folgt die Behauptung.

Insbesondere geht dann auch für jedes $(a, b) \in \mathcal{A}$:

$$\mathcal{P}_n(a, b) = \underbrace{\mathcal{P}_n((-\infty, b))}_{\overset{n\to\infty}{\longrightarrow} \mathcal{P}((-\infty, b))} - \underbrace{\mathcal{P}_n((-\infty, a])}_{\overset{n\to\infty}{\longrightarrow} \mathcal{P}((-\infty, a])} \overset{n\to\infty}{\longrightarrow} \mathcal{P}((a, b))$$

Nach Satz 4.4.5 folgt schließlich

$$\mathcal{P}_n \overset{n\to\infty}{\underset{\text{weak}}{\longrightarrow}} \mathcal{P}$$

□

Folgerung: Ist F stetig, so gilt

$$\mathcal{P}_n \overset{n\to\infty}{\underset{\text{weak}}{\longrightarrow}} \mathcal{P} \quad \Leftrightarrow \quad F_n \overset{n\to\infty}{\underset{\text{punktweise}}{\longrightarrow}} F$$

4.4.8 Satz über schwache Konvergenz und stetige Verteilungsfunktionen

Seien $\mathcal{P}_n,\ n \in \mathbb{N}$ und $\mathcal{P}$ Wahrscheinlichkeitsmaße auf $\mathbb{R}$ mit jeweils Verteilungsfunktionen $F_n,\ F$. Ist F stetig, so gilt sogar

$$\mathcal{P}_n \underset{\text{weak}}{\overset{n\to\infty}{\longrightarrow}} \mathcal{P} \quad \Leftrightarrow \quad F_n \underset{\text{gleichmäßig}}{\overset{n\to\infty}{\longrightarrow}} F$$

4.4.9 Definition: Lévy-Prokhorov Metrik

Sei (E, d) ein metrischer Raum und $\mathfrak{M}_1(E)$ die Menge aller Wahrscheinlichkeitsmaße auf $(E, \mathscr{B}(E))$. Zu $A \in \mathscr{B}(E)$ und $\varepsilon > 0$ bezeichne

$$A^\varepsilon := \{x \in E : \exists\ y \in A : d(x,y) < \varepsilon\} = \bigcup_{y \in A} B^o_\varepsilon(y)$$

Dann heißt die Abbildung $\pi : \mathfrak{M}_1(E) \times \mathfrak{M}_1(E) \to [0,\infty)$ definiert durch

$$\pi(\mathcal{P},\mathcal{Q}) := \inf\{\varepsilon > 0 : \mathcal{P}(A) \le \mathcal{Q}(A^\varepsilon) + \varepsilon \quad \wedge \quad \mathcal{Q}(A) \le \mathcal{P}(A^\varepsilon) + \varepsilon \quad \forall\ A \in \mathscr{B}(E)\} \tag{4.4.9.1}$$

$$\mathcal{P}, \mathcal{Q} \in \mathfrak{M}_1(E)$$

Lévy-Prokhorov Metrik auf $\mathfrak{M}_1(E)$.

Bemerkungen:

(i) $\pi : \mathfrak{M}_1(E) \times \mathfrak{M}_1(E) \to [0,\infty)$ ist tatsächlich eine Metrik.

(ii) Aus $\mathcal{P}(A) \le \mathcal{Q}(A^\varepsilon) + \varepsilon \quad \forall\ A \in \mathscr{B}(E)$ folgt auch $\mathcal{Q}(A) \le \mathcal{P}(A^\varepsilon) + \varepsilon \quad \forall\ A \in \mathscr{B}(E)$, so dass π auch die Darstellung

$$\pi(\mathcal{P},\mathcal{Q}) = \inf\{\varepsilon > 0 : \mathcal{P}(A) \le \mathcal{Q}(A^\varepsilon) + \varepsilon \quad \forall\ A \in \mathscr{B}(E)\} \tag{4.4.9.2}$$

besitzt.
Erläuterung: Wegen $((A^\varepsilon)^c)^\varepsilon \subseteq A^c$ (vgl. A.4.4(3)) folgt aus $\mathcal{P}(A) \le \mathcal{Q}(A^\varepsilon) + \varepsilon \ \ \forall\ A \in \mathscr{B}(E)$ auch

$$\underbrace{\mathcal{P}((A^\varepsilon)^c)}_{1-\mathcal{P}(A^\varepsilon)} \le \underbrace{\mathcal{Q}\left(((A^\varepsilon)^c)^\varepsilon\right)}_{\le \mathcal{Q}(A^c) = 1-\mathcal{Q}(A)} + \varepsilon$$

sprich

$$\mathcal{Q}(A) \le \mathcal{P}(A^\varepsilon) + \varepsilon$$

4.4.10 Charakterisierung der schwachen Konvergenz durch die Lévy-Prokhorov Metrik

Sei (E, d) ein separabler, metrischer Raum, dazu $\mathfrak{M}_1(E)$ der Raum aller Wahrscheinlichkeitsmaße ausgestattet mit der Lévy-Prokhorov Metrik π. Sind $\mathcal{P}_n, \mathcal{P} \in \mathfrak{M}_1(E)$ so gilt die Äquivalenz:

$$\mathcal{P}_n \underset{\text{weak}}{\overset{n\to\infty}{\longrightarrow}} \mathcal{P} \quad \Leftrightarrow \quad \pi(\mathcal{P}_n, \mathcal{P}) \overset{n\to\infty}{\longrightarrow} 0$$

4.5 Auswahlprinzip von Helly

4.5.1 Definition: Radonsches Wahrscheinlichkeitsmaß

Sei $(T, \mathscr{O})$ ein topologischer Raum und $\mathcal{P}$ ein Wahrscheinlichkeitsmaß auf $(T, \mathscr{B}(T))$. Dann heißt $\mathcal{P}$ *Radonsch* (*straff*, *inner regulär*, engl. *tight*) :⇔

$$\forall\ \varepsilon > 0 : \exists\ K \subseteq T \text{ kompakt } : \quad \mathcal{P}(K) \ge 1 - \varepsilon$$

4.5.2 Definition: Gleichmäßig Radonsche Wahrscheinlichkeitsmaße

Sei $(T, \mathscr{O})$ ein topologischer Raum und $(\mathcal{P}_i)_{i\in I}$ eine Familie von Wahrscheinlichkeitsmaßen auf $(T, \mathscr{B}(T))$. Dann heißen $(\mathcal{P}_i)$ *gleichmäßig Radonsch* (*gleichmäßig straff*, engl. *uniformly tight*) $:\Leftrightarrow$

$$\forall\, \varepsilon > 0 : \exists\, K \subseteq T \text{ kompakt } : \inf_{i\in I} \mathcal{P}_i(K) \geq 1 - \varepsilon$$

Beispiel: Die Familie von Wahrscheinlichkeitsmaßen $(\delta_n)_{n\in\mathbb{N}}$ ist **nicht** gleichmäßig Radonsch, obwohl jedes δ_n an sich Radonsch ist.

4.5.3 Permanenzeigenschaften gleichmäßig radonscher Maße

Sei $(T, \mathscr{O})$ ein topologischer Raum. Dann gilt:

1. Jede endliche Familie $\mathcal{P}_1, .., \mathcal{P}_n$ radonscher Maße auf $(T, \mathscr{B}(T))$ ist gleichmäßig radonsch.
2. Sind $(\mathcal{P}_k)_{k=1}^{\infty}$ und $(\mathcal{Q}_n)_{n=1}^{\infty}$ jeweils Familien gleichmäßig radonscher Maße auf $(T, \mathscr{B}(T))$, so ist es auch $(\mathcal{P}_k, \mathcal{Q}_n)_{k,n=1}^{\infty}$.
3. Ist (T, d) ein separabler metrischer Raum, so ist jedes Wahrscheinlichkeitsmaß $\mathcal{P} \in \mathfrak{M}_1(T)$ auf auf $(T, \mathscr{B}(T))$ radonsch.

Beweis:

1. Klar, da jede endliche Vereinigung kompakter Mengen auch kompakt ist.
2. Analog zu (1).
3. Siehe Literatur.

4.5.4 Theorem: Auswahlprinzip von Helly

Seien $(\mathcal{P}_n)_{n=1}^{\infty}$ Wahrscheinlichkeitsmaße auf $\mathbb{R}$ mit Verteilungsfunktionen F_n. Dann existiert Teilfolge $(\mathcal{P}_{n_k})_{k=1}^{\infty}$ und eine Funktion $F : \mathbb{R} \to [0, 1]$ mit:

- F ist monoton wachsend,
- F ist rechtsseitig stetig,
- Für jeden Stetigkeitspunkt $x \in \mathcal{C}(F)$ von F gilt

$$\lim_{k\to\infty} F_{n_k}(x) = F(x)$$

Beweis: Wähle $D \subseteq \mathbb{R}$ dicht, abzählbar. Dann existiert nach Satz A.4.5 eine Teilfolge $(F_{n_k})_{k=1}^{\infty}$ mit

$$\exists \lim_{k\to\infty} F_{n_k}(d) \ =: G(d) \quad \forall\, d \in D$$

Erhalten so die Abbildung $G : D \to [0, 1]$, wobei aus der Monotonie der $F_{n_k} : \mathbb{R} \to [0, 1]$ auch die entsprechende Monotonie von G folgt. Setzt man nun

$$F(x) := \inf \{G(d) : x \leq d \in D\} \quad , \quad x \in \mathbb{R} \tag{4.5.4.1}$$

so gilt:

- F ist wohldefinierte Abbildung $F : \mathbb{R} \to [0, 1]$
- Für $x_1 \leq x_2$ ist auch offensichtlich $F(x_1) \leq F(x_2)$

- F ist rechtsseitig stetig, denn zu $x \in \mathbb{R}$, $\varepsilon > 0$ existiert nach Konstruktion ein $d \in D$, $d > x$ mit $G(d) \leq F(x) + \varepsilon$. Sind nun $x_n \searrow x$, so existiert ein $n_0 \in \mathbb{N}$ mit $x \leq x_n \leq d \;\; \forall\, n \geq n_0$, sprich

$$F(x) \leq F(x_n) \leq G(d) \leq F(x) + \varepsilon \;\; \forall\, n \geq n_0$$

bzw. $F(x_n) \in B_\varepsilon(F(x)) \;\; \forall\, n \geq n_0$.

- Für jeden Stetigkeitspunkt $x \in \mathcal{C}(F)$ gehen $F_{n_k}(x) \overset{k\to\infty}{\longrightarrow} F(x)$. Tatsächlich: Für $\varepsilon > 0$ existiert zum einen Folge $(d_j) \subseteq D$ mit

$$d_j \searrow x \quad , \quad G(d_j) \searrow F(x)$$

(Def. von F) zum anderen Folge $(x_j) \subseteq \mathbb{R}$ mit

$$x_j \nearrow x \quad , \quad F(x_j) \nearrow F(x)$$

(Stetigkeit von F bei x) bzw. Folge $(d'_j) \subseteq D$ mit $x_j \leq d'_j \leq x$, sprich

$$d'_j \nearrow x \quad , \quad \underbrace{G(d'_j)}_{\substack{\geq F(x_j) \\ \leq F(x)}} \nearrow F(x)$$

Anderseits gehen

$$G(d'_j) = \lim_{k\to\infty} \underbrace{F_{n_k}(d'_j)}_{\leq F_{n_k}(x)} \leq \underline{\lim}_{k\to\infty}\; F_{n_k}(x) \leq \overline{\lim_{k\to\infty}}\; \underbrace{F_{n_k}(x)}_{\leq F_{n_k}(d_j)} \leq \lim_{k\to\infty} F_{n_k}(d_j) = G(d_j) \quad \forall\, j \in \mathbb{N}$$

sprich

$$F(x) \leq \underline{\lim}_{k\to\infty}\; F_{n_k}(x) \leq \overline{\lim_{k\to\infty}}\; F_{n_k}(x) \leq F(x)$$

was zu zeigen war.

□

Beachte: F muss nicht unbedingt Verteilungsfunktion eines Wahrscheinlichkeitsmaßes sein! So gehen z.B. für $\mathcal{P}_n := \delta_n$

$$F_n \underset{\text{punktweise}}{\overset{n\to\infty}{\longrightarrow}} 0$$

4.5.5 Theorem: 2. Auswahlprinzip von Helly im reellen

Seien $(\mathcal{P}_n)_{n=1}^\infty$ eine Folge gleichmäßig Radonscher Wahrscheinlichkeitsmaße auf $\mathbb{R}$. Dann existiert eine Teilfolge $(\mathcal{P}_{n_k})_{k=1}^\infty$ und Wahrscheinlichkeitsmaß $\mathcal{P}$ mit

$$\mathcal{P}_{n_k} \underset{\text{weak}}{\overset{k\to\infty}{\longrightarrow}} \mathcal{P}$$

Beweis: Seien F_n jeweils die Verteilungsfunktionen von $\mathcal{P}_n$. Dann existiert nach 4.5.4 Teilfolge $(F_{n_k})_{k=1}^\infty$ und monoton wachsende, rechtsseitig stetige $F : \mathbb{R} \to [0,1]$ mit

$$F_{n_k}(t) \overset{k\to\infty}{\longrightarrow} F(t) \;\; \forall\, t \in \mathcal{C}(F)$$

Beachte dass $\mathcal{C}(F)$ dicht in $\mathbb{R}$ liegt. Nach Charakterisierungssatz 4.4.7 genügt es zu zeigen, dass F Verteilungsfunktion eines Wahrscheinlichkeitsmaßes ist, sprich $\lim_{t\to\infty} F(t) = 1$ und $\lim_{t\to-\infty} F(t) = 0$. Tatsächlich: Zu $\varepsilon > 0$ existieren $0 < R_1, R_2$ mit $R_1, (-R_2) \in \mathcal{C}(F)$ und

$$\mathcal{P}_n([-R_1, R_1]) \geq 1 - \varepsilon$$

$$\mathcal{P}_n((-R_2, R_2)) \geq 1 - \varepsilon \;\; \forall\, n \in \mathbb{N}$$

Daraus lässt sich schließen

$$F(R_1) = \lim_{k\to\infty} F_{n_k}(R_1) \geq \varliminf_{k\to\infty} \mathcal{P}_{n_k}([-R_1, R_1]) \geq 1-\varepsilon$$

und

$$F(-R_2) = \lim_{k\to\infty} F_{n_k}(-R_2) = 1 - \varlimsup_{k\to\infty} \mathcal{P}_{n_k}((-R_2,\infty)) \leq 1 - \varliminf_{k\to\infty} \underbrace{\mathcal{P}_{n_k}((-R_2, R_2))}_{\geq 1-\varepsilon} \leq \varepsilon$$

sprich

$$F(x) \geq 1-\varepsilon \quad \forall\, x \geq R_1 \quad \wedge \quad F(x) \leq \varepsilon \quad \forall\, x \leq -R_2$$

□

4.5.6 Funktionalanalytische Vorbetrachtung

Wie schon in Abschnitt 4.4.2 angedeutet, sind die signierten, endlichen Borelmaße $\Lambda(K)$ auf einem kompakten metrischen Raum (K, d) eineindeutig mit dem Dualraum $\mathcal{C}(K)'$ des separablen Banachraumes $\mathcal{C}(K)$ identifizierbar. Dabei gilt

$$\|\mu\|_{\mathcal{C}(K)'} = |\mu|\,(K) \quad , \quad \mu \in \Lambda(K)$$

wobei $|\mu|$ die totale Variation des signierten Maßes μ sei (vgl. Hahn-Jordan-Zerlegung). Insbesondere befindet sich jedes Wahrscheinlichkeitsmaß $\mathcal{P}$ in der Einheitskugel

$$B_1 := \{\mu \in \mathcal{C}(K)' : \|\mu\| \leq 1\} \subseteq \mathcal{C}(K)'$$

die nach Banach-Anaoglu A.4.8 bzgl. der schwachen**schwach** Topologie folgenkompakt ist. Daher besitzt jede Folge $(\mathcal{P}_n) \in \Lambda(K)$ von Wahrscheinlichkeitsmaßen eine Teilfolge $(\mathcal{P}_{n_k})_{k=1}^\infty$ die schwach* gegen ein $\mathcal{P} \in B_1$ konvergiert:

$$\mathcal{P}_{n_k} \overset{n\to\infty}{\underset{\text{w*}}{\longrightarrow}} \mathcal{P}$$

Dabei entspricht die schwache* Konvergenz der Linearformen $\mathcal{P}_n \in \mathcal{C}(K)'$ genau der schwachen Konvergenz der Wahrscheinlichkeitsmaße $\mathcal{P}_n \in \Lambda(K)$. Offensichtlich ist $\mathcal{P}$ auch tatsächlich ein Wahrscheinlichkeitsmaß.

Obige Überlegungen lassen sich in folgendem Satz zusammenfassen.

4.5.7 Satz über Wahrscheinlichkeitsmaße auf kompakten Räumen

Sei (K, d) ein kompakter metrischer Raum, darauf die Folge $(\mathcal{P}_n)_{n\in\mathbb{N}}$ endlicher Maße. Diese seien dabei gleichmäßig beschränkt, sprich es existiere eine Konstante $R \geq 0$ mit $\mathcal{P}_n(K) \leq R \quad \forall\, n \in \mathbb{N}$. Dann existiert ein endliches Maß $\mathcal{P}$ mit

$$\mathcal{P}_n \overset{n\to\infty}{\underset{\text{weak}}{\longrightarrow}} \mathcal{P}$$

Sind $(\mathcal{P}_n)$ Wahrscheinlichkeitsmaße, so ist es auch $\mathcal{P}$.

Erläuterung: Die Maße $\mathcal{P}_n' := \mathcal{P}_n/R$ befinden sich alle in der Einheitskugel $B_1 \subseteq \Lambda(K)$ woraus sich die Behauptung nach obiger Vorbetrachtung ergibt.

4.5.8 Satz über Wahrscheinlichkeitsmaße auf $\mathbb{R}^n$

Seien $(\mathcal{P}_m)_{m\in\mathbb{N}}$ Wahrscheinlichkeitsmaße auf $\mathbb{R}^n$. Dann existiert eine Teilfolge $(\mathcal{P}_{m_k})_{k=1}^\infty$ und ein Maß μ mit $\mu(\mathbb{R}^n) \leq 1$ so dass

$$\int_{\mathbb{R}^n} f \; d\mathcal{P}_{m_k} \overset{k\to\infty}{\longrightarrow} \int_{\mathbb{R}^n} f \; d\mu$$

für alle stetigen $f : \mathbb{R}^n \to \mathbb{R}$ mit kompaktem Träger.

Beweis: Für die kompakten Mengen $K_m := [-m,m]^n$ gilt

$$K_1^0 \subseteq K_1 \subseteq K_2^o \subseteq K_2 \subseteq K_3^0 \subseteq \cdots \nearrow \mathbb{R}^n$$

Dann definieren die $(\mathcal{P}_m)$ auf natürlicher Weise Maße auf den Mengen K_j durch

$$\mathcal{P}_m^j(B \cap K_j) := \mathcal{P}_m(B \cap K_j) \quad \forall\, B \in \mathscr{B}(\mathbb{R}^n),\ m,j \in \mathbb{N}$$

Nach 4.5.7 existiert eine unendliche $N_1 \subseteq \mathbb{N}$ und Maß μ_1 auf K_1 mit

$$\lim_{\substack{m\to\infty\\ m\in N_1}} \underbrace{\int_{K_1} f\ d\mathcal{P}_m^1}_{\int_{K_1} f\ d\mathcal{P}_m} = \int_{K_1} f\ d\mu_1 \quad \forall\, f \in \mathcal{C}(K_1)$$

Ähnlich gibt es eine unendliche $N_2 \subseteq N_1$ mit

$$\lim_{\substack{m\to\infty\\ m\in N_2}} \underbrace{\int_{K_2} f\ d\mathcal{P}_m^2}_{\int_{K_2} f\ d\mathcal{P}_m} = \int_{K_2} f\ d\mu_2 \quad \forall\, f \in \mathcal{C}(K_2)$$

Durch entsprechende Fortsetzung und einen Diagonalisierungsschluss erhält man Teilfolge $(\mathcal{P}_{m_k})_{k=1}^\infty$ und Maße μ_j auf K_j so dass[17]

$$\lim_{k\to\infty} \int_{K_j} f\ d\mathcal{P}_{m_k} = \int_{K_j} f\ d\mu_j \quad \forall\, f \in \mathcal{C}(K_j),\ j \in \mathbb{N} \tag{4.5.8.1}$$

Behauptung: Es gilt $\mu_j(B) = \mu_{j+1}(B)$ für jede $B \subseteq K_j^o$.

Beweis: Sei zunächst $A \subseteq K_j^0$ abgeschlossen, dann existieren nach A.4.6 stetige, fallende Funktionen $g_j \searrow 1_A$ mit $\operatorname{supp}(g_j) \subseteq K_j$ und $0 \le g_j \le 1$. Für diese gilt

$$\int_{K_j} g_i\ d\mu_{j+1} = \int_{K_{j+1}} g_i\ d\mu_{j+1} \overset{(4.5.8.1)}{=} \lim_{k\to\infty} \int_{K_{j+1}} g_i\ d\mathcal{P}_{m_k} = \lim_{k\to\infty} \int_{K_j} g_i\ d\mathcal{P}_{m_k} \overset{(4.5.8.1)}{=} \int_{K_j} g_i\ d\mu_j \tag{4.5.8.2}$$

Nach Lebesgue also

$$\mu_{j+1}(A) = \int_{K_j} 1_A\ d\mu_{j+1} = \lim_{i\to\infty} \int_{K_j} g_i\ d\mu_{j+1} \overset{(4.5.8.2)}{=} \lim_{i\to\infty} \int_{K_j} g_i\ d\mu_j = \int_{K_j} 1_A\ d\mu_j = \mu_j(A) \tag{4.5.8.3}$$

Ist nun $B \subseteq K_j^o$ beliebig. Dann gilt bekanntlich

$$\mu_{j+1}(B) = \sup\{\mu_{j+1}(A) : A \subseteq B \text{ abgeschlossen}\} \overset{(4.5.8.3)}{=} \sup\{\mu_j(A) : A \subseteq B \text{ abgeschlossen}\} = \mu_j(B)$$

Setzen nun

$$\mu_j'(B) := \mu_j(B \cap K_j^o) \quad , \quad B \in \mathscr{B}(\mathbb{R}^n)$$

und

$$\mu(B) := \lim_{n\to\infty} \mu_j'(B) \quad , \quad B \in \mathscr{B}(\mathbb{R}^n)$$

Beachte dass nach voriger Behauptung $\mu_j' = \mu_{j+1}'\big|_{K_j^o}$.

Behauptung: μ ist tatsächlich wohldefiniert.

Beweis: Nach voriger Behauptung gilt

$$\mu_{j+1}(B \cap K_{j+1}^o) \ge \mu_{j+1}(B \cap K_j^o) = \mu_j(B \cap K_j^o)$$

das heißt die $\mu_j(B \cap K_j^o)$ sind wachsend in j und der Grenzwert daher wohldefiniert.

[17] Wähle n_k als den k-ten Index in N_k. Analog zu A.4.5.

Entsprechend leicht zu zeigen ist, dass μ ein Maß auf $\mathbb{R}^n$ definiert.
Behauptung: $\mu(\mathbb{R}^n) \le 1$.

