Wahrscheinlichkeitstheorie 2

Cornelius Hanel

Last updated: November 10, 2025

Studentische Mitschrift aus dem Sommersemester 2024. Fehler oder Ergänzungen gerne an
corneliush99@univie.ac.at.

Contents

8. Produkträume und Produktmaße

Seien im Folgenden $(\Omega_i, \mathcal{A}_i, \mu_i), i = 1, \ldots, n$ jeweils σ-endliche Maßräume. Setze $[n] := \{1, \ldots, n\}$.

Erzeugung von Produktmaßen

Wir möchten eine σ-Algebra $\mathcal{A}$ und Maß μ auf $\Omega = \Omega_1 \times \cdots \times \Omega_n$ finden, sodass

$$A_1 \times \cdots \times A_n \in \mathcal{A},\ \mu(A_1 \times \cdots \times A_n) = \prod_{i \in [n]} \mu_i(A_i)$$

für $A_i \in \mathcal{A}_i, i \in [n]$.

8.1. Definition: Setze

$$\mathcal{R} := \{A_1 \times \ldots \times A_n : A_i \in \mathcal{A}_i, i \in [n]\}$$

die Familie der messbare Rechtecke, i.e. eine Menge $A \in \mathcal{R}$ heißen messbares Rechteck. Man nennt

$$\bigotimes_{i \in [n]} \mathcal{A}_i := \sigma(\mathcal{R})$$

die Produkt-σ-Algebra auf Ω.

Bemerkung:

- Messbares Rechteck $\neq$ geometrisches Rechteck, e.g. $\mathbb{R}^2$ mit Borelmengen und $\mathbb{Q} \times \mathbb{Q}$.
- $\mathcal{R}$ ist ein π-System, da $\prod_{i \in [n]} A_i \cap \prod_{i \in [n]} B_i = \prod_{i \in [n]} (A_i \cap B_i)$ und $\mathcal{A}_i, i \in [n]$ als σ-Algebren durchschnittsstabil sind.
- $\bigotimes_{i \in [n]} \mathcal{A}_i \neq \prod_{i \in [n]} \mathcal{A}_i$.

8.2. Proposition: Betrachte die Koordinatenabbildungen

$$\Omega \ni \omega = (\omega_1, \ldots, \omega_n) \mapsto \omega_i = \pi_i(\omega) \in \Omega_i, i \in [n].$$

Dann gilt $\bigotimes_{i \in [n]} \mathcal{A}_n = \sigma(\pi_i, i \in [n])$.

Beweis: Mit Lemma 4.5 gilt

$$\sigma(\pi_i, i \in [n]) = \sigma\left(\left\{\bigcap_{i\in I} \pi_i^{-1}(A_i) : I \subseteq [n] \text{ endlich}, A_i \in \mathcal{A}_i, i \in I\right\}\right)$$
$$= \sigma\left(\left\{\bigcap_{i\in[n]} \pi_i^{-1}(A_i) : A_i \in \mathcal{A}_i, i \in [n]\right\}\right),$$

da $\bigcap_{i\in I} \pi_i^{-1}(A_i) = \left(\bigcap_{i\in I} \pi_i^{-1}(A_i)\right) \cap \left(\bigcap_{i\notin I} \pi_i^{-1}(\Omega_i)\right)$. Außerdem gilt $\bigcap_{i\in[n]} \pi_i^{-1}(A_i) = \prod_{i\in[n]} A_i$. □

8.3. Lemma Definiere

$$\mathcal{R} \ni A = A_1 \times \ldots \times A_n \mapsto \mu(A) := \prod_{i\in[n]} \mu_i(A_i) \in [0, \infty]$$

Dann ist μ eine σ-additive Mengenfunktion auf $\mathcal{R}$ und damit insbesondere auch endlich additiv.

Beweis: Seien $R_j \in \mathcal{R}, j \in [m]$ disjunkt, sodass auch $R = \bigcup_{j\in[m]} R_j \in \mathcal{R}$ (da $\mathcal{R}$ keine Algebra ist, ist dies nicht zwingend der Fall). Schreibe $R_j = A_j^{(1)} \times \ldots \times A_j^{(n)}$ und $R = A^{(1)} \times \ldots \times A^{(n)}$. Es gilt $\mathbb{1}_R(\omega) = \sum_{j\geqslant 1} \mathbb{1}_{R_j}(\omega)$ (Überlegung!) und damit

$$\mathbb{1}_R(\omega) = \prod_{i\in[n]} \mathbb{1}_{A^{(i)}}(\omega_i) = \sum_{j\geqslant 1} \prod_{i\in[n]} \mathbb{1}_{A_j^{(i)}}(\omega_i).$$

Integriere nun beide Seiten bezüglich $d\mu_n$ und erhalte mit MONK

$$\mu_n\left(A^{(n)}\right) \prod_{i\in[n-1]} \mathbb{1}_{A^{(i)}}(\omega_i) = \sum_{j\geqslant 1} \mu_n\left(A_j^{(n)}\right) \prod_{i\in[n-1]} \mathbb{1}_{A_j^{(i)}}(\omega_i).$$

Integriere induktiv bezüglich $\mu_{n-1}, \ldots, \mu_1$ und erhalte schließlich

$$\mu(R) = \prod_{i\in[n]} \mu_i\left(A^{(i)}\right) = \sum_{j\geqslant 1} \prod_{i\in[n]} \mu_i\left(A_j^{(i)}\right) = \sum_{j\geqslant 1} \mu(R_j)$$

wobei die erste und letzte Gleichung per Definition von μ folgen. □

8.4. Lemma: Die Mengenfamilie

$$\mathcal{R}^* := \left\{\bigcup_{m=1}^{M} R_m : M \in \mathbb{N}, R_m \in \mathcal{R}, m \in [M] \text{ disjunkt}\right\} \supseteq \mathcal{R}$$

ist eine Algebra.

Beweis:

- $\Omega \in \mathcal{R}^*$: folgt sofort aus $\Omega = \Omega_1 \times \Omega_2$ und $\Omega_1 \in \mathcal{A}_1, \Omega_2 \in \mathcal{A}_2$.

- $A, B \in \mathcal{R}^* \implies A \cap B \in \mathcal{R}^*$: Schreibe $A = \bigcup_{i=1}^k A_{i,1} \times \ldots \times A_{i,d}$ und $B = \bigcup_{j=1}^\ell B_{j,1} \times \ldots \times B_{j,d}$. Dann gilt

$$\begin{aligned} A \cap B &= \left(\bigcup_{i=1}^k A_{i,1} \times \ldots \times A_{i,d} \right) \cap \left(\bigcup_{j=1}^\ell B_{j,1} \times \ldots \times B_{j,d} \right) \\ &= \bigcup_{i=1}^k \bigcup_{j=1}^\ell (A_{i,1} \times \ldots \times A_{i,d} \cap B_{j,1} \times \ldots \times B_{j,d}) \end{aligned}$$

  Wir wissen, dass eine Vereinigung zweier messbarer Rechtecke wieder ein messbares Rechteck ist. Damit gilt $A \cap B \in \mathcal{R}^*$.

- $A \in \mathcal{R}^* \implies A^c \in \mathcal{R}^*$: Sei hier auch wieder $A = \bigcup_{i=1}^k A_{i,1} \times \ldots \times A_{i,d}$. Dann gilt mit de Morgan $A^c = \bigcap_{i=1}^k (A_{i,1} \times \ldots \times A_{i,d})^c$. Zeige also, dass das Komplement eines messbaren Rechtecks eine endliche Vereinigung messbarer Rechtecke ist (dann folgt mit Abgeschlossenheit bezüglich endlicher Durchschnitte die Aussage). Zeige also

$$A_1 \times \ldots \times A_d \in \mathcal{R} \implies (A_1 \times \ldots \times A_d)^c \in \mathcal{R}^*$$

  Definiere dazu für $i = 1, \ldots, d$ die Koordinatenabbildungen

$$\pi_i : \Omega \to \Omega_i, \ (\omega_1, \ldots, \omega_d) \mapsto \omega_i$$

  Dann gilt

$$A_1 \times \ldots \times A_d = \bigcap_{i=1}^d \{\pi_i \in A\}$$

  und es folgt

$$(A_1 \times \ldots \times A_d)^c = \bigcup_{j=1}^d \bigcup_{\substack{J \subseteq \{1,\ldots,d\} \\ |J|=j}} \left[\left(\bigcap_{i \in J} \{\pi_i \notin A_i\} \right) \cap \left(\bigcap_{i \notin J} \{\pi_i \in A_i\} \right) \right]$$

  wobei $\left(\bigcap_{i \in J} \{\pi_i \notin A_i\}\right) \cap \left(\bigcap_{i \notin J} \{\pi_i \in A_i\}\right) \in \mathcal{R}$ disjunkt sind. Da die Vereinigung endlich ist, folgt die Aussage. $\square$

Bemerkung: Bisher ist $\mu : \mathcal{R} \to [0, \infty]$ definiert als $\mu(A_1 \times \ldots \times A_d) := \mu(A_1) \cdot \ldots \cdot \mu(A_d)$. Wir erweitern μ nun zu einer Abbildung $\mu^* : \mathcal{R}^* \to [0, \infty]$ mit

$$\mu^* \left(\bigcup_{i=1}^k R_i \right) := \sum_{i=1}^k \mu(R_i)$$

für $R_i \in \mathcal{R}, i = 1, \ldots, k$ disjunkt.

8.5. Lemma: Für $R = \bigcup_{j=1}^m R_j \in \mathcal{R}^*$, definiere

$$\mu^*; \mathcal{R}^* \to [0, \infty], R \mapsto \sum_{j=1}^k \mu(R_j).$$

Dann ist μ^* wohldefiniert, d.h. unabhängig von der Darstellung von $R \in \mathcal{R}^*$.

Beweis: Sei $A = \bigcup_{i=1}^n R_i = \bigcup_{j=1}^m S_j$ mit $R_i, S_j \in \mathcal{R}$ für $i = 1, \ldots, n$ und $j = 1, \ldots, m$. Dann gilt

$$\begin{aligned}
\mu^*\left(\bigcup_{i=1}^n R_i\right) &\overset{\text{Def.}}{=} \sum_{i=1}^n \mu(R_i) \\
&= \sum_{i=1}^n \mu(R_i \cap A) = \sum_{i=1}^n \mu\left(R_i \cap \bigcup_{j=1}^m S_j\right) \\
&= \sum_{i=1}^n \mu\left(\bigcup_{j=1}^m (S_j \cap R_i)\right) = \sum_{i=1}^n \sum_{j=1}^m \mu\left(S_j \cap R_i\right) \\
&= \sum_{j=1}^m \mu\left(S_j \cap \bigcup_{i=1}^n R_i\right) = \sum_{j=1}^m \mu(S_j) \\
&= \mu^*\left(\bigcup_{j=1}^m S_j\right)
\end{aligned}$$

□

Bemerkung: Um μ^* nun zu einem Maß auf $(\prod_{i=1}^n \Omega_i, \sigma(\mathcal{R}^*))$ zu erweitern, sind mit dem Maßerweiterungssatz von Carathéodory folgende Voraussetzungen notwendig:

(i) $\mu^*(\emptyset) = 0$.

(ii) σ-Additivität: Für $A_i \in \mathcal{R}^*$ disjunkt mit $\bigcup_{i \geqslant 1} A_i \in \mathcal{R}^*$ gilt

$$\mu^*\left(\bigcup_{i \geqslant 1} A_i\right) = \sum_{i \geqslant 1} \mu^*(A_i)$$

(iii) σ-Endlichkeit: $\exists B_j \in \mathcal{R}^*, j \geqslant 1 : \Omega = \bigcup_{j \geqslant 1} B_j$ und $\forall j \geqslant 1 : \mu^*(B_j) < \infty$.

8.6. Lemma: μ^* erfüllt die Eigenschaften (i) bis (iii) aus der obigen Bemerkung.

Beweis:

(i) $\mu^*(\emptyset) = \mu^*\left(\prod_{i=1}^n \emptyset\right) = \prod_{i=1}^n \mu_i(\emptyset) = 0$.

(ii) Für $i \in [n]$, seien $(B_j^{(i)})_{j \geqslant 1}$ so, dass $\Omega_i = \bigcup_{j \geqslant 1} B_j^{(i)}$ und $\mu_i(B_j^{(i)}) < \infty$ für alle $j \geqslant 1$. Dann erfüllt $B_j := \prod_{i=1}^n B_j^{(i)}$ die gewünschten Eigenschaften.

(iii) Mit Lemma 3.10. genügt es hier zu zeigen, dass μ^* auf $\mathcal{R}$ σ-additiv ist, was sofort aus $\mu = \mu^*|_{\mathcal{R}}$ und Lemma 8.4. folgt. □

8.7. Satz: Seien $(\Omega_i, \mathcal{A}_i, \mu_i)$ jeweils σ-endliche Maßräume für $i = 1, \ldots, d$. Dann exisitert mit dem Maßerweiterungssatz von Carathéodory ein eindeutiges σ-endliches Maß μ auf dem Produktraum $(\prod_{i=1}^d \Omega_i, \bigotimes_{i=1}^d \mathcal{A}_i)$ mit der Eigenschaft, dass

$$\forall A_i \in \mathcal{A}_i, i = 1, \ldots, d : \mu(A_1 \times \ldots \times A_d) = \prod_{i=1}^d \mu_i(A_i)$$

Man nennt μ das Produktmaß.

Beweis: Folgt sofort aus dem Maßerweiterungssatz von Carathéodory.

8.8. Korollar: Seien $F_1, \dots, F_d$ Verteilungsfunktionen auf $\mathbb{R}$. Dann gibt es einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ und Zufallsvariablen $X_i : \Omega \to \mathbb{R}, i = 1, \dots, d$, sodass $X_i \sim F_i$ und die X_i unabhängig sind.

Beweis: Jedes F_i definiert ein Wahrscheinlichkeitsmaß $\mathbb{P}_i$ auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mit

$$\mathbb{P}_i((-\infty, t]) := F_i(t)$$

(cf. Satz 3.17). Definiere $\Omega := \mathbb{R}^d, \mathcal{A} := \bigotimes_{i=1}^d \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^d), \mathbb{P} := \bigotimes_{i=1}^d \mathbb{P}_i, X_i := \pi_i$. Dann sind die $X_i, i = 1, \dots, d$ alle messbar und für $t \in \mathbb{R}$ und $i = 1, \dots, d$ gilt

$$\mathbb{P}(X_i \leqslant t) = \mathbb{P}(\mathbb{R} \times \ldots \times (-\infty, t] \times \ldots \times \mathbb{R}) = 1 \cdot \ldots \cdot \mathbb{P}_i((-\infty, t]) \cdot \ldots \cdot 1 = F_i(t)$$

Schließlich gilt für $t \in \mathbb{R}^d$ und $i = 1, \dots, d$

$$\mathbb{P}(X_1 \leqslant t_1, \dots, X_d \leqslant t_d) = \mathbb{P}((\infty, t_1] \times \ldots \times (-\infty, d_t]) = \prod_{i=1}^d \mathbb{P}_i(-\infty, t_i]$$

sodass die X_i unabhängig sind. □

Produktmaß und Integral

Betrachte in diesem Abschnitt zwei σ-endliche Maßräume $(\Omega_i, \mathcal{A}_i, \mu_i), i = 1, 2$, den entsprechenden Produktaum $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2, \mu_1 \otimes \mu_2)$, sowie einen weiteren messbaren Raum $(\Omega', \mathcal{A}')$.

8.9. Definition: Sei $\omega_1 \in \Omega_1$ fixiert.

(i) Für $A \subseteq \Omega_1 \times \Omega_2$ sei

$$A_{\omega_1} := \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in A\}$$

der ω_1-Schnitt (ω_1-section) von A.

(ii) Für $f : \Omega_1 \times \Omega_2 \to \Omega'$ sei

$$f_{\omega_1} : \Omega_2 \to \Omega', \omega_2 \mapsto f(\omega_1, \omega_2)$$

der ω_1-Schnitt (ω_1 section) von f.

Bemerkung: Es gilt (einfacher Beweis, siehe Übung)

(i) $(\mathbb{1}_A)_{\omega_1} = \mathbb{1}_{A_{\omega_1}}$

(ii) Für $A' \subseteq \Omega'$ ist $f_{\omega_1}^{-1}(A') = (f^{-1}(A'))_{\omega_1}$

8.10. Proposition: Sei $\omega_1 \in \Omega_1$ fixiert. Dann gilt

(i) Ist $A \in \mathcal{A}_1 \otimes A_2$, dann ist $A_{\omega_1} \in \mathcal{A}_2$

(ii) Ist $f : \Omega_1 \times \Omega_2 \to \Omega'$ $\mathcal{A}_1 \otimes \mathcal{A}_2$–$\mathcal{A}'$-messbar, dann ist $f_{\omega_1} : \Omega_2 \to \Omega'$ $\mathcal{A}_2$–$\mathcal{A}'$-messbar.

Analoges gilt natürlich für die entsprechenden ω_2-Schnitte.

Beweis: Betrachte für $\omega_1 \in \Omega_1$ fixiert die Abbildung $g_{\omega_1} : \Omega_2 \to \Omega_1 \times \Omega_2, \omega_2 \mapsto (\omega_1, \omega_2)$. Dann gilt

$$\forall A = (A_1 \times A_2) \in \mathcal{R} : g_{\omega_1}^{-1}(A) = \begin{cases} A_2 & \text{falls } \omega_1 \in \mathcal{A}_1 \\ \emptyset & \text{falls } \omega_1 \in \mathcal{A}_1 \end{cases}$$

Damit ist g_{ω_1} $\mathcal{A}_2$–$\mathcal{A}_1 \otimes \mathcal{A}_2$-messbar (Da Messbarkeit im Erzeugendensystem eine hinreichende Bedingung ist). Damit folgt nun

(i) $A_{\omega_1} = \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in A\} = g_{\omega_1}^{-1}(A) \in \mathcal{A}_2$

(ii) $f_{\omega_1}(\omega_2) = f(\omega_1, \omega_2) = f(g(\omega_1, \omega_2)) = (f \circ g)(\omega_2)$

womit die Messbarkeit von f_{ω_1} aus der Messbarkeit von Zusammemsetzungen messbarer Funktionen folgt. $\square$

8.11. Satz (Tonelli's Theorem): Sei $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$ nicht-negativ und messbar. Dann ist die Abbildung

$$s_1 : \Omega_1 \to \overline{\mathbb{R}} \text{ mit } \omega_1 \mapsto \int_{\Omega_2} f_{\omega_1} \, d\mu_2$$

nicht-negativ und messbar und es gilt

$$\int_{\Omega_1 \times \Omega_2} f \, d(\mu_1 \otimes \mu_2) = \int_{\Omega_1} s_1 \, d\mu_1 = \int_{\Omega_1} \left(\int_{\Omega_2} f_{\omega_1} \, d\mu_2 \right) \, d\mu_1$$

Bemerkung: Die σ-Endlichkeit von μ_1 und μ_2 ist hier notwendig. Ein analoges Ergebnis gilt auch wenn die Reihenfolge der Integrale geändert wird.

Beweis: Betrachte zunächst den Fall wo μ_1 und μ_2 (und damit auch das Produktmaß $\mu_1 \otimes \mu_2$) endlich sind.

I. ***f* Indikatorfunktion auf messbarem Rechteck,** $f = \mathbb{1}_{A_1 \times A_2}$ **mit** $A_1 \times A_2 \in \mathcal{R}$
Hier gilt $f_{\omega_1}(\omega_2) = \mathbb{1}_{A_1}(\omega_1) \cdot \mathbb{1}_{A_2}(\omega_2)$ und damit

$$s_1(\omega_1) = \mathbb{1}_{A_1}(\omega_1) \int_{\Omega_2} \mathbb{1}_{A_2} \, d\mu_2 = \mathbb{1}_{A_1}(\omega_1) \cdot \mu_2(A_2) \geqslant 0$$

und als einfache Funktion auf einer $\mathcal{A}_1$-messbaren Menge auch $\mathcal{A}_1$–$\mathcal{B}(\overline{\mathbb{R}})$-messbar. Außerdem gilt

$$\begin{aligned} \int s_1 \, d\mu_1 &= \int \mathbb{1}_{A_1} \mu_2(A_2) \, d\mu_1 = \mu_1(A_1) \cdot \mu_2(A_2) \\ &= (\mu_1 \otimes \mu_2)(A_1 \times A_2) = \int \mathbb{1}_{A_1 \times A_2} \, d(\mu_1 \otimes \mu_") \end{aligned}$$

II. ***f* Indikatorfunktion auf endl. Vereinigung messbarer Rechtecke,** $f = \mathbb{1}_A$ **mit** $A \in \mathcal{R}^*$
Hier gilt $f_{\omega_1}(\omega_2) = (\mathbb{1}_A)_{\omega_1}(\omega_2) = \mathbb{1}_{A_{\omega_1}}(\omega_2)$ und da $A_{\omega_1} \in \mathcal{A}_2$, ist f_{ω_1} $\mathcal{A}_2$–$\mathcal{B}(\overline{\mathbb{R}})$-messbar. Es gilt

$$s_1(\omega_1) = \int \mathbb{1}_{A_{\omega_1}} \, d\mu_2 = \mu_2(A_{\omega_1}) \geqslant 0$$

Zeige nun die Messbarkeit von s_1: Definiere dazu

$$\mathcal{L} := \left\{ A \in \mathcal{A}_1 \otimes \mathcal{A}_2 : s_1(\cdot) = \int \mathbb{1}_A(\cdot, \omega_2)\ d\mu_2(\omega_2) \text{ ist } A_1\text{–}\mathcal{B}(\overline{\mathbb{R}})\text{-messbar} \right\}$$

und zeige $\mathcal{L} = \mathcal{A}_1 \otimes \mathcal{A}_2$. Es gilt natürlich $\mathcal{R} \subseteq \mathcal{L} \subseteq \mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma(\mathcal{R})$ wobei die erste Inklusion mit dem I. Fall und die zweite Inklusion laut Konstruktion gilt. Wir wissen, dass $\mathcal{R}$ ein π-System ist. Mit dem λ–π-Theorem genügt es also zu zeigen, dass $\mathcal{L}$ ein λ-System ist (einfache Überlegung). Es gilt also zu zeigen

- $\Omega_1 \times \Omega_2 \in \mathcal{L}$: Gilt, da $\Omega_1 \times \Omega_2 \in \mathcal{R} \subseteq \mathcal{L}$.
- $A, B \in \mathcal{L}, A \subseteq B \implies B \setminus A \in \mathcal{L}$: Hier gilt $\mathbb{1}_{B\setminus A} = \mathbb{1}_B - \mathbb{1}_A$, sodass

  $$\int \mathbb{1}_{B\setminus A}\ d\mu_2 = \int \mathbb{1}_B\ d\mu_2 - \int \mathbb{1}_A\ d\mu_2$$

  als Differenz zweier messbarer Funktionen (da $A, B \in \mathcal{L}$) wieder messbar ist.
- $A_1 \subseteq A_2 \subseteq \dots, A_i \in \mathcal{L}, \forall i \geqslant 1 \implies \bigcup_{i\geqslant 1} A_i \in \mathcal{L}$: Setze $A := \bigcup_{i\geqslant 1} A_i$, sodass $0 \leqslant \mathbb{1}_{A_i} \nearrow \mathbb{1}_A$ punktweise. Damit gilt

  $$\forall \omega_1 \in \Omega_1, \forall \omega_2 \in \Omega_2 : 0 \leqslant (\mathbb{1}_{A_i})_{\omega_1}(\omega_2) \nearrow (\mathbb{1}_A)_{\omega_1}(\omega_2)$$

  und mit MONK folgt

  $$\forall \omega_1 \in \Omega_1 : \int (\mathbb{1}_A)_{\omega_1}\ d\mu_2 = \lim_{i\to\infty} \int (\mathbb{1}_{A_i})_{\omega_1}\ d\mu_2$$

  Das Integral ist als Grenzwert messbarer Funktionen damit messbar (da der Grenzwert laut MONK auch existiert).