Beweis: Wegen $\mathcal{P}_m(K_j^o) \le 1 \;\forall\; m$ gilt $\mu_j(K_j^o) \le 1$ bzw.

$$\mu_i(K_j^o \cap K_i^o) = \mu_j(K_j^o) \le 1 \quad \forall\; i \ge j$$

und daher

$$\mu(K_j^o) = \lim_{i\to\infty} \mu_i(K_j^o \cap K_i^o) \le 1 \quad \forall\; j \in \mathbb{N}$$

Da $K_j^o \nearrow \mathbb{R}^n$ gilt auch

$$\mu(\mathbb{R}^n) \le 1$$

Sei nun $f \in \mathcal{C}(\mathbb{R}^n)$ mit kompaktem Träger, dann existiert ein $j \in \mathbb{N}$ mit $\operatorname{supp}(f) \subseteq K_j^o$, und daher

$$\lim_{k\to\infty} \int\limits_{\mathbb{R}^n} f\; d\mathcal{P}_{m_k} = \lim_{k\to\infty} \int\limits_{K_j^o} f\; d\mathcal{P}_{m_k} \overset{(4.5.8.1)}{=} \int\limits_{K_j^o} f\; d\mu_j = \int\limits_{K_j^o} f\; d\mu_j' \overset{\substack{\mu_j' = \mu_{j+1}'\big|_{K_j^o} \\ \mu_j' \nearrow \mu}}{=} \int\limits_{K_j^o} f\; d\mu = \int\limits_{\mathbb{R}^n} f\; d\mu$$

□

4.5.9 Theorem: 2. Auswahlprinzip von Helly im n-dimensionalen

Seien $(\mathcal{P}_m)_{m\in\mathbb{N}}$ gleichmäßig Radonsche Wahrscheinlichkeitsmaße auf $\mathbb{R}^n$. Dann existiert eine Teilfolge $(\mathcal{P}_{m_k})_{k=1}^\infty$ und Wahrscheinlichkeitsmaß $\mathcal{P}$ mit

$$\mathcal{P}_{m_k} \underset{\text{weak}}{\overset{k\to\infty}{\longrightarrow}} \mathcal{P}$$

Beweis: Nach 4.5.8 existiert ein Maß μ mit $\mu(\mathbb{R}^n) \le 1$ und Teilfolge $(\mathcal{P}_{m_k})_{k=1}^\infty$ mit

$$\int\limits_{\mathbb{R}^n} f\; d\mathcal{P}_{m_k} \overset{k\to\infty}{\longrightarrow} \int\limits_{\mathbb{R}^n} f\; d\mu \tag{4.5.9.1}$$

für jegliche stetige $f : \mathbb{R}^n \to \mathbb{R}$ mit kompaktem Träger.
Behauptung: $\mu(\mathbb{R}^n) = 1$.

Beweis: Zu $\varepsilon > 0$ existiert bekanntlich kompakte Menge $K \subseteq \mathbb{R}^n$ mit $\mathcal{P}_m(K) \ge 1 - \varepsilon \;\forall\; m$ und stetige Funktion $\psi : \mathbb{R}^n \to [0,1]$ mit $\psi\,|_K = 1$ und kompaktem Träger. Zudem gilt

$$\mu(\mathbb{R}^n) \ge \int\limits_{\mathbb{R}^n} \psi\; d\mu \overset{(4.5.9.1)}{=} \lim_{k\to\infty} \underbrace{\int\limits_{\mathbb{R}^n} \psi\; d\mathcal{P}_{m_k}}_{\ge \mathcal{P}_{m_k}(K)} \ge \varlimsup_{k\to\infty}\; \mathcal{P}_{m_k}(K) \ge 1 - \varepsilon$$

Da $\varepsilon > 0$ beliebig war folgt

$$\mu(\mathbb{R}^n) \ge 1$$

Behauptung: Beziehung (4.5.9.1) gilt auch für $f \in \mathcal{C}_b(\mathbb{R}^n)$ ohne kompakten Träger.

Beweis: Sei $f \in \mathcal{C}_b(\mathbb{R}^n)$ und $\varepsilon > 0$. Da auch μ Radonsch ist, existiert kompakte Menge $K \subseteq \mathbb{R}^n$ mit $\mu(K), \mathcal{P}_{m_k}(K) \ge 1 - \varepsilon$, dazu stetige Funktion $\psi : \mathbb{R}^n \to [0,1]$ mit kompaktem Träger und $\psi\big|_K = 1$. Können also abschätzen

$$\left| \int\limits_{\mathbb{R}^n} f\; d\mathcal{P}_{m_k} - \int\limits_{\mathbb{R}^n} f\; d\mu \right| \le \underbrace{\left| \int\limits_{\mathbb{R}^n} f(1-\psi)\; d\mathcal{P}_{m_k} \right|}_{\le 2\|f\|_\infty \mathcal{P}_{m_k}(K^c)} + \underbrace{\left| \int\limits_{\mathbb{R}^n} f(1-\psi)\; d\mu \right|}_{\le 2\|f\|_\infty \mu(K^c)} + \underbrace{\left| \int\limits_{\mathbb{R}^n} f\psi\; d\mathcal{P}_{m_k} - \int\limits_{\mathbb{R}^n} f\psi\; d\mu \right|}_{\underset{(4.5.9.1)}{\overset{k\to\infty}{\longrightarrow}} 0}$$

bzw.

$$\varlimsup_{k\to\infty} \left| \int\limits_{\mathbb{R}^n} f\; d\mathcal{P}_{m_k} - \int\limits_{\mathbb{R}^n} f\; d\mu \right| \le 4\,\|f\|_\infty\, \varepsilon$$

Da $\varepsilon > 0$ beliebig war, folgt die Behauptung.

□

4.6 Stetigkeitssatz von Lévy-Cramér

4.6.1 Lemma: Abschätzung von *Restmassen*

Sei $\mathcal{P}$ ein Wahrscheinlichkeitsmaß auf $\mathbb{R}^n$ mit charakteristischer Funktion $\widehat{\mathcal{P}}$. Dann gilt die Abschätzung

$$\mathcal{P}\left(\left\{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_\infty \geq \tfrac{1}{u}\right\}\right) \leq \frac{C}{u} \int_0^u \sum_{j=1}^n \left[1 - \Re\widehat{\mathcal{P}}(t\mathbf{e}_j)\right] \, dt \qquad \forall\, u > 0$$

für eine (n-unabhängige) Konstante $C > 1$.

Beweis: Behandeln zunächst den Fall $n = 1$. Dann lässt sich abschätzen

$$\frac{1}{u}\int_0^u \left[1 - \Re\widehat{\mathcal{P}}(t)\right] \, dt = \frac{1}{u^n}\int_0^u \left[1 - \Re\int_{\mathbb{R}^n} e^{itx} d\mathcal{P}(x)\right] \, dt = \frac{1}{u}\int_0^u \int_{\mathbb{R}} \underbrace{[1 - \cos(tx)]}_{\geq 0} \, d\mathcal{P}(x) \, dt$$

$$\overset{\text{Tornelli}}{=} \frac{1}{u}\int_{\mathbb{R}} \int_0^u \underbrace{[1 - \cos(tx)]}_{\geq 0} \, dt \, d\mathcal{P}(x) \geq \frac{1}{u}\int_{|x| \geq \frac{1}{u}} \int_0^u [1 - \cos(tx)] \, dt \, d\mathcal{P}(x)$$

$$= \frac{1}{u}\int_{|x| \geq \frac{1}{u}} \left[u - \frac{\sin(xu)}{x}\right] \, d\mathcal{P}(x) = \int_{|x| \geq \frac{1}{u}} \left[1 - \frac{\sin(xu)}{xu}\right] \, d\mathcal{P}(x)$$

$$\geq \underbrace{\left[1 - \sup_{|x| \geq 1} \frac{\sin x}{x}\right]}_{=: C^{-1}} \cdot \int_{|x| \geq \frac{1}{u}} d\mathcal{P}(x) = C^{-1} \cdot \mathcal{P}\left(x \in \mathbb{R} : |x| \geq \tfrac{1}{u}\right)$$

Der n-dimensionale Fall lässt sich dann wegen

$$\mathcal{P}\left(\mathbf{x} : |\langle \mathbf{x}, \mathbf{e}_j \rangle| \geq \tfrac{1}{u}\right) = \left(\mathcal{P} \circ \mathbf{e}_j^{-1}\right)\left(t : |t| \geq \tfrac{1}{u}\right) \leq \frac{C}{u}\int_0^u \left[1 - \Re \underbrace{\mathcal{F}\left(\mathcal{P} \circ \mathbf{e}_j^{-1}\right)(t)}_{\widehat{\mathcal{P}}(t\mathbf{e}_j)}\right] \, dt \tag{4.6.1.1}$$

und

$$\mathcal{P}\left(\mathbf{x} : \|\mathbf{x}\|_\infty \geq \tfrac{1}{u}\right) \leq \sum_{j=1}^n \mathcal{P}\left(\mathbf{x} : |\langle \mathbf{x}, \mathbf{e}_j \rangle| \geq \tfrac{1}{u}\right) \overset{(4.6.1.1)}{\leq} \frac{C}{u}\int_0^u \sum_{j=1}^n \left[1 - \Re\widehat{\mathcal{P}}(t\mathbf{e}_j)\right] \, dt$$

auf den 1-dim. Fall zurückführen.
□

4.6.2 Folgerung: Hinreichende Bedingung zur gleichmäßigen Straffheit

Seien $(\mathcal{P}_k)_{k=1}^\infty$ Wahrscheinlichkeitsmaße auf $\mathbb{R}^n$ mit charakteristischen Funktionen $\widehat{\mathcal{P}}_n$ und $h : \mathbb{R}^n \to \mathbb{C}$ eine in 0 stetige Funktion mit

$$\widehat{\mathcal{P}}_k(\mathbf{k}) \overset{k \to \infty}{\longrightarrow} h(\mathbf{k}) \quad \forall\, \mathbf{k} \in \mathbb{R}^n$$

Dann sind $(\mathcal{P}_k)$ gleichmäßig Radonsch.

Beweis: Wegen $\widehat{\mathcal{P}}_n(0) = 1 \ \ \forall\, n$ muss natürlich $h(0) = 1$ sein. Sei nun $\varepsilon > 0$ vorgegeben, dann sind nach 4.5.3 ein $R \geq 0$ und $k_0 \in \mathbb{N}$ zu finden mit

$$\mathcal{P}_k(\|x\|_\infty \geq R) \leq \varepsilon \quad \forall\, k \geq k_0$$

Da h stetig in 0 ist, existiert ein $\delta > 0$ mit

$$\Re h(\mathbf{k}) \geq 1 - \frac{\varepsilon}{2nC} \quad \forall\ \|\mathbf{k}\|_\infty \leq \delta$$

wobei $C > 1$ die Abschätzungskonstante aus 4.6.1 sei. Wähle dazu $R > \frac{1}{\delta}$. Aus

$$\left|1 - \Re\widehat{\mathcal{P}}(\mathbf{k})\right| \leq 2 \ \ , \ \ 1 - \Re\widehat{\mathcal{P}}_k(\mathbf{k}) \overset{k\to\infty}{\longrightarrow} 1 - \Re h(\mathbf{k})$$

folgt nach Lebesgue

$$\mathcal{P}_k\left(\|\mathbf{x}\|_\infty \geq R\right) \overset{4.6.1}{\leq} CR \int\limits_0^{\frac{1}{R}} \sum_{j=1}^n \left[1 - \Re\widehat{\mathcal{P}}_k(t\mathbf{e}_j)\right]\ dt$$

$$\underset{\text{Lebesgue}}{\overset{k\to\infty}{\longrightarrow}} CR \int\limits_0^{\frac{1}{R}} \sum_{j=1}^n \underbrace{\left[1 - \Re h(t\mathbf{e}_j)\right]}_{\leq \frac{\varepsilon}{2nC}}\ dt \leq CR \int\limits_0^{\frac{1}{R}} \frac{\varepsilon}{2C}\ dt = \frac{\varepsilon}{2}$$

Daher

$$\overline{\lim_{k\to\infty}}\ \mathcal{P}_k(\|\mathbf{x}\|_\infty \geq R) \leq \frac{\varepsilon}{2}$$

Insbesondere existiert ein $k_0 \in \mathbb{N}$ mit

$$\sup_{k\geq k_0} \mathcal{P}_k(\|\mathbf{x}\|_\infty \geq R) \leq \varepsilon$$

was zu zeigen war.
□

4.6.3 Lemma über die Konvergenz charakteristischer Funktionen

Seien $(\mathcal{P}_k)_{k=1}^\infty, \mathcal{P}$ Wahrscheinlichkeitsmaße auf $\mathbb{R}^n$ mit charakteristischen Funktionen $\widehat{\mathcal{P}}_k, \widehat{\mathcal{P}}$. Gehen $\mathcal{P}_k \underset{\text{weak}}{\overset{k\to\infty}{\longrightarrow}} \mathcal{P}$ so gehen auch

$$\widehat{\mathcal{P}}_k(\mathbf{k}) \overset{k\to\infty}{\longrightarrow} \widehat{\mathcal{P}}(\mathbf{k}) \quad \forall\ \mathbf{k} \in \mathbb{R}$$

Beweis: Die Funktionen $\mathbf{x} \mapsto \cos(\mathbf{kx})$ und $\mathbf{x} \mapsto \sin(\mathbf{kx})$ sind für jedes $\mathbf{k} \in \mathbb{R}^n$ auf $\mathbb{R}^n$ stetig, beschränkt. Nach Voraussetzung gehen daher

$$\widehat{\mathcal{P}}_k(\mathbf{k}) = \int\limits_{\mathbb{R}^n} \underbrace{e^{i\mathbf{kx}}}_{\cos(\mathbf{kx})+i\sin(\mathbf{kx})}\ d\mathcal{P}_n(\mathbf{x}) \overset{k\to\infty}{\longrightarrow} \int\limits_{\mathbb{R}^n} e^{i\mathbf{kx}} d\mathcal{P}(\mathbf{x}) = \widehat{\mathcal{P}}(\mathbf{k})$$

□

4.6.4 Lemma über gleichmäßig Radonsche Wahrscheinlichkeitsmaße

Seien $(\mathcal{P}_m)_{m=1}^\infty$ gleichmäßig Radonsche Wahrscheinlichkeitsmaße auf $\mathbb{R}^n$ mit charakteristischen Funktionen $\widehat{\mathcal{P}}_m$ und $h : \mathbb{R}^n \to \mathbb{C}$ eine Funktion mit

$$\lim_{m\to\infty} \widehat{\mathcal{P}}_m(\mathbf{k}) = h(\mathbf{k}) \quad \forall\ \mathbf{k} \in \mathbb{R}^n$$

Dann ist h charakteristische Funktion eines Wahrscheinlichkeitsmaßes $\mathcal{P}$ und es gehen

$$\mathcal{P}_m \underset{\text{weak}}{\overset{m\to\infty}{\longrightarrow}} \mathcal{P}$$

Beweis: Nach Helly 4.5.9 existiert ein Wahrscheinlichkeitsmaß $\mathcal{P}$ und Teilfolge $(\mathcal{P}_{m_k})_{k=1}^{\infty}$ mit $\mathcal{P}_{m_k} \underset{\text{weak}}{\overset{k\to\infty}{\longrightarrow}} \mathcal{P}$. Nach Lemma 4.6.3 gehen

$$h(\mathbf{k}) = \lim_{k\to\infty} \widehat{\mathcal{P}}_{m_k}(\mathbf{k}) = \widehat{\mathcal{P}}(\mathbf{k}) \quad \forall\ \mathbf{k} \in \mathbb{R}^n$$

woraus folgt $\widehat{\mathcal{P}} = h$, sprich h ist charakteristische Funktion von $\mathcal{P}$. Machen die <u>Annahme</u> $\mathcal{P}_n \underset{\text{weak}}{\overset{n\to\infty}{\not\longrightarrow}} \mathcal{P}$, dann existiert $f \in \mathcal{C}_b(\mathbb{R}^n)$ mit

$$\int_{\mathbb{R}} f\ d\mathcal{P}_m \overset{m\to\infty}{\not\longrightarrow} \int_{\mathbb{R}} f\ d\mathcal{P}$$

sprich es existiert ein $\varepsilon > 0$ und Teilfolge $(\mathcal{P}_{m'_k})_{k=1}^{\infty}$ mit

$$\left| \int_{\mathbb{R}^n} f\ d\mathcal{P}_{m'_k} - \int_{\mathbb{R}^n} f\ d\mathcal{P} \right| \geq \varepsilon \quad \forall\ k \in \mathbb{N} \tag{4.6.4.1}$$

Doch auch $(\mathcal{P}_{m'_k})$ sind gleichmäßig Radonsch, so dass nach Helly 4.5.9 ein Wahrscheinlichkeitsmaß $\mathcal{Q}$ und Teilfolge $(\mathcal{P}_{m'_{k_i}})_{i=1}^{\infty}$ mit $\mathcal{P}_{m'_{k_i}} \underset{\text{weak}}{\overset{i\to\infty}{\longrightarrow}} \mathcal{Q}$ existieren, sprich nach Lemma 4.6.3

$$\widehat{\mathcal{P}}_{m'_{k_i}}(\mathbf{k}) \overset{i\to\infty}{\longrightarrow} \widehat{\mathcal{Q}}(\mathbf{k}) \quad \forall\ \mathbf{k} \in \mathbb{R}$$

Doch dies impliziert auch $\widehat{\mathcal{Q}} = h$, und nach Theorem 4.3.6 $\mathcal{Q} = \mathcal{P}$, ein Widerspruch zu (4.6.4.1).
□

4.6.5 Stetigkeitssatz von Lévy-Cramér

Seien $(\mathcal{P}_n)_{n=1}^{\infty}$ Wahrscheinlichkeitsmaße auf $\mathbb{R}^n$ mit charakteristischen Funktionen $(\widehat{\mathcal{P}}_n)_{n=1}^{\infty}$ und $h : \mathbb{R}^n \to \mathbb{C}$ so dass

$$\widehat{\mathcal{P}}_n(\mathbf{k}) \overset{n\to\infty}{\longrightarrow} h(\mathbf{k}) \quad \forall\ \mathbf{k} \in \mathbb{R}^n$$

Dann sind folgende Aussagen äquivalent:

1. h ist stetig in der 0.
2. h ist stetig.
3. $(\mathcal{P}_n)_{n=1}^{\infty}$ sind gleichmäßig Radonsch.
4. h ist charakteristische Funktion.
5. $\mathcal{P}_n \underset{\text{weak}}{\overset{n\to\infty}{\longrightarrow}} \mathcal{P}$ für irgendein Wahrscheinlichkeitsmaß $\mathcal{P}$. Gegebenfalls ist dann $h = \widehat{\mathcal{P}}$.

Beweis:

(1)⇒(3): Folgt direkt aus 4.6.2.

(3)⇒(5): Folgt direkt aus 4.6.4.

(5)⇒(4): Nach Lemma 4.6.3 gehen

$$\widehat{\mathcal{P}}_n(\mathbf{k}) \overset{n\to\infty}{\longrightarrow} \widehat{\mathcal{P}}(\mathbf{k}) \quad \forall\ \mathbf{k} \in \mathbb{R}^n$$

Doch dies impliziert nach Voraussetzung $h = \widehat{\mathcal{P}}$.

(4)⇒(2): Nach 4.2.3 ist h sogar gleichmäßig stetig.

(2)⇒(1): Trivial.

□

4.6.6 Folgerung: Charakterisierung der schwachen Konvergenz (Lévy)

Seien $(\mathcal{P}_m)_{m=1}^{\infty}, \mathcal{P}$ Wahrscheinlichkeitsmaße auf $\mathbb{R}^n$ mit charakteristischen Funktionen $\widehat{\mathcal{P}}_m, \widehat{\mathcal{P}}$. Dann sind folgende äquivalent:

1. $\mathcal{P}_m \underset{\text{weak}}{\overset{m\to\infty}{\longrightarrow}} \mathcal{P}$
2. $\widehat{\mathcal{P}}_m(\mathbf{k}) \overset{m\to\infty}{\longrightarrow} \widehat{\mathcal{P}}(\mathbf{k}) \quad \forall\ \mathbf{k} \in (\mathbb{R}^n)'$
3. Für $\mathcal{P}_m \circ \mathbf{k}^{-1} \underset{\text{weak}}{\overset{m\to\infty}{\longrightarrow}} \mathcal{P} \circ \mathbf{k}^{-1} \quad \forall\ \mathbf{k} \in (\mathbb{R}^n)'$.

Beweis:

(1)⇒(2): Folgt aus Lemma 4.6.3

(2)⇒(1): Folgt aus 4.6.5 und Theorem 4.3.6

(1)⇔(3): Folgt aus

$$\mathcal{F}(\mathcal{P})(t\mathbf{k}) = \mathcal{F}\left(\mathcal{P} \circ \mathbf{k}^{-1}\right)(t) \quad \forall\ \mathbf{k} \in (\mathbb{R}^n)',\ t \in \mathbb{R}$$

(analog auch für $\mathcal{P}_m$) und der Äquivalenz von (1) & (2).

□

Bemerkung: Siehe auch A.2.4 für mögliche Probleme der Zurückführung des allgemeinen Falls auf den 1-dimensionalen.

Beispiele:

(i) Der Stetigkeitssatz 4.6.5 erweist sich als äußerst nützlich beim Nachweis dass so manche Funktionen $h : \mathbb{R} \to \mathbb{C}$ charakteristische Funktionen sind. Ist nämlich die Stetigkeit von h in der 0 gesichert, so genügt es eine Folge charakteristischer Funktionen $\widehat{\mathcal{P}}_n$ zu finden die $\widehat{\mathcal{P}}_n \underset{\text{punktweise}}{\overset{n\to\infty}{\longrightarrow}} h$ gehen.

(ii) Für $\mathcal{P}_n := \mathcal{N}_{0,\sigma_n^2}$ mit $\sigma_n^2 \searrow 0$ gehen

$$\widehat{\mathcal{P}}_n(t) = e^{-\sigma_n^2 t/2} \overset{n\to\infty}{\longrightarrow} 1 = \widehat{\delta}_0$$

sprich $\mathcal{N}_{0,\sigma_n^2} \underset{\text{weak}}{\overset{n\to\infty}{\longrightarrow}} \delta_0$.

(iii) Seien $\mathcal{P}_n$ jeweils gleichverteilt auf der Menge $\left\{0, \frac{1}{n}, \ldots, \frac{n-1}{n}\right\}$. Dann sind

$$\widehat{\mathcal{P}}_n(t) = \frac{1}{n}\sum_{k=1}^{n-1} e^{it\frac{k}{n}} = \frac{1 - e^{it}}{n\left(1 - e^{it/n}\right)} \overset{n\to\infty}{\longrightarrow} \frac{e^{it} - 1}{it} =: h(t)$$

und h ist stetig in 0, daher charakteristische Funktion. Tatsächlich ist h charakteristische Funktion der Gleichverteilung auf $[0,1]$, wohin die $\mathcal{P}_n$ schwach streben.

4.6.7 Definition: Schwache Konvergenz von Zufallsvariablen

Sei (E,d) ein metrischer Raum und $(X_n), X$ E-wertige, zufällige Größen auf dem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{S}, \mathcal{P})$. Dann konvergieren X_n *schwach* oder *in Verteilung* (*in distribution*) gegen X :⇔
die Bildmaße $\mathcal{P}_{X_n}$ konvergieren schwach gegen das Bildmaß $\mathcal{P}_X$, das heißt

$$\int_\Omega f(X_n)\ d\mathcal{P} \overset{n\to\infty}{\longrightarrow} \int_\Omega f(X)\ d\mathcal{P} \quad \forall\ f \in \mathcal{C}_b(E)$$

Man schreibt auch $X_n \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} X$.

Bemerkungen:

(i) Sind $X_n \overset{\text{d.}}{=} Y_n$ so gilt

$$X_n \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} X \quad \Leftrightarrow \quad Y_n \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} X$$

Sind $X \overset{\text{d.}}{=} Y$ so gilt

$$X_n \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} X \quad \Leftrightarrow \quad X_n \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} Y$$

Dies gibt Anlass zur alternativen Schreibweise $X_n \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} \mathcal{P}_X$.

(ii) Die schwache Konvergenz von Zufallsgrößen ist mit Vorsicht zu genießen! Z.B. folgt aus

$$X_n \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} X,\ Y_n \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} Y$$

nicht notwendigerweise $(X_n + Y_n) \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} (X+Y)$. Als Beispiel diene die Bernoulli Folge

$$\mathcal{P}(X_n = \pm 1) = \frac{1}{2} \quad , \quad \mathcal{P}(X = \pm 1) = \frac{1}{2}$$

Dann gehen zwar $(-X_n), X_n \underset{d}{\overset{n\to\infty}{\longrightarrow}} X$, doch

$$X_n - X_n = 0 \underset{\text{d}}{\overset{n\to\infty}{\not\longrightarrow}} 2X$$

4.6.8 Satz: Charakterisierung der schwachen Konvergenz von Zufallsvariablen

Sei (E,d) ein metrischer Raum, dazu die E-wertigen Zufallsvariablen X_n, X. Dann sind folgende Aussagen äquivalent:

1. $X_n \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} X$
2. $\overline{\lim\limits_{n\to\infty}}\ \mathcal{P}(X_n \in A) \leq \mathcal{P}(X \in A) \quad \forall\ A \subseteq E$ abgeschlossen
3. $\underline{\lim\limits_{n\to\infty}}\ \mathcal{P}(X_n \in G) \geq \mathcal{P}(X \in G) \quad \forall\ G \subseteq E$ offen
4. $\lim\limits_{n\to\infty} \mathcal{P}(X_n \in B) = \mathcal{P}(X \in B) \quad \forall\ B \in \mathscr{B}(E)$ mit $\mathcal{P}(X \in \partial B) = 0$.
5. Im Falle $E = \mathbb{R} : \mathcal{P}(X_n \leq t) \overset{n\to\infty}{\longrightarrow} \mathcal{P}(X \leq t) \ \ \forall\ t \in \mathbb{R}$ mit $\mathcal{P}(X = t) = 0$.
6. Im Falle $E = \mathbb{R}^m : \widehat{X}_n(\mathbf{k}) \overset{n\to\infty}{\longrightarrow} \widehat{X}(\mathbf{k}) \ \ \forall\ \mathbf{k} \in \mathbb{R}^m$
7. Im Falle $E = \mathbb{R}^m : \langle \mathbf{k}, X_n \rangle \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} \langle \mathbf{k}, X \rangle \ \ \forall\ \mathbf{k} \in \mathbb{R}^m$

Beweis: Siehe Theorem 4.4.4 für (1), (2), (3) und (4). Siehe 4.4.7 für (5). Siehe 4.6.6 für (6) & (7).

4.6.9 Satz: Übertragung der schwachen Konvergenz auf Summen

Seien $X, Y, X_k, Y_k,\ k \in \mathbb{N}\ \mathbb{R}^n$-wertige Zufallsgrößen mit X, Y unabhängig und jedes Paar X_k, Y_k unabhängig. Gehen

$$X_k \underset{\text{d}}{\overset{k\to\infty}{\longrightarrow}} X \quad , \quad Y_k \underset{\text{d}}{\overset{k\to\infty}{\longrightarrow}} Y$$

so gehen auch

$$(X_k + Y_k) \underset{\text{d}}{\overset{k\to\infty}{\longrightarrow}} (X+Y)$$

Beweis: Aus Charakterisierungssatz 4.6.8 und

$$\mathcal{F}(X_k + Y_k)(\mathbf{k}) = \widehat{X}_k(\mathbf{k}) \cdot \widehat{Y}_k(\mathbf{k}) \overset{k\to\infty}{\longrightarrow} \widehat{X}(\mathbf{k}) \cdot \widehat{Y}(\mathbf{k}) = \mathcal{F}(X+Y)(\mathbf{k}) \quad , \quad \mathbf{k} \in \mathbb{R}^n$$

folgt die Behauptung.
□

Beispiel: Betrachtet sei die Folge reeller, unabhängiger Zufallsgrößen X_k mit Verteilung

$$\mathcal{P}(X_k = \pm 1) = \frac{1}{2} \quad , \quad k \in \mathbb{N}$$

(Bernoulli-Folge) und charakteristischer Funktion

$$\widehat{X}_k(t) = \cos(t) \quad , \quad t \in \mathbb{R}$$

Dazu die Partialsummen

$$S_n := \sum_{k=1}^{n} X_k \quad , \quad n \in \mathbb{N}$$

Dann konvergieren

$$\widehat{S}_n(t) = \cos^n(t)$$

nicht, was erwartungsgemäß impliziert dass auch die S_n selber nirgends schwach konvergieren. Doch dafür gehen

$$\mathcal{F}\left(\frac{S_n}{\sqrt{n}}\right)(t) = \widehat{S}_n\left(\frac{t}{\sqrt{n}}\right) = \cos^n\left(\frac{t}{\sqrt{n}}\right) = \left[1 - \frac{t^2}{2n} + \mathcal{O}\left(\frac{1}{n^2}\right)\right]^n \underset{A.5.3}{\overset{n\to\infty}{\longrightarrow}} e^{-\frac{t^2}{2}} = \widehat{\mathcal{N}}_{0,1}$$

sprich $\frac{S_n}{\sqrt{n}} \underset{\mathrm{d}}{\overset{n\to\infty}{\longrightarrow}} \mathcal{N}_{0,1}$.

4.6.10 Weitere Permanenzeigenschaften der Konvergenz in Verteilung

1. Sind $X, X_1, X_2, ..$ & $Y_1, Y_2, ..$ reelle Zufallsvariablen und $c \in \mathbb{R}$ mit $X_n \underset{\mathrm{d}}{\overset{n\to\infty}{\longrightarrow}} X$ und $Y_n \underset{\mathrm{d}}{\overset{n\to\infty}{\longrightarrow}} c$, so gehen auch

$$(X_n + Y_n) \underset{\mathrm{d}}{\overset{n\to\infty}{\longrightarrow}} X + c$$

2. Sind $X, X_1, X_2, ...$ reelle Zufallsvariablen mit $X_n \underset{\mathrm{d}}{\overset{n\to\infty}{\longrightarrow}} X$ und $(a_n) \subseteq \mathbb{R}$ mit $a_n \overset{n\to\infty}{\longrightarrow} a$, so gehen auch

$$a_n X_n \underset{\mathrm{d}}{\overset{n\to\infty}{\longrightarrow}} aX$$

3. Sind (E_1, d_1), (E_2, d_2) metrische Räume, $f : E_1 \to E_2$ stetig und $X, X_1, X_2, ..$ E_1-wertige Zufallsvariablen mit $X_n \underset{\mathrm{d}}{\overset{n\to\infty}{\longrightarrow}} X$, so gehen auch

$$f \circ X_n \underset{\mathrm{d}}{\overset{n\to\infty}{\longrightarrow}} f \circ X$$

5 Grenzverteilungen

5.1 Poissonscher Grenzwertsatz

5.1.1 Motivationsbeispiel

Betrachtet sei eine Sammlung an (nicht fairen) Münzen $(M_n)_{n\in\mathbb{N}}$ die jeweils mit Wahrscheinlichkeit p_n bzw. $(1-p_n)$ den Wert 1 bzw. 0 ergeben. Dabei gehen $p_n \overset{n\to\infty}{\longrightarrow} 0$. Dies sei kompensiert dadurch, dass jede Münze M_n genau k_n mal geworfen wird, wobei $k_n p_n \overset{n\to\infty}{\longrightarrow} \lambda > 0$. Bezeichnet man mit $M_{n,j},\ j = 1,..,k_n$ die Ergebnisse des j-ten Wurfs der n-ten Münze, so sind die $M_{n,j}$ unabhängig, Binomialverteilt $M_{n,j} \sim \mathcal{B}_{1,p_n}$, bzw. die Summen

$$S_n := \sum_{j=1}^{k_n} M_{n,j} \quad , \quad n \in \mathbb{N}$$

unabhängig, Binomialverteilt $S_n \sim \mathcal{B}_{k_n,p_n}$, sprich

$$\mathcal{P}(S_n = k) = \binom{k_n}{k} p_n^k \cdot (1-p_n)^{k_n-k}$$

Nun stellt sich die Frage, wie sich die S_n für $n \to \infty$ verhalten.

5.1.2 Poissonscher Grenzwertsatz

Seien $(p_n)_{n\in\mathbb{N}} \subseteq [0,1]$ und $(k_n)_{n\in\mathbb{N}} \subseteq \mathbb{N}$ so dass

$$k_n \overset{n\to\infty}{\longrightarrow} \infty \quad , \quad k_n p_n \overset{n\to\infty}{\longrightarrow} \lambda \geq 0$$

und $\mathcal{B}_{k_n,p_n}$ die Binomialverteilungen mit Parametern k_n, p_n. Dann konvergieren

$$\mathcal{B}_{k_n,p_n} \underset{\text{weak}}{\overset{n\to\infty}{\longrightarrow}} \pi_\lambda$$

schwach gegen die Poissonverteilung π_λ mit Parameter λ.

Beweis: Sei $\widehat{\pi}_\lambda$ die charakteristische Funktion der Poissonverteilung mit Parameter λ, das heißt

$$\widehat{\pi}_\lambda(t) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} e^{ikt} = \exp\left[\lambda e^{it} - \lambda\right] \quad , \quad t \in \mathbb{R}$$

Dann gehen die charakteristischen Funktionen

$$\widehat{\mathcal{B}}_{k_n,p_n}(t) = \prod_{j=1}^{k_n} \widehat{\mathcal{B}}_{1,p_n}(t) = \left[(1-p_n) + p_n e^{it}\right]^{k_n} = \left[1 - \frac{1}{k_n} \cdot \underbrace{p_n k_n \left(1 - e^{it}\right)}_{\overset{n\to\infty}{\longrightarrow} \lambda(1-e^{it})}\right]^{k_n} \underset{A.5.3}{\overset{n\to\infty}{\longrightarrow}} \exp\left[\lambda\left(e^{it} - 1\right)\right] = \widehat{\pi}_\lambda(t)$$

punktweise gegen $\widehat{\pi}_\lambda$ und nach Lévy-Cramér 4.6.5

$$\mathcal{B}_{k_n,p_n} \underset{\text{weak}}{\overset{n\to\infty}{\longrightarrow}} \pi_\lambda$$

□

Beispiele:

(i) Eine Versicherung versichere $N \gg 1$ Hausbesitzer, denen jeweils eine geringe jährliche Schadenwahrscheinlichkeit $p \ll 1$ entspricht, $Np =: \lambda > 0$. Die Anzahl S an tatsächlich vorkommenden Schäden pro Jahr ist dann Binomialverteilt $S \sim \mathcal{B}_{N,p}$, und für hinreichend großes N *annähernd* Poissonverteilt $\sim \pi_\lambda$.

(ii) Die Lebensdauer $N \geq 1$ Atome sei exponentialverteilt $\sim \text{Exp}_\mu$ mit Parameter $0 < \mu \ll 1$, das heißt, die Wahrscheinlichkeit, dass ein bestimmtes Atom im Zeitintervall $[0,T]$ zerfällt, ist durch $p := 1 - e^{-\mu T} \approx \mu T$ gegeben.

Die Anzahl der in $[0,T]$ tatsächlich zerfallenen Atome ist dementsprechend Binomialverteilt $\sim \mathcal{B}_{N,p}$ bzw. für genügend großes N annähernd Poissonverteilt $\sim \pi_{Np}$. Die Wahrscheinlichkeit dass k Atome zerfallen ist insbesondere annähernd gegeben durch $\frac{(N\mu T)^k}{k!} e^{-N\mu T}$.

5.2 Der Zentrale Grenzwertsatz (CLT)

5.2.1 Vorbetrachtung

Für reelle, iid Zufallsvariablen $X_1, X_2, ..$ mit $\mathbb{E}|X_1|^2 < \infty$ sei gefragt nach dem asymptotischen Verhalten der Partialsummen

$$S_n := \sum_{k=1}^{n} X_k$$

für $n \to \infty$. Bekanntlich gehen im Fall $\mathbb{E}X_n > 0$

$$S_n \underset{\text{f.s.}}{\overset{n\to\infty}{\longrightarrow}} \infty$$

(vgl. Abschnitt 3.9.1).

Im Fall $\mathbb{E}X_1 = 0$ (o.B.d.A. $\mathcal{P}(X_1 = 0) < 1$) gilt zwar stets $\mathbb{E}S_n = 0$, dennoch gehen die Varianzen $\mathbb{V}S_n = n \cdot \mathbb{V}X_1 \overset{n\to\infty}{\longrightarrow} \infty$ und die Summen S_n nehmen fast-sicher beliebig große Werte an (vgl. (3.9.1.1)). Interpretationsgemäß *verbreitet sich die Verteilung der Summen S_n allmählich auf ganz* $\mathbb{R}$ (vgl. Abb. (15)).

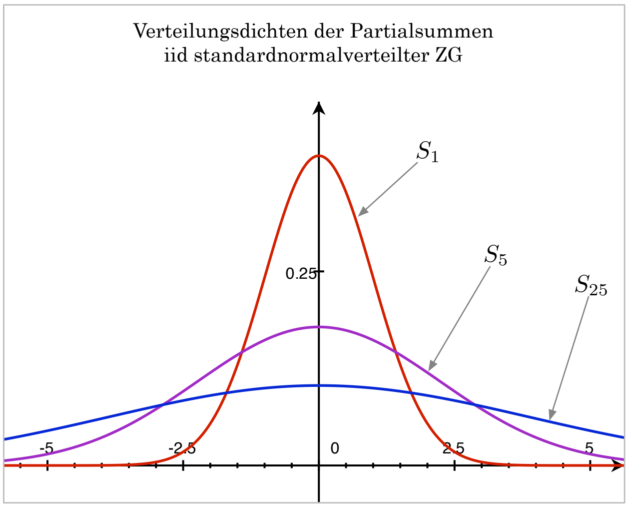

Abbildung 15: Verteilugsdichten der Partialsummen S_1, S_5 und S_{25} standardnormalverteilter Zufallsgrößen.

Es stellt sich nun die Frage, was passiert wenn letztere entsprechend *normiert* werden. Setzt man z.B.

$$Z_n := \frac{S_n - \mathbb{E}S_n}{\sqrt{\mathbb{V}S_n}} = \frac{S_n - n \cdot \mathbb{E}X_1}{\sqrt{n \cdot \mathbb{V}X_1}} \quad , \quad n \in \mathbb{N}$$

so sind $\mathbb{E}Z_n = 0$ und $\mathbb{V}Z_n = 1$, doch ist damit das exakte Verhalten der Z_n für $n \to \infty$ noch lange nicht geklärt.

5.2.2 Der Zentrale Grenzwertsatz für reelle, identisch verteilte Zufallsgrößen (CLT)

Seien $X_1, X_2, ..$ reelle, iid Zufallsvariablen mit $\mathbb{E}|X|^2 < \infty$, dazu $S_n := \sum_{k=1}^{n} X_k$. Dann konvergieren

$$\frac{S_n - \mathbb{E}S_n}{\sqrt{\mathbb{V}S_n}} = \frac{S_n - n \cdot \mathbb{E}X_1}{\sqrt{n \cdot \mathbb{V}X_1}} \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} \mathcal{N}_{0,1}$$

in Verteilung gegen die Standardnormalverteilung $\mathcal{N}_{0,1}$.

Beweis: Es genügt die Behauptung für $\mathbb{E}X_1 = 0$ und $\mathbb{V}X_1 = 1$ zu zeigen. Betrachten dazu die charakteristischen Funktionen

$$\mathcal{F}\left(\tfrac{S_n}{\sqrt{n}}\right)(t) = \widehat{S}_n\left(\tfrac{t}{\sqrt{n}}\right) = \left[\widehat{X}_1\left(\tfrac{t}{\sqrt{n}}\right)\right]^n \overset{4.2.6}{=} \left[1 + \frac{it}{\sqrt{n}}\underbrace{\mathbb{E}X_1}_{0} - \frac{t^2}{2n}\underbrace{\mathbb{E}\,|X_1|^2}_{1} + \frac{t^2}{2n}\cdot\Theta\left(\tfrac{t}{\sqrt{n}}\right)\right]^n \quad \Big| \quad \Theta \in o(1)$$

$$= \left[1 - \frac{1}{n}\frac{t^2}{2}\underbrace{\left(1 - \Theta\left(\tfrac{t}{\sqrt{n}}\right)\right)}_{\overset{n\to\infty}{\longrightarrow} 1}\right]^n \underset{A.5.3}{\overset{n\to\infty}{\longrightarrow}} e^{-\frac{t^2}{2}} = \mathcal{F}(\mathcal{N}_{0,1})(t) \quad , \quad t \in \mathbb{R}$$

Nach Lévy 4.6.6 folgt schließlich die Behauptung.
□

Folgerungen:

(i) Nach 4.6.10 gehen

$$\frac{1}{\sqrt{n}}\cdot(S_n - n\cdot\mathbb{E}X_1) \underset{d}{\overset{n\to\infty}{\longrightarrow}} \mathcal{N}_{0,\mathbb{V}X_1}$$

(ii) Ist Φ die Verteilungsfunktion der Standardnormalverteilung $\mathcal{N}_{0,1}$, sprich

$$\Phi(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-x^2/2}dx \quad , \quad t \in \mathbb{R}$$

so gehen nach dem CLT und 4.4.8

$$\sup_{t\in\mathbb{R}}\left|\mathcal{P}\left(\tfrac{S_n - n\mathbb{E}X_1}{\sqrt{n\mathbb{V}X_1}} \le t\right) - \Phi(t)\right| \overset{n\to\infty}{\longrightarrow} 0$$

Insbesondere gehen für Folgen $\alpha_n \le \beta_n,\ n \in \mathbb{N}$

$$\left|\mathcal{P}\left(\alpha_n \le \tfrac{S_n - n\mathbb{E}X_1}{\sqrt{n\mathbb{V}X_1}} \le \beta_n\right) - (\Phi(\beta_n) - \Phi(\alpha_n))\right| \overset{n\to\infty}{\longrightarrow} 0$$

Speziell für

$$\alpha_n := \frac{\gamma - n\mathbb{E}X_1}{\sqrt{n\mathbb{V}X_1}} \quad , \quad \beta_n := \frac{\delta - n\mathbb{E}X_1}{\sqrt{n\mathbb{V}X_1}}$$

gehen

$$\left|\mathcal{P}(\gamma \le S_n \le \delta) - \frac{1}{\sqrt{2\pi n\mathbb{V}X_1}}\int_{\gamma}^{\delta}\exp\left[-\frac{(x - n\mathbb{E}X_1)^2}{2n\mathbb{V}X_1}\right]dx\right| \overset{n\to\infty}{\longrightarrow} 0$$

Interpretationsgemäß ist S_n für genügend große n *annähernd* $\mathcal{N}_{n\mathbb{E}X_1,n\mathbb{V}X_1}$-verteilt!

(iii) Im Falle $\mathbb{E}X_1 = 0,\ \mathbb{V}X_1 < \infty$ geht

$$\lim_{n\to\infty}\mathcal{P}(S_n \le 0) = \lim_{n\to\infty}\mathcal{P}\left(\frac{S_n}{\sqrt{n\mathbb{V}X_1}} \le 0\right) = \Phi(0) = \frac{1}{2}$$

Beachte: Die Forderung $\mathbb{V}X_1 < \infty$ ist tatsächlich notwendig, da andernfalls Gegenbeispiele existieren!

Beispiel: Bei Banken werden nach-Komma-Stellen bei Cent gerundet. Sei X_k der k-te rundungsbedingte Verlust bzw. Gewinn der Bank und $S_n = \sum_{k=1}^n X_k$ der nach n solchen Rundungen Gesamt- Verlust bzw. Gewinn. Da die $X_1, X_2, ..$ unabhängig auf $[-0.5, 0.5]$ gleichverteilt sind mit $\mathbb{V}X_1 = 1/12$, gilt für $\alpha \le \beta$ und genügend großes $n \in \mathbb{N}$

$$\mathcal{P}\left(\alpha \le S_n\cdot\sqrt{\frac{12}{n}} \le \beta\right) \approx \frac{1}{\sqrt{2\pi}}\int_{\alpha}^{\beta} e^{-\frac{x^2}{2}}dx$$

So ist z.B für $n = 100$,

$$\mathcal{P}\left(|S_n| \leq 5.77\right) \approx \mathcal{P}\left(-2 \leq S_n \cdot \sqrt{\frac{12}{100}} \leq 2\right) \approx \frac{1}{\sqrt{2\pi}} \int\limits_{-2}^{2} e^{-\frac{x^2}{2}}\, dx \approx 0.9545$$

Anderseits würde sich nach Tschebyschow 1.8.3(ii) lediglich die Abschätzung

$$\mathcal{P}\left(|S_n| \leq 5.77\right) \geq 1 - \frac{\mathbb{V}S_n}{(5.77)^2} \approx 0.7435$$

ergeben.

5.2.3 Verallgemeinerung des CLT auf nicht-identisch verteilte Zufallsgrößen

Seien $(k_n) \subseteq \mathbb{N}$ mit $k_n \to \infty$ und für jedes $n \in \mathbb{N}$ $X_{n,1}, .., X_{n,k_n}$ reelle, unabhängige Zufallsgrößen. Dazu seien

$$S_n := \sum_{j=1}^{k_n} X_{n,j}$$

und

$$a_n := \mathbb{E}S_n = \sum_{j=1}^{k_n} \mathbb{E}X_{n,j} \quad , \quad \sigma_n^2 := \mathbb{V}S_n = \sum_{j=1}^{k_n} \mathbb{V}X_{n,j}$$

Dann stellen sich folgende Fragen:

- Wann konvergieren $\frac{S_n - a_n}{\sigma_n}$ in Verteilung?
- Was kann man gegebenfalls über die Verteilung des Grenzwertes sagen?
- Wann gilt der CLT, sprich, wann gehen $\frac{S_n - a_n}{\sigma_n} \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} \mathcal{N}_{0,1}$?

Tatsächlich existieren bestimmte Bedingungen, die den CLT implizieren [Feller-Lindeberg, Lyapunov].

Allgemein auszuschließen ist bei obigen Fragen, dass irgendein *Fehler* gegenüber der anderen *groß* ist. Speziell fordert man, dass das *Serienschema* (*Dreiecksschema*)

$$\begin{array}{l} X_{1,1} \dots X_{1,k_1} \\ X_{2,1} \dots\dots X_{2,k_2} \\ X_{3,1} \dots\dots\dots X_{3,k_3} \\ \vdots \qquad\qquad\qquad \ddots \end{array}$$

asymptotisch klein ist, das heißt

$$\lim_{n\to\infty} \max_{1\leq j\leq k_n} \mathcal{P}\left(|X_{n,j} - \mathbb{E}X_{n,j}| \geq \varepsilon\sigma_n\right) = 0 \quad \forall\, \varepsilon > 0$$

Ein triviales Beispiel eines nicht-asymptotisch-kleinen Schemas wären $X_{n,1} \neq 0,\ X_{n,2}, .., X_{n,k_n} = 0$.

Beispiel: Die unabhängigen Zufallsgrößen $X_{n,1}, .., X_{n,k_n} \sim \mathcal{B}_{1,p_n}$ seien Binomialverteilt, wobei $p_n \cdot k_n \overset{n\to\infty}{\longrightarrow} \lambda$. Dann ist wegen

$$\mathcal{P}\left(|X_{n,j} - p_n| \geq \varepsilon\sqrt{k_n p_n(1-p_n)}\right) = \mathcal{P}\left(|X_{n,j}| = 1\right) = p_n \overset{n\to\infty}{\longrightarrow} 0$$

das Dreiecksschema asymptotisch klein, doch nach Poisson 5.1.2 gehen

$$\frac{S_n - a_n}{\sigma_n} \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} \frac{S - \lambda}{\sqrt{\lambda}} \quad , \quad S \sim \pi_\lambda$$

Das Serienschema konvergiert also tatsächlich, jedoch **nicht** gegen die Normalverteilung!