Damit ist $\mathcal{L}$ ein λ-System. Zeige nun $\int \mathbb{1}_A\ d(\mu_1 \otimes \mu_2) = \int s_1\ d\mu_1$: Es gilt

$$\int \mathbb{1}_A\ d(\mu_1 \otimes \mu_2) = (\mu_1 \otimes \mu_2)(A)$$

Definiere

$$\nu(A) := \int \left(\int \mathbb{1}_A(\cdot, \omega_2)\ d\mu_2(\omega_2) \right) d\mu_1(\omega_1)$$

Wegen dem I. Fall wissen wir, dass $\nu(R) = (\mu_1 \otimes \mu_2)(R)$ für $R \in \mathcal{R}$ gilt. Falls ν ein Maß auf $\sigma(\mathcal{R}) = \mathcal{A}_1 \otimes \mathcal{A}_2$ ist, folgt mit Korollar 2.8, dass ν und $(\mu_1 \otimes \mu_2)$ auf $\mathcal{A}_1 \otimes \mathcal{A}_2 \supseteq \mathcal{R}^*$ übereinstimmen und damit die Aussage. $\nu(\emptyset)$ und $\nu \geqslant 0$ ergeben sich sofort aus den Eigenschaften vom Lebesgue-Integral. Zeige also die σ-Additivität:
Seien $A_i \in \mathcal{A}, i \geqslant 1$ disjunkt und definiere $B_n := \bigcup_{i=1}^n A_i$, $B := \bigcup_{i\geqslant 1} A_i$. Dann gilt $B_1 \subseteq B_2 \subseteq \ldots \subseteq B$ und damit $0 \leqslant \mathbb{1}_{B_n} \nearrow \mathbb{1}_B$. Folglich gilt auch $\forall \omega_1 \in \Omega_1 : 0 \leqslant (\mathbb{1}_{B_n})_{\omega_1} \nearrow (\mathbb{1}_B)_{\omega_1}$. Mit MONK folgt

$$\forall \omega_1 \in \Omega_1 : 0 \leqslant \int (\mathbb{1}_{B_n})_{\omega_1}\ d\mu_2 \xrightarrow[n\to\infty]{} \int (\mathbb{1}_B)_{\omega_1}\ d\mu_2$$

und (nochmal MONK)

$$0 \leqslant \int_{\Omega_1} \left(\int_{\Omega_2} (\mathbb{1}_{B_n})_{\omega_1}\ d\mu_2 \right)\ d\mu_1 \xrightarrow[n\to\infty]{} \int_{\Omega_1} \left(\int_{\Omega_2} (\mathbb{1}_B)_{\omega_1}\ d\mu_2 \right)\ d\mu_1$$

Also gilt

$$\begin{aligned}
\nu(B) = \nu\left(\bigcup_{i\geqslant 1} A_i\right) &= \lim_{n\to\infty} \nu(B_n) = \lim_{n\to\infty} \nu\left(\bigcup_{i=1}^{n} A_i\right) \\
&= \lim_{n\to\infty} \int_{\Omega_1} \left(\int_{\Omega_2} \mathbb{1}_{\bigcup_{i=1}^n A_i}\ d\mu_2\right)\ d\mu_1 \\
&= \lim_{n\to\infty} \int_{\Omega_1} \left(\int_{\Omega_2} \sum_{i=1}^{n} \mathbb{1}_{A_i}\ d\mu_2\right)\ d\mu_1 \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} \int_{\Omega_1} \left(\int_{\Omega_2} \mathbb{1}_{A_i}\ d\mu_2\right)\ d\mu_1 \\
&= \sum_{i\geqslant 1} \int_{\Omega_1} \left(\int_{\Omega_2} \mathbb{1}_{A_i}\ d\mu_2\right)\ d\mu_1 = \sum_{i\geqslant 1} \nu(A_i)
\end{aligned}$$

wobei die inneren Integrale jeweils über den ω_1-Schnitt der jeweiligen Funktionen zu verstehen sind.

III. **f einfache Funktion, $f = \sum_{i=1}^n \alpha_i \cdot \mathbb{1}_{A_i}$ mit $A_i \in \mathcal{A}_1 \otimes \mathcal{A}_2$ disjunkt**

$$s_1(\omega_1) = \int f_{\omega_1}\ d\mu_2 = \int \left(\sum_{i=1}^{n} \alpha_i \cdot \mathbb{1}_{A_i}\right)_{\omega_1} d\mu_2 = \sum_{i=1}^{n} \alpha_i \cdot \int \mathbb{1}_{A_i}\ d\mu_2 \geqslant 0$$

und s_1 ist als Linearkombination messbarer Funktionen wieder messbar. Weiters gilt

$$\begin{aligned}
\int f\ d(\mu_1 \otimes \mu_2) &= \int \sum_{i=1}^{n} \alpha_i \cdot \mathbb{1}_{A_i}\ d(\mu_1 \otimes \mu_2) \\
&= \sum_{i=1}^{n} \alpha_i \cdot \int \mathbb{1}_{A_i}\ d(\mu_1 \otimes \mu_2) \\
&= \sum_{i=1}^{n} \alpha_i \cdot (\mu_1 \otimes \mu_2)(A_i) \\
&= \sum_{i=1}^{n} \alpha_i \cdot \nu(A_i) \\
&= \sum_{i=1}^{n} \alpha_i \cdot \int_{\Omega_1} \left(\int_{\Omega_2} \mathbb{1}_{A_i}\ d\mu_2\right)\ d\mu_1 \\
&= \int_{\Omega_1} \left(\int_{\Omega_2} \sum_{i=1}^{n} \alpha_i \cdot \mathbb{1}_{A_i}\ d\mu_2\right)\ d\mu_1 \\
&= \int_{\Omega_1} \left(\int_{\Omega_2} f\ d\mu_2\right)\ d\mu_1
\end{aligned}$$

IV. **f nicht-negativ, messbar**
Wähle eine Folge einfacher Funktionen $f_n, n \geqslant 1$, sodass $0 \leqslant f_n \nearrow f$. Seien die f_n o.B.d.A. wie im III. Fall. Dann gilt $\forall \omega_1 \in \Omega_1 : 0 \leqslant (f_n)_{\omega_1} \nearrow f_{\omega_1}$ und mit MONK folgt

$$\forall \omega_1 \in \Omega_1 : 0 \leqslant \int_{\Omega_2} (f_n)_{\omega_1}\ d\mu_2 \xrightarrow[n\to\infty]{} \int_{\Omega_2} f_{\omega_1}\ d\mu_2$$

wobei die Integrale der f_n mit dem III. Fall messbar sind und der Grenzwert damit auch. Weiters gilt

$$\begin{aligned}\int f \, d(\mu_1 \otimes \mu_2) &\overset{\text{Def.}}{=} \lim_{n\to\infty} \int f_n \, d(\mu_1 \otimes \mu_2) \\ &\overset{\text{III.}}{=} \lim_{n\to\infty} \int_{\Omega_1} \left(\int_{\Omega_2} f_n \, d\mu_2 \right) \, d\mu_1 \\ &\overset{\text{MONK}}{=} \int_{\Omega_1} \left(\int_{\Omega_2} f \, d\mu_2 \right) \, d\mu_1 = \int_{\Omega_1} s_1 \, d\mu_1\end{aligned}$$

Betrachte nun den allgemeinen Fall, wo μ_1 und μ_2 beide σ-endlich sind.
Für $i = 1, 2$ gibt es Mengenfolgen $B_{i,n} \in \mathcal{A}_i, n \geqslant 1$, sodass $B_{i,1} \subseteq B_{i,2} \subseteq \ldots \subseteq \bigcup_{n\geqslant 1} B_{i,n}$ und $\forall n \geqslant 1 : \mu_i(B_{i,n}) < \infty$. Für $B_n := B_{1,n} \times B_{2,n} \in \mathcal{A}_1 \otimes \mathcal{A}_2$ ist $\Omega_1 \times \Omega_2 = \bigcup_{n\geqslant 1} B_n$ und $\forall n \geqslant 1 :$ $(\mu_1 \otimes \mu_2)(B_n) = \mu_1(B_{1,n}) \cdot \mu_2(B_{2,n}) < \infty$. Damit ist auch $\mu_1 \otimes \mu_2$ ein σ-endliches Maß. Definiere nun für $n \geqslant 1$ die folgenden Maße für messbare Mengen A:

$$\begin{aligned}\mu_{1,n}(A) &:= \mu_1(A \cap B_{1,n}) \\ \mu_{2,n}(A) &:= \mu_2(A \cap B_{2,n}) \\ \pi_n(A) &:= (\mu_1 \otimes \mu_2)(A \cap B_n)\end{aligned}$$

Dann gelten folgende Eigenschaften (leicht zu prüfen):

- $\mu_{2,n}$, $\mu_{1,n}$ und π_n sind endliche Maße für alle $n \geqslant 1$.
- $\pi_n = \mu_{1,n} \otimes \mu_{2,n}$
- Es gilt für $i = 1, 2$, dass
$$\int_{\Omega_i} f \, d\mu_{i,n} = \int_{\Omega_i} f \cdot \mathbb{1}_{B_{i,n}} \, d\mu_i$$
für $f : \Omega_i \to \overline{\mathbb{R}}$ nicht-negativ und messbar und
$$\int f \, d\pi_n = \int f \cdot \mathbb{1}_{B_n} \, d(\mu_1 \otimes \mu_2)$$
für $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$ nicht-negativ und messbar
- Der Satz von Tonelli gilt für $\mu_{1,n}$ und $\mu_{2,n}$ wie bereits bewiesen.

Sei also $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$ nicht-negativ und messbar. Dann gilt $0 \leqslant f \cdot \mathbb{1}_{B_n} \nearrow f$ und $\forall \omega_1 \in \Omega_1 :$ $0 \leqslant (f \cdot \mathbb{1}_{B_n})_{\omega_1} \nearrow f_{\omega_1}$. Mit MONK folgt also

$$\begin{aligned}s_1(\omega_1) &= \int_{\Omega_2} f_{\omega_1} \, d\mu_2 = \lim_{n\to\infty} \int_{\Omega_2} (f \cdot \mathbb{1}_{B_n})_{\omega_1} \, d\mu_2 \\ &= \begin{cases} \lim\limits_{n\to\infty} \int_{\Omega_2} f_{\omega_1} \, d\mu_{2,n} & \text{falls } \omega_1 \in B_{1,n} \\ 0 & \text{falls } \omega_1 \notin B_{1,n} \end{cases} \\ &= \mathbb{1}_{B_{1,n}}(\omega_1) \int_{\Omega_2} f_{\omega_1} \, d\mu_{2,n} \geqslant 0\end{aligned}$$

und messbar (IV. Fall und Produkt mit Indikatorfunktion auf messbarer Menge). Weiters gilt

$$\begin{aligned}
\int f \; d(\mu_1 \otimes \mu_2) &= \lim_{n\to\infty} \int \mathbb{1}_{B_n} \cdot f \; d(\mu_1 \otimes \mu_2) \\
&\overset{\text{s.o.}}{=} \lim_{n\to\infty} \int f \; d\pi_n \\
&= \lim_{n\to\infty} \int f \; d(\mu_{1,n} \otimes \mu_{2,n}) \\
&\overset{\text{IV.}}{=} \lim_{n\to\infty} \int_{\Omega_1} \left(\int_{\Omega_2} f_{\omega_1} \; d\mu_{2,n} \right) \; d\mu_{1,n} \\
&\overset{\text{s.o.}}{=} \lim_{n\to\infty} \int_{\Omega_1} \left(\mathbb{1}_{B_{1,n}} \int_{\Omega_2} f_{\omega_1} \cdot \mathbb{1}_{B_{2,n}} \; d\mu_2 \right) \; d\mu_1 \\
&\overset{\text{2x MONK}}{=} \int_{\Omega_1} \left(\int_{\Omega_2} f_{\omega_1} \; d\mu_2 \right) \; d\mu_1 = \int_{\Omega_1} s_1 \; d\mu_1
\end{aligned}$$

wobei der vorletzte Schritt mit den nicht-negativen monotonen Folgen $\left(f_{\omega_1} \cdot \mathbb{1}_{B_{2,n}}\right), n \geqslant 1$ und $\left(\mathbb{1}_{B_{1,n}} \cdot \int_{\Omega_2} f_{\omega_1} \mathbb{1}_{B_{2,n}} \; d\mu_2\right), n \geqslant 1$ und MONK folgt. □

8.12. Satz (Fubini's Theorem): Sei $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$ absolut $\mu_1 \otimes \mu_2$-integrierbar. Dann gibt es eine messbare Menge $N_1 \in \mathcal{A}_1$ mit folgenden Eigenschaften

(i) $\mu_1(N_1) = 0$ und $\forall \omega_1 \notin N_1 : f_{\omega_1} \in L^1(\mu_2)$

(ii) Die Abbildung $s_1 : \Omega_1 \to \overline{\mathbb{R}}$ mit

$$s_1(\omega_1) := \begin{cases} \displaystyle\int_{\Omega_2} f_{\omega_1} \; d\mu_2 & \text{falls } \omega_1 \notin N_1 \\ 0 & \text{falls } \omega_1 \in N_1 \end{cases}$$

ist absolut μ_1-integrierbar.

(iii)

$$\int f \; d(\mu_1 \otimes \mu_2) = \int_{\Omega_1} s_1 \; d\mu_1$$

Beweis:

(i) $|f|$ ist nicht-negativ und messbar. Laut Annahme und mit Tonelli (Satz 8.11) gilt

$$\int_{\Omega_1} \left(\int_{\Omega_2} |f| \; d\mu_2 \right) \; d\mu_1 = \int |f| \; d(\mu_1 \otimes \mu_2) < \infty$$

wobei das innere Integral auf der linken Seite für alle $\omega_a \in \Omega_1$ eine nicht-negative Funktion mit endlichem μ_1-Integral ist und damit $< \infty$ f.ü. ist. Definiere also

$$N_1 := \left\{ \omega_1 \in \Omega_1 : \int_{\Omega_2} |f|_{\omega_1} \; d\mu_2 = \infty \right\}$$

Dann gilt wegen der Messbarkeit von s_1 laut Tonelli $N_1 \in \mathcal{A}_1$ und trivial die gesuchten Eigenschaften (da $|f_{\omega_1}| = |f|_{\omega_1}$).

(ii) Schreibe $f = f_+ - f_-$ und wende jeweils Tonelli auf den Positiv- und Negativteil an. Schreibe

$$s_1(\omega_1) = \begin{cases} \int_{\Omega_2} (f_{\omega_1})_+ \, d\mu_2 - \int_{\Omega_2} (f_{\omega_1})_- \, d\mu_2 & \text{falls } \omega_1 \notin N_1 \\ 0 & \text{falls } \omega_1 \in N_1 \end{cases}$$
$$= \begin{cases} \int_{\Omega_2} (f_+)_{\omega_1} \, d\mu_2 - \int_{\Omega_2} (f_-)_{\omega_1} \, d\mu_2 & \text{falls } \omega_1 \notin N_1 \\ 0 & \text{falls } \omega_1 \in N_1 \end{cases}$$

Damit ist s_1 als Produkt und Summe messbarer Funktionen und mit Tonelli messbar. Weiters gilt

$$\begin{aligned}
\int_{\Omega_1} |s_1| \, d\mu_1 &= \int_{\Omega_1} \mathbb{1}_{N_1^c} \left| \int_{\Omega_2} (f_{\omega_1})_+ \, d\mu_2 - \int_{\Omega_2} (f_{\omega_1})_- \, d\mu_2 \right| \, d\mu_1 \\
&\overset{\text{DUG}}{\leqslant} \int_{\Omega_1} \mathbb{1}_{N_1^c} \left(\left| \int_{\Omega_2} (f_{\omega_1})_+ \, d\mu_2 \right| + \left| \int_{\Omega_2} (f_{\omega_1})_- \, d\mu_2 \right| \right) \, d\mu_1 \\
&= \int_{\Omega_1} \mathbb{1}_{N_1^c} \int_{\Omega_2} (f_{\omega_1})_+ \, d\mu_2 \, d\mu_1 + \int_{\Omega_1} \mathbb{1}_{N_1^c} \int_{\Omega_2} (f_{\omega_1})_- \, d\mu_2 \, d\mu_1 \\
&\leqslant \int_{\Omega_1} \int_{\Omega_2} (f_{\omega_1})_+ \, d\mu_2 \, d\mu_1 + \int_{\Omega_1} \int_{\Omega_2} (f_{\omega_1})_- \, d\mu_2 \, d\mu_1 \\
&\overset{\text{Tonelli}}{=} \int f_+ \, d(\mu_1 \otimes \mu_2) + \int f_- \, d(\mu_1 \otimes \mu_2) \\
&= \int f \, d(\mu_1 \otimes \mu_2) < \infty
\end{aligned}$$

Damit ist s_1 absolut μ_1-integrierbar.

(iii) Mit Tonelli für f_+ und f_- gilt

$$\begin{aligned}
\int_{\Omega_1} \, d\mu_1 &\overset{\text{Def.}}{=} \int_{\Omega_1} \mathbb{1}_{N_1^c} \left(\int_{\Omega_2} f_{\omega_1} \, d\mu_2 \right) \, d\mu_1 \\
&= \int_{\Omega_1} \mathbb{1}_{N_1^c} \left(\int_{\Omega_2} (f_+)_{\omega_1} \, d\mu_2 \right) \, d\mu_1 - \int_{\Omega_1} \mathbb{1}_{N_1^c} \left(\int_{\Omega_2} (f_-)_{\omega_1} \, d\mu_2 \right) \, d\mu_1 \\
&= \int_{\Omega_1} \left(\int_{\Omega_2} (f_+)_{\omega_1} \, d\mu_2 \right) \, d\mu_1 - \int_{\Omega_1} \left(\int_{\Omega_2} (f_-)_{\omega_1} \, d\mu_2 \right) \, d\mu_1 \\
&\overset{\text{Tonelli}}{=} \int f_+ \, d(\mu_1 \otimes \mu_2) - \int f_- \, d(\mu_1 \otimes \mu_2) = \int f \, d(\mu_1 \otimes \mu_2)
\end{aligned}$$

wobei der vorletzte Schritt aus $\mu_1(N_1) = 0$ und $f \overset{a.e.}{=} g \implies \int f \, d\mu = \int g \, d\mu$ folgt. $\square$.

8.13. Beispiel: Betrachte unabhängige Zufallsvariablen $X, Y : \Omega \to \mathbb{N}$ mit gemeinsamer pmf $p_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y)$, marginal pmfs $p_X = \mathbb{P}(X = x)$, $p_Y(y) = \mathbb{P}(Y = y)$ und eine messbare Funktion $f : \mathbb{N} \times \mathbb{N} \to \mathbb{Q}$ mit

$$(x, y) \mapsto \frac{1}{p_{X,Y}(x, y)} \frac{(-1)^{x+1}}{x + y} = \frac{1}{p_X(x) \cdot p_Y(y)} \frac{(-1)^{x+1}}{x + y}$$

Gesucht ist $\mathbb{E} f(X, Y)$. Betrachte dazu folgende Tabelle

↓ y / x →	1	2	3	4	…	Summe
1	1/2	-1/3	1/4	-1/5	…	$1-\log 2 =: c$
2	1/3	-1/4	1/5	-1/6	…	$-c+1/2$
3	1/4	-1/5	1/6	-1/7	…	$c-1/2+1/3$
4	1/5	-1/6	1/7	-1/8	…	$c+1/2-1/3+1/4$
⋮	⋮	⋮	⋮	⋮	⋱	⋮
Summe	$+\infty$	$-\infty$	$+\infty$	$-\infty$	…	

Damit gelten folgende Eigenschaften

(i) $\sum_{x\geqslant 1}\sum_{y\geqslant 1} f(x,y)\cdot p_{X,Y}(x,y) = \sum_{x\geqslant 1}(\infty-\infty+\infty-\ldots)$ existiert nicht!

(ii) $\sum_{y\geqslant 1}\sum_{x\geqslant 1} f(x,y)\cdot p_{X,Y}(x,y) = \sum_{y\geqslant 1}\left(\lim_{n\to\infty}\begin{cases}1/2+1/4+\ldots+1/n & \text{falls } n \text{ gerade}\\ c+1/3+1/5+\ldots+1/n & \text{falls } n \text{ ungerade}\end{cases}\right) = \infty$

Mit dem Kontrapositiv von Fubini (Satz 8.12) gilt also $\mathbb{E}|f(X,Y)| = \infty$.

Messbarkeit $\mathbb{R}^d$-wertiger Funktionen

Betrachte in diesem Abschnitt $\mathbb{R}^d$-wertige Abbildungen und die euklidische Norm $\|x\|_2 = \left(\sum_{i=1}^d x_i^2\right)^{1/2}$ für $x\in\mathbb{R}^d, x=(x_1,\ldots,x_d)'$.

8.14. Definition:

(i) Sei $\mathcal{O}_d := \{O\subseteq\mathbb{R}^d : O \text{ offen}\}$ die Familie aller offenen Mengen in $\mathbb{R}^d$. Die Borel-σ-Algebra auf $\mathbb{R}^d$ ist definiert als

$$\mathcal{B}(\mathbb{R}^d) := \sigma(\mathcal{O}_d)$$

(ii) Ist $(\Omega,\mathcal{A},\mathbb{P})$ ein Wahrscheinlichkeitsraum und $X:(\Omega,\mathcal{A})\to(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d))$ messbar, dann nennt man X einen d-dimensionalen Zufallsvektor. Man nennt die Funktion $F:\mathbb{R}^d\to[0,1]$ mit $t\to\mathbb{P}(X\leqslant t)$ (komponentenweise) die Verteilungsfunktion (cdf) von X.

8.15. Lemma Für $\mathbb{R}^d$ und $\mathbb{R}^\ell$ mit euklidischer Metrik gilt

$$f:\mathbb{R}^d\to\mathbb{R}^\ell \text{ stetig} \iff \forall O\in\mathcal{O}_\ell : f^{-1}(O)\in\mathcal{O}_d,$$

also ist f genau dann stetig, wenn Urbilder offener Mengen unter f offen sind.

Beweis: Übung! (cf. Höhere Analysis) □

8.16. Proposition: Sei $f:(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d))\to(\mathbb{R}^\ell,\mathcal{B}(\mathbb{R}^\ell))$ stetig. Dann ist f auch messbar.

Beweis: Mit Lemma 8.15 folgt unmittelbar die Messbarkeit im Erzeugendensystem und damit auch die Messbarkeit in $\mathcal{B}(\mathbb{R}^d)$. □

8.17. Proposition:

$$\mathcal{B}(\mathbb{R}^d) = \bigotimes_{i=1}^d \mathcal{B}(\mathbb{R})$$

Bemerkung: Allgemeiner gilt

$$\mathcal{B}(X^d) = \bigotimes_{i=1}^{d} \mathcal{B}(X)$$

für alle separablen metrischen Räume X (oder noch allgemeiner alle zweitabzählbaren topologischen Räume X).

Beweis: Um die Notation übersichtlich zu halten, betrachte hier nur den Fall $d = 2$. Betrachte zuerst die Richtung $\supseteq$, die auch ohne Annahmen an X auskommt. Setze

$$\mathcal{A} := \{A \subset X : A \times X \in \mathcal{B}(X^2)\},\ \mathcal{A}' := \{A \subset X : X \times A \in \mathcal{B}(X^2)\},$$

sodass $\mathcal{B}(X) \subseteq \mathcal{A}, \mathcal{A}'$ und $\mathcal{A}, \mathcal{A}'$ beide σ-Algebren sind (prüfe!). Nun gilt für $A, B \in \mathcal{B}(\mathbb{R})$, dass $A \times B = (A \times X) \cap (X \times B) \in \mathcal{B}(X^2)$ und damit

$$\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \sigma(\mathcal{R}) = \sigma\left(\{A \times B : A, B \in \mathcal{B}(\mathbb{R})\}\right) \subseteq \sigma(\mathcal{B}(X^2)) = \mathcal{B}(X^2).$$

Zeige nun die Richtung $\subseteq$. Beachte, dass X^2 in der Produktmetrik separabel ist und damit jede offene Menge $O \in \mathcal{O} := \{O \subseteq X^2 : O \text{ offen}\}$ von der Form

$$O = \bigcup_{i \geqslant 1} (O_i^{(1)} \times O_i^{(2)})$$

ist, wobei $O_i^{(1)}, O_i^{(2)}$ offen sind für alle $i \geqslant 1$. Nun gilt $O_i^{(1)}, O_i^{(2)} \in \mathcal{B}(\mathbb{R})$ und damit $O_i^{(1)} \times O_i^{(2)} \in \mathcal{R}$. Es folgt, dass $\mathcal{B}(X^2) = \sigma(\mathcal{O}) \subseteq \sigma(\mathcal{R}) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$. □

8.18 Korollar: Betrachte folgende Mengenfamilien

$$\begin{aligned}
\mathcal{J}_1 &:= \Big\{(-\infty, t_1] \times \ldots \times (-\infty, t_d] : t \in \mathbb{R}^d\Big\} \\
\mathcal{J}_2 &:= \Big\{(s_1, t_1) \times \ldots \times (s_d, t_d) : s, t \in \mathbb{R}^d\Big\} \\
\mathcal{J}_3 &:= \Big\{(s_1, t_1\rangle \times \ldots \times (s_d, t_d\rangle : s, t \in \overline{\mathbb{R}}^d, s_i \leqslant t_i \text{ für } i = 1, \ldots, d\Big\}
\end{aligned}$$

Dann gilt $\sigma(\mathcal{J}_1) = \sigma(\mathcal{J}_2) = \sigma(\mathcal{J}_3) = \mathcal{B}(\mathbb{R}^k)$

Beweis: Übung! Hinweis: Mit $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{O}_1)$ und $\mathbb{R} = \bigcup_{k \geqslant 1}(-k, k)$, wobei $(-k, k) \in \mathcal{O}_1$ gilt für $\mathcal{M} := \{A_1 \times \ldots \times A_n : A_i \in \mathcal{O}_1 \text{ für } i = 1, \ldots, d\}$, dass $\sigma(\mathcal{M}) = \bigotimes_{i=1}^d \sigma(\mathcal{O}_1) = \bigotimes_{i=1}^d \mathcal{B}(\mathbb{R})$. □

8.19. Korollar: Sei $(\Omega, \mathcal{A})$ ein messbarer Raum und $f : \Omega \to \mathbb{R}^d, \omega \mapsto (f_1(\omega), \ldots, f_d(\omega))'$ mit $f_i : \Omega \to \mathbb{R}$. Dann ist f genau dann $\mathcal{A}$–$\mathcal{B}(\mathbb{R}^d)$-messbar, wenn die Koordinatenfunktionen f_i für $i = 1, \ldots, d$ jeweils $\mathcal{A}$–$\mathcal{B}(\mathbb{R})$-messbar sind.