5.2.4 Lindeberg-Bedingung zum CLT

Seien $X_{n,1},..,X_{n,k_n}$ reelle, für jedes $n\in\mathbb{N}$ unabhängige Zufallsgrößen mit $k_n \overset{n\to\infty}{\longrightarrow} \infty$ und

$$S_n := \sum_{j=1}^{k_n} X_{n,j} \quad , \quad \sigma_n^2 := \mathbb{V}S_n$$

Für jedes $\delta > 0$ gelte

$$\lim_{n\to\infty} \frac{1}{\sigma_n^2} \cdot \sum_{j=1}^{k_n} \mathbb{E}\left[\left|X_{n,j} - \mathbb{E}X_{n,j}\right|^2 \cdot 1_{\{|X_{n,j}-\mathbb{E}X_{n,j}|\geq\delta\sigma_n\}}\right] = 0$$

Dann erfüllen die $X_{n,j}$ den CLT, d.h. es gilt

$$\frac{S_n - \mathbb{E}S_n}{\sqrt{\mathbb{V}S_n}} \underset{\mathrm{d}}{\overset{n\to\infty}{\longrightarrow}} \mathcal{N}_{0,1}$$

5.2.5 Der CLT und der empirische Mittelwert

Seien $x_1,..,x_n$ die unabhängigen Messwerte einer zufällig verteilten Größe, jeweils identisch verteilt gemäß $X_1,..,X_n$. Dann erweist sich der empirische Mittelwert

$$\widetilde{a} := \frac{1}{n}\sum_{k=1}^{n} x_k$$

der Stichprobe als *beste* Schätzung des Erwartungswertes $a := \mathbb{E}X_1$. Einerseits ist dieser verteilt gemäß

$$\widehat{a} \overset{\mathrm{d.}}{=} \frac{1}{n}\sum_{k=1}^{n} X_k =: \frac{S_n}{n}$$

mit $\mathbb{E}\widehat{a} = a$. Anderseits ist er für genügend große n annähernd gemäß $\mathcal{N}_{a,\mathbb{V}X_1/n}$ verteilt, wobei $\mathcal{N}_{a,\mathbb{V}X_1/n} \underset{\mathrm{d}}{\overset{n\to\infty}{\longrightarrow}} \delta_a$.

5.2.6 Fehlerabschätzung beim CLT

Die reellen Zufallsgrößen $X_1,X_2,..$ seien unabhängig identisch verteilt mit $\mathbb{E}\left|X_1\right|^3 < \infty$, dazu $S_n := \sum\limits_{k=1}^{n} X_k$. Dann existiert eine universelle (vom Verteilungsgesetzt unabhängige!) Konstante C_0 für die gilt

$$\sup_{t\in\mathbb{R}}\left|\mathcal{P}\left(\frac{S_n - \mathbb{E}S_n}{\sqrt{\mathbb{V}S_n}} \leq t\right) - \Phi(t)\right| \leq \frac{C_0}{\sqrt{n}} \cdot \frac{\mathbb{E}\left|X_1 - \mathbb{E}X_1\right|^3}{(\mathbb{V}X_1)^{\frac{3}{4}}}$$

[Berry 1941, Erséen 1945]. Beachtenswert ist dabei, dass der optimale (minimale) Wert von C_0 bisher noch nicht bekannt ist. Der aktuelle Stand liegt bei etwa $0.41 \leq C_0 \leq 0.80$. Anderseits ist $n^{-\frac{1}{2}}$ tatsächlich die optimalste Universellabschätzung[18]

Als Beispiel diene eine sogenannte Bernoullifolge $X_1,X_2,..$, sprich

$$\mathcal{P}(X_j = -1) = \mathcal{P}(X_j = 1) = \frac{1}{2}$$

dazu die Partialsummen $S_n := \sum_{k=1}^{n} X_k$ mit $\mathbb{E}S_n = 0,\ \mathbb{V}S_n = n$. Dann gilt

$$\mathcal{P}(S_{2n} = 0) = \binom{2n}{n} 2^{-2n} \overset{\text{Stirling}}{\sim} \sqrt{\frac{2}{\pi}} \cdot \frac{1}{\sqrt{n}}$$

Für $0 < \varepsilon < 1$ gilt dann

$$\mathcal{P}(S_{2n} = 0) = \mathcal{P}(S_{2n} < \varepsilon) - \mathcal{P}(S_{2n} < -\varepsilon)$$

[18] Für Spezialfälle existieren natürlich gegebenfalls optimalere Konvergenzgeschwindigkeiten.

und für genügend großes $n_0 \in \mathbb{N}$:

$$\frac{1}{2}\sqrt{\frac{2}{\pi}} \cdot \frac{1}{\sqrt{n}} \leq |\mathcal{P}(S_{2n} = 0) - \mathcal{P}(S_{2n} > -\varepsilon)| = \left|\mathcal{P}\left(\frac{S_{2n}}{\sqrt{2n}} < \frac{\varepsilon}{\sqrt{2n}}\right) - \mathcal{P}\left(\frac{S_{2n}}{\sqrt{2n}} < \frac{-\varepsilon}{\sqrt{2n}}\right)\right|$$

$$\leq \left|\mathcal{P}\left(\frac{S_{2n}}{\sqrt{2n}} < \frac{\varepsilon}{\sqrt{2n}}\right) - \Phi\left(\frac{\varepsilon}{\sqrt{2n}}\right)\right| + \left|\mathcal{P}\left(\frac{S_{2n}}{\sqrt{2n}} < \frac{-\varepsilon}{\sqrt{2n}}\right) - \Phi\left(\frac{-\varepsilon}{\sqrt{2n}}\right)\right| + \underbrace{\left|\Phi\left(\frac{\varepsilon}{\sqrt{2n}}\right) - \Phi\left(\frac{-\varepsilon}{\sqrt{2n}}\right)\right|}_{=:\delta(\varepsilon)}$$

$$\leq 2\sup_{t\in\mathbb{R}}\left|\mathcal{P}\left(\frac{S_{2n}}{\sqrt{2n}} \leq t\right) - \Phi(t)\right| + \underbrace{\delta(\varepsilon)}_{\stackrel{\varepsilon\to 0}{\longrightarrow} 0}$$

bzw.

$$\frac{1}{4}\sqrt{\frac{2}{\pi}} \cdot \frac{1}{\sqrt{n}} \leq \sup_{t\in\mathbb{R}}\left|\mathcal{P}\left(\frac{S_{2n}}{\sqrt{2n}} \leq t\right) - \Phi(t)\right|$$

Hieraus ist zu erkennen: $n^{-\frac{1}{2}}$ ist tatsächlich optimalste Schranke!

5.2.7 Spezialfall des CLT [Moivre, Laplace, 1733]

Betrachten die unabhängigen, Binomialverteilten $X_1, X_2, ... \sim \mathcal{B}_{1,p}$, dazu die Partialsummen

$$S_n := \sum_{k=1}^{n} X_k \ \sim \mathcal{B}_{n,p}$$

Dann gilt

$$\mathcal{P}\left(\alpha \leq \frac{S_n - \mathbb{E}S_n}{\sqrt{\mathbb{V}S_n}} \leq \beta\right) = \mathcal{P}\left(\alpha\sqrt{np(1-p)} + np \leq S_n \leq \beta\sqrt{np(1-p)} + np\right)$$

$$= \sum_{\substack{k\in\mathbb{N} \\ \alpha\sqrt{np(1-p)}+np\leq k\leq\beta\sqrt{np(1-p)}+np}} \binom{n}{k} p^k (1-p)^{n-k}$$

$$\underset{\text{M.,L.}}{\stackrel{n\to\infty}{\longrightarrow}} \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{x^2}{2}}\, dx$$

was genau der Aussage des CLT entspricht!

5.3 Der CLT für zufällige Vektoren

5.3.1 Definition: Mehrdimensionale Normalverteilung

Ein reeller, n-dimensionaler zufälliger Vektor $X = (X^1, .., X^n)$ heißt *Normalverteilt* falls für jedes $\mathbf{k} \in \mathbb{R}^n$ die Zufallsgröße $\langle X, \mathbf{k}\rangle$ Normalverteilt ist.

Bemerkungen: Die n- & m dimensionalen, reellen, zufälligen Vektoren X, Y besitzen jeweils Kovarianzmatrizen $R_X := \operatorname{Cov}(X)$, $R_Y := \operatorname{Cov}(Y)$ und Erwartungswerte $\boldsymbol{\mu}_X, \boldsymbol{\mu}_Y$.

(i) Ist X normalverteilt, so ist jede Komponente X^k normalverteilt.

(ii) Sind umgekehrt $X^1, .., X^n$ reell, unabhängig, normalverteilt, so ist der Vektor $X = (X^1, .., X^n)$ normalverteilt.

(iii) Ist X normalverteilt, so besitzt er charakteristische Funktion

$$\widehat{X}(\mathbf{k}) = \exp\left[-\frac{1}{2}\langle R_X\mathbf{k},\mathbf{k}\rangle + i\langle \mathbf{k},\boldsymbol{\mu}_X\rangle\right] \quad , \quad \mathbf{k}\in\mathbb{R}^n$$

(iv) Es gilt sogar: Ist $\boldsymbol{\mu}\in\mathbb{R}^n$ und $R\in\mathbb{R}^{n\times n}$ symmetrisch, nicht-negativ definit, so existiert ein eindeutig verteilter, normalverteilter Zufallsvektor mit Kovarianz R und Erwartungswert $\boldsymbol{\mu}$.
Gegebenfalls schreibt man $\mathcal{N}_{0,R}$ für die entsprechende Verteilung.

(v) Ist X normalverteilt und R_X sogar positiv definit ($R_X > 0$), so besitzt X bzgl. λ^n die Dichte

$$\rho(\mathbf{x}) = \frac{\exp\left[-\frac{1}{2}\left\langle R_X^{-1}(\mathbf{x}-\boldsymbol{\mu}_X),(\mathbf{x}-\boldsymbol{\mu}_X)\right\rangle\right]}{(2\pi)^{\frac{n}{2}}\sqrt{\det(R_X)}}$$

(vi) Sind X,Y unabhängig, normalverteilt so ist auch (X,Y) normalverteilt mit Kovarianzmatrix

$$\mathrm{Cov}((X,Y)) = \begin{pmatrix} R_X & 0 \\ 0 & R_Y \end{pmatrix}$$

(vii) Sind X,Y normalverteilt, so sind sie unkorreliert genau dann wenn sie unabhängig sind!
Insbesondere sind die Komponenten $X^1,..,X^n$ von X unabhängig, genau dann wenn R_X diagonal ist.

5.3.2 Vorbetrachtung des Problems

Gegeben seien die iid, reellen, n-dimensionalen Zufallsvektoren $X_1,X_2,...$, dazu die Partialsummen $S_m := \sum_{k=1}^{m} X_k$. Nach dem eindimensionalen CLT weis man dass

$$\frac{S_m-\mathbb{E}S_m}{\sqrt{m}} = \Bigg(\underbrace{\frac{S_m^1-\mathbb{E}S_m^1}{\sqrt{n}}}_{\xrightarrow[\mathrm{d}]{m\to\infty}\mathcal{N}_{0,\mathbb{V}X_1^1}},\dots,\underbrace{\frac{S_m^n-\mathbb{E}S_m^n}{\sqrt{m}}}_{\xrightarrow[\mathrm{d}]{m\to\infty}\mathcal{N}_{0,\mathbb{V}X_1^n}}\Bigg)$$

komponentenweise, schwach gegen $\mathcal{N}_{0,\mathbb{V}X^i}$ konvergieren. Doch ist dadurch nichts über die Gesamtverteilung des entsprechenden Grenzvektors gesagt.

5.3.3 Der mehrdimensionale Zentrale Grenzwertsatz (CLT)

Seien $X,X_1,X_2,..$ iid reelle, zufällige n-dim. Vektoren mit 2. Moment, dazu die Partialsummen $S_m := \sum_{k=1}^{m} X_k$.
Dann gehen

$$\frac{S_m-\mathbb{E}S_m}{\sqrt{m}} \xrightarrow[\mathrm{d}]{m\to\infty} \mathcal{N}_{0,\mathrm{Cov}(X)}$$

Beweis: Zurückführung auf den 1-dim Fall. Nach 4.6.8(7) genügt es zu zeigen: Für jedes $\mathbf{k}\in(\mathbb{R}^n)'$ gehen[19]

$$\left\langle \frac{S_m-\mathbb{E}S_m}{\sqrt{m}},\mathbf{k}\right\rangle \xrightarrow[\mathrm{d}]{m\to\infty} \mathcal{N}_{0,\mathrm{Cov}(X)}\circ\mathbf{k}^{-1} = \mathcal{N}_{0,\mathrm{Cov}\langle X,\mathbf{k}\rangle}$$

Tatsächlich gilt

$$\left\langle \frac{S_m-\mathbb{E}S_m}{\sqrt{m}},\mathbf{k}\right\rangle = \frac{\langle S_m,\mathbf{k}\rangle - m\langle\mathbb{E}X,\mathbf{k}\rangle}{\sqrt{m}} = \frac{\sum_{k=1}^m \overbrace{\langle X_k,\mathbf{k}\rangle}^{\substack{\text{iid}\\ \text{reell}}} - m\mathbb{E}\langle X,\mathbf{k}\rangle}{\sqrt{m}} \underset{(5.2.2)}{\xrightarrow[\text{weak}]{m\to\infty}} \mathcal{N}_{0,\mathbb{V}\langle X,\mathbf{k}\rangle}$$

□

[19] Beachte dass für $Z\sim\mathcal{N}_{0,\mathrm{Cov}(X)}$ folgt $\langle Z,\mathbf{k}\rangle\sim\mathcal{N}_{0,\mathbb{V}\langle Z,\mathbf{k}\rangle}$.

Bemerke: Sind die Komponenten $X^1, .., X^N$ von X unkorreliert, sprich $\mathrm{Cov}(X)$ diagonal, so sind die Komponenten der Grenzverteilung sogar unabhängig!

Beispiel: Der Vektor $X = (X^1, X^2)$ sei gleichverteilt auf der Einheitskugel $B_1^2 \subseteq \mathbb{R}^2$, sprich

$$\mathcal{P}_X(A) = \frac{\mathrm{Vol}_2(B_1^2 \cap A)}{\pi}$$

Dann besitzen X^1, X^2 jeweils die Verteilungsdichten

$$\rho(t) = \begin{cases} \frac{2}{\pi}\sqrt{1-t^2} & : |t| \leq 1 \\ 0 & : \text{sonst} \end{cases}$$

und der Vektor X Erwartungswert $\mathbb{E}X = 0$ und Kovarianzmatrix

$$\mathrm{Cov}(X) = \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{pmatrix}$$

Sind nun $X_1, X_2, ..$ iid Kopien von X, so gehen

$$\mathcal{P}\left(\frac{\sum_{k=1}^n X_k}{\sqrt{n}} \in B\right) \overset{n\to\infty}{\longrightarrow} \frac{4}{2\pi} \int\limits_B e^{-\frac{1}{2}\left(4t^2+4s^2\right)} dt\ ds$$

für jede $B \subseteq \mathbb{R}^2$ mit $\lambda^2(\partial B) = 0$.

6 Bedingte Verteilung

6.1 Motivation

6.1.1 Die bedingte Wahrscheinlichkeit

Betrachtet sei der Wahrscheinlichkeitsraum $(\Omega, \mathfrak{S}, \mathcal{P})$, dazu vorgegeben $C \in \mathfrak{S}$ mit $\mathcal{P}(C) > 0$. Dann definiert man typischerweise die bzgl. C bedingte Wahrscheinlichkeit von $A \in \mathfrak{S}$ durch

$$\mathcal{P}(A \mid C) := \frac{\mathcal{P}(A \cap C)}{\mathcal{P}(C)} \tag{6.1.1.1}$$

Interpretationsgemäß ist $\mathcal{P}(A \mid C)$ die Wahrscheinlichkeit des Ereignisses A, *unter der Kenntnis* das C eingetroffen ist.

Ist nun $\Omega = C_1 \dot\cup C_2 \dot\cup .. \dot\cup C_n$ eine disjunkte Überdeckung von Ω mit $C_k \in \mathfrak{S},\ \mathcal{P}(C_k) > 0\ \forall\ k$ so ist für jedes $A \in \mathfrak{S}$ die Abbildung

$$Y_A := \sum_{k=1}^{n} \mathcal{P}(A \mid C_k) \cdot 1_{C_k} \quad : \Omega \to \mathbb{R}$$

konstant auf jedem C_k und daher messbar bzgl. der unter-σ-Algebra

$$\mathscr{C} := \sigma(C_1, .., C_n) \subseteq \mathfrak{S}$$

Ferner gilt wegen

$$\int\limits_{C_k} Y_A\ d\mathcal{P} = \mathcal{P}(A \mid C_k) \cdot \mathcal{P}(C_k) = \mathcal{P}(A \cap C_k) = \int\limits_{C_k} 1_A\ d\mathcal{P} \quad \forall\ k = 1, .., n$$

allgemein

$$\int\limits_{C} Y_A\ d\mathcal{P} = \int\limits_{C} 1_A\ d\mathcal{P} \quad \forall\ C \in \mathscr{C}$$

Dabei liefert Y_A für jedes Elementarereignis $\omega \in \Omega$ die Bedingte Wahrscheinlichkeit $\mathcal{P}(A \mid C_k)$ für $\omega \in C_k$, sprich, die Wahrscheinlichkeit A einzutreffen unter der *groben* Kenntnis dass C_k eingetroffen ist.

6.1.2 Probleme dieser Definition der Bedingten Wahrscheinlichkeit

Eine Verallgemeinerung obiger Definition (6.1.1.1) der bedingten Wahrscheinlichkeit $\mathcal{P}(A \mid C)$ auf Nullmengen $C \in \mathfrak{S},\ \mathcal{P}(C) = 0$ ist nicht trivial. Zur Notwendigkeit wird dies jedoch zum Beispiel im Falle zweier Zufallsgrößen X, Y und Ausdrücken der Form

$$\mathcal{P}(X \in A \mid Y = s)$$

sprich, bei der Frage nach der Verteilung von X unter Kenntnis des Wertes von Y.

Besitzt (X, Y) als $\mathbb{R}^2$-wertige Zufallsgröße die Wahrscheinlichkeitsdichte ρ, dazu die Randdichte

$$\rho_r(v) := \int\limits_{\mathbb{R}} \rho(u, v)\ du$$

so ist für $\mathcal{P}_Y$-fast alle $s \in \mathbb{R}$ die Abbildung

$$\rho(u \mid s) := \frac{\rho(u, s)}{\rho_r(s)} \quad , \quad u \in \mathbb{R}$$

wohldefiniert, da ρ_r als Verteilungsdichte von Y $\mathcal{P}_Y$-fast überall ungleich 0 ist. Gegebenfalls ist $\rho(\cdot \mid s)$ eine Wahrscheinlichkeitsdichte, da

$$\int\limits_{\mathbb{R}} \rho(u \mid s)\ du = \frac{1}{\rho_r(s)} \underbrace{\int\limits_{\mathbb{R}} \rho(u, s)\ du}_{\rho_r(s)} = 1$$

Betrachtet man nun eine beliebig kleine Kugel $B_\varepsilon(s)$, erhält man im Kontext der in 6.1.1 formalen Definition der bedingten Wahrscheinlichkeit

$$\mathcal{P}\left(X \in A \mid Y \in B_\varepsilon(s)\right) = \frac{\mathcal{P}(X \in A,\ Y \in B_\varepsilon(s))}{\mathcal{P}(Y \in B_\varepsilon(s))} = \frac{\int\limits_A \int\limits_{B_\varepsilon(s)} \rho(u,v)\ dv\ du}{\int\limits_{B_\varepsilon(s)} \rho_r(v)\ dv}$$

$$= \int\limits_A \underbrace{\frac{\int\limits_{B_\varepsilon(s)} \rho(u,v)\ dv}{\int\limits_{B_\varepsilon(s)} \rho_r(v)\ dv}}_{\xrightarrow[\text{formal}]{\varepsilon\to 0} \frac{\rho(u,s)}{\rho_r(s)}} du \xrightarrow[\text{formal}]{\varepsilon\to 0^+} \int\limits_A \rho(u \mid s)\ du$$

sprich

$$''\ \mathcal{P}(X \in A \mid Y = s) = \int\limits_A \rho(u \mid s)\ du\ ''$$

Ziel des restlichen Kapitels ist die Formalisierung obiger Überlegungen.

6.2 Bedingter Erwartungswert bzgl. einer σ-Algebra

6.2.1 Definition: Bedingter Erwartungswert einer Zufallsgröße

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum, $X : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}^n$ eine Zufallsgröße mit $\mathbb{E}\,\|X\| < \infty$ und $\mathscr{C} \subseteq \mathfrak{S}$ eine Unter-σ-Algebra. Dann heißt eine $\mathscr{C}$-messbare Abbildung $Y : (\Omega, \mathscr{C}) \to \mathbb{R}^n$ *bedingter Erwartungswert von X, gegeben (bzgl.) $\mathscr{C}$*, falls gilt

$$\int\limits_C Y\ d\mathcal{P} = \int\limits_C X\ d\mathcal{P} \quad \forall\ C \in \mathscr{C}$$

Beachte dass Y eine Zufallsgröße auf $(\Omega, \mathscr{C}, \mathcal{P})$ ist.

6.2.2 Theorem: Existenz eines bedingten Erwartungswertes

Zu jeder Zufallsgröße $X : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}^n$ mit $\mathbb{E}\,\|X\| < \infty$ und Unter-σ-Algebra $\mathscr{C} \subseteq \mathfrak{S}$ existiert ein $\mathcal{P}$-fast-eindeutig bestimmter, bedingter Erwartungswert

$$\mathbb{E}(X \mid \mathscr{C}) : (\Omega, \mathscr{C}) \to \mathbb{R}^n$$

von X bzgl. $\mathscr{C}$, sprich

$$\int\limits_C \mathbb{E}(X \mid \mathscr{C})\ d\mathcal{P} = \int\limits_C X\ d\mathcal{P} \quad \forall\ C \in \mathscr{C}$$

Beweisidee: Es genügt die Aussage für den reellen Fall zu zeigen. Ist zunächst $X \geq 0$ so definiert

$$\mu(C) := \int\limits_C X\ d\mathcal{P} \quad , \quad C \in \mathscr{C}$$

ein bzgl. der Einschränkung $\mathcal{P}\big|_{\mathscr{C}}$ absolut stetiges, Maß auf $(\Omega, \mathscr{C})$. Nach Radon-Nikodym existiert eine $\mathcal{P}\big|_{\mathscr{C}}$-fast-eindeutige, $\mathscr{C}$-messbare Funktion $Y : \Omega \to \mathbb{R}$ mit

$$\mu(C) = \int\limits_C Y\ d\mathcal{P}\big|_{\mathscr{C}} = \int\limits_C Y\ d\mathcal{P}$$

was genau dem bedingten Erwartungswert von X bzgl. $\mathscr{C}$ entspricht. Mit Hilfe des Zerlegung $X =: X^+ - X^-$ in positiven bzw. negativen Anteil lässt sich auch der allgemeinere Fall zeigen, wobei benötigt wird $\mathbb{E}\,|X| < \infty$.
$\square$

Bemerkung: Aus dem Beweis wird ersichtlich, dass die Existenz (und fast-Eindeutigkeit) von $\mathbb{E}(X \mid \mathscr{C})$ auch für $\mathbb{E}|X| = \infty$ gesichert ist, insofern $X \geq 0$ fast-sicher!

Beispiele: Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $X : \Omega \to \mathbb{R}^n$ eine Zufallsgröße mit $\mathbb{E}\|X\| < \infty$.

(i) Ist $\mathscr{C} := \sigma(C_1, .., C_n)$ für disjunkte $C_k \in \mathfrak{S}$ mit $\Omega = \bigcup_{k=1}^{n} C_k$, so besitzt X den bedingten Erwartungswert

$$\mathbb{E}(X \mid \mathscr{C}) = \sum_{\substack{k=1 \\ \mathcal{P}(C_k)\neq 0}}^{n} \frac{1_{C_k}}{\mathcal{P}(C_k)} \cdot \int_{C_k} X \, d\mathcal{P} \tag{6.2.2.1}$$

(ii) Ist $\Omega = [0,1]^2,\ \mathfrak{S} = \mathscr{B}([0,1]^2),\ \mathcal{P} = \lambda^2$ und

$$\mathscr{C} := \mathscr{B}([0,1]) \times [0,1] := \{B \times [0,1] : B \in \mathscr{B}([0,1])\}$$

so besitzt X den bedingten Erwartungswert

$$\mathbb{E}(X \mid \mathscr{C})(\mathbf{x}) = \int_0^1 X(x^1, s)\ ds$$

hängt also insbesondere nur von der ersten Komponente ab. Tatsächlich ist $\mathbf{x} \mapsto \int_0^1 X(x^1, s)\ ds$ $\mathscr{C}$-messbar und es gilt

$$\int_{B\times[0,1]} \mathbb{E}(X \mid \mathscr{C})(\mathbf{x})\ d\mathbf{x} = \int_B \int_0^1 \int_0^1 X(x^1, s)\ ds\ dx^2\ dx^1 = \int_B \int_0^1 X(x^1, s)\ ds\ dx^1 = \int_{B\times[0,1]} X\ d\mathbf{x}$$

6.2.3 Eigenschaften des bedingten Erwartungswertes

Seien $X, X_k : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}$ Zufallsgrößen mit $\mathbb{E}|X|, \mathbb{E}|X_k| < \infty,\ k \in \mathbb{N}$ und $\mathscr{C} \subseteq \mathfrak{S}$ eine Unter-σ-Algebra. Dann gilt:

1. $0 \leq X$ f.s. $\Rightarrow$ $\mathbb{E}(X \mid \mathscr{C}) \geq 0$ f.s.

2. $X_1 \leq X_2$ f.s. $\Rightarrow$ $\mathbb{E}(X_1 \mid \mathscr{C}) \leq \mathbb{E}(X_2 \mid \mathscr{C})$ f.s.

3. Für $\alpha_1, \alpha_2 \in \mathbb{R}$ gilt
$$\mathbb{E}(\alpha_1 X_1 + \alpha_2 X_2 \mid \mathscr{C}) \overset{\text{f.s.}}{=} \alpha_1 \mathbb{E}(X_1 \mid \mathscr{C}) + \alpha_2 \mathbb{E}(X_2 \mid \mathscr{C})$$

4. $|\mathbb{E}(X \mid \mathscr{C})| \leq \mathbb{E}(|X| \mid \mathscr{C})$

5. Ist X bereits $\mathscr{C}$-messbar, so gilt
$$\mathbb{E}(X \mid \mathscr{C}) \overset{\text{f.s.}}{=} X$$

6. Sind $0 \leq X_1 \leq X_2 \leq .. \underset{\text{f.s.}}{\longrightarrow} X$ so gehen auch
$$\mathbb{E}(X_k \mid \mathscr{C}) \underset{\text{f.s}}{\overset{k\to\infty}{\longrightarrow}} \mathbb{E}(X \mid \mathscr{C})$$

7. Sind die σ-Algebren $\sigma(X)$ und $\mathscr{C}$ unabhängig, so gilt
$$\mathbb{E}(X \mid \mathscr{C}) \overset{\text{f.s.}}{=} \mathbb{E}X$$

8. Es gilt
$$\mathbb{E}\left[\mathbb{E}(X \mid \mathscr{C})\right] = \mathbb{E}X$$

9. Sind $\mathscr{C}_1 \subseteq \mathscr{C}_2 \subseteq \mathfrak{S}$ σ-Algebren, so gilt
$$\mathbb{E}\left[\mathbb{E}(X \mid \mathscr{C}_2) \mid \mathscr{C}_1\right] = \mathbb{E}(X \mid \mathscr{C}_1)$$

Beweis:

1. Klar (vgl. Beweis von 6.2.2).
2. Analog.
3. Klar.
4. Folgt aus $-|X| \leq X \leq |X|$ und Eigenschaften (2) & (3).
5. Klar.
6. Setzt man $Y_k := \mathbb{E}(X_k \mid \mathscr{C})$ so gilt

$$0 \leq Y_1 \leq Y_2 \leq \ldots \underset{\text{f.s.}}{\overset{k\to\infty}{\longrightarrow}} \sup_{k\in\mathbb{N}} Y_k =: Y$$

Dabei ist Y $\mathscr{C}$-messbar und nach dem Satz über monotone Konvergenz gilt

$$\int_C X \; d\mathcal{P} \overset{\text{M.K.}}{=} \lim_{k\to\infty} \int_C X_k \; d\mathcal{P} = \lim_{k\to\infty} \int_C Y_k \; d\mathcal{P} \overset{\text{M.K.}}{=} \int_C Y \; d\mathcal{P} \quad \forall\, C \in \mathscr{C}$$

das heißt

$$Y \overset{\text{f.s.}}{=} \mathbb{E}(X \mid \mathscr{C})$$

7. Wegen

$$\int_C X \; d\mathcal{P} = \mathbb{E}(\underbrace{1_C \cdot X}_{\text{unabhängig}}) = \mathbb{E}(1_C) \cdot \mathbb{E}(X) = \int_C (\mathbb{E}X) \; d\mathcal{P} \quad \forall\, C \in \mathscr{C}$$

ist

$$\mathbb{E}X \overset{\text{f.s.}}{=} \mathbb{E}(X \mid \mathscr{C})$$

8. Klar.
9. Klar.

□

6.2.4 Die Jensensche Ungleichung

Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $X : (\Omega, \mathfrak{S}, \mathbb{R}) \to I \subseteq \mathbb{R}$ eine Zufallsgröße mit $\mathbb{E}\,|X| < \infty$, $\mathscr{C} \subseteq \mathfrak{S}$ eine Unter-σ-Algebra und $\varphi : I \to \mathbb{R}$ konvex mit $\mathbb{E}\,|\varphi(X)| < \infty$. Dann gilt:

$$\varphi\left[\mathbb{E}\left(X \mid \mathscr{C}\right)\right] \leq \mathbb{E}\left(\varphi(X) \mid \mathscr{C}\right)$$

Beweis: Für jede affine Abbildung

$$\psi : \mathbb{R} \to \mathbb{R},\; \psi(t) := \alpha + \beta t \quad , \quad \alpha, \beta \in \mathbb{R}$$

gilt

$$\psi(\mathbb{E}(X \mid \mathscr{C})) = \mathbb{E}(\psi(X) \mid \mathscr{C})$$

Da φ konvex ist, gilt

$$\varphi(t) = \sup\left\{\psi(t) : \psi \leq \varphi \text{ affin}\right\} \quad \forall\, t \in I \tag{6.2.4.1}$$

und natürlich

$$\mathbb{E}\left(\varphi(X) \mid \mathscr{C}\right) \geq \mathbb{E}\left(\psi(X) \mid \mathscr{C}\right) = \psi(\mathbb{E}(X \mid \mathscr{C}))$$

für jegliche affine $\psi \leq \varphi$. Folglich auch

$$\mathbb{E}\left(\varphi(X) \mid \mathscr{C}\right) \geq \sup_{\substack{\psi\leq\varphi \\ \text{affin}}} \psi\left[\mathbb{E}(X \mid \mathscr{C})\right] \overset{(6.2.4.1)}{=} \varphi\left[\mathbb{E}(X \mid \mathscr{C})\right]$$

□

6.2.5 Folgerungen der Jensenschen Ungleichung

Ist $I \subseteq \mathbb{R}$ ein offenes Intervall, $X : (\Omega, \mathfrak{S}, \mathcal{P}) \to I$ eine Zufallsgröße mit $\mathbb{E}\,|X| < \infty$, $\varphi : I \to \mathbb{R}$ konvex und $\mathbb{E}\,|\varphi(X)| < \infty$, so gilt:

1. $\varphi(\mathbb{E}X) \leq \mathbb{E}(\varphi(X))$
2. $\mathbb{E}\left[\varphi\left(\mathbb{E}(X \mid \mathscr{C})\right)\right] \leq \mathbb{E}\varphi(X)$

Beweis:

1. Setzt man $\mathscr{C} := \{\emptyset, \Omega\}$ so sind sowohl $\sigma(X)$ als auch $\sigma(\varphi(X))$ unabhängig von $\mathscr{C}$. Nach 6.2.3(7) und obiger Ungleichung folgt schließlich die Behauptung.
2. Aus 6.2.3(8) und der Jensenschen Ungleichung folgt

$$\mathbb{E}\left[\varphi\left(\mathbb{E}(X \mid \mathscr{C})\right)\right] \leq \mathbb{E}\left[\mathbb{E}(\varphi(X) \mid \mathscr{C})\right] = \mathbb{E}\varphi(X)$$

□

6.2.6 Lemma über den bedingten Erwartungswert von Produkten

Seien $X, Y : (\Omega, \mathfrak{S}, \mathcal{P}) \to \mathbb{R}$ zufällige Größen mit $\mathbb{E}\,|XY|\,, \mathbb{E}\,|X| < \infty$ und Y $\mathscr{C}$-messbar. Dann gilt:

$$\mathbb{E}(XY \mid \mathscr{C}) \overset{\text{f.s.}}{=} Y \cdot \mathbb{E}(X \mid \mathscr{C})$$

Beweis: Sei zunächst $Y = 1_{C_0}$ für irgendein $C_0 \in \mathscr{C}$. Dann gilt

$$\int\limits_C 1_{C_0} \cdot \mathbb{E}(X \mid \mathscr{C})\; d\mathcal{P} = \int\limits_{C \cap C_0} \mathbb{E}(X \mid \mathscr{C})\; d\mathcal{P} = \int\limits_{C \cap C_0} X\; d\mathcal{P} = \int\limits_C 1_{C_0} X\; d\mathcal{P} \quad \forall\, C \in \mathscr{C}$$

das heißt

$$\mathbb{E}(1_{C_0} \cdot X \mid \mathscr{C}) \overset{\text{f.s.}}{=} 1_{C_0} \cdot \mathbb{E}(X \mid \mathscr{C})$$

Aufgrund der Linearität überträgt sich das Ergebnis auch auf beliebige Treppenfunktionen Y und nach 6.2.3(6) auf beliebige nicht-negative $Y \geq 0$. Durch die Separierung $Y = Y^+ - Y^-$ erhält man schließlich die allgemeine Behauptung.
□

6.2.7 Korollar: Charakterisierung des bedingten Erwartungswertes

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum, $\mathscr{C} \subseteq \mathfrak{S}$ eine Unter-σ-Algebra und $X : \Omega \to \mathbb{R}$ eine Zufallsgröße mit $\mathbb{E}\,|X| < \infty$. Dann ist $\mathbb{E}(X \mid \mathscr{C})$ genau die (fast-eindeutig bestimmte) $\mathscr{C}$-messbare Abbildung, für die gilt:

$$\mathbb{E}\,(XY) = \mathbb{E}\left[\mathbb{E}\,(X \mid \mathscr{C}) \cdot Y\right] \quad \forall\, Y : (\Omega, \mathscr{C}) \to \mathbb{R} \text{ messbar mit } \mathbb{E}\,|XY| < \infty$$

Erläuterung: Richtung "$\Leftarrow$" ergibt sich direkt aus den Fällen $Y = 1_C,\ C \in \mathscr{C}$.
Richtung "$\Rightarrow$" ergibt sich aus Lemma 6.2.6 gemäß

$$\mathbb{E}\left[\mathbb{E}\,(X \mid \mathscr{C}) \cdot Y\right] \overset{6.2.6}{=} \mathbb{E}\left[\mathbb{E}\,(XY \mid \mathscr{C})\right] = \mathbb{E}(XY)$$

6.2.8 Der bedingte Erwartungswert als Projektor

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum, $\mathscr{C} \subseteq \mathfrak{S}$ eine Unter-σ-Algebra und $1 \leq p < \infty$. Durch Betrachtung der konvexen Funktion $\varphi(t) := |t|^p$ erhält man nach Jensen 6.2.5

$$\mathbb{E}\,|\mathbb{E}(X \mid \mathscr{C})|^p \leq \mathbb{E}\,|X|^p \tag{6.2.8.1}$$

das heißt für $X \in L_p(\Omega, \mathfrak{S}, \mathcal{P})$ ist $\mathbb{E}(X \mid \mathscr{C}) \in L_p(\Omega, \mathscr{C}, \mathcal{P})$. Die damit zusammenhängende lineare Abbildung

$$\Pi : L_p(\Omega, \mathfrak{S}, \mathcal{P}) \to L_p(\Omega, \mathscr{C}, \mathcal{P}) \quad , \quad \Pi(X) := \mathbb{E}(X \mid \mathscr{C})$$

ist nach 6.2.3(9) ein Projektor in den abgeschlossenen Unterraum $L_p(\Omega, \mathscr{C}, \mathcal{P}) \subseteq L_p(\Omega, \mathfrak{S}, \mathcal{P})$, sprich $\Pi^2 = \Pi$. Nach (6.2.8.1) ist

$$\|\Pi\| := \sup_{\|X\|_p \leq 1} \|\Pi X\| \leq 1$$

also eigentlich sogar[20]

$$\|\Pi\| = 1$$

Im Fall $p = 2$ ist Π sogar ein Orthogonalprojektor im Hilbertraum $L_2(\Omega, \mathfrak{S}, \mathcal{P})$ (siehe unten).

6.2.9 Der bedingte Erwartungswert als Orthogonalprojektor

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum, $\mathscr{C} \subseteq \mathfrak{S}$ eine Unter-σ-Algebra und

$$\Pi : L_2(\Omega, \mathfrak{S}, \mathcal{P}) \to L_2(\Omega, \mathscr{C}, \mathcal{P})$$

der Projektor definiert durch

$$\Pi(X) := \mathbb{E}(X \mid \mathscr{C})$$

Dann ist Π ein Orthogonalprojektor, das heißt

$$X - \Pi X \perp L_2(\Omega, \mathscr{C}, \mathcal{P})$$

Beweis: Wegen $\Pi Y = Y \quad \forall\, Y \in L_2(\Omega, \mathscr{C}, \mathcal{P})$ genügt es nach A.4.9 zu zeigen dass

$$X - \Pi X \perp \Pi X \quad \forall\, X \in L_2(\Omega, \mathfrak{S}, \mathcal{P})$$

sprich

$$\langle X, \Pi X \rangle = \langle \Pi X, \Pi X \rangle$$

Wegen

$$\mathbb{E}\,|X \cdot \Pi X| = \|X \cdot \Pi X\|_1 \leq \|X\|_2 \cdot \underbrace{\|\Pi X\|_2}_{\leq \|X\|_2} \leq \|X\|_2^2 < \infty$$

gilt nach Lemma 6.2.6

$$\|\Pi X\|_2^2 = \mathbb{E}\,[\Pi X \cdot \mathbb{E}(X \mid \mathscr{C})] \overset{6.2.7}{=} \mathbb{E}\,(\Pi X \cdot X) = \langle X, \Pi X \rangle$$

was zu zeigen war.
□

Interpretation: Als Orthogonalprojektion ist ΠX die *beste* Approximation von $X \in L_2(\Omega, \mathfrak{S}, \mathcal{P})$ im Unterraum $L_2(\Omega, \mathscr{C}, \mathcal{P})$, das heißt

$$\|X - \mathbb{E}(X \mid \mathscr{C})\|_2 = \inf\left\{\|X - Y\|_2 : Y \in L_2(\Omega, \mathscr{C}, \mathcal{P})\right\}$$

$\mathbb{E}(X \mid \mathscr{C})$ ist sozusagen die beste Voraussage für X bei Kenntnis von $\mathscr{C}$.

[20] Für jeden Projektor Π gilt $\|\Pi\| \geq 1$.

6.2.10 Definition: Bedingte Wahrscheinlichkeit einer messbaren Menge

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum, $\mathscr{C} \subseteq \mathfrak{S}$ eine Unter-σ-Algebra. Zu Menge $A \in \mathfrak{S}$ heißt

$$\mathcal{P}(A \mid \mathscr{C}) := \mathbb{E}(1_A \mid \mathscr{C}) \quad : (\Omega, \mathscr{C}) \to [0,1]$$

bedingte Wahrscheinlichkeit von A, gegeben $\mathscr{C}$.

Bemerke: $\mathbb{E}(A \mid \mathscr{C})$ ist die (fast-eindeutig bestimmte) $\mathscr{C}$-messbare Abbildung für die gilt

$$\int_C \mathcal{P}(A \mid \mathscr{C})\, d\mathcal{P} = \mathcal{P}(A \cap C) \quad \forall\, C \in \mathscr{C}$$

6.2.11 Satz über die bedingte Wahrscheinlichkeit von Mengen

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum, $\mathscr{C} \subseteq \mathfrak{S}$ eine Unter-σ-Algebra. Dann gilt:

1. $\mathcal{P}(\Omega \mid \mathscr{C}) \overset{\text{f.s.}}{=} 1$
2. Zu disjunkten $A_1, A_2, .. \in \mathfrak{S}$ existiert eine Nullmenge $N \in \mathfrak{S}$ mit

$$\mathcal{P}\left(\bigcup_{j=1}^{\infty} A_j \middle| \mathscr{C}\right)(\omega) = \sum_{j=1}^{\infty} \mathcal{P}(A_j \mid \mathscr{C})(\omega) \quad \forall\, \omega \notin N$$

Beweis:

1. Folgt direkt aus

$$\int_C 1\, d\mathcal{P} = \mathcal{P}(C) = \mathcal{P}(\Omega \cap C) \quad \forall\, C \in \mathscr{C}$$

2. Folgt aus

$$\sum_{j=1}^{\infty} \mathcal{P}(A_j \mid \mathscr{C}) = \lim_{n\to\infty} \mathcal{P}\left(\sum_{j=1}^{n} 1_{A_j} \mid \mathscr{C}\right) \underset{\text{f.s.}}{\overset{6.2.3(6)}{=}} \mathcal{P}\left(\sum_{j=1}^{\infty} 1_{A_j} \mid \mathscr{C}\right) = \mathcal{P}\left(\bigcup_{j=1}^{\infty} A_j \mid \mathscr{C}\right)$$

□

Bemerkung: Es stellt sich angesichts obiger Eigenschaften die Frage, ob denn solch eine Nullmenge N existiert, die unabhängig von den vorliegenden $A_1, A_2, ..$ ist, sprich, ob $\mathcal{P}(\cdot \mid \mathscr{C})$ nicht doch fast-sicher σ-additiv (und somit ein Wahrscheinlichkeitsmaß) ist. Leider ist die Antwort darauf negativ!

6.2.12 Definition: Reguläre bedingte Wahrscheinlichkeit

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $\mathscr{C} \subseteq \mathfrak{S}$ eine Unter-σ-Algebra. Dann heißt die bedingte Wahrscheinlichkeit $\mathcal{P}(\cdot \mid \mathscr{C})$ *regulär*, falls

$$A \mapsto \mathcal{P}(A \mid \mathscr{C})(\omega) \quad , \quad A \in \mathfrak{S}$$

für alle $\omega \in \Omega$ ein Wahrscheinlichkeitsmaß ist.

6.2.13 Satz: Integration über die bedingte Wahrscheinlichkeit

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum, $\mathscr{C} \subseteq \mathfrak{S}$ eine Unter-σ-Algebra, dazu die bedingte Wahrscheinlichkeit $\mathcal{P}(\cdot \mid \mathscr{C})$. Ist $\mathcal{P}(\cdot \mid \mathscr{C})(\omega)$ ein Wahrscheinlichkeitsmaß, so gilt für jede Zufallsvariable $X : (\Omega, \mathfrak{S}) \to \mathbb{R}$

$$\mathbb{E}(X \mid \mathscr{C})(\omega) = \int_\Omega X(\omega')\, \mathcal{P}(d\omega' \mid \mathscr{C})(\omega)$$

Beweis: Es genügt die Aussage für $X \geq 0$ zu zeigen. Nach dem Satz über monotone Konvergenz und 6.2.3(6) genügt es die Aussage für Treppenfunktionen zu zeigen. Aufgrund der Linearität von $\mathbb{E}(\cdot \mid \mathscr{C})$ und des Integrals genügt es schließlich die Aussage für $Y = 1_A,\ A \in \mathfrak{S}$ zu zeigen. Doch tatsächlich:

$$\int\limits_{\Omega} 1_A(\omega')\ \mathcal{P}(d\omega' \mid \mathscr{C})(\omega) = \mathcal{P}(A \mid \mathscr{C})(\omega) = \mathbb{E}(1_A \mid \mathscr{C})(\omega)$$

□

6.3 Bedingter Erwartungswert bzgl. einer Zufallsvariablen

6.3.1 Definition: Bedingter Erwartungswert bzgl. einer Zufallsvariablen

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum, $(S, \mathcal{S})$ ein messbarer Raum und $Z : \Omega \to S$ eine Zufallsvariable, dazu die erzeugte σ-Algebra $\sigma(Z) := Z^{-1}(\mathcal{S})$. Dann heißt für jede weitere Zufallsvariable $X : \Omega \to \mathbb{R}$ mit $\mathbb{E}\,|X| < \infty$ die Abbildung

$$\mathbb{E}(X \mid Z) := \mathbb{E}(X \mid \sigma(Z))$$

bedingter Erwartungswert von X, gegeben Z. Sie ist die (fast-eindeutig bestimmte) $\mathscr{C}$-messbare Abbildung für die gilt

$$\int\limits_{Z^{-1}(B)} \mathbb{E}(X \mid Z)\ d\mathcal{P} = \int\limits_{Z^{-1}(B)} X\ d\mathcal{P} \quad \forall\ B \in \mathcal{S}$$

Allgemeiner heißt für Zufallsvariablen $Z_i : \Omega \to (S_i, \mathcal{S}_i),\ i \in I$

$$\mathbb{E}(X \mid Z_i : i \in I) := \mathbb{E}\left(X \mid \sigma(Z_i : i \in I)\right)$$

bedingter Erwartungswert von X, gegeben $(Z_i)_{i \in I}$. Analog heißt für messbare $A \in \mathfrak{S}$

$$\mathcal{P}(A \mid Z) := \mathcal{P}(A \mid \sigma(Z))$$

bedingte Wahrscheinlichkeit von A, gegeben Z.

Bemerke: Nach 2.3.6 erzeugt die Zufallsvariable

$$\mathbf{Z} : \Omega \to \left(\bigtimes_{i \in I} S_i, \bigotimes_{i \in I} \mathcal{S}_i \right)$$

die gleiche σ-Algebra $\sigma(\mathbf{Z}) = \sigma(Z_i : i \in I)$, so dass wir im Folgenden in der Regel o.B.d.A. nur bedingte Erwartungswerte bzgl. einzelner Zufallsgrößen Z betrachten werden.

6.3.2 Eigenschaften von $\mathbb{E}(X \mid Z)$

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum, und $Z : \Omega \to (S, \mathcal{S})$ eine Zufallsvariable in den messbaren Raum $(S, \mathcal{S})$. Dann gilt für jegliche Zufallsvariable $X : \Omega \to \mathbb{R}$ mit $\mathbb{E}\,|X| < \infty$:

1. $\mathbb{E}(f(Z) \cdot X \mid Z) \overset{\text{f.s.}}{=} f(Z) \cdot \mathbb{E}(X \mid Z)$ für jede messbare $f : S \to \mathbb{R}$, insofern $\mathbb{E}\,|f(Z) \cdot X| < \infty$.

2. Sind X, Z unabhängig, so gilt $\mathbb{E}(X \mid Z) \overset{\text{f.s.}}{=} \mathbb{E}X$.

Erläuterung: Diese und ähnliche Eigenschaften folgen aus den allgemeineren Eigenschaften 6.2.3 bedingter Erwartungswerte.