Beweis:

I. $\Longrightarrow$
   Die Koordinatenprojektionen $\pi_i : \mathbb{R}^d \to \mathbb{R}$ sind stetig und damit $\mathcal{B}(\mathbb{R}^d)$–$\mathcal{B}(\mathbb{R})$-messbar. Damit ist $f_i = (\pi_i \circ f)$ auch messbar.

II. $\Longleftarrow$

Mit Korollar 8.18 genügt es die Messbarkeit für Urbilder unter f aus $\mathcal{J}_\infty$ zu zeigen. Es gilt

$$f^{-1}((-\infty, t_1] \times \ldots \times (-\infty, t_d]) = \bigcap_{i=1}^{d} \{f_i \leqslant t_i\} \in \mathcal{A}$$

da ein endlicher Durchschnitt messbarer Mengen wieder messbar ist. □

8.20. Korollar: Sei $(\Omega, \mathcal{A})$ ein messbarer Raum und seien $f, g : (\Omega, \mathcal{A}) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ und $h : \Omega \to \mathbb{R}$ messbar. Sei weiters $M \in \mathbb{R}^{\ell \times d}$ eine deterministische Matrix. Dann gilt

(i) $f + g$ ist $\mathcal{A}$–$\mathcal{B}(\mathbb{R}^d)$-messbar.

(ii) $\langle f, g\rangle = \sum_{i=1}^{d} f_i g_i$ ist $\mathcal{A}$–$\mathcal{B}(\mathbb{R})$-messbar.

(iii) Mf ist $\mathcal{A}$–$\mathcal{B}(\mathbb{R}^\ell)$-messbar.

(iv) hf ist $\mathcal{A}$–$\mathcal{B}(\mathbb{R}^d)$-messbar.

Beweis: Folgt aus der Messbarkeit der Komponentenfunktionen und der Messbarkeit von Summen und Produkten messbarer Funktionen. □

8.21. Proposition: Die Verteilungsfunktion $F : \mathbb{R}^d \to [0,1]$ einer Zufallsvariable $X : (\Omega, \mathcal{A}) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ hat folgende Eigenschaften:

(i) Für eine Folge $t_n = (t_n^{(1)}, \ldots, t_n^{(d)})' \in \mathbb{R}^d, n \geqslant 1$ mit $\min\limits_{1 \leqslant i \leqslant d} t_n^{(i)} \searrow -\infty$ gilt $\lim\limits_{n\to\infty} F(t_n) = 0$ und für eine Folge $s_n = (s_n^{(1)}, \ldots, s_n^{(d)})' \in \mathbb{R}^d, n \geqslant 1$ mit $\min\limits_{1 \leqslant i \leqslant d} s_n^{(i)} \nearrow \infty$ gilt $\lim\limits_{n\to\infty} F(s_n) = 1$.

(ii) Für $t_n \in \mathbb{R}^d, n \geqslant 1$ mit $t_n \searrow t_0$ komponentenweise gilt $\lim\limits_{n\to\infty} F(t_n) = F(t_0)$. **(Rechtsstetigkeit)**

(iii) Für $t_n \in \mathbb{R}^d, n \geqslant 1$ mit $t_n \nearrow t_0$ komponentenweise existiert $\lim\limits_{n\to\infty} F(t_n)$.

(iv) Für $s, t \in \mathbb{R}^d$ mit $s \leqslant d$ komponentenweise gilt

$$0 \leqslant \sum_{k=0}^{d} (-1)^k \sum_{x \in I_k} F(x)$$

wobei $I_k = \left\{ x \in \mathbb{R}^d : \sum_{i=1}^{d} \delta_{x_i, s_i} = k, \sum_{i=1}^{d} \delta_{x_i, t_i} = d - k \right\}$, also die Menge der der x, sodass x in k Komponenten mit s übereinstimmt und in den restlichen Komponenten mit t übereinstimmt.

Beweis:

(i) Sei $m_n := \min\limits_{1 \leqslant i \leqslant d} t_n^{(i)}$ und sei i_n so, dass $m_n = t_n^{(i_n)}$ (also, dass das Minimum in der i_n-ten Koordinate angenommen wird).

- Falls $m_n \searrow -\infty$

$$\begin{aligned} F(t_n) &= \mathbb{P}(X_1 \leqslant t_n^{(1)}, \dots, X_d \leqslant t_n^{(d)}) \\ &\leqslant \mathbb{P}(X_{i_n} \leqslant m_n) \\ &\leqslant \sum_{i=1}^{d} \mathbb{P}(X_i \leqslant m_n) \xrightarrow[n\to\infty]{} 0 \end{aligned}$$

- Falls $m_n \nearrow \infty$

$$\begin{aligned} F(t_n) &= \mathbb{P}(X_1 \leqslant t_n^{(1)}, \dots, X_d \leqslant t_n^{(d)}) \\ &\geqslant \mathbb{P}(X_1 \leqslant m_n, \dots, X_d \leqslant m_n) \xrightarrow[n\to\infty]{\text{S.V.U.}} \mathbb{P}(X \in \mathbb{R}^d) = 1 \end{aligned}$$

  wobei die Stetigkeit von unten mit den Mengen $(-\infty, t_n^{(i)}] \supseteq (-\infty, m_n], i = 1, \dots, d$ gilt.

(ii) Setze $A_n := \{X \leqslant t_n\}$. Dann gilt $A_1 \supseteq A_2 \supseteq \dots \supseteq \bigcap_{n\geqslant 1} A_n$. Mit der Stetigkeit von oben folgt $\mathbb{P}(A) = \lim_{n\to\infty} \mathbb{P}(A_n)$ und damit $F(t_0) = \lim_{n\to\infty} F(t_n)$.

(iii) Für A_n wie in (ii) gilt hier $A_1 \subseteq A_2 \subseteq \dots \subseteq \bigcup_{n\geqslant 1} A_n =: A$. Mit der Stetigkeit von unten folgt $\lim_{n\to\infty} F(t_n) = \mathbb{P}(A) \in [0,1]$, sodass der Grenzwert exisitiert.

(iv) Es gilt

$$\begin{aligned} 0 &\leqslant \mathbb{P}(X \in (s_1, t_1] \times \dots \times (s_d, t_d]) \\ &= \mathbb{P}(\{X \leqslant t\} \setminus \{\exists i \leqslant d : X_i \leqslant s_i\}) \\ &= \mathbb{P}(\{X \leqslant t\} \setminus \{\forall i \leqslant d : X_i \leqslant t_i, \exists i \leqslant d : X_i \leqslant s_i\}) =: (*) \end{aligned}$$

Setze $A_j := \{X \leqslant t, X_j \leqslant s_j\}$ für $j = 1, \dots, d$. Dann gilt

$$\begin{aligned} (*) &= F(t) - \mathbb{P}\left(\bigcup_{j=1}^{d} A_j\right) \overset{\text{In-Ex}}{=} F(t) + \sum_{\ell=1}^{d} (-1)^\ell \sum_{\substack{I\subseteq\{1,\dots,d\} \\ |I|=\ell}} \mathbb{P}\left(\bigcap_{j\in I} A_j\right) \\ &= F(t) + \sum_{\ell=1}^{d} (-1)^\ell \sum_{x\in I_k} F(x) = \sum_{\ell=0}^{d} (-1)^\ell \sum_{x\in I_k} F(x) \end{aligned}$$

□

8.22. Proposition: Sei $F : \mathbb{R}^d \to [0,1]$ eine Funktion mit den Eigenschaften (i)-(iv) aus Proposition 8.21. Dann gibt es einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ und einen Zufallsvektor $X : \Omega \to \mathbb{R}^d$, sodass F die cf von X ist.

Beweis: Nur Beweisidee: Konstruiere ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, sodass

$$\mathbb{P}((-\infty, t_1] \times \dots \times (-\infty, t_d]) = F(t)$$

und setze $X(\omega) := \omega$ für $\omega \in \mathbb{R}^d$. □

8.23. Definition: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $f : (\Omega, \mathcal{A}) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ messbar. Sind alle Komponentenfunktionen $f_i : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})), i = 1, \ldots, d$ (quasi-)integrierbar, dann nennt man f (quasi-)integrierbar und setzt

$$\int f \, d\mu := \left(\int f_1 \, d\mu, \ldots, \int f_d \, d\mu \right)'$$

8.24. Definition: Sei V ein Vektorraum über einen Körper K. Eine Abbildung $\|\cdot\| : V \to [0, \infty)$ ist eine Norm auf V, falls für $x, y \in V$ und $\lambda \in K$ gilt:

- $\|x\| = 0 \iff x = 0$
- $\|\lambda x\| = |\lambda| \|x\|$
- $\|x + y\| \leqslant \|x\| + \|y\|$

Bemerkung: Eine Norm $\|\cdot\|$ erfüllt auch die umgekehrte Dreiecksungleichung

$$|\|x\| - \|y\|| \leqslant \|x - y\|$$

8.25. Lemma: Alle Normen auf $\mathbb{R}^d$ sind Lipschitz-äquivalent, i.e. für eine beliebige Norm $\|\cdot\|$ auf $\mathbb{R}^d$ gibt es Konstanten $\alpha, \beta > 0$, sodass

$$\forall x \in \mathbb{R}^d : \alpha \|x\|_\infty \leqslant \|x\| \leqslant \beta \|x\|_\infty$$

für $\|x\|_\infty = \max\limits_{1 \leqslant i \leqslant d} x_i$. Insbesondere sind damit alle Normen auf $\mathbb{R}^d$ topologisch äquivalent und erzeugen dieselben offenen Mengen.

Beweis: Betrachte die kanonische Basis $\{e_1, \ldots, e_d\}$ und setze $\mu := \max\limits_{1 \leqslant i \leqslant d} \|e_i\|$. Dann gilt

$$\|x\| = \left\| \sum_{i=1}^d x_i e_i \right\| \leqslant \sum_{i=1}^d \|x_i e_i\| = \sum_{i=1}^d |x_i| \|e_i\| \leqslant \mu \sum_{i=1}^d |x_i| \leqslant k\mu \|x\|_\infty =: \beta \|x\|_\infty$$

Damit ist die Abbildung $f : \mathbb{R}^d \to [0, \infty)$ mit $x \mapsto \|x\|$ stetig, denn

$$|\|x\| - \|y\|| \leqslant \|x - y\| \leqslant \beta \|x - y\|_\infty$$

und für $\varepsilon > 0$, setze $\delta := \varepsilon / \beta$. Betrachte nun $S := \left\{ s \in \mathbb{R}^d : \|x\|_\infty = 1 \right\}$. Dann ist S mit Heine–Borel kompakt und für f gilt der Extremwertsatz. Sei also $p \in \arg\min_{x \in S} \|x\|$. Dann gilt $\|p\| \neq 0$ und für alle $x \neq 0$ gilt

$$\|x\| = \left\| \|x\|_\infty \cdot \frac{x}{\|x\|_\infty} \right\| = \|x\|_\infty \cdot \left\| \frac{x}{\|x\|_\infty} \right\| \geqslant \|x\|_\infty \cdot \|p\| =: \alpha \|x\|_\infty$$

□

Bemerkung: Sei $\|\cdot\|$ eine beliebige Norm auf $\mathbb{R}^d$. Dann gilt

$$f \in L^1 \iff \|f\| \in L^1$$

da für $i = 1, \ldots, d$ gilt

$$|f_i| \leqslant \|f\|_2 \leqslant \sum_{j=1}^d |f_j|$$

und alle Normen auf $\mathbb{R}^d$ Lipschitz-äquivalent sind (cf. Lemma 8.25).

Unendliche Produkträume

Dieser Abschnitt behandelt die Konstruktion unendlicher Folgen von Zufallsvariablen und wird in Wahrscheinlichkeitstheorie 2 nicht behandelt, kann aber vor allem für die LV Stochastische Prozesse interessant sein.

8.26. Definition: Sei $\Omega := \{(\omega_1, \omega_2, \dots) : \omega_i \in \mathbb{R}, i \geqslant 1\}$. Sei $n \in \mathbb{N}$ und $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$. Man nennt Mengen der Form

$$E = \left(\prod_{i=1}^{n} B_i\right) \times \left(\prod_{i>n} \mathbb{R}\right)$$

Zylindermengen. Sei $\mathcal{C}$ die Familie aller Zylindermengen E. Dann nennt man $\mathcal{A} := \sigma(\mathcal{C})$ die Zylinder-σ-Algebra. Für $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$, setze außerdem

$$\mathrm{cyl}(B_1, \dots, B_n) := \left(\prod_{i=1}^{n} B_i\right) \times \left(\prod_{i>n} \mathbb{R}\right) \in \mathcal{C}.$$

8.27. Satz (Erweiterungssatz von Kolmogorov): Sei $(\Omega, \mathcal{A})$ wie oben. Für $n \geqslant 1$ sei ρ_n ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, sodass

$$\forall B_1, \dots, B_n \in \mathcal{B}(\mathbb{R}) : \rho_{n+1}(B_1 \times \ldots \times B_n \times \mathbb{R}) = \rho_n(B_1 \times \ldots \times B_n).$$

Dann existiert ein eindeutiges Wahrscheinlichkeitsmaß ρ auf $(\Omega, \mathcal{A})$, sodass

$$\forall n \geqslant 1, \forall B_1, \dots, B_n \in \mathcal{B}(\mathbb{R}) : \rho\left(\mathrm{cyl}(B_1, \dots, B_n)\right) = \rho_n(B_1 \times \ldots \times B_n).$$

Beweis: siehe z.B. P. Billingsley, *Probability and Measure* (2nd Ed.), Theorems 36.1, 36.2 $\square$

8.28. Korollar: Seien $\mathbb{P}_1, \mathbb{P}_2, \dots$ Wahrscheinlichkeitsmaße auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Dann existiert ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ mit u.a. Zufallsvariablen $X_i : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})), i \geqslant 1$, sodass $X_i \sim \mathbb{P}_i$ für alle $i \geqslant 1$.

Beweis: Setze $\rho_n := \bigotimes_{i=1}^n \mathbb{P}_i$ für alle $i \geqslant 1$. Dann gilt

$$\begin{aligned}
\rho_{n+1}(B_1 \times \ldots \times B_n \times \mathbb{R}) &= \mathbb{P}_1(B_1) \cdot \ldots \cdot \mathbb{P}_n(B_n) \cdot \mathbb{P}_{n+1}(\mathbb{R}) \\
&= \prod_{i=1}^{n} \mathbb{P}_i(B_i) \\
&= \rho_n(B_1 \times \ldots \times B_n)
\end{aligned}$$

für alle $n \geqslant 1$ und $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$. Sei also $\mathbb{P} := \rho$ das Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{A})$ aus Satz 8.27. Dann gilt

$$\mathbb{P}(\mathrm{cyl}(B_1, \dots, B_n)) = \prod_{i=1}^{n} \mathbb{P}_i(B_i)$$

für alle $n \geqslant 1$ und $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$. Setze nun $X_i := \pi_i$ die i-te Koordinatenabbildung, sodass $X_i^{-1} B = \{\omega \in \Omega : \omega_i \in B\} \in \mathcal{C} \subseteq \mathcal{A}$, i.e. X_i ist eine Zufallsvariable für alle $i \geqslant 1$. Sei nun für $n \geqslant 1$

beliebig $i : [n] \to \mathbb{N}$ eine monoton steigende Funktion. Für $k \notin i([n])$, setze $B_k := \mathbb{R}$ und sonst seien $B_{i(j)} \in \mathcal{B}(\mathbb{R}), j \in [n]$ beliebig. Dann gilt

$$\begin{aligned}
\mathbb{P}(X_{i(1)} \in B_{i(1)}, \dots, X_{i(n)} \in B_{i(n)}) &= \mathbb{P}(\operatorname{cyl}(B_1, \dots, B_{i(n)})) \\
&= \prod_{k=1}^{i(n)} \mathbb{P}_k(B_k) \\
&= \left(\prod_{k \in i([n])} \mathbb{P}_k(B_k) \right) \cdot \left(\prod_{k \notin i([n])} \mathbb{P}_k(\mathbb{R}) \right) \\
&= \prod_{j=1}^{n} \mathbb{P}_{i(j)} \left(B_{i(j)} \right) = \prod_{j=1}^{n} \mathbb{P} \left(X_{i(j)} \in B_{i(j)} \right),
\end{aligned}$$

womit $X_i, i \geqslant 1$ u.a. sind. $\square$

Bemerkung: Die obige Konstruktion lässt sich auf den Fall mit Zufallsvariablen Werte in einem separablem, vollständigen metrischen Raum annehmen und in einer geordneten Menge indexiert sind, also z.B. euklidische Prozesse $X_t : (\Omega, \mathcal{A}) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)), t \in \mathbb{R}$.

9. Konvergenz von messbaren Abbildungen

Sei in diesem Kapitel $(\Omega, \mathcal{A}, \mu)$ immer ein generischer Maßraum.

Konvergenz von $\overline{\mathbb{R}}$-wertigen Funktionen

Seien in diesem Abschnitt $f_n, f, g_n, g : (\Omega, \mathcal{A}) \to (\overline{\mathbb{R}}, \mathbb{B}(\overline{\mathbb{R}}))$ messbare Funktionen. Weiters setze hier $\infty - \infty = -\infty + \infty := 0$.

9.1. Definition: Eine Funktionenfolge $f_n, n \geqslant 1$ konvergiert μ-fast-überall (kurz f.ü.) gegen f, falls

$$\mu\left(\left\{\omega \in \Omega : \lim_{n\to\infty} f_n(\omega) = f(\omega)\right\}^c\right) = 0$$

Wir schreiben dann $f_n \xrightarrow[n\to\infty]{a.e.} f$ (almost everywhere) oder im Falle eines Wahrscheinlichkeitsraumes $a.s.$ (amost surely).

9.2. Lemma: Es gilt $f_n \xrightarrow[n\to\infty]{a.e.} f$ genau dann, wenn

(i) $\forall \varepsilon > 0 : \mu\left(\limsup_{n\to\infty} \{|f_n - f| \geqslant \varepsilon\}\right) = 0$

(ii) Falls μ endlich ist, dann ist (i) äquivalent zu $\forall \varepsilon > 0 : \lim_{N\to\infty} \mu\left(|f_n - f| > \varepsilon \text{ für alle } n \geqslant N\right) = 0$

Beweis: Mit der archimedischen Eigenschaft von $\mathbb{R}$ genügt es jeweils den Fall $\varepsilon = 1/k$ für alle $k \geqslant 1$ zu betrachten.

I. $f_n \xrightarrow[n\to\infty]{a.e.} f \implies$ (i):

$$\begin{aligned}\left\{\lim_{n\to\infty} f_n = f\right\} &= \left\{\forall k \geqslant 1 \exists N \geqslant 1 \forall n \geqslant N : |f_n - f| < \frac{1}{k}\right\} \\ &= \bigcap_{k\geqslant 1} \bigcup_{N\geqslant 1} \bigcap_{n\geqslant N} \left\{|f_n - f| < \frac{1}{k}\right\} \\ &= \bigcap_{k\geqslant 1} \liminf_{n\to\infty} \left\{|f_n - f| > \frac{1}{k}\right\}\end{aligned}$$

Also gilt mit de Morgan und den Gesetzen zu lim sup und lim inf von Mengen

$$\left\{\lim_{n\to\infty} f_n = f\right\}^c = \bigcup_{k\geqslant 1} \limsup_{n\to\infty} \left\{|f_n - f| \geqslant \frac{1}{k}\right\}$$

und laut Annahme damit

$$\mu\left(\bigcup_{k\geqslant 1} \limsup_{n\to\infty} \left\{|f_n - f| \geqslant \frac{1}{k}\right\}\right) = 0$$

und damit insbesondere

$$\mu\left(\limsup_{n\to\infty} \left\{|f_n - f| \geqslant \frac{1}{k}\right\}\right) = 0$$

für jedes $k \geqslant 1$ (da $A_k \subseteq \cup_{k\geqslant 1} A_k$ für alle $k \geqslant 1$).

II. (i) $\Longrightarrow$ $f_n \xrightarrow[n\to\infty]{a.e.} f$:

$$\begin{aligned}\mu\left(\lim_{n\to\infty} f_n \neq f\right) &\overset{\text{s.o.}}{=} \mu\left(\bigcup_{k\geqslant 1} \limsup_{n\to\infty} \left\{|f_n - f| \geqslant \frac{1}{k}\right\}\right)\\ &\overset{\sigma\text{-Subadd.}}{\leqslant} \sum_{k\geqslant 1} \mu\left(\limsup_{n\to\infty} \left\{|f_n - f| \geqslant \frac{1}{k}\right\}\right) = 0\end{aligned}$$

wobei der letzte Schritt aus der Annahme folgt.

III. μ endlich $\Longrightarrow$ ((i) $\Longleftrightarrow$ (ii)):

$$\limsup_{n\to\infty} \left\{|f_n - f| \geqslant \frac{1}{k}\right\} = \bigcap_{N\geqslant 1} \bigcup_{n\geqslant N} \left\{|f_n - f| \geqslant \frac{1}{k}\right\} =: \bigcap_{N\geqslant 1} A_N$$

Dann gilt $A_1 \supseteq A_2 \supseteq \ldots \supseteq \bigcap_{N\geqslant 1} A_N$ und mit der Stetigkeit von oben gilt

$$\begin{aligned}\mu\left(\limsup_{n\to\infty} \left\{|f_n - f| \geqslant \frac{1}{k}\right\}\right) &= \lim_{N\to\infty} \mu(A_N)\\ &= \lim_{N\to\infty} \mu\left(|f_n - f| > \frac{1}{k} \text{ für alle } n \geqslant N\right) = 0\end{aligned}$$

für alle $k \geqslant 1$. □

9.3. Lemma:

$$\forall \varepsilon > 0 : \sum_{n\geqslant 1} \mu\left(|f_n - f| > \varepsilon\right) < \infty \implies f_n \xrightarrow[n\to\infty]{a.e.} f$$

Beweis: Mit dem 1. Borel–Cantelli-Lemma für allgemeine Maße (Lemma 7.8, Bemerkung 2) gilt

$$\mu\left(\limsup_{n\to\infty} |f_n - f| > \varepsilon\right) = 0$$

und mit Lemma 9.2 folgt die Behauptung. □

9.4. Definition: Eine Funktionenfolge $f_n, n \geqslant 1$ konvergiert im Maß μ gegen f, falls

$$\forall \varepsilon > 0 : \lim_{n\to\infty} \mu\left(|f_n - f| > \varepsilon\right) = 0$$

Wir schreiben dann $f_n \xrightarrow[n\to\infty]{\mu} f$.

9.5. Proposition: Ist μ endlich, dann gilt $f_n \xrightarrow[n\to\infty]{a.e.} f \implies f_n \xrightarrow[n\to\infty]{\mu} f$.