Beispiele:

(i) Zum n-fachen Würfelexperiment seien $X_1, .., X_n$ die gewürfelten Werte, dazu $S_n := X_1 + \cdots + X_m$. Dann gilt für $m \leq n$:

$$\mathbb{E}\left(S_n \mid X_1, \ldots, X_m\right) \overset{\text{f.s}}{=} \underbrace{\mathbb{E}\left(X_1 + \cdots + X_m \mid X_1, \ldots, X_m\right)}_{X_1 + \cdots + X_m} + \underbrace{\mathbb{E}\left(X_{m+1} + \cdots + X_n \mid X_1, \ldots, X_m\right)}_{\mathbb{E}(X_{m+1} + \cdots + X_n)}$$

$$\overset{\text{f.s}}{=} X_1 + \cdots + X_m + \frac{7}{2}(n - m)$$

(ii) Die Zufallsgröße $Z : (\Omega, \mathfrak{S}, \mathcal{P}) \to (S, \mathcal{S})$ nehme nur die Werte $s_1, .., s_n \in S$ an (mit $\{s_i\} \in \mathcal{S}$). Dann gilt

$$\mathbb{E}\left(X \mid Z\right) \overset{\text{f.s}}{=} \sum_{j=1}^{n} \int_{Z^{-1}(s_j)} X \, d\mathcal{P} \cdot \frac{1_{Z^{-1}(s_j)}}{\mathcal{P}(Z^{-1}(s_j))}$$

für jegliche Zufallsgröße $X : \Omega \to \mathbb{R}$ (vgl. (6.2.2.1)).

6.3.3 Spezialfälle für Folge von Zufallsvariablen

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $X_1, X_2, .. : \Omega \to \mathbb{R}$ reelle Zufallsgrößen (z.B. mittlere Temperatur pro Tag) mit $\mathbb{E}\,|X_n| < \infty$, dazu betrachtet die bedingten Erwartungswerte

$$\mathbb{E}(X_{n+1} \mid X_1, .., X_n)$$

Folgende Fälle sind von besonderem Interesse:

Fall: X_{n+1} ist $\sigma(X_1, .., X_n)$-messbar. Dann ist nach Faktorisierungslemma A.1.3

$$X_{n+1} = h(X_1, .., X_n)$$

für geeignete, messbare $h : \mathbb{R}^n \to \mathbb{R}$ und

$$\mathbb{E}(X_{n+1} \mid X_1, .., X_n) \overset{\text{f.s}}{=} X_{n+1}$$

Fall: X_{n+1} ist unabhängig von $X_1, .., X_n$. Dann gilt

$$\mathbb{E}(X_{n+1} \mid X_1, .., X_n) \overset{\text{f.s}}{=} \mathbb{E}X_{n+1}$$

Fall: $\mathbb{E}(X_{n+1} \mid X_1, .., X_n) \overset{\text{f.s}}{=} \mathbb{E}(X_{n+1} \mid X_n)$ (Markoveigenschaft)

Fall: $\mathbb{E}(X_{n+1} - X_n \mid X_1, .., X_n) \overset{\text{f.s}}{=} 0$ (Martingaleigenschaft)

6.3.4 Definition: Bedingter Erwartungswert bzgl. einzelner Werte

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum, $(S, \mathcal{S})$ ein messbarer Raum, $Z : \Omega \to (S, \mathcal{S})$ eine Zufallsvariable. Dann ist für jede Zufallsvariable $X : \Omega \to \mathbb{R}^n$ mit $\mathbb{E}\,\|X\| < \infty$ der bedingte Erwartungswert $\mathbb{E}(X \mid Z)$ messbar bzgl. $\sigma(Z)$, so dass nach Faktorisierungslemma A.1.3 eine $\mathcal{P}_Z$-fast eindeutige, messbare Funktion $h_X : (S, \mathcal{S}) \to \mathbb{R}^n$ existiert[21] mit

$$\mathbb{E}(X \mid Z) = h_X \circ Z$$

Zu $s \in S$ heißt nun der Wert

$$\mathbb{E}(X \mid Z = s) := h_X(s)$$

[21]Genauer: Es existiert eine Menge $N \subseteq S$ mit $\mathcal{P}(Z^{-1}(N)) = 0$, so dass $h_X(s)$ für alle $s \notin N$ eindeutig festgelegt ist. Dabei kann N unabhängig von X gewählt werden, z.B. $N := Z(\Omega)^c$ (nicht unbedingt messbar!).

bedingter Erwartungswert von X, gegeben $Z = s$. Insbesondere gilt

$$\int_B \mathbb{E}(X \mid Z = s)\, d\mathcal{P}_Z(s) = \int_{Z^{-1}(B)} h_X(Z)\, d\mathcal{P} = \int_{Z^{-1}(B)} X\, d\mathcal{P} \quad \forall\, B \in \mathcal{S}$$

Ähnlich nennt man für messbare $A \in \mathfrak{S}$ den Wert[22]

$$\mathcal{P}(A \mid Z = s) := \mathbb{E}(1_A \mid Z = s)$$

bedingte Wahrscheinlichkeit von A, gegeben $Z = s$. Analog gilt auch hier

$$\int_B \mathcal{P}(A \mid Z = s)\, d\mathcal{P}_Z(s) = \mathcal{P}(A \cap Z^{-1}(B)) \quad \forall\, B \in \mathcal{S}$$

Speziell gilt

$$\mathcal{P}(A \mid Z = s) \cdot \mathcal{P}(Z = s) = \mathcal{P}\left(A \cap \{Z = s\}\right)$$

(vgl. 6.1.1) und

$$\mathcal{P}(A) = \int_S \mathcal{P}(A \mid Z = s)\, d\mathcal{P}_Z(s)$$

(*Formel über die totale Wahrscheinlichkeit*). $\mathcal{P}(\cdot \mid Z = s)$ heißt *reguläre bedingte Wahrscheinlichkeit*, falls die Abbildung

$$A \mapsto \mathcal{P}(A \mid Z = s) \quad , \quad A \in \mathfrak{S}$$

für jedes $s \in S$ ein Wahrscheinlichkeitsmaß ist. Gegebenfalls ist dann auch bedingte Wahrscheinlichkeit

$$\mathcal{P}(\cdot \mid Z)(\omega) := \mathcal{P}(\cdot \mid Z = Z(\omega))$$

regulär.

Ist nun $Y : \Omega \to (M, \mathcal{M})$ eine weitere Zufallsvariable in den messbaren Raum, so kann man

$$B \mapsto \mathcal{P}(Y \in B \mid Z = s) \quad , \quad B \in \mathcal{M}$$

als Einschränkung von $\mathcal{P}(\cdot \mid Z = s)$ auf $\sigma(Y)$ betrachten. Ist $\mathcal{P}(\cdot \mid Z = s)$ Wahrscheinlichkeitsmaß, so ist

$$\mathcal{P}(Y^{-1}(\cdot) \mid Z = s)$$

einfach das Bildmaß von $\mathcal{P}(\cdot \mid Z = s)$ unter der Abbildung Y.

6.3.5 Theorem: Existenz einer regulären bedingten Wahrscheinlichkeit

Sei (E, d) ein vollständiger, separabler metrischer Raum, $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum, $(S, \mathcal{S})$ ein messbarer Raum und $Y : \Omega \to E,\ Z : \Omega \to S$ zufällige Größen. Dann existiert stets eine reguläre Version von $\mathcal{P}(Y \in \cdot \mid Z = s)$, das heißt die Abbildung

$$B \mapsto \mathcal{P}(Y \in B \mid Z = s) \quad , \quad B \in \mathscr{B}(E)$$

ist für jedes $s \in S$ ein Wahrscheinlichkeitsmaß auf $(E, \mathscr{B}(E))$.

[22] Beachte: Die Mengen-Abbildung $\mathcal{P}(\cdot \mid Z = s) : \mathfrak{S} \to \mathbb{R}$ ist für $\mathcal{P}_Z$-fast alle s eindeutig festgelegt.

6.3.6 Bedingte Dichten

Sei $(\Omega, \mathfrak{S}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $X, Z : \Omega \to \mathbb{R}$ Zufallsgrößen, so dass (X, Z) die Dichte ρ besitzt, sprich

$$\mathcal{P}((X,Z) \in B) = \int_B \rho(x,z)\ d(x,z) \quad \forall\ B \in \mathscr{B}(\mathbb{R}^2)$$

Dazu sei definiert die Randdichte

$$\rho_r(z) := \int_{\mathbb{R}} \rho(x,z)\ dx \quad , \quad z \in \mathbb{R}$$

(Wahrscheinlichkeitsdichte von Z) und die für $\mathcal{P}_Z$-fast alle z wohldefinierte Funktion

$$\rho(x \mid z) := \frac{\rho(x,z)}{\rho_r(z)} \quad , \quad x \in \mathbb{R}$$

Dann lässt sich zeigen: Für $\mathcal{P}_Z$-fast alle $s \in \mathbb{R}$ gilt

$$\mathbb{E}\left(f(X,Z) \mid Z = s\right) = \int_{\mathbb{R}} f(x,s)\ \rho(x \mid s)\ dx \tag{6.3.6.1}$$

für jede messbare $f : \mathbb{R}^2 \to \mathbb{R}$. Gegebenfalls gilt dann für $A \in \mathscr{B}(\mathbb{R}),\ f(x,z) := 1_A(x)$:

$$\mathcal{P}(X \in A \mid Z = s) \overset{\text{def}}{=} \underbrace{\mathbb{E}\left(1_{X^{-1}(A)} \mid Z = s\right)}_{\mathbb{E}(1_A(X)|Z=s)} \overset{(6.3.6.1)}{=} \int_A \rho(x \mid s)\ dx$$

(vgl. mit entsprechenden Ausdruck in 6.1.2) und $\mathcal{P}(X^{-1}(\cdot) \mid Z = s)$ ist reguläre, bedingte Wahrscheinlichkeit auf $\mathbb{R}$.

A Anhang

A.1 Bemerkungen zu Messbarkeit

A.1.1 Satz über Schnittfunktionen

Seien $(\Omega_1, \mathfrak{S}_1), \ldots, (\Omega_n, \mathfrak{S}_n)$ messbare Räume, dazu die Produkträume

$$\left(\Omega^k, \mathfrak{S}^k\right) := \left(\mathop{\times}_{i=1}^{k} \Omega_i,\ \bigotimes_{i=1}^{k} \mathfrak{S}_i\right) \quad , \quad k \le n$$

und die messbare Abbildung

$$f : (\Omega^n, \mathfrak{S}^n) \to (S, \mathcal{S})$$

in den messbaren Raum $(S, \mathcal{S})$. Dann ist für beliebige (feste)

$$y_{k+1} \in \Omega_{k+1}, .., y_n \in \Omega_n,\ k \le n$$

die *Schnittfunktion*

$$f_{\mathrm{s}} : (\Omega^k, \mathfrak{S}^k) \to (S, \mathcal{S}) \quad , \quad f_{\mathrm{s}}(x_1, .., x_k) := f(x_1, .., x_k, y_{k+1}, .., y_n)$$

messbar.

Beweis: Wegen

$$f_{\mathrm{s}} = f \circ \mathrm{Id}_{\mathrm{s}} \quad , \quad \mathrm{Id}_{\mathrm{s}}(x_1, .., x_k) := (x_1, .., x_k, y_{k+1}, .., y_n)$$

genügt es die Behauptung für $f := \mathrm{Id}$ und $(S, \mathcal{S}) = (\Omega^n, \mathfrak{S}^n)$ zu zeigen. Für $A = \mathop{\times}_{i=1}^{n} A_i \in \mathop{\times}_{i=1}^{n} \mathfrak{S}_i$ ist

$$\mathrm{Id}_{\mathrm{s}}^{-1}(A) = \left\{(x_1, .., x_k) \in \Omega^k : (x_1, .., x_k, y_{k+1}, .., y_n) \in A\right\} = \underbrace{\begin{cases} \emptyset & : (y_{k+1}, .., y_n) \notin \mathop{\times}_{i=k+1}^{n} A_i \\ \mathop{\times}_{i=1}^{k} A_i & : (y_{k+1}, .., y_n) \in \mathop{\times}_{i=k+1}^{n} A_i \end{cases}}_{\text{messbar}}$$

Wegen $\mathfrak{S}^n = \sigma\left(\mathop{\times}_{i=1}^{n} \mathfrak{S}_i\right)$ ist Id_{s} messbar.
□

Beispiel: Ist $(G, \circ, \mathcal{G})$ eine messbare Gruppe, $B \in \mathcal{G}$ und $y \in G$, so ist auch $B \circ y^{-1} \in \mathcal{G}$.
Beweis: Die Abbildung

$$g : G \to G \quad , \quad g(x) := x \circ y$$

ist nach Satz (A.1.1) messbar und

$$B \circ y^{-1} = g^{-1}(B)$$

A.1.2 Satz: Messbarkeit der n-fachen Gruppenverknüpfung

Sei $(G, \circ, \mathcal{G})$ eine messbare Gruppe, dazu der n-fache Produktraum

$$(G^n, \mathcal{G}^n) := \left(\mathop{\times}_{i=1}^{n} G,\ \bigotimes_{i=1}^{n} \mathcal{G}\right)$$

Dann ist die Abbildung

$$g_n : (G^n, \mathcal{G}^n) \to (G, \mathcal{G}) \quad , \quad g_n(x_1, .., x_n) = x_1 \circ \cdots \circ x_n$$

messbar.

Beweis: Sei

$$P_{i_1,..,i_m} : G^n \to G^m \quad , \quad P_{i_1,..,i_m}(x_1,..,x_n) = (x_{i_1},..,x_{i_m}) \quad , \quad m \leq n$$

die (messbare) *Projektion* von G^n in den Unterproduktraum G^m. Der Fall $n = 2$ ist nach Voraussetzung an die Gruppe gegeben. Gilt nun die Behauptung für $n-1$, so ist

$$g_n(x_1,..,x_n) = g_{n-1}(x_1,..,x_{n-1}) \circ x_n \quad \forall\ x_i \in G$$

das heißt

$$g_n = g_2 \circ (g_{n-1} \circ P_{1,..,n-1},\ P_n)$$

Dabei ist $g_{n-1} \circ P_{1,..,n-1} : G^n \to G$ messbar, weshalb auch der Vektor $(g_{n-1} \circ P_{1,..,n-1},\ P_n) : G^n \to G \times G$ bzw. g_n messbar ist.
$\square$

A.1.3 Faktorisierungslemma

Sei Ω eine nicht-leere Menge, $(S, \mathcal{S})$ ein messbarer Raum und $X : \Omega \to S$. Ist die Abbildung $Y : \Omega \to \mathbb{R}^n$ messbar bzgl. $\sigma(X) := X^{-1}(\mathcal{S})$, so existiert eine messbare Abbildung $h : (S, \mathcal{S}) \to \mathbb{R}^n$ mit

$$Y = h \circ X$$

Diese ist auf dem Bild $X(\Omega)$ eindeutig festgelegt. Genauer gesagt, gilt

$$X^{-1}\left(\{s \in S : h(s) \text{ eindeutig festgelegt}\}\right) = \Omega$$

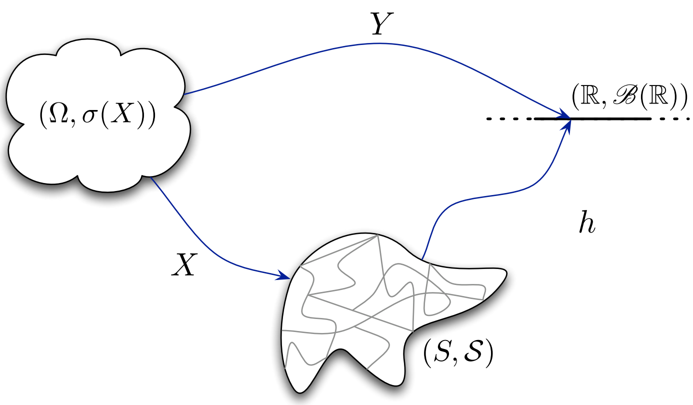

Abbildung 16: Zum Faktorisierungslemma: Zusammenhang zwischen X und bzgl. $\sigma(X)$ messbaren Abbildungen.

Beweisidee: Da die Messbarkeit eines Vektors $Y : \Omega \to \mathbb{R}^n$ äquivalent ist zur Messbarkeit jeder seiner Komponenten, genügt es die Behauptung für den 1-dimensionalen Fall zu zeigen.

Zunächst wird die Aussage für Treppenfunktionen $Y = \sum_{j=1}^n \alpha_j \cdot 1_{A_j},\ A_j \in \sigma(X)$ disjunkt, bewiesen. Tatsächlich existieren $B_j \in \mathcal{S}$ mit $A_j = Z^{-1}(B_j)$ und

$$h := \sum_{j=1}^n \alpha_j \cdot 1_{B_j}$$

erfüllt die Behauptung. Der allgemeinere Fall wird mit entsprechender Approximation durch Treppenfunktionen gezeigt. Die Behauptungen zur Eindeutigkeit sind klar.

A.2 Konvergenz von Maßen und Zufallsvariablen

A.2.1 Hilfslemma über die Stetigkeit der Metrik

Sei (E, d) ein metrischer Raum und

$$d(x, A) := \inf_{y \in A} d(x, y) \quad , \quad x \in E, A \subseteq E \tag{A.2.1.1}$$

Dann ist für jedes $A \subseteq E$ die Abbildung $d(\cdot, A) : x \mapsto d(x, A)$ stetig.

Beweis: Es gilt die Dreiecksungleichung

$$d(x, y) + d(y, A) = d(x, y) + \inf_{z \in A} d(y, z) = \inf_{z \in A} \underbrace{(d(x, y) + d(y, z))}_{\geq d(x,z)} \geq \inf_{z \in A} d(x, z) = d(x, A)$$

Dementsprechend gilt für Folge $x_n \overset{n\to\infty}{\longrightarrow} x \in E$:

$$\underbrace{d(x, A) - d(x_n, x)}_{\overset{n\to\infty}{\longrightarrow} d(x,A)} \leq d(x_n, A) \leq \underbrace{d(x_n, x) + d(x, A)}_{\overset{n\to\infty}{\longrightarrow} d(x,A)}$$

das heißt $d(x_n, A) \overset{n\to\infty}{\longrightarrow} d(x, A)$.

□

Bemerke: Ist $A \subseteq E$ abgeschlossen, so ist $d(x, A) > 0$ für $x \notin A$, denn sonst gäbe es eine Folge $(y_n) \subseteq A$ mit $y_n \overset{n\to\infty}{\longrightarrow} x$.

A.2.2 Lemma: Approximation von Indikatorfunktionen in metrischen Räumen

Sei (E, d) ein metrischer Raum. Dann gilt:

1. Zu abgeschlossener $A \subseteq E$ existieren $g_n \in \mathcal{C}_b(E)$ mit

$$g_n \underset{n\to\infty}{\searrow} 1_A \text{ punktweise}$$

   und $0 \leq g_n \leq 1$.

2. Zu offener $G \subseteq E$ existieren $f_n \in \mathcal{C}_b(E)$ mit

$$f_n \underset{n\to\infty}{\nearrow} 1_G \text{ punktweise}$$

   und $0 \leq f_n \leq 1$.

Beweis:

1. Definiere die Funktionen

$$g_n : E \to \mathbb{R} \quad , \quad g_n(x) := 1 - \min\{1, n \cdot d(x, A)\} \quad , \quad n \in \mathbb{N}$$

   wobei $d(\cdot, A) : E \to [0, \infty)$ die Abstandsfunktion aus (A.2.1.1) sei. Dann gilt:

   - Da $d(\cdot, A)$ nach A.2.1 stetig ist, sind auch alle g_n stetig.
   - Alle g_n sind beschränkt: $0 \leq g_n \leq 1$
   - Die g_n sind punktweise monoton fallend in n, das heißt $g_1(x) \geq g_2(x) \geq \ldots$
   - Ist $x \in A$ so sind $g_n(x) = 1$. Ist $x \notin A$, sprich $d(x, A) > 0$, so gehen $g_n(x) \overset{n\to\infty}{\longrightarrow} 0$. Daher gehen $g_n \underset{\text{punktweise}}{\overset{n\to\infty}{\longrightarrow}} 1_A$.

2. Da $G^c \subseteq E$ abgeschlossen ist, existieren nach Teil (1) monoton fallende $g_n \in \mathcal{C}_b(E)$ mit $g_n \underset{\text{punktweise}}{\overset{n\to\infty}{\longrightarrow}} 1_{G^c}$. Die

$$f_n := 1 - g_n$$

   erfüllen dann alle Behauptungen.

   □

A.2.3 Satz: Eindeutigkeit von Wahrscheinlichkeitsmaßen

Sei (E, d) ein metrischer Raum und μ, ν zwei endliche Maße auf $(E, \mathscr{B}(E))$ für die gilt

$$\int_E f\ d\mu = \int_E f\ d\nu$$

für jede beschränkte, stetige $f : E \to \mathbb{R}$. Dann ist $\mu = \nu$.

Beweis: Sei $G \subseteq E$ offen, dann existieren nach A.2.2 monoton wachsende Funktionen $f_n \in \mathcal{C}_b(E)$ mit $f_n \underset{\text{punktweise}}{\overset{n\to\infty}{\longrightarrow}} 1_G$ und $0 \leq f_n$. Nach Satz über monotone Konvergenz (Beppo-Levi)

$$\mathcal{P}(G) = \int_E 1_G\ d\mathcal{P} \overset{\text{B.L.}}{=} \lim_{n\to\infty} \int_E f_n\ d\mathcal{P} = \lim_{n\to\infty} \int_E f_n\ d\mathcal{Q} \overset{\text{B.L.}}{=} \int_E 1_G\ d\mathcal{Q} = \mathcal{Q}(G)$$

Da die offenen Mengen $\cap$-stabiler Erzeuger von $\mathscr{B}(E)$ sind und E offen, gilt $\mathcal{P} = \mathcal{Q}$.
$\square$

A.2.4 Probleme bei der Zurückführung des Satzes von Lévy auf den 1-dim. Fall

Es sei die Gültigkeit des Satzes von Lévy 4.6.6 für den 1-dimensionalen Fall angenommen, das heißt für Wahrscheinlichkeitsmaße $(\mathcal{P}_m), \mathcal{P}$ auf $\mathbb{R}$ gilt

$$\mathcal{P}_m \underset{\text{weak}}{\overset{m\to\infty}{\longrightarrow}} \mathcal{P} \quad \Leftrightarrow \quad \widehat{\mathcal{P}}_m \underset{\text{punktweise}}{\overset{m\to\infty}{\longrightarrow}} \widehat{\mathcal{P}}$$

Als direkte Folgerung des 1-dim. Satzes von Lévy-Cramér 4.6.5, der wiederum über Lemma 4.6.4 die 1-dim Version des Auswahlprinzips von Helly 4.5.9 benötigt, kann dieser ohne jegliche Funktionalanalytische Betrachtungen (4.5.6) bewiesen werden.

Zu zeigen wäre, dass für Wahrscheinlichkeitsmaße $\mathcal{P}_n, \mathcal{P}$ auf $\mathbb{R}^n$ mit

$$\widehat{\mathcal{P}}_n(\mathbf{k}) \overset{n\to\infty}{\longrightarrow} \widehat{\mathcal{P}}(\mathbf{k}) \quad \forall\ \mathbf{k} \in (\mathbb{R}^n)'$$

ebenso folgt $\mathcal{P}_n \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} \mathcal{P}$. Für $t \in \mathbb{R}$ gilt auf jeden Fall

$$\widehat{\mathcal{P}}_n(t\mathbf{k}) = \int_{\mathbb{R}^n} e^{it\langle \mathbf{x}, \mathbf{k}\rangle} d\mathcal{P}_n(\mathbf{x}) = \int_{\mathbb{R}} e^{its} d\left(\mathcal{P}_n \circ \mathbf{k}^{-1}\right)(s) = \mathcal{F}\left(\mathcal{P}_n \circ \mathbf{k}^{-1}\right)(t) \overset{n\to\infty}{\longrightarrow} \mathcal{F}\left(\mathcal{P} \circ \mathbf{k}^{-1}\right)(t)$$

das heißt nach dem 1-dim. Fall von Lévy 4.6.6

$$\underbrace{\mathcal{P}_n \circ \mathbf{k}^{-1}}_{\substack{\text{W-Maße} \\ \text{auf } \mathbb{R}}} \underset{\text{d}}{\overset{n\to\infty}{\longrightarrow}} \mathcal{P} \circ \mathbf{k}^{-1} \quad \forall\ \mathbf{k} \in (\mathbb{R}^n)'$$

Äquivalent dazu gilt

$$\varlimsup_{n\to\infty} \mathcal{P}_n\left(\{\mathbf{x} : \langle \mathbf{x}, \mathbf{k}\rangle \in A\}\right) \leq \mathcal{P}\left(\{\mathbf{x} : \langle \mathbf{x}, \mathbf{k}\rangle \in A\}\right)$$

für jede abgeschlossene $A \subseteq \mathbb{R}^n$ und $\mathbf{k} \in (\mathbb{R}^n)'$. Benötigt wird aber die allgemeiner Form

$$\varlimsup_{n\to\infty} \mathcal{P}_n(B) \leq \mathcal{P}(B) \quad \forall\ B \subseteq \mathbb{R}^n \text{ abgeschlossen}$$

worin auch das Problem liegt!

A.3 Charakteristische Funktion

A.3.1 Differenzierungslemma von Lebesgue

Sei $U \subseteq \mathbb{R}^n$ offen, $(\Omega, \mathfrak{S}, \mu)$ ein Maßraum und $f : U \times \Omega \to \mathbb{R}$ so dass $f(\mathbf{k}, \cdot)$ messbar ist für jedes $\mathbf{k} \in U$. Weiterhin gelte:

1. $\int\limits_\Omega |f(\mathbf{k},\omega)| \; d\mu(\omega) < \infty \quad \forall\, \mathbf{k} \in U$

2. Für fast alle $\omega \in \Omega$ existieren $\dfrac{\partial f}{\partial k^i}(\mathbf{k},\omega), \; i = 1, .., n.$

3. $|\partial_{k^i} f(\mathbf{k},\omega)| \leq h(\omega) \quad \forall\, \mathbf{k} \in U$ für irgendeine integrable $h : \Omega \to [0, \infty)$.

Dann gilt:

$$\frac{\partial}{\partial k^i} \int\limits_\Omega f(\mathbf{k},\omega) \; d\mu(\omega) = \int\limits_\Omega \frac{\partial f}{\partial k^i}(\mathbf{k},\omega) \; d\mu(\omega)$$

Beweis: Setze

$$F(\mathbf{k}) := \int\limits_\Omega f(\mathbf{k},\omega) \; d\mu(\omega)$$

Dann ist

$$\frac{F(\mathbf{k} + h\mathbf{e}_i) - F(\mathbf{k})}{h} = \int\limits_\Omega \underbrace{\frac{f(\mathbf{k} + h\mathbf{e}_i) - f(\mathbf{k})}{h}}_{\overset{h\to 0}{\longrightarrow} \partial_{k^i} f(\mathbf{k},\omega)} \; d\mu(\omega)$$

Anderseits kann man für ein geeignetes $\vartheta(\mathbf{k}, h, \omega)$ abschätzen

$$\left| \frac{f(\mathbf{k} + h\mathbf{e}_i) - f(\mathbf{k})}{h} \right| = |\partial_{k^i} f(\mathbf{k} + \vartheta(\mathbf{k}, h, \omega) \cdot h, \; \omega)| \leq h(\omega)$$

(Majorante) so dass nach Lebesgue

$$\frac{\partial F}{\partial k^i}(\mathbf{k}) = \lim_{h\to 0} \frac{F(\mathbf{k} + h\mathbf{e}_i) - F(\mathbf{k})}{h} = \int\limits_\Omega \frac{\partial f}{\partial k^i}(\mathbf{k},\omega) \; d\mu(\omega)$$

$\square$

A.3.2 Lemma: Charakterisierung von Dichten

Sei $\mathcal{P}$ ein Wahrscheinlichkeitsmaß auf $\mathbb{R}$, $D \subseteq \mathbb{R}$ dicht und $\rho : \mathbb{R} \to \mathbb{R}$ stetig mit

$$\int\limits_a^b \rho(x) \; dx = \mathcal{P}\left((a,b)\right) \quad \forall\; a < b \in D$$

und

$$\int\limits_\mathbb{R} |\rho(x)| \; dx < \infty$$

Dann ist $\mathcal{P} \ll \lambda^1$ und besitzt die Dichte ρ.

Beweis: Für beliebige $a < b \in \mathbb{R}$ (der Fall $a = b$ ist trivial) wähle Folgen $(a_n), (b_n) \subseteq D$ mit

$$a_n \underset{n\to\infty}{\searrow} a \quad , \quad b_n \underset{n\to\infty}{\nearrow} b \quad , \quad a_n < b_n$$

Dann gilt

$$\int\limits_a^b \rho(x)\ dx \overset{\text{Lebesgue}}{=} \lim_{n\to\infty} \int\limits_{a_n}^{b_n} \rho(x)\ dx = \lim_{n\to\infty} \mathcal{P}\left([a_n, b_n]\right) = \mathcal{P}\left([a, b]\right)$$

Insbesondere muss $\rho \geq 0$ sein, da sonst aufgrund der Stetigkeit eine Umgebung $\emptyset \neq B_\varepsilon(x_0)$ existieren würde mit $\rho\big|_{B_\varepsilon(x_0)} < 0$ und daher

$$\int\limits_{x_0-\varepsilon}^{x_0+\varepsilon} \rho(x)\ dx < 0$$

Daher definiert ρ durch

$$\mathcal{Q}(A) := \int\limits_A \rho(x)\ dx \quad , \quad A \in \mathscr{B}(A)$$

ein endliches Maß auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ mit $\mathcal{Q} \ll \lambda^1$, das nach obiger Überlegung auf allen Intervallen $[a, b]$ mit $\mathcal{P}$ übereinstimmt. Bekanntlich sind dann $\mathcal{Q} = \mathcal{P}$.
□

A.3.3 Satz über die Konzentration von Maßen

Sei μ ein σ-endliches Maß auf dem messbaren Raum $(\Omega, \mathfrak{S})$ und $(A_i)_{i\in I} \subseteq \mathfrak{S}$ eine Familie disjunkter Mengen mit $\mathcal{P}(A_i) > 0\ \forall\ i \in I$. Dann ist I höchstens abzählbar.

Beweis für endliche Maße: Sei zunächst μ endlich. Für jedes $\varepsilon > 0$ ist dann

$$\{i \in I : \mathcal{P}(A_i) > \varepsilon\}$$

offensichtlich endlich! Daher ist

$$I_0 := \bigcup_{n\in\mathbb{N}} \left\{i \in I : \mathcal{P}(A_i) > \tfrac{1}{n}\right\}$$

höchstens abzählbar. Doch tatsächlich ist $I_0 = I$.

Beweis für σ-endliche Maße: Seien $(\Omega_n) \subseteq \mathfrak{S}$ mit $\Omega_n \underset{n\to\infty}{\nearrow} \Omega$ und $\mu(\Omega_n) < \infty$. Dann gilt

$$\forall\ i \in I : \exists\ n \in \mathbb{N} : \mu(A_i \cap \Omega_n) > \frac{\mu(A_i)}{2}$$

bzw. für jedes $\varepsilon > 0$:

$$\{i \in I : \mu(A_i) > \varepsilon\} \subseteq \underbrace{\bigcup_{n\in\mathbb{N}} \underbrace{\left\{i \in I : \mu(A_i \cap \Omega_n) > \tfrac{\varepsilon}{2}\right\}}_{\substack{\text{endlich} \\ \text{da } \mu(\Omega_n)<\infty}}}_{\text{abzählbar}}$$

Analog zu oben ist dann

$$I_0 := \bigcup_{n\in\mathbb{N}} \underbrace{\left\{i \in I : \mu(A_i) > \tfrac{1}{n}\right\}}_{\text{abzählbar}}$$

auch höchstens abzählbar, da abzählbare Vereinigung abzählbarer Mengen. Doch tatsächlich ist $I_0 = I$.
□

Beispiel: Ist μ σ-endlich, so ist der Unstetigkeitsbereich $\mathcal{C}(\mu)^c$ höchstens abzählbar.

A.3.4 Satz über Quader mit Masselosem Rand

Seien $(\mu_j)_{j\in J}$ abzählbar viele, σ-endliche Maße auf $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$. Dann gibt es zu jedem Quader der Form

$$Q(\mathbf{a},\mathbf{b}) := \bigtimes_{i=1}^{n} (a_i, b_i) \;\subseteq \mathbb{R}^n \quad , \quad \mathbf{a},\mathbf{b} \in \mathbb{R},\; a_i < b_i \tag{A.3.4.1}$$

eine Folge von Quadern $Q(\mathbf{a}^k, \mathbf{b}^k) \subseteq \mathbb{R}^n,\; a_i^k < b_i^k$ mit

$$Q(\mathbf{a}^k, \mathbf{b}^k) \underset{k\to\infty}{\nearrow} Q(\mathbf{a},\mathbf{b})$$

und $\mu_j\left(\partial Q(\mathbf{a}^k, \mathbf{b}^k)\right) = 0 \quad \forall\, j \in J$.

Beweis: Zu $t \in \mathbb{R},\; 1 \le i \le n$ setze

$$H_i(t) := \{\mathbf{x} \in \mathbb{R}^n : x_i = t\}$$

Dann gibt es nach A.3.3 zu jedem $i \in \mathbb{N}$ und $j \in J$ höchstens abzählbar viele $t \in \mathbb{R}$ mit $\mu_j(H_i(t)) > 0$ und

$$A_i := \{t \in \mathbb{R} : \mu_j(H_i(t)) = 0 \quad \forall\, j \in J\} = \Bigg[\underbrace{\bigcup_{j\in J} \{t \in \mathbb{R} : \mu_j(H_i(t)) > 0\}}_{\text{abzählbar}}\Bigg]^c$$

ist dicht in $\mathbb{R}$.
Ist nun $Q(\mathbf{a},\mathbf{b}) \subseteq \mathbb{R}^n$ ein beliebiger Quader der Form (A.3.4.1), so gibt es für jedes $1 \le i \le n$ Folgen $(a_i^k)_k,\; (b_i^k)_k \subseteq A_i$ mit

$$a_i^k \underset{k\to\infty}{\searrow} a_i \quad , \quad b_i^k \underset{k\to\infty}{\nearrow} b_i \quad , \quad a_i^k < b_i^k$$

bzw.

$$Q(\mathbf{a}^k, \mathbf{b}^k) \underset{k\to\infty}{\nearrow} Q(\mathbf{a},\mathbf{b})$$

Wegen

$$\partial Q(\mathbf{a}^k, \mathbf{b}^k) \subseteq \bigcup_{i=1}^{n} \underbrace{H_i(a_i^k) \cup H_i(b_i^k)}_{\mu_j-\text{0-Mengen } \forall\ j\in J}$$

ist auch

$$\mu_j\left(\partial Q(\mathbf{a}^k, \mathbf{b}^k)\right) = 0 \quad \forall\, j \in J$$

□

A.4 Topologie & Funktionalanalysis

A.4.1 Satz: Offene Vereinigungen auf metrischen Räumen

Sei (E,d) ein separabler metrischer Raum und $(G_i)_{i\in I} \subseteq \mathscr{P}(E),\; I \neq \emptyset$ Familie offener Mengen. Dann existieren abzählbare $i_1, i_2, .. \in I$ mit

$$\bigcup_{i\in I} G_i = \bigcup_{k=1}^{\infty} G_{i_k}$$

Beweis: Sei $D \subseteq E$ abzählbar und dicht in E. Zu $n \in \mathbb{N}$ setze

$$D_n := \underbrace{\left\{x \in D : \exists\ i_{x,n} \in I : B_{\frac{1}{n}}(x) \subseteq G_{i_{x,n}}\right\}}_{\text{höchstens abzählbar}}$$

Dann gilt

$$\bigcup_{i\in I} G_i = \underbrace{\bigcup_{n\in\mathbb{N}} \bigcup_{x\in D_n} G_{i_{x,n}}}_{\substack{\text{abzählbare}\\ \text{Vereinigung}}}$$

denn: Zu $y \in G_i$ existiert ein $n \in \mathbb{N}$ mit $B_{\frac{2}{n}}(y) \subseteq G_i$, da G_i offen. Da D dicht in E liegt, existiert ferner ein $x \in D \cap B_{\frac{1}{n}}(y)$, also

$$B_{\frac{1}{n}}(x) \overset{\triangle\text{-Ungl.}}{\subseteq} B_{\frac{2}{n}}(y) \subseteq G_i$$

bzw. $x \in D_n$, woraus folgt

$$y \in B_{\frac{1}{n}}(x) \overset{x \in D_n}{\subseteq} G_{i_{x,n}}$$

Die andere Inklusionsrichtung ist trivial.
□

A.4.2 Korollar: Erzeugung offener Mengen

Ist (E,d) ein separabler metrischer Raum, so kann jede offene Menge $\emptyset \neq G \subseteq E$ durch abzählbar viele offene Kugeln erzeugt werden, sprich

$$\exists\ x_1, x_2, .. \in E,\ \varepsilon_1, \varepsilon_2, .. > 0 : G = \bigcup_{k=1}^{\infty} B_{\varepsilon_k}(x_k)$$

Beweis: Zu jedem $x \in G$ existiert ein $\varepsilon_x > 0$ mit $B_{\varepsilon_x}(x) \subseteq G$, sprich

$$G = \bigcup_{x \in G} B_{\varepsilon_x}(x)$$

Nach Satz A.4.1 folgt die Behauptung.
□

A.4.3 Korollar: Erzeugung offener Mengen auf $\mathbb{R}$

Sei $D \subseteq \mathbb{R}$ dicht in $\mathbb{R}$ und $\emptyset \neq G \subseteq \mathbb{R}$ offen. Dann existieren (offene) Intervalle $I_1, I_2, .. \subseteq \mathbb{R}$ mit $\partial I_k \subseteq D$ und

$$G = \bigcup_{k=1}^{\infty} I_k$$

Beweis: Zu jedem $x \in G$ existiert ein $\varepsilon_x > 0$ mit $B_{\varepsilon_x}(x) \subseteq G$. Da D dicht in $\mathbb{R}$ liegt, existieren $a_x, b_x \in B_{\varepsilon_x}(x)$ mit $a_x < x < b_x$, sprich $x \in (a_x, b_x)$ mit $\partial(a_x, b_x) \subseteq D$. Wegen

$$G = \bigcup_{x \in G} (a_x, b_x)$$

folgt nach Satz A.4.1 die Behauptung.
□

A.4.4 Satz über ε-Umgebungen von Mengen

Sei (E,d) ein metrischer Raum und $\varepsilon > 0$. Zu Menge $A \subseteq E$ bezeichne

$$A^{\varepsilon} := \{x \in E : \exists\ y \in A : d(x,y) < \varepsilon\} = \bigcup_{y \in A} B^o_{\varepsilon}(y) \tag{A.4.4.1}$$

und

$$A_{\varepsilon} := \{x \in A : \forall\ y \in E : d(x,y) < \varepsilon\ \Rightarrow\ y \in A\} = \{x \in E : B^o_{\varepsilon}(x) \subseteq A\} \tag{A.4.4.2}$$

Dann gilt für $A \subseteq E$ und $\varepsilon, \delta > 0$:

1. $A \subseteq (A^{\varepsilon})_{\varepsilon}$ und $(A_{\varepsilon})^{\varepsilon} \subseteq A$
2. $(A^{\varepsilon})^c = (A^c)_{\varepsilon}$ und $(A_{\varepsilon})^c = (A^c)^{\varepsilon}$

3. $((A^\varepsilon)^c)^\varepsilon \subseteq A^c \subseteq ((A_\varepsilon)^c)_\varepsilon$
4. $(A^\varepsilon)^\delta \subseteq A^{\varepsilon+\delta}$
5. $(A_\varepsilon)_\delta \supseteq A_{\varepsilon+\delta}$

Beweis:

1. Ist $x \in A$, so ist natürlich $B^o_\varepsilon(x) \subseteq A^\varepsilon$, sprich $x \in (A^\varepsilon)_\varepsilon$.
   Ist $x \in (A_\varepsilon)^\varepsilon$, so existiert ein $y \in A_\varepsilon$ mit $x \in B^o_\varepsilon(y) \subseteq A$, sprich $x \in A$.
2. Ist $x \in (A^\varepsilon)^c$, das heißt $d(x,y) \geq \varepsilon \ \forall\ y \in A$, so gilt $B^o_\varepsilon(x) \cap A = \emptyset$ bzw. $x \in (A^c)_\varepsilon$.
   Ist $x \in (A^c)_\varepsilon$, sprich $B^o_\varepsilon(x) \subseteq A^c$, so gilt $d(x,y) \geq \varepsilon \quad \forall\ y \in A$ bzw. $\nexists\ y \in A : x \in B^o_\varepsilon(y)$. Daher $x \notin A^\varepsilon$.
   Ist $x \in (A_\varepsilon)^c$, sprich $B^o_\varepsilon(x) \not\subseteq A$, so $\exists\ y \in A^c \cap B^o_\varepsilon(x)$ und daher $x \in B^o_\varepsilon(y)$ bzw. $x \in (A^c)^\varepsilon$.
   Ist $x \in (A^c)^\varepsilon$, sprich $x \in B^o_\varepsilon(y)$ für irgendein $y \in A^c$, so gilt $y \in B^o_\varepsilon(x) \cap A^c \neq \emptyset$, daher $x \notin A_\varepsilon$.
3. $((A^\varepsilon)^c)^\varepsilon \overset{(2)}{=} ((A^c)_\varepsilon)^\varepsilon \overset{(1)}{\subseteq} A^c$.
   $((A_\varepsilon)^c)_\varepsilon) \overset{(2)}{=} ((A_\varepsilon)^\varepsilon)^c \overset{(1)}{\supseteq} A^c$.
4. Sei $x \in (A^\varepsilon)^\delta$, sprich $x \in B^o_\delta(y)$ für irgendein $y \in A^\varepsilon$. Da auch $y \in B^o_\varepsilon(z)$ für irgendein $z \in A$, gilt $x \in B^o_{\varepsilon+\delta}(z)$, sprich $x \in A^{\varepsilon+\delta}$.
5. Ist $x \in A_{\varepsilon+\delta}$, sprich $B^o_{\varepsilon+\delta}(x) \subset A$. Ist nun $y \in B^o_\delta(x)$ und $z \in B^o_\varepsilon(y)$, so gilt $z \in B^o_{\varepsilon+\delta}(x) \subseteq A$, das heißt $B^o_\varepsilon(y) \subseteq A$ und daher $y \in A_\varepsilon$. Daher $B^o_\delta(x) \subseteq A_\varepsilon$, also $x \in (A_\varepsilon)_\delta$.

   □

A.4.5 Satz zur Folgenkompaktheit von Funktionenmengen

Sei $(T, \mathscr{O})$ ein kompakter, topologischer Hausdorff Raum der das erste Abzählbarkeitsaxiom erfüllt, $X = \{x_1, x_2, ..\}$ abzählbar und

$$T^X := \{f : X \to T\}$$

die Menge aller Abbildungen von X nach T, ausgestattet mit dem Konvergenzbegriff

$$f_n \overset{n\to\infty}{\longrightarrow} f \quad :\Leftrightarrow \quad f_n(x) \overset{n\to\infty}{\longrightarrow} f(x) \ \ \forall\ x \in X,\ (f_n)_{n=1}^\infty, f \in T^X$$

Dann besitzt jede Folge $(f_n) \subseteq T^X$ eine in T^X konvergente Teilfolge.

Beweis: Da T kompakt ist und das 1. Abzählbarkeitsaxiom erfüllt, existiert eine unendliche Teilmenge $N_1 \subseteq \mathbb{N}$ mit

$$\exists \lim_{\substack{n\in N_1 \\ n\to\infty}} f_n(x_1) =: f(x_1)$$

(konvergente Teilfolge). Beachte: Da T Hausdorffsch ist, ist dieser Grenzwert eindeutig. Doch aus gleichem Grunde existiert auch eine unendliche Teilmenge $N_2 \subseteq N_1$ mit

$$\exists \lim_{\substack{n\in N_2 \\ n\to\infty}} f_n(x_2) =: f(x_2)$$

Induktiv fortgesetzt, erhält man eine fallende Folge unendlicher Mengen $\mathbb{N} \supseteq N_1 \supseteq N_2 \supseteq \ldots$ so dass

$$\exists \lim_{\substack{n\in N_k \\ n\to\infty}} f_n(x_k) =: f(x_k)$$

Assoziieren die Indizierung der Elemente $x_k \in X$ mit einer entsprechenden Ordnung und setzen n_j als die j-te Zahl aus N_j. Erhalten dadurch eine Teilfolge $(f_{n_j})_{j=1}^\infty$ mit

$$f_{n_j}(x_k) \overset{j\to\infty}{\longrightarrow} f(x_k) \quad , \quad x_k \in X$$

Beachte dass für $j \geq k$ stets $n_j \in N_k$ gilt.
□

A.4.6 Satz: Existenz abklingender Funktionen

Seien $\Omega_1, \Omega \subseteq \mathbb{R}^n$ offen und $\overline{\Omega}_1 \subseteq \Omega_2$ kompakt. Dann existieren stetige, reellwertige Funktionen $\varphi_j \in \mathcal{C}(\mathbb{R}^n)$ mit

$$\operatorname{supp}(\varphi_j) \subseteq \Omega_2 \quad , \quad \varphi_j \searrow 1_{\overline{\Omega}_1} \quad , \quad 0 \leq \varphi_j \leq 1$$

A.4.7 Schwache*-Topologie

Sei $(E, \|\cdot\|)$ ein normierter Raum, dazu der Dualraum E'. Dann definiert man auf E' die *schwache*-Konvergenz* gemäß:

$$a_n \underset{\mathrm{w}^*}{\overset{n\to\infty}{\longrightarrow}} a \quad :\Leftrightarrow \quad a_n(x) \underset{|\cdot|}{\overset{n\to\infty}{\longrightarrow}} a(x) \quad \forall\, x \in E$$

für Dualformen $a_n, a \in E'$. Definiert man für $x_1, .., x_n \in E$ die Halbnormen

$$p_{x_1,..,x_n}(a) := \sup_{1\leq i\leq n} \|a(x_i)\| \quad , \quad a \in E'$$

so gilt die Äquivalenz

$$a_n \underset{\mathrm{w}^*}{\overset{n\to\infty}{\longrightarrow}} a \quad \Leftrightarrow \quad p_{x_1,..,x_n}(a_n - a) \overset{n\to\infty}{\longrightarrow} 0 \;\; \forall\, x_1, .., x_n \in E,\; n \in \mathbb{N}$$

Die Halbnormen $p_{x_1,..,x_n}$ definieren auf E' die *schwache*-Topologie*: Eine Menge $U \subseteq E'$ heißt *schwach*-offen* falls zu jeder Linearform $a \in U$ Punkte $x_1, .., x_n \in E$ und $\varepsilon > 0$ existieren so dass

$$B^o_{x_1,..,x_n,\varepsilon}(a) := \{b \in E' : p_{x_1,..,x_n}(b - a) < \varepsilon\} \subseteq U$$

Dabei stimmt die schwache*-Topologie mit der durch die schwache Konvergenz erzeugten überein, das heißt:

$$a_n \underset{\mathrm{w}^*}{\overset{n\to\infty}{\longrightarrow}} a \quad \Leftrightarrow \quad \forall\, a \in U \subseteq E' \text{ schwach*-offen} : a_n \in U \text{ schließlich immer}$$

Offensichtlich impliziert die Konvergenz $a_n \underset{\|\cdot\|}{\overset{n\to\infty}{\longrightarrow}} a$ stets die schwache*-Konvergenz.

A.4.8 Satz von Banach-Alaoglu

Sei $(E, \|\cdot\|)$ ein normierter Raum, dazu der Dualraum E'. Dann ist die abgeschlossene Einheitskugel

$$B_1(0_{E'}) := \{a \in E' : \|a\| \leq 1\}$$

bzgl. des schwachen*-Topologie kompakt.