Beweis: Es gelte $f_n \xrightarrow[n\to\infty]{a.e.} f$. Mit Lemma 9.2 (ii) folgt

$$\forall \varepsilon > 0 \lim_{N\to\infty} \mu\left(|f_n - f| > \varepsilon \text{ für alle } n \geqslant N\right) = 0$$

Aber $\{|f_n - f| > \varepsilon\} \subseteq \{|f_n - f| > \varepsilon \text{ für alle } n \geqslant N\}$ und damit folgt per Definition von Konvergenz im Maß die Aussage. □

9.6. Proposition: Sei μ endlich. Dann ist $f_n \xrightarrow[n\to\infty]{\mu} f$ äquivalent zu folgender Aussage: Jede Teilfolge $f_{n_k}, k \geqslant 1$ von $f_n, n \geqslant 1$ enthält eine weitere Teilfolge $f_{n_{k_j}}, j \geqslant 1$, sodass

$$f_{n_{k_j}} \xrightarrow[j\to\infty]{a.e.} f$$

Beweis:

I. $\implies$
Sei $0 < \varepsilon < 1$. Da $\mu(|f_n - f| > \varepsilon) \xrightarrow[n\to\infty]{} 0$, kann man o.B.d.A. annehmen, dass $\mu(|f_n - f| > \varepsilon) \leqslant 1$ für alle $n \geqslant 1$ (wähle einfach einen Index $N \geqslant 1$ für den die Aussage wahr ist, das Argument ändert sich dadurch nicht). Sei nun eine beliebige Teilfolge $f_{n_k}, k \geqslant 1$ gegeben. Wähle nun eine weitere Teilfolge $f_{n_{k_j}}, j \geqslant 1$ (einfache Überlegung), sodass

$$\mu\left(|f_{n_{k_j}} - f| > \frac{1}{j}\right) < 2^{-j}$$

Dann gilt

$$\sum_{j\geqslant 1} \mu(|f_{n_{k_j}} - f| > \varepsilon) = \sum_{\substack{j\geqslant 1\\ j\leqslant 1/\varepsilon}} \mu(|f_{n_{k_j}} - f| > \varepsilon) + \sum_{\substack{j\geqslant 1\\ j\geqslant 1/\varepsilon}} \mu(|f_{n_{k_j}} - f| > \varepsilon)$$

Die erste Summe ist endlich, und jeder Term ist nach oben durch 1 beschränkt. Eine Abschätzung der Terme in der zweiten Summe erfolgt mit obigem Argument. Damit gilt

$$\sum_{j\geqslant 1} \mu(|f_{n_{k_j}} - f| > \varepsilon) \leqslant \frac{1}{\varepsilon} + \sum_{j\geqslant 1} 2^{-j} = \frac{1}{\varepsilon} + 2 < \infty$$

und mit Lemma 9.3 folgt $f_{n_{k_j}} \xrightarrow[j\to\infty]{a.e.} f$.

II. $\Longleftarrow$

Angenommen $f_n \not\xrightarrow[n\to\infty]{\mu} f$. Dann gibt es $\varepsilon > 0$, sodass $\mu(|f_n - f| > \varepsilon)$ nicht gegen 0 geht. Da μ aber endlich ist, ist die Folge $(\mu(|f_n - f| > \varepsilon))_{n\geqslant 1}$ beschränkt und mit Bolzano–Weierstraß existiert eine Teilfolge $f_{n_k}, k \geqslant 1$, die konvergiert, also

$$\lim_{k\to\infty} \mu(|f_{n_k} - f| > \varepsilon) = \alpha > 0.$$

Da alle Teilfolgen von konvergenten Folgen gegen denselben Grenzwert konvergieren, folgt für die Teilfolge $f_{n_{k_j}}, j \geqslant 1$ aus der Annahme

$$\lim_{j\to\infty} \mu(|f_{n_{k_j}} - f| > \varepsilon) = \alpha$$

Mit Proposition 9.5 gilt aber $f_{n_{k_j}} \xrightarrow[j\to\infty]{\mu} f$ und damit erhalten wir einen Widerspruch. □

Bemerkung: Die Richtung $\Longrightarrow$ gilt sogar für allgemeine Maßräume, da Lemma 9.3. keine Annahmen an Endlichkeit macht.

9.7. Definition: Sei $p \geqslant 1$. Eine Funktionenfolge $f_n, n \geqslant 1$ konvergiert in L^p (bzw. im p-ten Mittel) gegen f, falls

$$\int |f_n - f|^p \, d\mu \xrightarrow[n\to\infty]{} 0$$

Wir schreiben dann $f_n \xrightarrow[n\to\infty]{L^p} f$

9.8. Proposition:

$$f_n \xrightarrow[n\to\infty]{L^p} f \implies f_n \xrightarrow[n\to\infty]{\mu} f$$

Beweis: Sei $\varepsilon > 0$. Es gilt

$$\begin{aligned}\int |f_n - f|^p \, d\mu &\geqslant \int |f_n - f|^p \cdot \mathbb{1}_{\{|f_n - f|^p > \varepsilon^p\}} \, d\mu \\ &\geqslant \int \varepsilon^p \cdot \mathbb{1}_{\{|f_n - f|^p > \varepsilon^p\}} \, d\mu \\ &= \varepsilon^p \cdot \mu(|f_n - f|^p > \varepsilon^p)\end{aligned}$$

Teile beide Seiten durch ε^p und die linke Seite konvergiert noch immer gegen 0, und damit auch die rechte Seite. Damit folgt per Definiton von Konvergenz im Maß die Aussage. □

Bemerkung: Der Beweis liefert auch eine allgemeine Form der Markov-Ungleichung: Für $g \geqslant 0$ und $\varepsilon > 0$ gilt

$$\mu(g \geqslant \varepsilon) \leqslant \frac{1}{\varepsilon} \int f \, d\mu$$

9.9. Proposition: Für $1 \leqslant p \leqslant q < \infty$ gilt

$$f_n \xrightarrow[n\to\infty]{L^q} f \implies f_n \xrightarrow[n\to\infty]{L^p} f$$

Beweis: Mit der Ljapunov-Ungleichung gilt

$$\left(\int |f_n - f|^p \, d\mu\right)^{1/p} \leqslant \left(\int |f_n - f|^q \, d\mu\right)^{1/q}$$

wobei die rechte Seite laut Annahme gegen 0 konvergiert. Die Aussage folgt mit dem continuous mapping theorem für konvergente Folgen reeller Zahlen. □

9.10. Proposition: Sei $f_n, n \geqslant 1$ eine Funktionenfolge, sodass $f_n \xrightarrow[n\to\infty]{L^1} f$ für eine absolut integrierbare Funktion f, i.e. $f \in L^1$. Dann folgt

$$\int f_n \, d\mu \xrightarrow[n\to\infty]{} \int f \, d\mu$$

Beweis: Es gilt $f_n \in L^1$ für hinreichend große n (also $\exists N \geqslant 1 \forall n \geqslant N : f_n \in L^1$), denn mit der Dreiecksungleichung gilt

$$\limsup_{n\to\infty} \int |f_n| \, d\mu \leqslant \limsup_{n\to\infty} \left(\int |f_n - f| \, d\mu + \int |f| \, d\mu\right) < \infty$$

Weiters ist

$$\limsup_{n\to\infty} \left| \int f_n \, d\mu - \int f \, d\mu \right| \leqslant \limsup_{n\to\infty} \int |f_n - f| \, d\mu = 0$$

womit die Aussage folgt. □

9.12. Proposition: Sei μ ein endliches Maß und f_n, g_n, f, g alle reellwertig für alle $n \geqslant 1$, sodass $f_n \xrightarrow[n\to\infty]{\mu/a.e.} f$ und $g_n \xrightarrow[n\to\infty]{\mu/a.e.} g$. Dann gilt

(i) $f_n \pm g_n \xrightarrow[n\to\infty]{\mu/a.e.} f \pm g$

(ii) $f_n \cdot g_n \xrightarrow[n\to\infty]{\mu/a.e.} f \cdot g$

(iii) Falls $\mu(g = 0)$, dann $\dfrac{f_n}{g_n} \xrightarrow[n\to\infty]{\mu/a.e.} \dfrac{f}{g}$

Beweis: Der Fall für Konvergenz f.ü. folgt sofort aus der Tatsache, dass die Vereinigung zweier Nullmengen wieder eine Nullmenge ist. Zeige also die Aussage für Konvergenz im Maß.

Sei $n_k, k \geqslant 1$ eine Teilfolge von $n, n \geqslant 1$. Weil $f_n \xrightarrow[n\to\infty]{\mu} f$, gibt es eine wegen Proposition 9.6 eine weitere Teilfolge $n_{k_j}, j \geqslant 1$ von $n_k, k \geqslant 1$, sodass $f_{n_{k_j}} \xrightarrow[j\to\infty]{a.e.} f$. Da $n_{k_j}, j \geqslant 1$ aber auch eine Teilfolge der ursprünglichen Folge $n, n \geqslant 1$ ist, gibt es eine weitere Teilfolge $n_{k_{j_\ell}}, \ell \geqslant 1$, sodass $g_{n_{k_{j_\ell}}} \xrightarrow[\ell\to\infty]{a.e.} g$. Es gilt aber auch $f_{n_{k_{j_\ell}}} \xrightarrow[\ell\to\infty]{a.e.} f$ und damit $f_{n_{k_{j_\ell}}} \pm g_{n_{k_{j_\ell}}} \xrightarrow[\ell\to\infty]{a.e.} f \pm g$ und $f_{n_{k_{j_\ell}}} \cdot g_{n_{k_{j_\ell}}} \xrightarrow[\ell\to\infty]{a.e.} f \cdot g$. Damit gibt es für jede Teilfolge $n_k, k \geqslant 1$ von $n, n \geqslant 1$ eine weitere Teilfolge $n_{k_{j_\ell}}, \ell \geqslant 1$, sodass Summe/Differenz/Produkt konvergieren und mit Proposition 9.6 folgt die Aussage für Konvergenz im Maß. Für (iii) siehe Übung! □

9.13. Satz (Continuous Mapping Theorem, CMT): Sei μ endlich und sei $f_n, n \geqslant 1$ eine reellwertige Funktionenfolge, sodass $f_n \xrightarrow[n\to\infty]{\mu/a.e.} f$. Sei weiters $h : \mathbb{R} \to \mathbb{R}$ eine Abbildung, die stetig auf einer Menge $H \subseteq N$, mit $\mu(f \notin N) = 0$ ist. Dann gilt

$$h(f_n) \xrightarrow[n\to\infty]{\mu/a.e.} h(f)$$

Beweis: Zeige den Fall mit Konvergenz f.ü. Sei $A := \{\lim_{n\to\infty} f_n \neq f\} \cup \{f \notin N\}$. Dann gilt mit der σ-Subadditivität $\mu(A) = 0$ und für $\omega \notin A$ gilt $f_n(\omega) \xrightarrow[n\to\infty]{} f(\omega)$ und $f(\omega) \in H$. Mit dem Continuous Mapping Theorem für punktweise Konvergenz folgt $h(f_n(\omega)) \xrightarrow[n\to\infty]{} h(f(\omega))$. Damit folgt

$$h(f_n) \xrightarrow[n\to\infty]{a.e.} h(f)$$

Die Aussage für Konvergenz im Maß folgt mit Proposition 9.6. $\square$

Konvergenz von $\mathbb{R}^d$-wertigen Funktionen

In diesem Abschnitt seien $f_n, f : \Omega \to \mathbb{R}^d$ $\mathcal{A}$–$\mathcal{B}(R^d)$-messbar und $\|\cdot\|$ die euklidische Norm auf $\mathbb{R}^d$.

9.14. Definition:

$$f_n \xrightarrow[n\to\infty]{\mu/a.e./L^p} f \iff \|f_n - f\| \xrightarrow[n\to\infty]{\mu/a.e./L^p} 0$$

9.15. Proposition: Für $f_n = \left(f_n^{(1)}, \dots, f_n^{(d)}\right)'$ und $f = \left(f^{(1)}, \dots, f^{(d)}\right)'$ gilt

$$f_n \xrightarrow[n\to\infty]{\mu/a.e./L^p} f \iff \forall i = 1, \dots, d : f_n^{(i)} \xrightarrow[n\to\infty]{\mu/a.e./L^p} f^{(i)}$$

Beweis: Für $x = (x_1, \dots, x_d)'$ gilt die folgende Ungleichung für alle $j = 1, \dots, d$

$$|x_j| = \sqrt{x_j^2} \leqslant \sqrt{\sum_{j=1}^d x_j^2} = \|x\| \leqslant \sqrt{d \cdot \max_{1\leqslant j\leqslant d} x_j^2} = \sqrt{d} \max_{1\leqslant j\leqslant d} |x_j| \leqslant \sqrt{d} \sum_{j=1}^d |x_j|$$

und damit

$$|f_n^{(i)} - f^{(i)}| \leqslant \|f_n - f\| \leqslant \sqrt{d} \sum_{j=1}^d |f_n^{(j)} - f^{(j)}|$$

womit die Behauptung folgt. $\square$

Bermekung: Damit gelten die Resultate aus dem vorherigen Abschnitt auch für vektorwertige rationale Operationen, soweit diese definiert sind.

Konvergenz von Integralen

Seien in diesem Abschnitt $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $f_n, f : (\Omega, \mathcal{A}) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ messbar.

9.16. Lemma (von Fatou, 1.Version): Sei $f_n, n \geqslant 1$ eine Folge nicht-negativer Funktionen. Dann gilt

$$\int \liminf_{n\to\infty} f_n \; d\mu \leqslant \liminf_{n\to\infty} \int f_n \; d\mu$$

Beweis: Setze $g_n := \inf_{k\geqslant n} f_k$, sodass $\liminf_{n\to\infty} = \lim_{n\to\infty} g_n$ und $0 \leqslant g_1 \leqslant \ldots \leqslant \lim_{n\to\infty} g_n$. Dann gilt mit MONK

$$\int \lim_{n\to\infty} g_n = \int \liminf_{n\to\infty} f_n \; d\mu = \lim_{n\to\infty} \int g_n \; d\mu$$

Da aber $g_n = \inf_{k\geqslant n} f_k \leqslant f_n$ gilt für alle $n \geqslant 1 : \int g_n \; d\mu \leqslant \int f_n \; d\mu$ und damit

$$\int \liminf_{n\to\infty} f_n \; d\mu = \lim_{n\to\infty} \int g_n \; d\mu \leqslant \liminf_{n\to\infty} \int f_n \; d\mu$$

$\square$

9.17. Lemma (von Fatou, 2.Version): Sei $f_n, n \geqslant 1$ eine Folge von Funktionen, sodass $g \leqslant f_n$ für alle $n \geqslant 1$ und $g_- \in L^1$. Sei weiters f_n quasi-integrierbar für alle $n \geqslant 1$. Dann gilt

$$\int \liminf_{n\to\infty} f_n \; d\mu \leqslant \liminf_{n\to\infty} \int f_n \; d\mu$$

Beweis:

I. Fall ($\int g_+ \; d\mu = \infty$)
Weil $f \leqslant f_n$ für alle $n \geqslant 1$ und damit $g \leqslant \liminf_{n\to\infty} f_n$, folgt $\int f_n \; d\mu = \infty$ für alle $n \geqslant 1$. Damit gilt auch $\liminf_{n\to\infty} \int f_n \; d\mu = \infty$ und die Aussage folgt trivial.

II. Fall ($\int g_+ \; d\mu < \infty$)
Damit gilt laut Voraussetzung $g \in L^1$ und damit $g < \infty$ f.ü., sodass für alle $n \geqslant 1$ $(f_n - g)$ f.ü. wohldefiniert und f.ü. nicht-negativ ist. Mit Lemma 9.16 folgt

$$\int \liminf_{n\to\infty} (f_n - g) \; d\mu \leqslant \liminf_{n\to\infty} \int f_n - g \; d\mu$$

und mit der Linearität des Integrals und der Tatsache, dass g nicht von n abhängt folgt die Aussage. $\square$

9.18. Satz (Dominated Convergence Theorem, DOMK): Sei $f_n, n \geqslant 1$ eine Funktionenfolge, sodas $f_n \xrightarrow[n\to\infty]{\mu} f$ und $|f_n| \leqslant g$ für alle $n \geqslant 1$ und eine integrierbare Funktion g. Dann folgt

(i) $f \in L^1$

(ii) $\displaystyle\int f_n \; d\mu \xrightarrow[n\to\infty]{} \int f \; d\mu$

(iii) $f_n \xrightarrow[n\to\infty]{L^1} f$

Beweis: In zwei Schritten: Zeige das Resultat zunaächst unter der stärkeren Annahme $f_n \xrightarrow[n\to\infty]{a.e.} f$ und erweitere den Beweis dann um ein Teil-Teilfolgen-Argument. Beachte, dass $|f| \leqslant g$ f.ü. und mit der Monotonie $f_n, f \in L^1$ für alle $n \geqslant 1$. Es gilt f.ü. $f_n \geqslant -g \in L^1$ und Lemma 9.17 liefert

$$\int f \; d\mu = \int \liminf_{n\to\infty} f_n \; d\mu \leqslant \liminf_{n\to\infty} \int f_n \; d\mu.$$

Ebenso gilt $-f_n \geqslant -g$ f.ü und mit Lemma 9.17 folgt

$$-\int f \; d\mu = \int \liminf_{n\to\infty} (-f_n) \; d\mu \leqslant \liminf_{n\to\infty} \int (-f_n) \; d\mu = -\limsup_{n\to\infty} \int f_n \; d\mu$$

und damit

$$\int f \; d\mu \geqslant \limsup_{n\to\infty} \int f_n \; d\mu.$$

Zusammengefasst

$$\limsup_{n\to\infty} \int f_n \; d\mu \leqslant \int f \; d\mu \leqslant \liminf_{n\to\infty} \int f_n \; d\mu$$

und damit $\int f \; d\mu = \lim_{n\to\infty} \int f_n \; d\mu$. Es gelte nun die schwächere Annahme $f_n \xrightarrow[n\to\infty]{\mu} f$. Setze

$$I_n := \int f_n \; d\mu, \; I := \int f \; d\mu.$$

Sei $(f_{n_k})_{k\geqslant 1}$ eine Teilfolge von $(f_n)_{n\geqslant 1}$. Dann gibt es eine Teil-Teilfolge $(f_{n_{k_j}})_{j\geqslant 1}$ von $(f_{n_k})_{k\geqslant 1}$, sodass $f_{n_{k_j}} \xrightarrow[n\to\infty]{a.e.} f$ und da $|f_{n_{k_j}}| \leqslant g$ f.ü., folgt mit dem ersten Teil $I_{n_{k_j}} \xrightarrow[j\to\infty]{} I$ und insbesondere $|I_n| \leqslant \int |g| \; d\mu < \infty$ für alle $n \geqslant 1$. Angenommen I_n konvergiert nicht. Da I_n beschränkt ist, folgt

$$-\infty < m := \liminf_{n\to\infty} I_n < \limsup_{n\to\infty} =: M < \infty.$$

Damit existieren Teilfolgen $(I_{n_{k^{(m)}}})_{k^{(m)}\geqslant 1}$ und $(I_{n_{k^{(M)}}})_{k^{(M)}\geqslant 1}$, sodass

$$I_{n_{k^{(m)}}} \xrightarrow[k^{(m)}\to\infty]{} m, \; I_{n_{k^{(M)}}} \xrightarrow[k^{(M)}\to\infty]{} M,$$

ein Widerspruch, da jede Teil-Teilfolge denselben Grenzwert I haben muss. Ebenso lässt sich zeigen, dass I_n nur gegen I konvergieren kann. $\square$

9.19. Korollar (Bounded Convergence Theorem): Sei μ ein endliches Maß und $f_n, n \geqslant 1$ eine Funktionenfolge, sodass $f_n \xrightarrow[n\to\infty]{\mu} f$ und $|f_n| \leqslant K$ für ein $K \in [0, \infty)$. Dann folgt (i), (ii) und (iii) aus Satz 9.18.

Beweis: Folgt sofort aus DOMK (Satz 9.18) und $\int K \; d\mu = K \cdot \mu(\Omega) < \infty$ für endliche Maße. $\square$

9.20. Lemma (Scheffé's Lemma): Seien $f_n, n \geqslant 1, f$ nicht-negative, integrierbare Funktionen, sodass $f_n \xrightarrow[n\to\infty]{\mu} f$. Falls zusätzlich $\int f_n \; d\mu \xrightarrow[n\to\infty]{} \int f \; d\mu$, dann folgt $f_n \xrightarrow[n\to\infty]{L^1} f$.

Beweis: Setze $h_n := f - f_n$ und wende DOMK (Satz 9.18) auf die Folgen $(h_n)_+, n \geqslant 1$ und $(h_n)_-, n \geqslant 1$ an. $\square$

9.21. Proposition: Falls $f_n \xrightarrow[n\to\infty]{L^p} f$ und $f \in L^p$ für ein $p \geqslant 1$, dann folgt

(i) $\displaystyle\int |f_n|^p \, d\mu \xrightarrow[n\to\infty]{} \int |f|^p \, d\mu$

(ii) $\displaystyle\int f_n \, d\mu \xrightarrow[n\to\infty]{} \int f \, d\mu$

Beweis:

(i) Mit der Minkowski-Ungleichung gilt

$$\left(\int |f_n|^p \, d\mu\right)^{1/p} = \left(\int |f_n - f + f|^p \, d\mu\right)^{1/p} \leqslant \left(\int |f_n - f|^p \, d\mu\right)^{1/p} + \left(\int |f|^p \, d\mu\right)^{1/p}$$

wobei der erste Summand laut Annahme gegen 0 konvergiert (und damit insbesondere ab einem Index $N \geqslant 1$ endlich ist) und der zweite Summand laut Annahme endlich ist. Es gilt also $f_n \in L^p$ für hinreichend große n. Weiters folgt

$$\limsup_{n\to\infty} \int |f_n|^p \, d\mu \leqslant \int |f|^p \, d\mu$$

da der erste Summand nicht-negativ ist. Aber mit der Minkowski-Ungleichung gilt auch

$$\left(\int |f|^p \, d\mu\right)^{1/p} = \left(\int |f - f_n + f_n|^p \, d\mu\right)^{1/p} \leqslant \left(\int |f - f_n|^p \, d\mu\right)^{1/p} + \left(\int |f_n|^p \, d\mu\right)^{1/p}$$

sodass

$$\liminf_{n\to\infty} \int |f_n|^p \, d\mu \geqslant \int |f|^p \, d\mu$$

und daher

$$\lim_{n\to\infty} \int |f_n|^p \, d\mu = \int |f| \, d\mu$$

(ii) Es gilt $|(f_n)_+ - f_+| \leqslant |f_n - f|$ und $|(f_n)_- - f_-| \leqslant |f_n - f|$ (einfache Überlegung). Laut Annahme gilt $f_n \xrightarrow[n\to\infty]{L^p} f$ und mit Proposition 9.9 auch $f_n \xrightarrow[n\to\infty]{L^1} f$. Mit den beiden Ungleichung oben folgt also

$$(f_n)_+ \xrightarrow[n\to\infty]{L^1} f_+ \text{ und } (f_n)_- \xrightarrow[n\to\infty]{L^1} f_-$$

Da $f \in L^p$ gilt auch $f \in L^1$ und damit $f_+, f_- \in L^1$. Mit Teil (i) folgt

$$\int (f_n)_+ \, d\mu \xrightarrow[n\to\infty]{} \int f_+ \, d\mu \text{ und } \int (f_n)_- \, d\mu \xrightarrow[n\to\infty]{} \int f_- \, d\mu$$

und mit den Rechenregeln für Konvergenz von Folgen reeller Zahlen die Aussage. □

Gleichgradige Integrierbarkeit

Seien in diesem Abschnitt $f_n : (\Omega, \mathcal{A}) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ messbare Funktionen für alle $n \geqslant 1$.

9.22. Definition: Sei μ endlich. Eine Folge messbarer Funktionen $f_n, n \geqslant 1$ ist gleichgradig integrierbar ("uniformly integrable", g.i.), falls

$$\lim_{\alpha\to\infty} \limsup_{n\to\infty} \int\limits_{|f_n|\geqslant\alpha} |f_n| \; d\mu = 0$$

Bemerkung: Eine konstante Folge absolut integrierbarer, reellwertiger Funktionen ist gleichgradig integrierbar, da

$$\lim_{\alpha\to\infty} \int\limits_{\{|f|\geqslant\alpha\}} |f| \; d\mu = \lim_{\alpha\to\infty} \int |f| \cdot \mathbb{1}_{\{|f|\geqslant\alpha\}} \; d\mu$$

wobei $|f| < \infty$ f.ü., da $f \in L^1$ und damit

$$|f| \cdot \mathbb{1}_{\{|f|\geqslant\alpha\}} \xrightarrow[\alpha\to\infty]{a.e.} |f| \cdot 0 \overset{a.e.}{=} \mathbb{1}_{\{|f|=\infty\}}$$

Da $|f| \cdot \mathbb{1}_{\{|f|\geqslant\alpha\}} \leqslant f \in L^1$ folgt mit DOMK

$$\int\limits_{\{|f|\geqslant\alpha\}} |f| \; d\mu \xrightarrow[\alpha\to\infty]{} \mu(|f| = \infty) = 0$$

9.23. Lemma: Sei μ endlich und $f_n, n \geqslant 1$ gleichgradig integrierbar. Dann folgt

$$\limsup_{n\to\infty} \int |f_n| \; d\mu < \infty$$

i.e. $f_n \in L^1$ für hinreichend große n.

Beweis: Wähle $\alpha > 0$, sodass $\limsup_{n\to\infty} \int_{\{|f_n|\geqslant\alpha\}} |f_n| < \infty$. Dann gilt

$$\limsup_{n\to\infty} \int |f_n| \; d\mu = \limsup_{n\to\infty} \left(\int\limits_{\{|f_n|\geqslant\alpha\}} |f_n| \; d\mu + \int\limits_{\{|f_n|\geqslant\alpha\}} |f_n| \; d\mu \right) < \infty$$

wobei der erste Summand laut Annahme endlich ist und der zweite Summand $\leqslant \alpha \cdot \mu(\Omega)$ ist. □

9.24. Lemma: Sei μ endlich und seien $f_n, n \geqslant 1$ und $g_n, n \geqslant 1$ gleichgradig integrierbar. Dann ist $f_n + g_n$ f.ü. wohldefiniert für hinreichend große n und $f_n + g_n, n \geqslant 1$ ist gleichgradig integrierbar.

Beweis: Es ist $\int |f_n| \; d\mu < \infty$ für $n \geqslant n_f$ und $\int |g_n| \; d\mu < \infty$ für $n \geqslant n_g$. Damit ist

$$\int f_n + g_n \; d\mu \leqslant \int |f_n| + |g_n| \; d\mu < \infty$$

und damit $f_n + g_n < \infty$ fast überall für $n \geqslant \max(n_f, n_g)$.

Setze nun $h_n := \max(|f_n|, |g_n|), n \geqslant 1$. Dann gilt $|f_n + g_n| \leqslant 2h_n$ und

$$\begin{aligned} h_n \cdot \mathbb{1}_{\{h_n\geqslant\alpha/2\}} &= h_n \cdot \mathbb{1}_{\{|f_n|\geqslant\alpha/2\}} \cdot \mathbb{1}_{\{f_n\geqslant g_n\}} + h_n \cdot \mathbb{1}_{\{|g_n|\geqslant\alpha/2\}} \cdot \mathbb{1}_{\{f_n<g_n\}} \\ &\leqslant |f_n| \cdot \mathbb{1}_{\{|f_n|\geqslant\alpha/2\}} + |g_n| \cdot \mathbb{1}_{\{|g_n|\geqslant\alpha/2\}} \end{aligned}$$

und damit für $\alpha > 0$

$$\begin{aligned}\int\limits_{\{|f_n+g_n|\geqslant\alpha\}} |f_n + g_n| \; d\mu &\leqslant 2 \int\limits_{\{h_n\geqslant\alpha/2\}} h_n \; d\mu \\ &\leqslant 2 \int\limits_{\{|f_n|\geqslant\alpha/2\}} |f_n| \; d\mu + 2 \int\limits_{\{|g_n|\geqslant\alpha/2\}} |g_n| \; d\mu\end{aligned}$$

wobei der lim sup für $n \to \infty$ beider Summanden für $\alpha \to \infty$ gegen 0 geht, womit $f_n + g_n, n \geqslant 1$ gleichgradig integrierbar sind. $\square$

9.25. Satz: Sei μ endlich. Dann ist folgendes äquivalent

(i) $f_n \xrightarrow[n\to\infty]{L^1} f$ und $f \in L^1$

(ii) $f_n \xrightarrow[n\to\infty]{\mu} f$ und $f_n, n \geqslant 1$ gleichgradig integrierbar

Bemerkung: Aus (i) folgt mit Proposition 9.10, dass

$$\int f_n \; d\mu \xrightarrow[n\to\infty]{} \int f \; d\mu$$

Beweis:

I. (i) $\Longrightarrow$ (ii)
Aus (i) folgt mit Proposition 9.8, dass $f_n \xrightarrow[n\to\infty]{\mu} f$. Weiters ist für $\alpha > 0$

$$0 = \lim_{n\to\infty} \int |f_n - f| \; d\mu \geqslant \int\limits_{\{|f_n-f|\geqslant\alpha\}} |f_n - f| \; d\mu \geqslant 0$$

Damit ist $f_n - f, n \geqslant 1$ gleichgradig integrierbar. Da $f \in L^1$ ist die konstante Folge $f, n \geqslant 1$ gleichgradig integrierbar und mit Satz 9.24 auch die Summe $f_n = (f_n - f) + f, n \geqslant 1$

II. (ii) $\Longrightarrow$ (i)
Mit Proposition 9.6 genügt es, die Aussage unter der stärkeren Annahme $f_n \xrightarrow[n\to\infty]{a.e.} f$ zu zeigen. Mit Fatou I gilt

$$\int |f| \; d\mu = \int \liminf_{n\to\infty} f_n \; d\mu \leqslant \liminf_{n\to\infty} \int |f_n| \; d\mu \leqslant \limsup_{n\to\infty} \int |f_n| \; d\mu < \infty$$

wobei die Endlichkeit des Integrals für große n aus der Annahme der gleichgradigen Integrierbarkeit folgt. Damit gilt $f \in L^1$. Für $\alpha > 0$ gilt

$$\int |f_n - f| \; d\mu = \int\limits_{\{|f_n-f|\geqslant\alpha\}} |f_n - f| \; d\mu + \int\limits_{\{|f_n-f|<\alpha\}} |f_n - f| \; d\mu$$

wobei mit Korollar 9.19 gilt

$$\int\limits_{\{|f_n-f|<\alpha\}} |f_n - f| \; d\mu \xrightarrow[n\to\infty]{} 0$$

und damit

$$\limsup_{n\to\infty} \int |f_n - f| \; d\mu = \limsup_{n\to\infty} \int\limits_{\{|f_n - f| \geqslant \alpha\}} |f_n - f| \; d\mu$$

Da aber $f \in L^1$ ist die konstante Folge $f, n \geqslant 1$ gleichgradig integrierbar, womit die Aussage mit der Dreiecksungleichung für $\alpha \searrow 0$ folgt. $\square$

9.26. Proposition: Sei μ endlich. Angenommen $f_n \xrightarrow[n\to\infty]{\mu} f$ und $\limsup_{n\to\infty} \int |f_n|^{1+\delta} \; d\mu < \infty$ für ein $\delta > 0$. Dann folgt

(i) $f \in L^1$

(ii) $\displaystyle\int f_n \; d\mu \xrightarrow[n\to\infty]{} \int f \; d\mu$

(iii) $f_n \xrightarrow[n\to\infty]{L^1} f$

Beweis: Unter der Annahme $f_n \xrightarrow[n\to\infty]{\mu}$ genügt es mit Satz 9.25 zu zeigen, dass $f_n, n \geqslant 1$ gleichgradig integrierbar ist.

$$\begin{aligned}
\limsup_{n\to\infty} \int\limits_{\{|f_n| \geqslant \alpha\}} |f_n| \; d\mu &= \limsup_{n\to\infty} \int\limits_{\{|f_n|^\delta \geqslant \alpha^\delta\}} |f_n| \cdot \frac{\alpha^\delta}{\alpha^\delta} \; d\mu \\
&\leqslant \frac{1}{\alpha^\delta} \limsup_{n\to\infty} \int\limits_{\{|f_n|^\delta \geqslant \alpha^\delta\}} |f_n||f_n|^\delta \; d\mu \\
&= \frac{1}{\alpha^\delta} \limsup_{n\to\infty} \int\limits_{\{|f_n| \geqslant \alpha\}} |f_n||f_n|^\delta \; d\mu \xrightarrow[\alpha\to\infty]{} 0
\end{aligned}$$

da der $\limsup$ im letzten Ausdruck endlich ist und $\alpha^{-\delta} \xrightarrow[\alpha\to\infty]{} 0$. $\square$

10. Konvergenz von Zufallsvariablen

Gesetze der großen Zahlen

Sei im folgenden Kapitel $(\Omega, \mathcal{A}, \mathbb{P})$ jeweils ein Wahrscheinlichkeitsraum und $X_n, X \in L^1, n \geqslant 1$ jeweils $\overline{\mathbb{R}}$-wertige Zufallsvariablen.

10.1. Proposition: Sei $X \in L^2$. Dann ist folgendes äquivalent:

1. $\displaystyle \frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}X_i) \xrightarrow[n\to\infty]{L^2} 0$
2. $\displaystyle \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\operatorname{Cov}(X_i, X_j) \xrightarrow[n\to\infty]{} 0$

Beweis:

$$\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}X_i\right)^2\right] = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{E}\left[(X_i - \mathbb{E}X_i)(X_j - \mathbb{E}X_j)\right]$$

Bermerkung: Für $X_i \in L^2$ unkorreliert ist 2. äquivalent zu

$$\frac{1}{n^2}\sum_{i=1}^{n}\operatorname{Var}(X_i) \longrightarrow 0$$

Falls die $X_i \in L^2$ und i.i.d. sind, gilt 1. und 2. trivial.

10.2. Satz (Schwaches Gesetz der großen Zahlen, WLLN): Seien $X_i \in L^1$ i.i.d.. Dann gilt:

$$\frac{1}{n}\sum_{i=1}^{n}X_i \xrightarrow[n\to\infty]{P} \mathbb{E}X_1$$

Beweis: Sei $M > 0$. Dann gilt

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}X_i) &= \frac{1}{n}\sum_{i=1}^{n}\left(X_i \cdot \mathbb{1}_{\{|X_i|\leqslant M\}} - \mathbb{E}[X_i \cdot \mathbb{1}_{\{|X_i|\leqslant M\}}]\right)\\ &+ \frac{1}{n}\sum_{i=1}^{n}\left(X_i \cdot \mathbb{1}_{\{|X_i|> M\}} - \mathbb{E}[X_i \cdot \mathbb{1}_{\{|X_i|> M\}}]\right)\end{aligned}$$

I. Zum ersten Summanden:
Die Zufallsvariablen $X_i \cdot \mathbb{1}_{\{|X_i| \leqslant M\}}$ sind unabhängig und beschränkt (also $\in L^\infty$ und damit insbesondere $\in L^2$). Damit ist

$$\operatorname{Var}\left(X_i \cdot \mathbb{1}_{\{|X_i| \leqslant M\}}\right) = \mathbb{E}\left[X_i^2 \cdot \mathbb{1}_{\{|X_i| \leqslant M\}}\right] - \left(\mathbb{E}\left[X_i \cdot \mathbb{1}_{\{|X_i| \leqslant M\}}\right]\right)^2 \leqslant 2M^2$$

Mit Bemerkung (ii) zu Proposition 10.1 folgt damit, dass der erste Summand $\xrightarrow[n\to\infty]{L^2}$ und damit $\xrightarrow[n\to\infty]{P} 0$.

II. Zum zweiten Summanden:

$$\begin{aligned}
\mathbb{E}\left|\frac{1}{n}\sum_{i=1}^n \left(X_i \cdot \mathbb{1}_{\{|X_i|>M\}} - \mathbb{E}[X_i \cdot \mathbb{1}_{\{|X_i|>M\}}]\right)\right| &\overset{\text{2x DUG}}{\leqslant} \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n \left|X_i \cdot \mathbb{1}_{\{|X_i|>M\}}\right| + \left|\mathbb{E}[X_i \cdot \mathbb{1}_{\{|X_i|>M\}}]\right|\right] \\
&\leqslant \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left|X_i \cdot \mathbb{1}_{\{|X_i|>M\}}\right| + \mathbb{E}[|X_i| \cdot \mathbb{1}_{\{|X_i|>M\}}] \\
&= 2 \cdot \mathbb{E}\left[|X_1| \cdot \mathbb{1}_{\{|X_i|>M\}}\right] \xrightarrow[M\to\infty]{} 0
\end{aligned}$$

Sei also $\varepsilon > 0$. Dann ist

$$\begin{aligned}
\limsup_{n\to\infty} \mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^n (X_i - \mathbb{E}X_i)\right| > \varepsilon\right) &= \limsup_{n\to\infty} \mathbb{P}(|(1)+(2)| > \varepsilon) \\
&\leqslant \limsup_{n\to\infty} \mathbb{P}\left(|(1)| > \frac{\varepsilon}{2}\right) + \mathbb{P}\left(|(2)| > \frac{\varepsilon}{2}\right) \\
&= \limsup_{n\to\infty} \mathbb{P}\left(|(2)| > \frac{\varepsilon}{2}\right) \\
&\overset{\text{Markov}}{\leqslant} \limsup_{n\to\infty} \frac{2 \cdot \mathbb{E}|(2)|}{\varepsilon} \\
&\leqslant \frac{4 \cdot \mathbb{E}\left[|X_1| \cdot \mathbb{1}_{\{|X_i|>M\}}\right]}{\varepsilon}
\end{aligned}$$

für jedes $M > 0$. Wähle nun für $\varepsilon > 0$ ein $M > 0$ so, sodass dieser Ausdruck $< \delta$ ist. Das funktioniert wegen II. für jedes $\delta > 0$ und daher folgt

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^n (X_i - \mathbb{E}X_i)\right| > \varepsilon\right) \xrightarrow[n\to\infty]{} 0$$

und damit die Aussage. □

10.3. Satz (Starkes Gesetz der großen Zahlen, SLLN): Seien $X_i \in L^1$ i.i.d.. Dann gilt sogar die stärkere Aussage:

$$\frac{1}{n}\sum_{i=1}^n X_i \xrightarrow[n\to\infty]{a.s.} \mathbb{E}X_1$$

Beweis: siehe z.B.: P. Billingsley, *Probability and Measure* (2nd Ed.), Theorem 6.1 □

10.4. Proposition (Momentenmethode): Betrachte i.i.d. Zufallsvariablen $X_i, i \geqslant 1$, sodass $X_i^d \in L^1$ für ein $d \geqslant 1$. Sei außerdem $f : \mathbb{R}^d \to \mathbb{R}^\ell$ und setze $\theta := f(\mathbb{E}X_1^1, \ldots, \mathbb{E}X_1^d)$ (θ könnte z.B. als Funktion der ersten d Momente die Verteilung von X_1 parametrisieren). Falls f stetig im Punkt $(\mathbb{E}X_1^1, \ldots, \mathbb{E}X_1^d)'$ ist, dann gilt für $\hat{\theta}_n := f\left(\frac{1}{n}\sum_{i=1}^n X_i^1, \ldots, \frac{1}{n}\sum_{i=1}^n X_i^d\right)$, dass $\hat{\theta}_n \xrightarrow[n\to\infty]{P} \theta$.

Beweis: Mit der Ljapunov-Unggleichung gilt $X_i^j \in L^1$ für $j = 1, \ldots, d$. Weiters sind X_i^j i.i.d. Mit dem SLLN gilt damit für $j = 1, \ldots, d$

$$\frac{1}{n}\sum_{i=1}^n X_i^j \xrightarrow[n\to\infty]{a.s.} \mathbb{E}X_1^j$$

und mit Proposition 9.15 folgt

$$\left(\frac{1}{n}\sum_{i=1}^n X_i^1, \ldots, \frac{1}{n}\sum_{i=1}^n X_i^d\right)' \xrightarrow[n\to\infty]{a.s.} (\mathbb{E}X_1^1, \ldots, \mathbb{E}X_1^d)'$$

Die Aussage folgt schließlich mit dem CMT (quasi Satz 9.13. für $f : \mathbb{R}^d \to \mathbb{R}^\ell$). $\square$

11. Schwache Konvergenz

Betrachte in diesem Kapitel einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$, reelwertige Zufallsvariablen $X, X_n, n \geqslant 1$ mit entsprechenden cdfs $F, F_n, n \geqslant 1$ bzw. den entsprechenden induzierten Wahrscheinlichkeitsmaßen $\mathbb{P}, \mathbb{P}_n, n \geqslant 1$ auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

11.1. Definition: $\mathbb{P}_n, n \geqslant 1$ konvergieren schwach/in Verteilung gegen $\mathbb{P}$, wenn

$$\forall t \in \mathbb{R} \text{ mit } F(\cdot) \text{ stetig in t} : F_n(t) \xrightarrow[n\to\infty]{} F(t)$$

Kurz: $\mathbb{P}_n \xrightarrow[n\to\infty]{d} \mathbb{P}$, oder $F_n \xrightarrow[n\to\infty]{d} F$, oder $X_n \xrightarrow[n\to\infty]{d} X$.

11.2. Proposition: $X_n \xrightarrow[n\to\infty]{P} X \implies X_n \xrightarrow[n\to\infty]{d} X$

Beweis: Sei $\varepsilon > 0$ beliebig.

$$\begin{aligned} F_n(t) &= \mathbb{P}(X_n \leqslant t) \\ &= \mathbb{P}(X_n \leqslant t, |X_n - X| \leqslant \varepsilon) + \mathbb{P}(X_n \leqslant t, |X_n - X| > \varepsilon) \\ &\leqslant \mathbb{P}(X - \varepsilon \leqslant t, |X_n - X| \leqslant \varepsilon) + \mathbb{P}(|X_n - X| > \varepsilon) \\ &\leqslant \mathbb{P}(X - \varepsilon \leqslant t) + \mathbb{P}(|X_n - X| > \varepsilon) \end{aligned}$$

Damit folgt $\forall t \in \mathbb{R}$:

$$\limsup_{n\to\infty} F_n(t) \leqslant F(t+\varepsilon)$$

Gleichzeitig gilt:

$$\begin{aligned} F(t-\varepsilon) &= \mathbb{P}(X \leqslant t - \varepsilon) \\ &= \mathbb{P}(X \leqslant t - \varepsilon, |X_n - X| \leqslant \varepsilon) + \mathbb{P}(X \leqslant t - \varepsilon, |X_n - X| > \varepsilon) \\ &\leqslant \mathbb{P}(X_n - \varepsilon \leqslant t - \varepsilon, |X_n - X| \leqslant \varepsilon) + \mathbb{P}(|X_n - X| > \varepsilon) \\ &= \mathbb{P}(X_n \leqslant t) + \mathbb{P}(|X_n - X| > \varepsilon) \end{aligned}$$

Damit folgt

$$\liminf_{n\to\infty} F_n(t) \geqslant F(t-\varepsilon)$$

Wenn $F(\cdot)$ nun stetig im Punkt t ist, dann gilt

$$\lim_{\varepsilon \searrow 0} F(t-\varepsilon) = \lim_{\varepsilon \searrow 0} F(t+\varepsilon) = F(t)$$

und damit

$$F(t) \leqslant \liminf_{n\to\infty} F_n(t) \leqslant \limsup_{n\to\infty} F_n(t) \leqslant F(t)$$

und es folgt für alle $t \in \mathcal{C}(F)$:

$$F_n(t) \xrightarrow[n\to\infty]{} F(t)$$

$\square$

11.3. Proposition: Es gelte $X_n \xrightarrow[n\to\infty]{d} X$ mit $\mathbb{P}(X = c) = 1$ für ein $c \in \mathbb{R}$. Dann folgt $X_n \xrightarrow[n\to\infty]{P} c$.

Beweis:

$$F(t) = \begin{cases} 1, & \text{if } t \geqslant c \\ 0, & \text{if } t < c \end{cases}$$

ist stetig auf $\mathbb{R} \setminus \{c\}$ und damit stetig in $c \pm \varepsilon$ für alle $\varepsilon > 0$. Also gilt

$$\begin{aligned} \mathbb{P}(|X_n - X| > \varepsilon) &= \mathbb{P}(|X_n - c| > \varepsilon) \\ &= \mathbb{P}(X_n - c \notin [-\varepsilon, \varepsilon]) \\ &= 1 - \mathbb{P}(X_n - c \in [-\varepsilon, \varepsilon]) \\ &\leqslant 1 - F_n(c+\varepsilon) + F_n(c-\varepsilon) \xrightarrow[n\to\infty]{} 1 - F(c+\varepsilon) + F(c-\varepsilon) = 0 \end{aligned}$$

Damit folgt $X_n \xrightarrow[n\to\infty]{P} X \overset{a.s.}{=} c$. $\square$

11.4. Satz (von Glivenko–Cantelli): Seien $X_n, n \geqslant 1$ i.i.d.reellwertige Zufallsvariablen mit cdf F. Definiere die empirical cdf:

$$\hat{F}_n(t) := \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{X_i \leqslant t\}}$$

für alle $\omega \in \Omega$. Dann gilt:

$$\sup_{t\in\mathbb{R}} |\hat{F}_n(t) - F(t)| \xrightarrow[n\to\infty]{a.s.} 0$$

Beweis: Zeige zuerst, dass $\sup_{t\in\mathbb{R}} |\hat{F}_n(t) - F(t)|$ messbar ist. Wähle dazu $t_k \in \mathbb{R}, k \geqslant 1$, sodass

$$|\hat{F}_n(t_k) - F(t_k)| \geqslant \sup_{t\in\mathbb{R}} |\hat{F}_n(t) - F(t)| - \frac{1}{k}$$

Beachte, dass t_k von ω abhängt! Wähle nun $q_k \in \mathbb{Q}, k \geqslant 1$, sodass

$$|\hat{F}_n(q_k) - F(q_k)| \geqslant |\hat{F}_n(t_k) - F(t_k)| - \frac{1}{k}$$

Dieser Schritt funktioniert wegen der rechtsseitigen Stetitgkeit von $F, \hat{F}_n, n \geqslant 1$. Nun folgt aber

$$|\hat{F}_n(q_k) - F(q_k)| \geqslant \sup_{t\in\mathbb{R}} |\hat{F}_n(t) - F(t)| - \frac{2}{k}$$

und damit

$$|\hat{F}_n(q_k) - F(q_k)| \xrightarrow[k\to\infty]{} \sup_{t\in\mathbb{R}} |\hat{F}_n(t) - F(t)|$$

Also
$$\sup_{q\in\mathbb{Q}} |\hat{F}_n(q) - F(q)| = \sup_{t\in\mathbb{R}} |\hat{F}_n(t) - F(t)|$$
wobei die linke Seite als abzählbares Supremem messbarer Funktionen messbar ist.
Setze nun $F(-\infty) = \hat{F}_n(-\infty) = 0$ und $F(\infty) = \hat{F}_n(\infty) = 1$ und wähle ein Mesh
$$-\infty = t_0 < t_1 < \ldots < t_{k-1} < t_k = +\infty$$
sodass
$$F(t_{j-}) - F(t_{j-1}) \leqslant \varepsilon \tag{1}$$
für $1 \leqslant j \leqslant k$. Dazu verfährt man folgendermaßen:
Beginne mit $t_0 = -\infty$. Angenommen wir haben schon $j-1 \geqslant 0$ Punkte gewählt, sodass (1) für alle $i \leqslant j-1$ gilt und $t_{j-1} < \infty$. Setze dann
$$t_j := \sup\{t > t_{j-1} : F(t_-) - F(t_{j-1}) < \varepsilon\}$$
Mit der Stetigkeit von oben gilt dann $F(t_-) - F(t_{j-1}) = \mathbb{P}(t_{j-1} < X < t) \xrightarrow[t\searrow t_{j-1}]{} 0$. Damit ist t_j wohldefiniert. Falls $t_j = \infty$ sind wir fertig. Andernfalls wiederholt man die Prozedur für $j+1$. Wir haben am Schluss also Punkte $(t_j)_{0\leqslant j\leqslant k}$ für die gilt:
$$\begin{aligned} F(t_{j-}) - F(t_{j-1}) &\leqslant \varepsilon \\ F(t_j) - F(t_{j-1}) &\geqslant \varepsilon \end{aligned}$$
Da $F(\infty) - F(-\infty) = 1$, gilt $k\varepsilon \leqslant 1$ und damit $k < \left\lfloor \frac{1}{\varepsilon} \right\rfloor$, also endet die Prozedur immer in endlich vielen Schritten.
Zeige nun, dass die empirische cdf gleichmäßig in $t \in \mathbb{R}$ gegen die tatsächliche cdf konvergiert. Mit dem SLLN gilt
$$\hat{F}_n(t_j) \xrightarrow[n\to\infty]{a.s.} F(t_j) \tag{2}$$
$$\hat{F}_n(t_{j-}) \xrightarrow[n\to\infty]{a.s.} F(t_{j-}) \tag{3}$$
Sei also N_ε, so dass für alle $\omega \in N_\varepsilon^c$ (2) und (3) gelten. Für jedes $t \in \mathbb{R}$ gibt es $1 \leqslant j \leqslant k$, sodass $t \in [t_{j-1}, t_j)$ und es gilt
$$\begin{aligned} \hat{F}_n(t) - F(t) &\overset{\text{Monotonie}}{\leqslant} \hat{F}_n(t_{j-}) - F(t_{j-1}) = \\ &= \hat{F}_n(t_{j-}) - F(t_{j-}) + F(t_{j-}) - F(t_{j-1}) \leqslant \\ &\leqslant \hat{F}_n(t_{j-}) - F(t_{j-}) + \varepsilon \end{aligned}$$
und
$$\begin{aligned} \hat{F}_n(t) - F(t) &\overset{\text{Monotonie}}{\geqslant} \hat{F}_n(t_{j-1}) - F(t_{j-}) = \\ &= \hat{F}_n(t_{j-1}) - F(t_{j-1}) + F(t_{j-1}) - F(t_{j-}) \geqslant \\ &\geqslant \hat{F}_n(t_{j-1}) - F(t_{j-1}) - \varepsilon \end{aligned}$$
Damit folgt
$$\begin{aligned} \limsup_{n\to\infty} \sup_{t\in\mathbb{R}} \left(\hat{F}_n(t_{j-}) - F(t)\right) &\leqslant \limsup_{n\to\infty} \max\left\{\hat{F}_n(t_{j-1}) - F(t_{j-1} \leqslant j \leqslant k)\right\} + \varepsilon \\ \liminf_{n\to\infty} \inf_{t\in\mathbb{R}} \left(\hat{F}_n(t_{j-}) - F(t)\right) &\geqslant \liminf_{n\to\infty} \min\left\{\hat{F}_n(t_{j-1}) - F(t_{j-1} \leqslant j \leqslant k)\right\} - \varepsilon \end{aligned}$$

Weil $\varepsilon > 0$ beliebig war, folgt

$$\liminf_{n\to\infty} \inf_{t\in\mathbb{R}} \left(\hat{F}_n(t_{j-}) - F(t)\right) = \limsup_{n\to\infty} \sup_{t\in\mathbb{R}} \left(\hat{F}_n(t_{j-}) - F(t)\right) = 0$$

und damit

$$\lim_{n\to\infty} \sup_{t\in\mathbb{R}} |\hat{F}_n(t) - F(t)| = 0$$

für alle $\omega \in \bigcup_{\varepsilon>0} N_\varepsilon^c = \bigcup_{n\geqslant 1} N_n^c$ mit Maß 0. $\square$

Bemerkung: Für eine Klasse von Funktionen $\mathcal{F}$, sei

$$\forall f \in \mathcal{F}: \ \hat{F}_n(f) := n^{-1} \sum_{i=1}^{n} f(X_i)$$
$$F(f) := \mathbb{E}[f(X_1)]$$

Im Fall von Satz 11.4 war z.B.$\mathcal{F} = \left\{\mathbb{1}_{(-\infty,t]}(\cdot), t \in \mathbb{R}\right\}$. Im allgemeinen Fall geben Glivenko-Cantelli Theoreme Bedinungen an die Klasse $\mathcal{F}$, sodass

$$\sup_{f\in\mathcal{F}} |\hat{F}_n(f) - F(f)| \xrightarrow[n\to\infty]{a.e.} 0$$

11.5. Satz (Portemanteau Theorem 1): Es gilt $X_n \xrightarrow[n\to\infty]{d} X$ genau dann, wenn eine der folgenden Bedingungen erfüllt ist:

(i) $\mathbb{E}f(X_n) \xrightarrow[n\to\infty]{} \mathbb{E}f(X)$ für alle f stetig mit kompaktem Träger.