Bemerkungen: Sei E separabel. Dann gilt:

(i) Die schwache*-Topologie ist auf $B_1(0_{E'})$ metrisierbar.[5]

(ii) Als kompakter, metrisierbarer Raum ist $B_1(0_{E'})$ (in der schwachen*-Topologie) separabel und erfüllt das 2e Abzählbarkeitsaxiom.

(iii) Insbesondere ist $B_1(0_{E'})$ auch folgenkompakt. Jede Folge $(a_n) \subseteq B_1(0_{E'})$ besitzt also eine Teilfolge $(a_{n_k})_{k=1}^{\infty}$ mit $a_{n_k} \underset{\mathrm{w}^*}{\overset{n\to\infty}{\longrightarrow}} a$ für irgendein $a \in B_1(0_{E'})$.

A.4.9 Satz: Hinreichende Bedingung für Orthogonalprojektionen

Sei $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ ein Hilbertraum, $U \subseteq \mathcal{H}$ ein abgeschlossener Unterraum und $P : \mathcal{H} \to U$ linear mit folgenden Eigenschaften:

1. $Py = y \;\; \forall\, y \in U$

2. $x - Px \perp Px \;\; \forall\, x \in \mathcal{H}$

Dann ist P ein Orthogonalprojektor auf U.

Beweis: Zu zeigen wäre $x - Px \perp U \quad \forall\, x \in \mathcal{H}$. Tatsächlich gilt für $y \in U$:

$$Px + y = Px + \overbrace{Py}^{y} = P(x+y) \perp (x+y) - P(x+y) = x - Px + y - \underbrace{Py}_{y} = x - Px$$

Wegen $Px \perp x - Px$ muss schließlich auch $y \perp x - Px$ sein.
□

A.5 Allgemeine Hilfsaussagen

A.5.1 Lemma von Kronecker

Seien $(x_k)_{k=1}^{\infty} \subseteq \mathbb{R}$ und $0 < b_1 \leq b_2 \leq \dots \in \mathbb{R}$ mit

$$\exists \sum_{k=1}^{\infty} x_k \in \mathbb{R} \quad , \quad b_k \overset{k\to\infty}{\longrightarrow} \infty$$

Dann gilt:

$$\frac{1}{b_n} \sum_{k=1}^{n} b_k x_k \overset{n\to\infty}{\longrightarrow} 0$$

Beweis: Setze

$$s := \sum_{k=1}^{\infty} x_k \quad , \quad r_n := \sum_{k=n+1}^{\infty} x_k, \ n \in \mathbb{N} \quad , \quad r_0 := s$$

Dann sind $r_{n-1} - r_n = x_n$ und es gehen $r_n \overset{n\to\infty}{\longrightarrow} 0$. Ferner gilt

$$\sum_{k=1}^{n} b_k x_k = \sum_{k=1}^{n} b_k (r_{k-1} - r_k) = \sum_{k=0}^{n-1} b_{k+1} r_k - \sum_{k=1}^{n} b_k r_k = b_1 r_0 - b_n r_n + \sum_{k=1}^{n-1} r_k \cdot (b_{k+1} - b_k)$$

bzw.

$$\frac{1}{n} \left| \sum_{k=1}^{n} b_k x_k \right| \leq \frac{b_1 \,|s|}{b_n} + |r_n| + \sum_{k=1}^{n-1} |r_k| \underbrace{\frac{(b_{k+1} - b_k)}{b_n}}_{\geq 0}$$

Sei nun $\varepsilon > 0$ fest, dann existiert ein $N \in \mathbb{N}$ mit $|r_k| \leq \varepsilon \ \forall\, n \geq N$. Mit $c := \sup_{k\in\mathbb{N}} |r_k|$ kann man daher für $n \geq N$ die Abschätzung machen

$$\sum_{k=1}^{n-1} |r_k| \underbrace{\frac{(b_{k+1} - b_k)}{b_n}}_{\geq 0} = \sum_{k=1}^{N-1} |r_k| \frac{b_{k+1} - b_k}{b_n} + \sum_{k=N}^{n-1} |r_k| \frac{b_{k+1} - b_k}{b_n}$$

$$\leq \sum_{k=1}^{N-1} c \cdot \frac{b_{k+1} - b_k}{b_n} + \sum_{k=N}^{n-1} \varepsilon \frac{b_{k+1} - b_k}{b_n} = \underbrace{\frac{c}{b_n}(b_N - b_1)}_{\overset{n\to\infty}{\longrightarrow} 0} + \underbrace{\varepsilon \cdot \frac{b_n - b_N}{b_n}}_{\overset{n\to\infty}{\longrightarrow} \varepsilon}$$

das heißt

$$\overline{\lim_{n\to\infty}} \frac{1}{b_n} \left| \sum_{k=1}^{n} b_k x_k \right| \leq \underbrace{\overline{\lim_{n\to\infty}} \frac{b_1 \,|s|}{b_n}}_{0} + \underbrace{\lim_{n\to\infty} |r_n|}_{0} + \underbrace{\overline{\lim_{n\to\infty}} \sum_{k=1}^{n-1} |r_k| \cdot \frac{b_{k+1} - b_k}{b_n}}_{\leq \varepsilon} \leq \varepsilon \quad \forall\, \varepsilon > 0$$

bzw.

$$\overline{\lim_{n\to\infty}} \frac{1}{b_n} \left| \sum_{k=1}^{n} b_k x_k \right| = 0$$

□

A.5.2 Folgerungen des Lemmas von Kronecker

Seien $(x_k)_{k=1}^{\infty} \subseteq \mathbb{R}$.

1. Sind $0 < b_1 \leq b_2 \leq \cdots \to \infty$ so dass
$$\sum_{k=1}^{\infty} \frac{x_k}{b_k}$$
in $\mathbb{R}$ konvergiert, so gilt
$$\frac{1}{b_n} \sum_{k=1}^{n} x_k \overset{n\to\infty}{\longrightarrow} 0$$

2. Konvergieren $x_k \overset{k\to\infty}{\longrightarrow} x \in \mathbb{R}$, so konvergiert auch
$$\frac{1}{n} \sum_{k=1}^{n} x_k \overset{n\to\infty}{\longrightarrow} x$$

Beachte: Die Umkehrung von (1) gilt im allgemeinen nicht. So ist z.B.
$$\sum_{k=3}^{\infty} \frac{1}{k(\ln k)(\ln\ln k)} = \infty$$
obwohl
$$\frac{1}{(\ln n)(\ln\ln n)} \underbrace{\sum_{k=1}^{n} \frac{1}{k}}_{\sim \ln n} \overset{n\to\infty}{\longrightarrow} 0$$

Beweis:

1. Wende Lemma von Kronecker A.5.1 auf die Folge $y_k := x_k/b_k$ an.

2. Setze $y_k := x_k - x_{k-1},\ x_0 := 0$, dann gilt
$$\sum_{k=1}^{\infty} y_k = x$$
bzw. nach Kronecker A.5.1
$$\frac{1}{n} \sum_{k=1}^{n} k y_k \overset{n\to\infty}{\longrightarrow} 0 \tag{A.5.2.1}$$
Doch anderseits ist
$$\frac{1}{n} \sum_{k=1}^{n} k y_k = \frac{1}{n} \left[1(x_1 - x_0) + 2(x_2 - x_1) + 3(x_3 - x_2) + \cdots + n(x_n - x_{n-1})\right]$$
$$= \frac{1}{n}\left[n x_n - \sum_{k=1}^{n} x_k\right] = \underbrace{x_n}_{\overset{n\to\infty}{\longrightarrow} x} - \frac{1}{n} \sum_{k=1}^{n} x_k$$
so dass mit (A.5.2.1) folgt
$$\frac{1}{n} \sum_{k=1}^{n} x_k \overset{n\to\infty}{\longrightarrow} x$$

□

Beispiel: Für $\alpha > 1$ existiert

$$\sum_{k=1}^{\infty} \frac{1}{k\,[\ln(k+1)]^{\alpha}}$$

in $\mathbb{R}$, so dass

$$\frac{1}{[\ln(n+1)]^{\alpha}} \sum_{k=1}^{n} \frac{1}{k} \overset{n\to\infty}{\longrightarrow} 0$$

geht.

A.5.3 Lemma: Stetigkeit der Darstellung der Exponentialfunktion

Seien $z_n, z \in \mathbb{C}$ mit $z_n \overset{n\to\infty}{\longrightarrow} z$. Dann gilt:

$$\left(1 + \frac{z_n}{n}\right)^n \overset{n\to\infty}{\longrightarrow} e^z$$

Beweis: Bekanntlich geht

$$\left(1 + \frac{x}{n}\right)^n \overset{n\to\infty}{\longrightarrow} e^x \quad \forall\, x \in \mathbb{C}$$

Es genügt also zu zeigen, dass

$$\left|\left(1 + \frac{z}{n}\right)^n - \left(1 + \frac{z_n}{n}\right)^n\right| \overset{n\to\infty}{\longrightarrow} 0$$

Tatsächlich folgt durch die allgemeine Abschätzung

$$|a^n - b^n| = |a - b| \cdot \left|a^{n-1} + a^{n-2}b + \ldots ab^{n-2} + b^{n-1}\right| \le n \cdot |a - b| \cdot \max\left\{|a|, |b|\right\}^{n-1}$$

mit $a := 1 + \frac{z_n}{n}$, $b := 1 + \frac{z}{n}$ und die Abschätzung

$$|a| \le 1 + \frac{|z_n|}{n} \le 1 + \frac{|z| + 1}{n} \quad \forall\, n \ge n_0$$

für genügend großes $n_0 \in \mathbb{N}$ auch die Abschätzung

$$\left|\left(1 + \frac{z}{n}\right)^n - \left(1 + \frac{z_n}{n}\right)^n\right| \le n \cdot \frac{|z_n - z|}{n} \cdot \underbrace{\left(1 + \frac{|z| + 1}{n}\right)^{n-1}}_{\overset{n\to\infty}{\longrightarrow} e^{|z|+1}} \overset{n\to\infty}{\longrightarrow} 0$$

□

A.5.4 Satz über n-mal stetig differenzierbare Funktionen (Taylor)

Sei $f : \mathbb{R} \to \mathbb{C}$ n-mal stetig differenzierbar. Dann gilt

$$f(t) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} \cdot t^k + \frac{t^n}{n!} \cdot \Theta(t) \quad , \quad t \in \mathbb{R}$$

für geeignetes $\Theta : \mathbb{R} \to \mathbb{C}$ mit

$$\lim_{t\to 0} \Theta(t) = 0$$

Beweis: Nach Taylor ist bekanntlich

$$f(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} \cdot t^k + R_{n-1}(t)$$

mit

$$R_{n-1}(t) = \frac{1}{(n-1)!} \int_0^t (t - u)^{n-1} f^{(n)}(u)\; du$$

als Restglied. Daher

$$f(t) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} \cdot t^k + \underbrace{\frac{1}{(n-1)!} \int\limits_0^t (t-u)^{n-1} \left[f^{(n)}(u) - f^{(n)}(0) \right] \, du}_{=: \frac{t^n}{n!} \cdot \Theta(t)}$$

wobei man abschätzen kann

$$\frac{|t|^n}{n!} \cdot |\Theta(t)| \leq \underbrace{\sup_{0 \leq \vartheta \leq 1} \left| f^{(n)}(\vartheta t) - f^{(n)}(0) \right|}_{\substack{\xrightarrow{t \to 0} 0 \\ \text{da } f^{(n)} \text{ stetig}}} \cdot \underbrace{\frac{1}{(n-1)!} \int\limits_0^t |t-u|^{n-1} \, du}_{\frac{|t|^n}{n!}}$$

sprich $\Theta(t) \stackrel{t \to 0}{\longrightarrow} 0$.

A.5.5 Differentiationstheorem von Lebesgue

Sei $f \in L_1(\mathbb{R}^n)$. Dann gilt für fast-alle $x \in \mathbb{R}^n$

$$\lim_{\varepsilon \to 0^+} \frac{1}{\lambda^n(B_\varepsilon(x))} \int\limits_{B_\varepsilon(x)} f(y) \, dy = f(x)$$

[4]

B Symbol-Referenz

$\mathbb{R}_+$: $= [0, \infty)$.

$\mathbb{C}_+$: $\mathbb{R}_+ + i\mathbb{R}_+$.

$\mathbf{e}_k$: Standard-Einheitsvektor in $\mathbb{R}^n$.

iid: **I**dentically, **I**ntependently **D**istributed.

z^*: Komplex-Konjugierte zu $z \in \mathbb{C}$.

$\|\mathbf{x}\|_\infty$: Supremumsnorm für Vektoren $\mathbf{x} \in \mathbb{C}^n$: $\|\mathbf{x}\|_\infty := \sup_{1 \le j \le n} \left|x^j\right|$.

$\mathcal{C}(M)$: Raum aller stetigen Funktionen $M \to \mathbb{C}$.

$\mathcal{C}(M, N)$: Raum aller stetigen Funktionen $M \to N$.

$\mathcal{C}_b(M)$: Raum aller stetigen, beschränkten Funktionen $M \to \mathbb{C}$. Konventionell ausgestattet mit der Norm $\|f\|_\infty := \sup_{t \in M} |f(t)|$.

$\mathcal{C}_b(M, N)$: Raum aller stetigen, beschränkten Funktionen $M \to N$.

1_A: Indikatorfunktion für Menge A.

$\mathscr{P}(M)$: Potenzmenge von M.

$\sigma(\mathscr{A})$: Für $\mathscr{A} \subset \mathscr{P}(M)$, kleinste σ-Algebra die $\mathscr{A}$ enthält.

$\mathfrak{S} \otimes \mathfrak{M}$: Für σ-Algebren $\mathfrak{S}, \mathfrak{M}$, produkt-$\sigma$-Algebra: $\mathfrak{S} \otimes \mathfrak{M} := \sigma(\{A \times B : A \in \mathfrak{S},\ B \in \mathfrak{M}\})$.

$\mathfrak{S}^n$: Für σ-Algebra $\mathfrak{S}$: $\mathfrak{S}^n := \underbrace{\mathfrak{S} \otimes \mathfrak{S} \otimes \cdots \otimes \mathfrak{S}}_{\times n}$.

$\mathscr{Z}(\Omega)$: Für Produktraum Ω: Menge aller Zylindermengen auf Ω. Siehe 2.1.1.

Π_J: Projektion aus Produktraum Ω in die Komponenten $J = \{j_1, .., j_n\}$. Siehe 2.1.5.

$(E, \|\cdot\|)$: Normierter Raum.

E': Für normierten Raum E, der Dualraum.

(E, d): Metrischer Raum.

$B_r(x)$: Abgeschlossene Kugel um x mit Radius r.

$B_r^o(x)$: Offene Kugel um x mit Radius r.

B_1^n: n-dimensionale Einheitskugel: $B_n^1 := B_1(0)$.

$d(A, B)$: Für Untermengen $A, B \subset E$ eines metrischen Raumes XE: $d(A, B) := \inf \{d(a, b) : a \in A,\ b \in B\}$.

$d(x, A)$: Für Untermenge $A \subset E$ eines metrischen Raumes und $x \in E : d(x, A) := d(\{x\}, A)$. Siehe A.2.1.

$(T, \mathscr{O})$: Topologischer Raum mit Topologie (offene Mengen) $\mathscr{O}$.

$\overline{A}$: Abschluss einer Untermenge $A \subseteq T$ eines topologischen Raumes T: $\mathrm{cl}(A) = \overline{A}$.

A^o: Inneres einer Untermenge $A \subseteq T$ eines topologischen Raumes T: $\mathrm{int}(A) = A^o$.

∂B: Rand der Untermenge $B \subseteq T$ eines topologischen Raumes: $\partial B := \mathrm{cl}(B) \setminus \mathrm{int}(B)$

$\mathscr{B}(T)$: Borel-σ-Algebra des topologischen Raumes T: $\mathscr{B}(T) := \sigma(\mathscr{O}(T))$.

$(T, \mathscr{B}(T))$: Messbarer Raum über Grundmenge T (topologischen Raum) und Borel-σ-Algebra $\mathscr{B}(T)$.

$\mathfrak{M}(\Omega)$: Für messbaren Raum $(\Omega, \mathfrak{S})$, die Menge aller Maße auf $(\Omega, \mathfrak{S})$.

$\mathfrak{M}_1(\Omega)$: Für messbaren Raum $(\Omega, \mathfrak{S})$ die Menge aller Wahrscheinlichkeitsmaße auf $(\Omega, \mathfrak{S})$.

$\Lambda(T)$: Die Menge aller signierten Borel-Maßen auf dem topologischen Raum T. Siehe 4.4.2.

$\mu * \nu$: Faltung zweier Maße μ, ν.

$(\Omega, \mathfrak{S}, \mu)$: Maßraum über Grundmenge Ω, mit σ-Algebra $\mathfrak{S}$ und Maß μ.

μ_X: Bildmaß von Maß μ unter Abbildung X: $\mu_X := \mu \circ X^{-1}$.

$\mu * \nu$: Faltung der Maße μ, ν auf $(\Omega, \mathfrak{S}, \mu)$: $\mu * \nu(B) := \int\limits_{\Omega} \mu(B - x)\ d\nu(x)$.

λ: Lebesgue-Maß auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$.

λ^n: Lebesgue-Maß auf $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$.

$\mathcal{U}$: Auf $[0,1]$ gleichverteilte Zufallsvariable.

$\mathcal{N}_{x_0,s}$: Normalverteilung mit Erwartungswert x_0 und Varianz s.

$\mathcal{N}_{\mu,R}$: Verteilung eines Gaußvektors $X : \Omega \to \mathbb{R}^n$ mit Erwartungswert $\mathbb{E}(X) = \mu$ und Kovarianzmatrix $R = (\mathrm{Cov}(X_i, X_j))_{i,j=1}^n$.

$\mathcal{B}_{n,p}$: Binomialverteilung mit Versuchs-Erfolgswahrscheinlichkeit p, über n Versuche.

π_λ: Poissonverteilung mit Parameter (Intensität) $\lambda > 0$.

Exp_λ: Exponentialverteilung mit Parameter λ: $\mathrm{Exp}_\lambda((t,\infty)) = e^{-\lambda t},\ t \geq 0$.

$L_p(\Omega)$: Funktionenraum der p-integrierbaren, komplexwertigen Funktionen im Maßraum $(\Omega, \mathfrak{S}, \mu)$.

$\mathcal{F}$: Fouriertransformation.

$\mathrm{supp}(f) := \mathrm{cl}\{x \in X : f(x) \neq 0\}$: Träger von $f : X \to \mathbb{C}$.

S^T: Für Mengen $S, T \neq \emptyset$: $S^T := \{f : T \to S\}$.

$\mathfrak{B}(\Omega)$: Raum aller beschränkten, komplexwertigen, gegebenfalls messbaren, Funktionen auf Ω. Dabei ist $\mathfrak{B}(\Omega)$ mit der $\|\cdot\|_\infty$-Norm ausgestattet.

$X \overset{\mathbf{d.}}{=} Y$: Zufallsvariablen X, Y identisch verteilt, das heißt $\mathcal{P}_X = \mathcal{P}_Y$.

$A^\dagger$: Adjungierte Operator zu A. Für Matrix $A \in \mathbb{C}^{n\times n}$ ist $A^\dagger = (A^T)^*$.

$\mathbb{E}(X)$: Erwartungswert der Zufallsvariable X.

$\mathrm{Cov}(X,Y)$: Kovarianz zweier Zufallsvektoren X, Y, $\mathrm{Cov}(XY) := \mathbb{E}\left[(X - \mathbb{E}X) \otimes (Y - \mathbb{E}Y)\right]$. Siehe 1.8.2.

$\mathrm{Cov}(X)$: Kovarianzmatrix des Zufallsvektors X, $\mathrm{Cov}(X) := \mathrm{Cov}(X,X)$. Siehe 1.8.2.

$M(X)$: Median der reellen Zufallsvariablen X. Siehe 3.5.1.

$\widetilde{\mathcal{P}}$: (auch $\mathcal{F}(\mathcal{P})$) Fouriertransformierte des Maßes $\mathcal{P}$ auf $\mathscr{B}(\mathbb{R}^n)$.

$\mathrm{supp}(\mu)$: Träger des Maßes μ. Siehe 1.1.2.

$\mathcal{C}(\mu)$: Stetigkeitsbereich des Maßes μ. Siehe 1.1.2.

$\mathcal{C}(F)$: Stetigkeitsbereich der Abbildung F, das heißt $x \in \mathcal{C}(F) \ \Leftrightarrow\ F$ stetig in x.

$\widetilde{X}$: (auch $\mathcal{F}(X)$) Fouriertransformierte der Zufallsvariable X, $\widetilde{X} := \widetilde{\mathcal{P}_X}$.

$\mu_n \underset{\mathbf{weak}}{\overset{n\to\infty}{\longrightarrow}} \mu$: Die Maß-Folge (μ_n) strebt schwach gegen μ. Kurz: $\mu_n \underset{\mathrm{w}}{\overset{n\to\infty}{\longrightarrow}} \mu$.

$X^n \underset{\mathbf{p}}{\overset{n\to\infty}{\longrightarrow}} X$: Die Zufallsvariablen $(X^n)_{n\in\mathbb{N}}$ gehen stochastisch (in Wahrscheinlichkeit) gegen X.

$X^n \underset{\mathbf{d}}{\overset{n\to\infty}{\longrightarrow}} X$: Die Zufallsvariablen $(X^n)_{n\in\mathbb{N}}$ gehen in Verteilung (schwach) gegen X.

$X^n \underset{\mathbf{f.s.}}{\overset{n\to\infty}{\longrightarrow}} X$**:** Die Zufallsvariablen $(X^n)_{n\in\mathbb{N}}$ gehen in fast-sicher (fast-überall) gegen X.

$a_n \underset{\mathbf{w*}}{\overset{n\to\infty}{\longrightarrow}} a$**:** Schwache* Konvergenz der Linearformen a_n. Siehe A.4.7.

$\mathbb{E}(X \mid \mathscr{C})$**:** Bedingter Erwartungswert der Zufallsgröße X, gegeben die σ-Algebra $\mathscr{C}$. Siehe 6.2.1.

$\mathcal{P}(A \mid \mathscr{C})$**:** Bedingte Wahrscheinlichkeit der messbaren Menge A, gegeben die σ-Algebra $\mathscr{C}$. Siehe 6.2.10.

$\mathbb{E}(X \mid Z)$**:** Bedingter Erwartungswert der Zufallsgröße X, gegeben die Zufallsgröße Z. Siehe 6.3.1.

$\mathcal{P}(A \mid Z)$**:** Bedingte Wahrscheinlichkeit der messbaren Menge A, gegeben die Zufallsgröße Z. Siehe 6.3.1.

$\mathbb{E}(X \mid Z = s)$**:** Bedingter Erwartungswert der Zufallsgröße X, gegeben $Z = s$. Siehe 6.3.4.

$\mathcal{P}(A \mid Z = s)$**:** Bedingte Wahrscheinlichkeit der messbaren Menge, gegeben $Z = s$. Siehe 6.3.4.

$\rho(x \mid z)$**:** Bedingte Dichte der Zufallsgröße X an der Stelle x, gegeben $Z = z$. Siehe 6.3.6.

Literatur

[1] *Wahrscheinlichkeitstheorie*, H. Bauer
Walter de Gruyter, 2001

[2] *Grundbegriffe der Wahrscheinlichkeitsrechnung*, A.N. Kolmogorov
Springer, 1933

[3] *Lecture Notes on Ergodicity*, Terence Tao
http://terrytao.wordpress.com/2008/02/04/254a-lecture-9-ergodicity/ (18.02.2010)

[4] *Sur l'intégration des fonctions discontinues*, H. Lebesgue
Annales scientifiques de l'École Normale Supérieure, Sér. 3, 27 (1910), p. 361-450

[5] *The weak dual topology*, Gabriel Nagy
Kansas State University
http://www.math.ksu.edu/ñagy/real-an/2-04-weak-dual.pdf (21.02.2010)

Index