(ii) $\mathbb{E}f(X_n) \xrightarrow[n\to\infty]{} \mathbb{E}f(X)$ für alle f stetig und beschränkt.

Bemerkung: Ist f stetig und $\overline{\operatorname{supp} f}$ kompakt, dann ist f auch beschränkt. Damit gilt trivial (ii) $\Longrightarrow$ (i).

Beweis:

I. $X_n \xrightarrow[n\to\infty]{d} X \implies$ (i)
Sei f wie in (i). Dann ist f auf $\operatorname{supp} f$ auch gleichmäßig stetig. Für $\varepsilon > 0$ wähle ein Mesh

$$a_0 < a_1 < \ldots < a_{k-1} < a_k$$

sodass

$$\forall x \notin (a_0, a_k] : f(x) = 0$$
$$\forall x \in (a_{i-1}, a_i] : |f(x) - f(a_i)| < \varepsilon \quad \text{für } 1 \leqslant i \leqslant k$$

und so, dass $a_o, \ldots, a_k$ alle Stetigkeitsstellen von F sind (davon gibt es höchstens abzählbar viele, also ist das ohne Probleme möglich).
Definiere

$$g_\varepsilon(x) := \sum_{i=1}^{k} f(a_i)\mathbb{1}_{(a_{i-1},a_i]}(x)$$

sodass

$$\sup_{x\in\mathbb{R}} |f(x) - g_\varepsilon(x)| \leqslant \varepsilon \tag{4}$$

Dann gilt

$$\begin{aligned}\mathbb{E}g_\varepsilon(X_n) &= \sum_{i=1}^{k} f(a_i)\mathbb{P}(a_{i-1} < X_n \leqslant a_i) = \\ &= \sum_{i=1}^{k} f(a_i)(F_n(a_i) - F_n(a_{i-1})) \xrightarrow[n\to\infty]{} \sum_{i=1}^{k} f(a_i)(F(a_i) - F(a_{i-1})) = \mathbb{E}g_\varepsilon(X)\end{aligned}$$

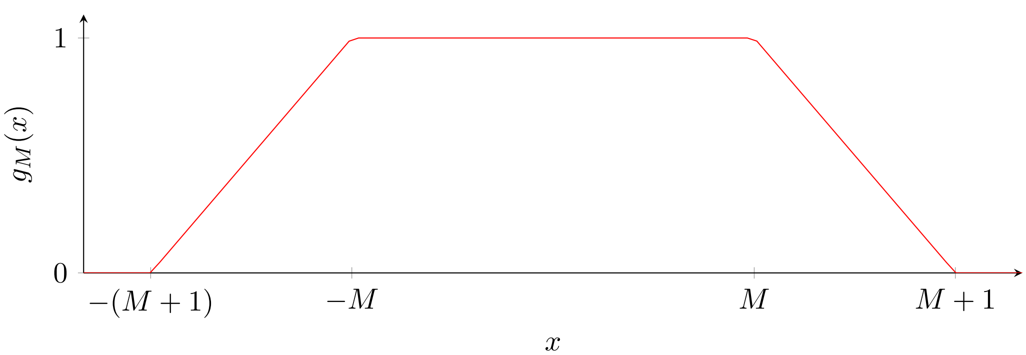
und mit (4) folgt $\forall\varepsilon > 0 : f \leqslant g_\varepsilon + \varepsilon$ und $g_\varepsilon \leqslant f + \varepsilon$ und damit

$$\begin{aligned}\limsup_{n\to\infty} \mathbb{E}f(X_n) \leqslant \limsup_{n\to\infty} \mathbb{E}g_\varepsilon(X_n) + \varepsilon = \mathbb{E}g_\varepsilon(X) + \varepsilon \leqslant \mathbb{E}f(X) + 2\varepsilon \\ \liminf_{n\to\infty} \mathbb{E}f(X_n) \geqslant \limsup_{n\to\infty} \mathbb{E}g_\varepsilon(X_n) - \varepsilon = \mathbb{E}g_\varepsilon(X) - \varepsilon \geqslant \mathbb{E}f(X) - 2\varepsilon\end{aligned}$$

und für $\varepsilon \to 0$ folgt, dass

$$\mathbb{E}f(X_n) \xrightarrow[n\to\infty]{} \mathbb{E}f(X)$$

II. (i) $\Longrightarrow$ (ii)
Sei f wie in (ii). Für $M > 1$ definiere eine stetige Funktion g_M mit kompaktem Träger, wie folgt:



Also $g_M(x) := \begin{cases} 0 \text{ if } x \notin (-M-1, M+1] \\ 1 \text{ if } x \in (-M, M] \\ x + (M+1) \text{ if } x \in (-M-1, -M] \\ (M+1) - x \text{ if } x \in (M, M+1] \end{cases}$

Wähle nun M groß genug, sodass $|1 - \mathbb{E}g_M(X)| < \varepsilon$ (möglich wegen $g_M(x) \xrightarrow[M\to\infty]{} 1$ und MONK). Die Funktion $f_M := f \cdot g_M$ ist dann stetig, beschränkt und hat kompakten Träger (einfache überlegung). Weiters folgt

$$\begin{aligned}\big|\mathbb{E}f(X_n) - \mathbb{E}f(X)\big| &= \big|\mathbb{E}\left[f(X_n) - f_M(X_n) + f_M(X_n) - f_M(X) + f_M(X) - f(X)\right]\big| \leqslant \\ &\leqslant \big|\mathbb{E}f(X_n) - \mathbb{E}f_M(X_n)\big| + \big|\mathbb{E}f_M(X_n) - \mathbb{E}f_M(X)\big| + \big|\mathbb{E}f_M(X) - f(X)\big|\end{aligned}$$

Betrachte nun der Reihe nach alle drei Summanden

$$\big|\mathbb{E}f(X_n) - f_M(X_n)\big| \leqslant \|f\|_\infty \big|\mathbb{E}[1 - g_M(X_n)]\big| \xrightarrow[n\to\infty]{} \|f\|_\infty \big|\mathbb{E}[1 - g_M(X)]\big| \leqslant \varepsilon \cdot \|f\|_\infty$$

$$\big|\mathbb{E}[f_M(X_n) - f_M(X)]\big| \xrightarrow[n\to\infty]{\text{Ann.}} 0$$

$$\big|\mathbb{E}f(X) - f_M(X)\big| \overset{\text{s.o.}}{\leqslant} \varepsilon \cdot \|f\|_\infty$$

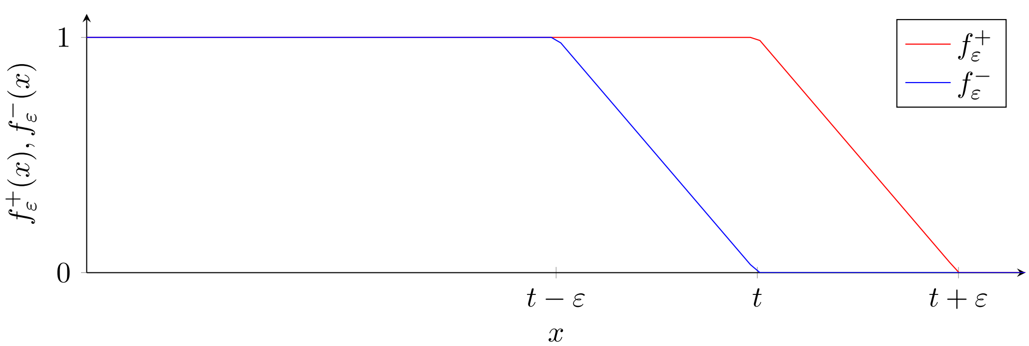
Damit folgt

$$0 \leqslant \limsup_{n\to\infty} \big|\mathbb{E}[f(X_n) - f(X)]\big| \leqslant 2\varepsilon$$

und da $\varepsilon > 0$ beliebig war auch die Aussage.

III. (ii) $\implies X_n \xrightarrow[n\to\infty]{d} X$

Für $t \in \mathbb{R}$ definiere stetige, beschränkte Funktionen f_ε^- und f_ε^+ wie folgt:



Dann ist $f_\varepsilon^- \leqslant \mathbb{1}_{(-\infty,t]} \leqslant f_\varepsilon^+$ und damit

$$\mathbb{E}f_\varepsilon^-(X_n) \leqslant F_n(t) \leqslant \mathbb{E}f_\varepsilon^+(X_n)$$

für alle $n \geqslant 1$. Es folgt mit DOMK

$$\mathbb{E}f_\varepsilon^-(X) \leqslant \liminf_{n\to\infty} F_n(t) \leqslant \limsup_{n\to\infty} F_n(t) \leqslant \mathbb{E}f_\varepsilon^+(X)$$

Aber

$$\mathbb{E}f_\varepsilon^-(X) \geqslant \mathbb{E}[\mathbb{1}_{(-\infty,t-\varepsilon]}(X)] = F(t-\varepsilon)$$
$$\mathbb{E}f_\varepsilon^-(X) \leqslant \mathbb{E}[\mathbb{1}_{(-\infty,t+\varepsilon]}(X)] = F(t+\varepsilon)$$

Wenn F also stetig im Punkt t ist, folgt $F_n(t) \xrightarrow[n\to\infty]{} F(t)$. □

11.6. Satz (Portemanteau Theorem 2): Es gilt $X_n \xrightarrow[n\to\infty]{d} X$ genau dann, wenn eine der folgenden Bedingungen zutrifft:

(i) Für $O \subseteq \mathbb{R}$ offen ist $\mathbb{P}(X \in O) \leqslant \liminf\limits_{n\to\infty} \mathbb{P}(X_n \in O)$.

(ii) Füpr $A \subseteq \mathbb{R}$ abgeschlossen ist $\mathbb{P}(X \in A) \geqslant \limsup\limits_{n\to\infty} \mathbb{P}(X_n \in A)$.

(iii) Für $B \in \mathcal{B}(\mathbb{R})$ mit $\mathbb{P}(X \in \partial B) = 0$ gilt $\mathbb{P}(X_n \in B) \xrightarrow[n\to\infty]{} \mathbb{P}(X \in B)$.

Beweis: (i) $\iff$ (ii) folgt sofort aus O offen $\iff$ O^c abgeschlossen.

I. $X_n \xrightarrow[n\to\infty]{} X \implies$ (i)
Sei $O \subseteq \mathbb{R}$ offen und wähle stetige und beschränkte Funktionen (einfache Überlegung, Hinweis: Jede offene Teilmenge von $\mathbb{R}$ ist eine abzählbare Vereinigung offener Intervalle) $f_m, m \geqslant 1$, sodass

$$0 \leqslant f_1 \leqslant \ldots \leqslant \lim_{m\to\infty} f_m = \mathbb{1}_O$$

Es folgt mit Portemanteau 1 (Satz 11.5)

$$\forall m \geqslant 1 : \mathbb{E} f_m(X) = \lim_{n\to\infty} \mathbb{E} f_m(X_n) \leqslant \liminf_{n\to\infty} \mathbb{P}(X_n \in O)$$

und damit

$$\mathbb{P}(X \in O) = \mathbb{E}\mathbb{1}_O(X) \overset{\text{MONK}}{=} \lim_{m\to\infty} \mathbb{E} f_m(X) \overset{\text{s.o.}}{\leqslant} \lim_{m\to\infty} \liminf_{n\to\infty} \mathbb{P}(X_n \in O) = \liminf_{n\to\infty} \mathbb{P}(X_n \in O)$$

II. (i),(ii) $\implies$ (iii)
Sei $B \in \mathcal{B}(\mathbb{R})$ mit $\mathbb{P}(X \in \partial B) = 0$. Dann ist

$$\mathbb{P}(X \in B \setminus \partial B) = \mathbb{P}(X \in \operatorname{int}(S)) = \mathbb{P}(X \in \operatorname{clos}(S)) = \mathbb{P}(X \in S)$$

und damit

$$\mathbb{P}(X \in B) = \mathbb{P}(X \in \operatorname{clos}(B)) \overset{\text{(ii)}}{\geqslant} \limsup_{n\to\infty} \mathbb{P}(X_n \in \operatorname{clos}(B)) \geqslant \limsup_{n\to\infty} \mathbb{P}(X_n \in B)$$
$$\mathbb{P}(X \in B) = \mathbb{P}(X \in \operatorname{int}(B)) \overset{\text{(i)}}{\leqslant} \liminf_{n\to\infty} \mathbb{P}(X_n \in \operatorname{int}(B)) \leqslant \liminf_{n\to\infty} \mathbb{P}(X_n \in B)$$

und damit

$$\mathbb{P}(X_n \in B) \xrightarrow[n\to\infty]{} \mathbb{P}(X \in B)$$

III. (iii) $\implies$ $X_n \xrightarrow[n\to\infty]{d} X$
Für $B = (-\infty, t]$ ist $\mathbb{P}(X_n \in B) = F_n(t)$ und $\partial B = \{t\}$. Wenn F also stetig in t ist, folgt die Aussage. $\square$

11.7. Satz (Slutsky's Theorem): Wenn $X_n \xrightarrow[n\to\infty]{d} X$ und $Z_n \xrightarrow[n\to\infty]{d} Z$, mit $\mathbb{P}(Z = c) = 1$, dann gilt:

(i) $X_n + Z_n \xrightarrow[n\to\infty]{d} X + c$

(ii) $X_n Z_b \xrightarrow[n\to\infty]{d} Xc$

(iii) Falls $c \neq 0$: $\dfrac{X_n}{Z_n} \xrightarrow[n\to\infty]{d} \dfrac{X}{c}$

Beweis:

(i) Wenn t Stetigkeitspunkt der Verteilung von $X + c$ ist, dann ist $t - c$ Stetigkeitspunkt der Verteilung von X, also

$$\lim_{x \nearrow t} \mathbb{P}(X \leqslant x - c) = \lim_{x \nearrow c} \mathbb{P}(X + c \leqslant x) \overset{\text{Ann.}}{=} \mathbb{P}(X + c \leqslant t) = \mathbb{P}(X \leqslant t - c)$$

Damit folgt $X_n + c \xrightarrow[n\to\infty]{d} X + c$ für $c \in \mathbb{R}$.
Sei also $f : \mathbb{R} \to \mathbb{R}$ stetig mit kompaktem Träger. Dann ist f beschränkt und gleichmäßg stetig. Sei $\varepsilon > 0$ und wähle $\delta > 0$, sodass

$$\forall x, y \in \mathbb{R} : |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$$

Dann gilt

$$\begin{aligned}\left|\mathbb{E}f(X_n + Z_n) - \mathbb{E}f(X_n + c)\right| &\leqslant \mathbb{E}\left|f(X_n + Z_n) - f(X_n + c)\right| = \\ &= \mathbb{E}\left[\left|f(X_n + Z_n) - f(X_n + c)\right| \mathbb{1}_{|Z_n - c| \geqslant \delta}\right] \\ &+ \mathbb{E}\left[\left|f(X_n + Z_n) - f(X_n + c)\right| \mathbb{1}_{|Z_n - c| < \delta}\right] \leqslant \\ &\leqslant \varepsilon + 2\|f\|_\infty \cdot \mathbb{P}(|Z_n - c| \geqslant \delta) \xrightarrow[n\to\infty]{} \varepsilon\end{aligned}$$

Für $\varepsilon \searrow 0$ und mit $X_n + c \xrightarrow[n\to\infty]{d} X + c$ folgt die Aussage mit Portememanteau 1 (Satz 11.5).

(ii) Zeige zuerst den Fall wo $c = 0$:
Hier genügt es mit Proposition 11.3 zu zeigen, dass $X_n Z_n \xrightarrow[n\to\infty]{P} 0$. Für $\varepsilon > 0$ gilt

$$\begin{aligned}&\mathbb{P}(|X_n Z_n| > \varepsilon) \\ &= \mathbb{P}(|X_n Z_n| > \varepsilon, |Z_n| > \delta) + \mathbb{P}(|X_n Z_n| > \varepsilon, |Z_n| \leqslant \delta) \\ &\leqslant \mathbb{P}(|X_n Z_n| > \varepsilon, |Z_n| \leqslant \delta) + \mathbb{P}(|Z_n| > \delta) \\ &= \mathbb{P}(X_n < -\varepsilon/\delta) + [1 - \mathbb{P}(X_n \leqslant \varepsilon/\delta)] + \mathbb{P}(|Z_n| > \delta) \\ &\leqslant F_n(-\varepsilon/\delta) + 1 - F_n(\varepsilon/\delta) + \mathbb{P}(|Z_n| > \delta)\end{aligned}$$

Wir können nun $\delta > 0$ so wählen, dass F in den Punkten $\pm\varepsilon/\delta$ stetig ist. Dann folgt

$$\limsup_{n\to\infty} \mathbb{P}(|X_n Z_n| > \varepsilon) \leqslant F(-\varepsilon/\delta) + [1 - F(\varepsilon/\delta)]$$

und für $\delta \searrow 0$ folgt die gewünschte Aussage.
Zeige nun den Fall wo $c \neq 0$:
Schreibe dazu $X_n Z_n = X_n(Z_n - c) + X_n c$. Es gilt $X_n c \xrightarrow[n\to\infty]{d} Xc$ (einfach zu prüfen) und $(Z_n - c) \xrightarrow{P} 0$. Mit dem ersten Fall folgt

$$X_n(Z_n - c) \xrightarrow{P/d} 0$$

Mit (i) folgt also

$$X_n Z_n = X_n(Z_n - c) + X_n c \xrightarrow[n\to\infty]{d} Xc$$

(iii) Die Abbildung $t \mapsto 1/t$ ist steig auf $\mathbb{R} \setminus \{0\}$ und damit gilt

$$\frac{1}{Z_n} \xrightarrow[n\to\infty]{P} \frac{1}{c}$$

Mit (ii) folgt damit die Aussage. □

11.8. Satz (Continuous Mapping Theorem, CMT): Seien $X_n, n \geqslant 1$ und X Zufallsvariablen, sodass $X_n \xrightarrow[n\to\infty]{d} X$ und sei $h : \mathbb{R} \to \mathbb{R}$ eine Abbildung, sodass es $H \in \mathcal{B}(\mathbb{R})$ gibt, mit $H \subseteq C(h)$ und $\mathbb{P}(X \in H) = 1$. Dann folgt

$$h(X_n) \xrightarrow[n\to\infty]{d} h(X)$$

Beweis: Sei $A \subseteq \mathbb{R}$ abgeschlossen. Dann gilt

$$h^{-1}A \subseteq \operatorname{clos}(h^{-1}A) \overset{\dagger}{\subseteq} h^{-1}A \cup H^c$$

† folgt mit einer kurzen überlegung aus der Tatsache, dass Urbilder abgeschlossener Mengen unter stetigen Funktionen wieder abgeschlossen sind.

Damit folgt nun

$$\begin{aligned}
\limsup_{n\to\infty} \mathbb{P}(h(X_n) \in A) &= \limsup_{n\to\infty} \mathbb{P}(X_n \in h^{-1}A) \\
&\leqslant \limsup_{n\to\infty} \mathbb{P}(X_n \in \operatorname{clos}(h^{-1}A)) \\
&\overset{\text{PMT2}}{\leqslant} \mathbb{P}(X \in h^{-1}A) \\
&\leqslant \mathbb{P}(X \in h^{-1}A \cup H^c) \\
&\leqslant \mathbb{P}(X \in h^{-1}A) + \mathbb{P}(X \in H^c) \\
&= \mathbb{P}(X \in h^{-1}A) = \mathbb{P}(h(X) \in A)
\end{aligned}$$

und da A eine beliebige abgeschlossene Teilmenge von $\mathbb{R}$ war folgt mit PMT2 (Satz 11.6), dass

$$h(X_n) \xrightarrow[n\to\infty]{d} h(X)$$

$\square$

11.9. Proposition (δ-Methode): Für eine Folge von Zufallsvariablen $X_n, n \geqslant 1$ mit

$$\sqrt{n}(X_n - \mu) \xrightarrow[n\to\infty]{d} \mathcal{N}(\prime, \sigma^{\in})$$

und eine Abbildung $f : \mathbb{R} \to \mathbb{R}$, die stetig im Punkt μ ist, gilt

$$\sqrt{n}(f(X_n) - f(\mu)) \xrightarrow[n\to\infty]{d} \mathcal{N}(0, (f'(\mu))^2\sigma^2)$$

Eine hinreichende Bedingung an die $X_n, n \geqslant 1$ wäre z.B., dass sie einen zentralen Grenzwertsatz (siehe Sätze 11.25, 11.26, 11.27) erfüllen.

Beweis: Betrachte die Abbildung $g : \mathbb{R} \to \mathbb{R}$ mit

$$x \mapsto \begin{cases} f'(\mu) - \dfrac{f(x) - f(\mu)}{x - \mu} & \text{if } x \neq \mu \\ 0 & \text{if } x = \mu \end{cases}$$

Da $f'(\mu)$ existiert ist g stetig im Punkt μ (Definition der Ableitung). Dann gilt

$$f(x) - f(\mu) = f'(\mu)(x-\mu) - g(x)(x-\mu)$$

und damit

$$\sqrt{n}(f(X_n) - f(\mu)) = f'(\mu)\sqrt{n}(X_n - \mu) - g(X_n)\sqrt{n}(X_n - \mu) =: A_n - B_n$$

Mit Slutsky's Theorem (Satz 11.7. (ii)) und dem Reproduktionssatz folgt sofort $A_n \xrightarrow[n\to\infty]{d} \mathcal{N}(0, (f'(\mu))^2\sigma^2)$. Es genügt mit Slutky's Theorem (Satz 11.7. (i)) also zu zeigen, dass $B_n \xrightarrow[n\to\infty]{P} 0$.

Sei $Z \sim \mathcal{N}(0,\sigma^2)$. Dann gilt (einfache überlegung) $X_n - Z \xrightarrow[n\to\infty]{P} 0$. Da g stetig in μ ist, gilt mit dem Continuous Mapping Theorem (Satz 11.8) $g(X_n) \xrightarrow[n\to\infty]{P} g(\mu) = 0$. Dann folgt erneut mit Slutsky's Theorem $B_n \xrightarrow[n\to\infty]{P/d} 0$ und damit die Behauptung. □

11.10. Satz Seien $X_n, n \geqslant 1$ gleichgradig integrierbar und $X_n \xrightarrow[n\to\infty]{d} X$. Dann folgt $X \in L^1$ und $\mathbb{E}X_n \xrightarrow[n\to\infty]{} \mathbb{E}X$.

Beweis:

I. $X_n, n \geqslant 0$ und X alle nicht-negativ
Aus der gleichgradigen Integrierbarkeit folgt mit Lemma 9.2

$$\limsup_{n\to\infty} \mathbb{E}|X_n| = B < \infty$$

Für alle $M > 1$ definiere $g_M(\cdot)$ wie im Beweis von Satz 11.5 (PMT1, Teil II). Dann ist g_M stetig, beschränkt und hat kompakten Träger. Also ist die Abbildung mit $t \mapsto t \cdot g_M(t)$ stetig mit kompaktem Träger und es gilt $\forall m \geqslant 1$

$$0 \leqslant \mathbb{E}[Xg_M(X)] \overset{\text{PMT1}}{=} \lim_{n\to\infty} \mathbb{E}[X_n g_M(X_n)] \overset{g_M \leqslant 1}{\leqslant} \limsup_{n\to\infty} \mathbb{E}X_n \leqslant \limsup_{n\to\infty} \mathbb{E}|X_n| = B < \infty$$

Gleichzeitig gilt aber

$$0 \leqslant Xg_1(X) \leqslant \ldots \leqslant \lim_{M\to\infty} Xg_M(X) = X$$

und mit MONK

$$0 \leqslant \mathbb{E}X = \lim_{M\to\infty} \mathbb{E}[Xg_M(X)] \leqslant B$$

und da $\mathbb{E}X = \mathbb{E}|X|$, folgt $X \in L^1$.

Zeige nun die Konvergenz:

$$\begin{aligned} \big|\mathbb{E}X_n - \mathbb{E}X\big| &= \big|\mathbb{E}[X_n g_M(X_n) - X + X_n - X_n g_M(X_n)]\big| \\ &\leqslant \big|\mathbb{E}[X_n g_M(X_n)] - \mathbb{E}X\big| + \mathbb{E}|X_n|\big|1 - g_M(X_n)\big| \\ &\leqslant \big|\mathbb{E}[X_n g_M(X_n)] - \mathbb{E}X\big| + \mathbb{E}|X_n| \cdot \mathbb{1}_{\{|X_n| \geqslant M\}} \end{aligned}$$

Damit folgt mit PMT1

$$\limsup_{n\to\infty} |\mathbb{E}X_n - \mathbb{E}X| \leqslant \left|\mathbb{E}[X g_M(X)] - \mathbb{E}X\right| + \limsup_{n\to\infty} \mathbb{E}|X_n| \cdot \mathbb{1}_{\{|X_n|\geqslant M\}}$$

und für $M \to \infty$

$$\limsup_{n\to\infty} |\mathbb{E}X_n - \mathbb{E}X| \leqslant \limsup_{M\to\infty} \left|\mathbb{E}X g_M(X) - \mathbb{E}X\right| + \limsup_{M\to\infty} \limsup_{n\to\infty} \mathbb{E}|X_n| \cdot \mathbb{1}_{\{|X_n|\geqslant M\}} = 0$$

wobei die letzte Gleichung mit MONK und der Def. von gleichgradiger Integrierbarkeit folgt.

II. allgemeiner Fall
Die Abbildung mit $t \mapsto \max(t, 0)$ ist stetig auf $\mathbb{R}$, sodass mit dem Continuous Mapping Theorem (Satz 11.8) für $X^+ := \max(X, 0)$ und $X_n^+ := \max(X_n, 0)$ gilt

$$X_n^+ \xrightarrow[n\to\infty]{d} X^+$$

Da $0 \leqslant X_n^+ \leqslant |X_n^+|$ und die $X_n, n \geqslant 1$ gleichgradig integrierbar sind, sind auch die $X_n^+, n \geqslant 1$ gleichgradig integrierbar. Mit dem I. Fall folgt

$$\mathbb{E}X_n^+ \xrightarrow[n\to\infty]{} \mathbb{E}X^+$$

Dasselbe gilt für den Negativteil und es folgt

$$\mathbb{E}X_n \xrightarrow[n\to\infty]{} \mathbb{E}X$$

mit den Rechenregeln für Konvergenz von Folgen reeler Zahlen. □

Schwach konvergente Teilfolgen

Erinnerung Analysis: Jede Folge reeller Zahlen $a_n, n \geqslant 1$ enthält eine Teilfolge $a_{n_i}, i \geqslant 1$, die gegen einen Grenzwert $\alpha \in \overline{\mathbb{R}}$ konvergiert, i.e. $\alpha = \lim_{i\to\infty} a_{n_i} \in \overline{\mathbb{R}}$. Falls $a_n, n \geqslant 1$ zusätzlich beschränkt ist, dann gilt $\alpha \in \mathbb{R}$. ähnliches gilt für Wahrscheinlichkeitsmaße.
Reimnder: $b_n, n \geqslant 1$ ist eine Teilfolge von $a_n, n \geqslant 1$, falls eine bijektive, streng monoton steigende Abbildung $f : \mathbb{N} \to \mathbb{N}$ existiert, sodass $b_n = a_{f(n)}$ für alle $n \geqslant 1$.

11.11. Satz (Helly's Selection Theorem): Sei $F_n, n \geqslant 1$ eine Folge von Verteilungsfunktionen (cdfs). Dann existiert eine Teilfolge $F_{n_i}, i \geqslant 1$ und eine monoton-nichtfallende rechtsstetige Funktion $F : \mathbb{R} \to \mathbb{R}$ mit linksseitigen Grenzwerten (cádlág), sodass

$$\forall t \in C(F) : F_{n_i}(t) \xrightarrow[i\to\infty]{} F(t)$$

Dabei ist aber nicht garantiert, dass F auch eine cdf ist (also $\lim\limits_{t\to\infty} F(t) = 1, \lim\limits_{t\to-\infty} F(t) = 0$)!

Beweis: Ordne $\mathbb{Q} = \{q_1, q_2, \ldots\}$ und wähle aus der vollen Folge $n = 1, 2, 3, \ldots$ eine erste Teilfolge $n_i(1), i \geqslant 1$, sodass $F_{n_{i(1)}}(q_1)$ für $i \to \infty$ kovergiert (möglich da $F_n \leqslant 1$, cf. Anmerkung unter der Unterüberschrift).
Wähle nun eine weitere (Teil-)Teilfolge $n_i(2), i \geqslant 1$, sodass $F_{n_{i(2)}}(q_2)$ für $i \to \infty$ konvergiert.

$\vdots$

Für die k-te Teilfolge $n_i(k), i \geqslant 1$ existieren dann die Grenzwerte

$$\lim_{i\to\infty} F_{n_i(k)}(q_\ell),\ \ \ell = 1, \dots, k$$

Setze nun $n_i := n_i(i)$ für $i \geqslant 1$. Dann konvergiert $F_{n_i}(q_\ell)$ für jedes $\ell \geqslant 1$, da n_i ab dem Index $i = \ell$ eine Teilfolge von $n_\ell(\ell)$ ist. Setze nun für $q \in \mathbb{Q}$

$$G(q) := \lim_{i\to\infty} F_{n_i}(q)$$

Dann ist $G : \mathbb{Q} \to [0, 1]$ monoton-nichtfallend auf $\mathbb{Q}$, da jedes Element der Folge $F_{n_i}(q)$ monoton-nichtfallend ist. Für $t \in \mathbb{R}$ definiere nun

$$F(t) := \inf\{G(q) : q \geqslant t, q \in \mathbb{Q}\}$$

Als infimum einer nichtleeren Menge ist $F(t)$ damit für alle $t \in \mathbb{R}$ wohldefiniert und

1. F ist monoton-nichtfallend auf $\mathbb{R}$, da

$$s \leqslant t \implies \{G(q) : q \geqslant s, q \in \mathbb{Q}\} \supseteq \{G(q) : q \geqslant t, q \in \mathbb{Q}\}$$

2. F ist rechtsstetig auf $\mathbb{R}$:
Sei $\varepsilon > 0, t \in \mathbb{R}$. Dann gibt es mit der Greatest Lower Bound Property $q \in \mathbb{Q}, q \geqslant t$, sodass

$$F(t) \leqslant G(q) \leqslant F(t) + \varepsilon$$

da $\mathbb{Q}$ dicht in $\mathbb{R}$ liegt. Damit gilt

$$\forall s \in [t, q] : F(t) \leqslant F(s) \leqslant F(q) = G(q) \leqslant F(t) + \varepsilon$$

und für $\varepsilon \searrow 0$ folgt die Aussage.

3. $F_{n_i} \xrightarrow[n\to\infty]{} F$ auf $C(F)$:
Sei F stetig in t, und sei $\varepsilon > 0$. Wähle nun $\delta > 0$, sodass

$$|t - s| < \delta \implies |F(t) - F(s)| < \varepsilon$$

Wähle nun $r, q \in \mathbb{Q}$, sodass

$$t - \delta < r < t < q < t + \delta$$

und somit

$$F(t) - \varepsilon < F(r) \leqslant F(t) \leqslant F(q) < F(t) + \varepsilon$$

Mit der Definition von G und F folgt nun

$$F_{n_i}(r) \leqslant F_{n_i}(t) \leqslant F_{n_i}(q)$$

wobei die linke und rechte Schranke gegen $G(r)$ bzw. $G(q)$ konvergieren, also

$$F(t) - \varepsilon < G(r) \leqslant \liminf_{i\to\infty} F_{n_i}(t) \leqslant \limsup_{i\to\infty} F_{n_i}(t) \leqslant G(q) < F(t) + \varepsilon$$

und für $\varepsilon \searrow 0$ folgt die Aussage. $\square$

11.12. Definition: Eine Folge von Zufallsvariablen $X_n, n \geqslant 1$ ist *stochastisch beschränkt*, wenn gilt:

$$\sup_{n\geqslant 1} \mathbb{P}(|X_n| > M) \xrightarrow[M\to\infty]{} 0$$

Die entsprechende Folge der Maße/cdfs nennt man dann *straff*.

11.13. Lemma: $X_n, n \geqslant 1$ ist genau dann stochastisch beschränkt, wenn gilt:

$$\forall \varepsilon > 0 \exists g : \mathbb{R} \to [0,\infty) \text{ messbar} : \limsup_{n\to\infty} \mathbb{E}g(X_n) \leqslant \varepsilon \cdot \liminf_{|t|\to\infty} g(t) \tag{5}$$

Beweis:

I. $\Longrightarrow$
Wähle $a_0 < a_1 < \ldots$ mit $a_k \xrightarrow[k\to\infty]{} \infty$, sodass mit der stochastischen Beschränktheit

$$\forall n \geqslant 1 : \mathbb{P}(|X_n| > a_k) < \left(\frac{1}{4}\right)^k$$

(Setze $\varepsilon = 4^{-k}$ und setze a_k dann einfach groß genug). Setze nun

$$g(t) := \sum_{k\geqslant 0} 2^k \mathbb{1}_{(a_k, a_{k+1}]}(|t|)$$

Dann ist g messbar (einfaches Argument) und $\liminf_{|t|\to\infty} g(t) = \infty$ (folgt z.B. aus $\liminf \geqslant \inf$).
Weiters gilt

$$\begin{aligned}\forall n \geqslant 1 : \mathbb{E}g(X_n) &= \sum_{k\geqslant 0} 2^k \mathbb{P}(a_k < |X_n| \leqslant a_{k+1}) \\ &\leqslant \sum_{k\geqslant 0} 2^k \mathbb{P}(|X_n| > a_k) \\ &\leqslant \sum_{k\geqslant 0} 2^k \frac{1}{4^k} = 2 < \infty\end{aligned}$$

Es gilt also $\forall \varepsilon > 0$

$$\limsup_{n\to\infty} \mathbb{E}g(X_n) \leqslant 2 \leqslant \infty = \varepsilon \cdot \liminf_{|t|\to\infty} g(t)$$

II. $\Longleftarrow$
Sei $\varepsilon > 0$ und wähle g, sodass (5) gilt.

$$\liminf_{|t|\to\infty} g(t) = \lim_{T\to\infty} \inf_{|t|>T} g(t)$$

wobei $\inf_{|t|>T} g(t)$ monoton nicht-fallend in T ist. Mit der Annahme können wir $T > 0$ wählen, sodass

$$\limsup_{n\to\infty} \mathbb{E}g(X_n) \leqslant \varepsilon \cdot \inf_{|t|>T} g(t) \leqslant \varepsilon \cdot \liminf_{|t|\to\infty} g(t)$$

Wähle nun $c \in \mathbb{R}$, sodass

$$\limsup_{n\to\infty} \mathbb{E}g(X_n) < c <\leqslant \varepsilon \cdot \inf_{|t|>T} g(t) \leqslant \varepsilon \cdot \liminf_{|t|\to\infty} g(t)$$

Für $|t| > T$ ist dann $c < \varepsilon g(t)$, bzw. $1 < \varepsilon g(t)/c$ und damit

$$\begin{aligned}\limsup_{n\to\infty} \mathbb{P}(|X_n| > T) &= \limsup_{n\to\infty} \mathbb{E}\left[1 \cdot \mathbb{1}_{\{|X_n|>T\}}\right] \\ &\leqslant \limsup_{n\to\infty} \mathbb{E}\left[\frac{\varepsilon g(X_n)}{c} \mathbb{1}_{\{|X_n|>T\}}\right] \\ &\leqslant \frac{\varepsilon}{c} \cdot \limsup_{n\to\infty} \mathbb{E} g(X_n) < \varepsilon\end{aligned}$$

Damit gibt es ein $n_0 \in \mathbb{N}$, sodass für $n \geqslant n_0 : \mathbb{P}(|X_n| > T) < \varepsilon$. Wähle nun $S > 0$ groß genug, sodass für $n = 1, \ldots, n_0 - 1 : \mathbb{P}(|X_n| > S) < \varepsilon$ und setze $M := \max(S, T)$. Damit gibt es für jedes $\varepsilon > 0$ ein $M > 0$, sodass

$$\forall n \geqslant 1 : \mathbb{P}(|X_n| > M) < \varepsilon$$

und insbesondere $\sup_{n\geqslant 1} \mathbb{P}(|X_n| > M) \xrightarrow[M\to\infty]{} 0$. □

11.14. Satz (Prokorov Theorem): Betrachte eine Folge von cdfs $F_n, n \geqslant 1$. Dann sind folgende Aussagen äquivalent:

(i) Die $F_n, n \geqslant 1$ sind straff.

(ii) Jede Teilfolge $F_{n_i}, i \geqslant 1$ enthält eine weitere Teilfolge $F_{n_{i_k}}, k \geqslant 1$, sodass

$$F_{n_{i_k}} \xrightarrow[k\to\infty]{d} F$$

für eine cdf F.

Beweis:

I. (i) $\Longrightarrow$ (ii)
Helly's Selection Theorem (Satz 11.11) liefert uns eine Teilfolge $F_{n_{i_k}}, k \geqslant 1$ und eine monoton nicht-fallende cádlág Funktion F, sodass $F_{n_{i_k}} \xrightarrow[k\to\infty]{d} F$. Zeige also

$$F(t) \xrightarrow[t\to\infty]{} 1 \text{ und } F(t) \xrightarrow[t\to-\infty]{} 0$$

Da die $F_n, n \geqslant 1$ straff sind, kann man $M_\ell, \ell \geqslant 1$ wählen, sodass

$$\sup_{n\geqslant 1} [F_n(-M_\ell) + 1 - F_n(M_\ell)] \leqslant \sup_{n\geqslant 1} \mathbb{P}(|X_n| > M_\ell) < \frac{1}{\ell}$$

O.B.d.A. seien $\pm M_\ell$ Stetigskeitspunkte und $M_\ell + 1 < M_{\ell+1} \xrightarrow[\ell\to\infty]{} \infty$ für alle $\ell \geqslant 1$. Dann gilt $F_{n_{i_k}}(M_\ell) = 1 - [1 - F_{n_{i_k}}(M_\ell)] > 1 - 1/\ell$ und damit

$$\lim_{t\to\infty} F_n(t) = \lim_{\ell\to\infty} F(M_\ell) = \lim_{\ell\to\infty} \left(\lim_{k\to\infty} F_{n_{i_k}}(M_\ell)\right) \geqslant 1$$

Da aber für alle $k \geqslant 1$ $F_{n_{i_k}}(M_\ell) \leqslant 1$ sind, folgt

$$\lim_{t\to\infty} F(t) = 1$$

Ein ähnliches Argument funktionert für $\lim_{t\to-\infty} F(t)$.

II. (ii) $\implies$ (i)

Angenommen die $F_n, n \geqslant 1$ sind nicht straff

$$\implies \exists \varepsilon > 0 \forall M > 0 : \sup_{n \geqslant 1} \mathbb{P}(|X_n| > M) \geqslant \varepsilon \tag{6}$$

Wähle nun $n_i, i \geqslant 1$ so, dass

$$F_{n_i}(-i) + 1 - F_{n_i}(i) > \frac{\varepsilon}{2}$$

Wegen (6) können die $n_i, i \geqslant 1$ monoton steigend gewählt werden. Laut Voraussetzung existiert nun eine weitere Teilfolge $F_{n_{i_k}}, k \geqslant 1$ und eine cdf F, sodass

$$F_{n_{i_k}} \xrightarrow[k \to \infty]{d} F$$

Weil F eine cdf ist, gibt es $M > 0$, sodass

$$F(-M) + 1 - F(M) < \frac{\varepsilon}{2}$$

Seien o.B.d.A. $\pm M$ Stetigkeitspunkte von F. Dann gilt

$$F(-M) + 1 - F(M) = \lim_{k \to \infty} \left[F_{n_{i_k}}(-M) + 1 - F_{n_{i_K}}(M) \right] < \frac{\varepsilon}{2}$$

Da aber

$$F_{n_{i_k}}(-M) + 1 - F_{n_{i_k}}(M) \geqslant F_{n_{i_k}}(-i_k) + 1 - F_{n_{i_k}}(i_k) \geqslant \varepsilon > \frac{\varepsilon}{2}$$

für hinreichend große k (sodass $i_k \geqslant M$) ergibt sich hier ein Widerspruch zur Annahme (6). □

Charakteristische Funktionen

$(\mathbb{C} = \{a + ib\, a, b \in \mathbb{R}\}, d_{|\cdot|})$ und $(\mathbb{R}^2, d_{\|\cdot\|})$ sind als metrische Räume via $\mathbb{C} \ni z = a + ib \mapsto (a, b) \in \mathbb{R}^2$ isometrisch isomoprh. Damit ergibt sich eine ähnliche Korrespondenz zwischen Borelmengen in $\mathcal{C}$ und $\mathbb{R}^2$. Eine Funktion $f : \Omega \to \mathbb{C}$ ist also genau dann messbar, wenn ihr Realteil und ihr Imaginärteil messbar sind, analog zur koordinatenweise Messbarkeit von $\mathbb{R}^2$-wertigen Funktionen. In diesem Kapitel sei $(\Omega, \mathcal{A}, \mathbb{P})$ wieder ein generischer Wahrscheinlichkeitsraum.

11.15. Definition: Sei $X : \Omega \to \mathbb{C}$ eine Zufallsvariable und schreibe $X = \Re(X) + i\Im(X)$ ($\Re(z)$ ist der Realteil und $\Im(z)$ der Imaginärteil von $z \in \mathbb{C}$). Sind $\Re(X)$ und $\Im(X)$ beide integrierbar bzgl. $\mathbb{P}$, dann nennt man X integrierbar und setzt

$$\int X \, d\mathbb{P} := \int \Re(X) \, d\mathbb{P} + i \int \Im(X) \, d\mathbb{P}$$

Bemerkung: Es gilt $X \in L^1 \iff |X| \in L^1$, denn

$$\max(\Re(X), \Im(X)) \leqslant |(\Re(X))^2 + (\Im(X))^2|^{1/2} = |X| \leqslant |\Re(X)| + |\Im(X)|$$

11.16. Lemma: Das oben definierte Lebesgue-Integral ist linear und für $X \in L^1$ gilt die Dreiecksungleichung

$$|\mathbb{E}X| \leqslant \mathbb{E}|X|$$

Beweis: Schreibe $f = \Re(X), g = \Im(X)$ und $z = a + ib$. Dann gilt

$$\begin{aligned}
\int zX \; d\mathbb{P} &= \int (a+ib)(f+ig) \; d\mathbb{P} \\
&= a\int f \; d\mathbb{P} + ia\int g \; d\mathbb{P} + ib\int f \; d\mathbb{P} - b\int g \; d\mathbb{P} \\
&= (a+ib)\left(\int f \; d\mathbb{P} + i\int g \; d\mathbb{P}\right) = z\int X \; d\mathbb{P}.
\end{aligned}$$

Die Identität $\int X + Y \; d\mathbb{P} = \int X \; d\mathbb{P} + \int Y \; d\mathbb{P}$ folgt trivial aus der Linearität des reellen Integrals. Falls $\mathbb{E}X = 0$, gilt die Ungleichung trivial. Sei also $\mathbb{E}X \neq 0$ und setze $w := |\mathbb{E}X|^{-1}\overline{\mathbb{E}X} \in \mathbb{C}$, sodass $|w| = 1$ Dann gilt

$$\mathbb{E}[wX] = w \cdot \mathbb{E}X = \frac{\mathbb{E}X\overline{\mathbb{E}X}}{|\mathbb{E}X|} = |\mathbb{E}X| \in \mathbb{R}$$

und somit $wX \in \mathbb{R}$ und $wX \leqslant |wX|$. Mit Cauchy–Schwarz folgt nun

$$|\mathbb{E}X| = \mathbb{E}[wX] \leqslant \mathbb{E}|wX| = |w| \cdot \mathbb{E}|X| = \mathbb{E}|X|.$$

□

Bemerkung: Es gilt

$$\begin{aligned}
e^{ix} &= \cos(x) + i\sin(x) \\
\frac{\partial}{\partial x}e^{ix} &= ie^{ix}
\end{aligned}$$

11.17. Definition: Die charakteristische Funktion (cf) einer Zufallsvariable $X : \Omega \to \mathbb{R}$ ist definiert als

$$\varphi_X(t) := \mathbb{E}\left[e^{itX}\right], \forall t \in \mathbb{R}$$

$\varphi : \mathbb{R} \to \mathbb{C}$ ist mit DOMK immer wohldefiniert, da $\left|e^{itX}\right| \leqslant 1$.

11.18. Proposition: Seien X, Y reellwertige, unabhängige Zufallsvariablen. Dann gilt

$$\varphi_{X+Y} = \varphi_X(t)\varphi_Y(t)$$

Beweis:

$$\begin{aligned}
\varphi_{X+Y}(t) &= \mathbb{E}\left[e^{itX}e^{itY}\right] \\
&= \mathbb{E}\left[(\cos(tX) + i\sin(tX))(\cos(tY) + i\sin(tY))\right] \\
&= \mathbb{E}\left[\cos(tX)\cos(tY) + i\cos(tX)\sin(tY) + i\sin(tX)\cos(tY) - \sin(tX)\sin(tY)\right] \\
&\overset{\in\mathbb{R}}{=} \mathbb{E}\left[\cos(tX) + i\sin(tX)\right]\mathbb{E}\left[\cos(tY) + i\sin(tY)\right] = \varphi_X(t)\varphi_Y(t)
\end{aligned}$$

□

11.19. Lemma: Die momenterzeugende Funktion (mgf) einer reellwertigen Zufallsvariable X

$$M_X(t) := \mathbb{E}\left[e^{tX}\right]$$

sei wohldefiniert in einer offenen Umgebung von 0, i.e.

$$\exists t_0 > 0 : M_X(t) \text{ exisitert für } t \in (-t_0, t_0)$$

Dann folgt

$$\forall k \geqslant 0 : X \in L^k \text{ und } \mathbb{E}X^k = \frac{\partial^k}{\partial t^k} M_X(t)\bigg|_{t=0}$$

Beweis: Sei $M_X(t)$ wohldefiniert für $|t| < t_0$ mit $t_0 > 0$. Dann folgt mit

$$0 \leqslant e^{t|X|} \leqslant e^{tX} + e^{-tX}$$

dass

$$\forall |t| < t_0 : e^{t|X|} \in L^1$$

Sei also $t < t_0$. Per Definition ist

$$e^{tX} = \sum_{k \geqslant 0} \frac{(tX)^k}{k!}$$

wobei die Partialsummen im Betrag durch $e^{|tX|} \in L^1$ beschränkt sind. Mit DOMK gilt damit

$$M_X(t) = E\left[e^{tX}\right] = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{t^k \mathbb{E}X^k}{k!}$$

wobei die Reihe absolut konvergiert, da

$$\sum_{k \geqslant 0} \frac{|t^k \mathbb{E}X^k|}{k!} \leqslant \sum_{k \geqslant 0} \frac{|t^k| \mathbb{E}|X|^k}{k!} \overset{\text{MONK}}{=} \mathbb{E}\left[\sum_{k \geqslant 0} \frac{|t^k| X^k}{k!}\right] = \mathbb{E}\left[e^{|tX|}\right] < \infty$$

Differenzieren liefert dann

$$\frac{\partial^\ell}{\partial t^\ell} M_X(t) = \sum_{k \geqslant \ell} \frac{k(k-1)\dots(k-\ell+1)}{k!} t^{k-\ell} \mathbb{E}X^k$$

und mit $t = 0$ das gewünschte Ergebnis. □

11.20. Satz (Inversions- und Eindeutigkeitssatz): Sei φ die cf und F die cdf einer reelwertigen Zufallsvariable X. Dann gilt

$$F(t) = \lim_{n\to\infty} \frac{1}{2\pi} \int_{-\infty}^{t} \int_{-\infty}^{+\infty} \exp\left(-\frac{w^2}{2n^2} + iwv\right) \varphi(-w)\, dw\, dv$$

für jedes t mit $\mathbb{P}(x = t) = 0$. Wegen der rechtsstetigkeit von F ist $F(t)$ damit für jedes $t \in \mathbb{R}$ eindeutig durch φ definiert.

Bermerkung: In der Literatur wird oft

$$\mathbb{P}(a \leqslant X < b) = \lim_{T\to\infty} \frac{1}{2\pi} \int\limits_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t))\ dt$$

angegeben für $\mathbb{P}(X = a) = \mathbb{P}(x = b) = 0$.

Beweis: Die Idee ist Folgende:
Sei $Z \sim \mathcal{N}(0,1)$ u.a. von X und setze $X_n := X + Z/n$. Dann gilt $X_n \xrightarrow[n\to\infty]{a.s.} X$ und als Konsequenz auch $F_n \xrightarrow[n\to\infty]{d} F$. Für t mit $\mathbb{P}(X = t) = 0$, folgt dann $F_n(t) \xrightarrow[n\to\infty]{} F(t)$. Zeige nun:

I. X_n hat eine Dichte f_n bzgl. dem Lebesgue-Maß $\lambda = \mathrm{vol}$

II. Diese Dichte ist von der Form

$$f_n(v) = \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} \exp\left(-\frac{w^2}{2n^2} + iwv\right) \varphi(-w)\ dw$$

I. Z/n hat eine Riemann-integrierbare Dichte $\phi_{0,1/n^2} =: \phi$ bzgl. λ und somit folgt

$$\begin{aligned}
F_n(t) &= \mathbb{P}(X + Z/n \leqslant t) \\
&\overset{\text{Tonelli, u.a.}}{=} \int\limits_{(X)} \int\limits_{(Z)} \mathbb{1}_{(-\infty,t]}(x+v)\ d\mathbb{P}_{Z/n}(v)\ d\mathbb{P}_X(x) \\
&= \int\limits_{(X)} \int\limits_{-\infty}^{t-x} \phi(v)\ dv\ d\mathbb{P}_X(x) \\
&\overset{w:=v+x}{=} \int\limits_{X} \int\limits_{-\infty}^{t} \phi(w-x)\ dw\ d\mathbb{P}_X(x) \\
&= \int\limits_{(X)} \int\limits_{(W)} \mathbb{1}_{(-\infty,t]}(w)\phi(w-x)\ d\lambda(w)\ d\mathbb{P}_X(x) \\
&\overset{\text{Tonelli}}{=} \int\limits_{(W)} \int\limits_{(X)} \mathbb{1}_{(-\infty,t]}(w)\phi(w-x)\ d\mathbb{P}_X(x)\ d\lambda(w) \\
&= \int\limits_{(W)} \mathbb{1}_{(-\infty,t]}(w) \left[\int\limits_{(X)} \phi(w-x)\ d\mathbb{P}_X(x) \right] d\lambda(w) \\
&= \int\limits_{-\infty}^{t} \mathbb{E}_X\left[\phi(w-X)\right] d\lambda(w)
\end{aligned}$$

Damit ist $f_n(v) := \mathbb{E}_X\left[\phi(v-X)\right]$ eine Dichte von X_n.

II. Es gilt

$$\phi(t) = \frac{n}{\sqrt{2\pi}} \exp\left(-\frac{t^2 n^2}{2}\right) = \frac{n}{\sqrt{2\pi}} \varphi_Z(nt)$$

und damit

$$\begin{aligned}
f_n(v) &= \mathbb{E}_X\left[\phi(v - X)\right] \\
&= \mathbb{E}_X\left[\frac{n}{\sqrt{2\pi}} \varphi_Z[n(v - X)]\right] \\
&= \mathbb{E}_X\left[\frac{n}{\sqrt{2\pi}} \mathbb{E}_Z\left[e^{invZ - inXZ}\right]\right] \\
&\overset{\text{Fubini}}{=} \frac{n}{\sqrt{2\pi}} \mathbb{E}_Z\left[\mathbb{E}_X\left[e^{invZ} e^{-inXZ}\right]\right] \\
&= \frac{n}{\sqrt{2\pi}} \mathbb{E}_Z\left[e^{invZ} \mathbb{E}_X\left[e^{-inXZ}\right]\right] \\
&= \frac{n}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} \frac{1}{n\sqrt{2\pi}} \exp\left(-\frac{w^2}{2n^2}\right) \exp\left(iwv\right) \mathbb{E}_X\left[e^{-iwX}\right]\, dw \\
&= \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} \exp\left(-\frac{w^2}{2n^2} + iwv\right) \varphi(-w)\, dw
\end{aligned}$$

wobei der vorletzte Schritt mit der Erwartung von $nZ \sim \mathcal{N}(0, n^2)$ folgt. □

11.21. Satz (Lévy Continuity Theorem/Stetigkeitssatz): Betrachte eine Folge von reellwertigen Zufallsvariablen $X_n, n \geqslant 1$ mit entsprechenden cdfs $F_n, n \geqslant 1$ und cfs $\varphi_n, n \geqslant 1$. Wenn

(i) $\forall t \in \mathbb{R} : \exists \lim_{n\to\infty} \varphi_n(t) =: \gamma(t) \in \mathbb{C}$

(ii) $\gamma(t)$ stetig im Punkt $t = 0$

dann folgt, dass γ die cf einer cdf F ist, und

$$F_n \xrightarrow[n\to\infty]{d} F$$

Beweis: Zeige zuerst, dass $F_n, n \geqslant 1$ straff sind. Beachte

$$\int_{-u}^{u} (1 - e^{itx})\, dt = 2u - \int_{-u}^{u} \cos(tx) + i\sin(tx)\, dt = 2u - \frac{2\sin(ux)}{x},$$

sodas mit Erwartungswert auf beiden Seiten und dem Satz von Fubini folgt

$$\frac{1}{u} \int_{-u}^{u} (1 - \varphi_n(t))\, dt = 2 \int 1 - \frac{\sin(uX_n)}{uX_n}\, d\mathbb{P}.$$

Mit der Ungleichung $|\sin(x)| \leqslant |x|$ für alle $x \in \mathbb{R}$ folgt, dass der Integrand im rechten Integral nicht-negativ ist. Für $|x| \in [2/u, u]$ gilt $|\sin(ux)| \leqslant 1$ und damit

$$\frac{1}{u} \int_{-u}^{u} (1 - \varphi_n(t))\, dt \geqslant 2 \int \left(1 - \frac{1}{|ux|}\right) \mathbb{1}_{\{|X_n| \geqslant 2/u\}}\, d\mathbb{P} \geqslant \mathbb{P}(|X_n| > 2/u)$$

Da γ stetig in $t = 0$ ist, gilt

$$\frac{1}{u}\int_{-u}^{u}(1-\gamma(t))\;dt \xrightarrow[u\to 0]{} 0$$

Sei nun $\varepsilon > 0$ und u_ε, sodass der obige Ausdruck kleiner als ε ist. Da $\varphi_n \xrightarrow[n\to\infty]{} \gamma$ folgt mit dem DOMK (für komplexwertige Integrale, da $|\varphi_n| \leqslant 1$), dass für alle $u \geqslant 0$

$$\frac{1}{u}\int_{-u}^{u}(1-\varphi_n(t))\;dt \xrightarrow[n\to\infty]{} \frac{1}{u}\int_{-u}^{u}(1-\gamma(t))\;dt.$$

Wähle nun $N \geqslant 1$ so, dass

$$\left|\frac{1}{u}\int_{-u}^{u}(1-\varphi_n(t))\;dt - \frac{1}{u}\int_{-u}^{u}(1-\gamma(t))\;dt\right| < \varepsilon$$

für alle $n \geqslant N$. Damit folgt

$$\mathbb{P}(|X_n| > 2/u_\varepsilon) \leqslant \frac{1}{u_\varepsilon}\int_{-u_\varepsilon}^{u_\varepsilon}(1-\varphi_n(t))\;dt < 2\varepsilon$$

für alle $n \geqslant N$. Außerdem gilt

$$\sup_{n\geqslant 1}\mathbb{P}(|X_n| > M) = \max\left\{\sup_{n\geqslant N}\mathbb{P}(|X_n| > M), \max_{n<N}\mathbb{P}(|X_n| > M)\right\}$$

und damit folgt, dass $F_n, n \geqslant 1$ straff sind. Mit Satz 11.14 hat jede Teilfolge $F_{n_k}, k \geqslant 1$ eine weitere Teilfolge $F_{n_{k_j}}, j \geqslant 1$, sodass $F_{n_{k_j}} \xrightarrow[j\to\infty]{d} F$ für eine cdf F einer Zufallsvariable X. Da die Funktionen $\sin, \cos$ auf $\mathbb{R}$ stetig und beschränkt sind folgt mit Satz 11.5

$$\forall t \in \mathbb{R} : \varphi_{n_{k_j}}(t) \xrightarrow[j\to\infty]{} \varphi_X(t)$$

für φ_X die cf von X. Wegen Annahme (i) und weil der Grenzwert von $\varphi_{n_{k_j}}(t)$ derselbe wie von $\varphi_n(t)$ sein muss (cf. Analysis: wenn eine Folge konvergiert, dann konvergiert jede Teilfolge zum selben Grenzwert), folgt damit

$$\forall t \in \mathbb{R} : \varphi_n(t) \xrightarrow[n\to\infty]{} \varphi_X(t)$$

Zeige nun $F_n \xrightarrow[n\to\infty]{d} F$. Angenommen nicht:

$$\Longrightarrow \exists \text{Teilfolge}(F_{n_k})_{k\geqslant 1} : F_{n_k} \cancel{\xrightarrow[k\to\infty]{d}} F$$

Aber $F_{n_k}, k \geqslant 1$ ist straff, da $F_n, n \geqslant 1$ straff ist und enthält mit Helly (Satz 11.11) und Prokorov (Satz 11.14) daher eine weitere Teilfolge $F_{n_{k_\ell}}, \ell \geqslant 1$ mit

$$F_{n_{k_\ell}} \xrightarrow[\ell\to\infty]{d} G$$

für eine cdf G. Damit gilt

$$\forall t \in \mathbb{R} : \varphi_{n_{k_\ell}}(t) \xrightarrow[\ell\to\infty]{} \varphi_G(t)$$

und damit (siehe oben)

$$\forall t \in \mathbb{R}\varphi_n(t) \xrightarrow[n\to\infty]{} \varphi_G(t)$$

Wegen (ii) gilt aber $\varphi_X(t) = \varphi_G(t)$ und mit dem Inversion Theorem (Satz 11.20) folgt

$$G \stackrel{d}{=} F$$

ein Widerspruch. $\square$

11.22. Satz Betrachte eine Folge von Zufallsvariablen $X_n, n \geqslant 1$ mit cdfs $F_n, n \geqslant 1$ und cfs $\varphi_n, n \geqslant 1$. Es gilt

$$X_n \xrightarrow[n\to\infty]{d} X \iff \varphi_n \xrightarrow[n\to\infty]{} \varphi$$

Beweis: Die Implikation $\implies$ folgt trivial mit PMT1, da $\cos, \sin$ stetig und beschränkt auf $\mathbb{R}$ sind. Zeige also die Implikation $\impliedby$
Dazu genügt es zu zeigen, dass $\varphi(\cdot)$ stetig im Punkt $t = 0$ ist (dann folgt mit dem Lévy Continuity Theorem, dass $F_n \xrightarrow[n\to\infty]{d} F$). Es gilt

$$\left|e^{itX}\right| \leqslant 1 \text{ und } e^{itX} \xrightarrow[t\to 0]{a.s.} 1 \text{ (sogar punktweise)}$$

Mit DOMK folgt daher

$$\varphi(t) = \mathbb{E}\left[e^{itX}\right] \xrightarrow[t\to 0]{} 1 = \varphi(0)$$

und damit die Aussage. □

Zentrale Grenzwertsätze

11.23. Lemma: Seien $z_i, w_i \in \mathbb{C}$ mit $|z_i|, |w_i| \leqslant 1, i = 1, \ldots, n$. dann gilt folgende Ungleichung

$$|z_1 \cdot \ldots \cdot z_n - w_1 \cdot \ldots \cdot w_n| \leqslant \sum_{i=1}^{n} |z_i - w_i|$$

Beweis: Induktiv nach n.

1. Base ($P(1)$): $|z_1 - w_1| \leqslant |z_1 - w_1|$
2. Induktionsschritt ($P(n) \implies P(n+1)$):

$$\begin{aligned}
&|z_1 \cdot \ldots \cdot z_n \cdot z_{n+1} - w_1 \cdot \ldots \cdot w_n \cdot w_{n+1}| \\
&= |z_1 \cdot \ldots \cdot z_n \cdot z_{n+1} - z_1 \cdot \ldots \cdot z_n w_{n+1} + z_1 \cdot \ldots \cdot z_n w_{n+1} - w_1 \cdot \ldots \cdot w_n \cdot w_{n+1}| \\
&= |z_1 \cdot \ldots \cdot z_n \cdot (z_{n+1} - w_{n+1}) - w_{n+1} \cdot (z_1 \cdot \ldots \cdot z_n - w_1 \cdot \ldots \cdot w_n)| \\
&\leqslant |z_1 \cdot \ldots \cdot z_n| \cdot |z_{n+1} - w_{n+1}| + |w_{n+1}||z_1 \cdot \ldots \cdot z_n - w_1 \cdot \ldots \cdot w_n| \\
&\leqslant \sum_{i=1}^{n+1} |z_i - w_i|
\end{aligned}$$

□

11.24. Lemma: Für $x \in \mathbb{R}$ und $n \in \mathbb{N}_0$ ist

$$e^{ix} = \sum_{k=0}^{n} \frac{(ix)^k}{k!} + R_n(x)$$

mit Rest

$$R_n(x) = \frac{i^{n+1}}{n!} \int_0^x (x-s)^n e^{is} \, ds$$

wobei insbesondere gilt

$$R_n(x) \leqslant \min\left(\frac{|x|^{n+1}}{(n+1)!}, \frac{2|x|^n}{n!}\right)$$

(wobei uns später im Beweis des Lindeberg-Lévy CLT der Fall $n = 2$ interessiert).

Beweis: Induktiv nach n.

1. Base (P(0)):

$$\begin{aligned}
\int_0^x e^{is}\ ds &= i^{-1}e^{is}\Big|_{s=0}^{s=x} \\
&= -ie^{is}\Big|_{s=0}^{s=x} \\
&= i - ie^{ix}
\end{aligned}$$

$$\implies ie^{ix} = i - \int_0^x e^{is}$$

$$\implies e^{ix} = 1 + i\int_0^x e^{is}\ ds = \sum_{k=0}^{0}\frac{(ix)^k}{k!} + \frac{i^{0+1}}{0!}\int_0^x (x-s)^0 e^{is}\ ds$$

2. Induktionsschritt ($P(n) \implies P(n+1)$): Laut Voraussetzung gilt

$$e^{ix} = \sum_{k=0}^{n}\frac{(ix)^k}{k!} + R_n(X)$$

wobei

$$R_n(x) = \frac{i^{n+1}}{n!}\int_0^x (x-s)^n e^{is}\ ds$$

Zeige also

$$R_n(x) = \frac{(ix)^{n+1}}{(n+1)!} + R_{n+1}(x)$$

Mit partieller Integration ($f(s) = e^{is}, g'(s) = -\frac{(x-s)^{n+1}}{n+1}$) gilt

$$\begin{aligned}
R_n(x) &= \frac{i^{n+1}}{n!} \int_0^x (x-s)^n e^{is}\ ds \\
&= \frac{i^{n+1}}{n!} \left(\left[-e^{is} \frac{(x-s)^{n+1}}{n+1} \right]_{s=0}^{s=x} + \int_0^x ie^{is} \frac{(x-s)^{n+1}}{n+1}\ ds \right) \\
&= \frac{i^{n+1}}{n!} \left(\frac{x^{n+1}}{n+1} + \frac{i}{n+1} \int_0^x (x-s)^{n+1} e^{is}\ ds \right) \\
&= \frac{(ix)^{n+1}}{(n+1)!} + \frac{i^{n+2}}{(n+1)!} \int_0^x (x-s)^{n+1} e^{is}\ ds \\
&= \frac{(ix)^{n+1}}{(n+1)!} + R_{n+1}(x)
\end{aligned}$$

Zur ersten Ungleichung:
Für $x \geqslant 0$ und mit $|e^{ix}| \leqslant 1, |i^\ell| = 1, \forall \ell \geqslant 1$ ist

$$\begin{aligned}
|R_n(x)| &\leqslant \frac{1}{n!} \int_0^x |x-s|^n\ ds \\
&= \frac{1}{n!} \int_0^x (x-s)^n\ ds \\
&= \frac{x^{n+1}}{(n+1)!} \overset{x \geqq 0}{=} \frac{|x|^{n+1}}{(n+1)!}
\end{aligned}$$

Für $x < 0$ ist

$$\begin{aligned}
|R_n(x)| &\leqslant \frac{1}{n!} \left| \int_0^x (x-s)^n e^{is}\ ds \right| \\
&\frac{1}{n!} \left| \int_x^0 (x-s)^n e^{is}\ ds \right| \\
&= \frac{1}{n!} \int_x^0 (s-x)^n\ ds = \frac{(-x)^{n+1}}{(n+1)!} \overset{x \leq 0}{=} \frac{|x|^{n+1}}{(n+1)!}
\end{aligned}$$

womit die erste Ungleichung für alle $x \in \mathbb{R}$ gilt.

Zur zweiten Ungleichung:

$$\begin{aligned}
R_n(x) &= \frac{i^{n+1}}{n!}\int\limits_0^x (x-s)^n e^{is}\ ds \\
&= \frac{i^{n+1}}{n!}\left(\left[-ie^{is}(x-s)^n\right]_{s=0}^{s=x} - in\int\limits_0^x (x-s)^{n-1}e^{is}\ ds\right) \\
&= \frac{i^{n+2}x^n}{n!} - \frac{i^{n+2}}{(n-1)!}\int\limits_0^x (x-s)e^{is}\ ds
\end{aligned}$$

wobei der erste Schritt wieder mit partieller Integration für $f(s) = (x-s)^n$ und $g'(s) = e^{is}$ folgt. Ähnlich wie für die erste Ungleichung folgt damit

$$|R_n(x)| \leqslant \frac{|x|^n}{n!} + \frac{1}{(n+1)!}\left|\int\limits_0^x (x-s)^{n-1}e^{is}\right| \leqslant \frac{2|x|^n}{n!}$$

$\square$

11.25. Satz (Lindeberg–Lévy CLT): Es seien $X_i \overset{\text{i.i.d.}}{\sim} X, i \geqslant 1$ quadratisch integrierbare Zufallsvariablen mit $\mathbb{E}X = \mu, \operatorname{Var}(X) = \sigma^2 \in (0,\infty)$. Für $\overline{X}_n := \sum_{i=1}^n X_i$ gilt dann

$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1)$$

Beweis: Seien o.B.d.A. $\mu = 0, \sigma = 1$. Sei im Folgenden ψ die cf von X und φ_n die cf von $\sqrt{n}\overline{X}_n$. Zu zeigen ist dann

$$\varphi_n(t) \xrightarrow[n\to\infty]{} e^{-t^2/2}$$

Mit Lemma 11.23 gilt $e^{itX} = 1 + itX - \frac{(tX)^2}{2} + R_2(tX)$ wobei $|R_2(tX)| \leqslant \min\left(\frac{|tX|^3}{6}, t^2X^2\right)$. Damit gilt also

$$\psi(t) = 1 + \mathbb{E}[itX] - \frac{t^2\mathbb{E}X^2}{2} + \mathbb{E}R_2(tX)$$

Sei $t > 0$. Mit der zweiten Ungleichung aus Lemma 11.23 gilt $\left|\frac{R_2(tX)}{t^2}\right| \leqslant X^2 \in L^1$ und mit der ersten Ungleichung gilt $\left|\frac{R_2(tX)}{t^2}\right| \xrightarrow[t\searrow 0]{a.s.} 0$. Mit DOMK folgt schließlich

$$\mathbb{E}\left|\frac{R_2(tX)}{t^2}\right| \xrightarrow[t\searrow 0]{} 0$$

Nun gilt

$$\varphi_n(t) = \mathbb{E}\left[e^{it\sqrt{n}\overline{X}_n}\right] \overset{\text{i.i.d.}}{=} \left[\psi\left(\frac{t}{\sqrt{n}}\right)\right]^n = \left(1 - \frac{t^2}{2n} + \mathbb{E}[R_2(tX)]\right)^n$$

Für $t = 0$ gilt also $\varphi_n(0) = 1 = e^{-0^2/2}$.

Für $t \neq 0$ ist

$$\begin{aligned}
\left|\varphi_n(t) - e^{-t^2/2}\right| &= \left|\varphi_n(t) - \left(1-\frac{t^2}{2n}\right)^n + \left(1-\frac{t^2}{2n}\right)^n - e^{-t^2/2}\right| \\
&\leqslant \left|\varphi_n(t) - \left(1-\frac{t^2}{2n}\right)^n\right| + \left|\left(1-\frac{t^2}{2n}\right)^n - e^{-t^2/2}\right| \\
&= \left|\left[\psi\left(\frac{t}{\sqrt{n}}\right)\right]^n - \left(1-\frac{t^2}{2n}\right)^n\right| + \mathcal{O}(1) \\
&\overset{11.22}{\leqslant} \sum_{i=1}^{n} \left|\psi\left(\frac{t}{\sqrt{n}}\right) - \left(1-\frac{t^2}{2n}\right)\right| \\
&= n\left|\psi\left(\frac{t}{\sqrt{n}}\right) - \left(1-\frac{t^2}{2n}\right)\right| \\
&= n\left|\mathbb{E}\left[R_2\left(\frac{tX}{\sqrt{n}}\right)\right]\right| \\
&= \frac{nt^2}{n}\left|\frac{\mathbb{E}\left[R_2\left(\frac{tX}{\sqrt{n}}\right)\right]}{t^2/n}\right| \xrightarrow[n\to\infty]{} 0
\end{aligned}$$

□

Lindeberg- und Ljapunov-Bedingung

Betrachte nun allgemeiner ein trianguläres Feld von quadratisch integrierbaren Zufallsvariablen $X_{n,i}, i = 1, \dots, r_n, n \geqslant 1$, sodass für $n \geqslant 1$ die $X_{n,1}, \dots, X_{n,r_n}$ unabhängig sind. Definiere weiters

$$\sigma_{n_i}^2 := \mathrm{Var}(X_{n,i}) \text{ und } s_n^2 := \sum_{i=1}^{r_n} \sigma_{n,i}^2$$

Eine Folge von quadratisch integrierbaren, unabhängigen Zufallsvariablen $Y_k, k \geqslant 1$ erfüllt diese Eigenschaften natürlich, da wir dann $X_{n,i} \sim Y_i$ für alle $n \geqslant 1$ und $i = 1, \dots, r_n$ haben.

11.26. Definition (Lindeberg-Bedingung): Wir sagen, dass die Lindeberg-Bedingung für $X_{n,i}, i = 1, \dots, r_n, n \geqslant 1$ gilt, falls

$$\forall \varepsilon > 0 : \frac{1}{s_n^2} \sum_{i=1}^{r_n} \mathbb{E}\left[(X_{n,i} - \mathbb{E}X_{n,i})^2 \mathbb{1}_{\{|X_{n,i} - \mathbb{E}X_{n,i}| > \varepsilon \cdot s_n\}}\right] \xrightarrow[n\to\infty]{} 0$$

11.27. Satz (Lindeberg CLT): Unter der Lindeberg-Bedingung gilt für $\overline{X}_n = n^{-1}\sum_{i=1}^{r_n} X_{n,i}$, dass

$$\sqrt{n}\frac{\overline{X}_n - \mathbb{E}\overline{X}_n}{s_n} = \frac{\overline{X}_n - \mathbb{E}\overline{X}_n}{\sqrt{\mathrm{Var}(\overline{X}_n)}} \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1)$$

Beweis: siehe z.B. P. Billingsley, *Probability and Measure* (3rd Ed.), p.360. □

11.28. Definition (Ljapunov-Bedingung): Wir sagen, dass die Ljapunov-Bedingung für $X_{n,i}, i = 1, \ldots, r_n, n \geqslant 1$ gilt, falls

$$\exists \delta > 0 : \frac{1}{s_n^{2+\delta}} \sum_{i=1}^{r_n} \mathbb{E}\,|X_{n,i} - \mathbb{E}X_{n,i}|^{2+\delta} \xrightarrow[n\to\infty]{} 0$$

11.29. Lemma: Die Ljapunov-Bedingung impliziert die Lindeberg-Bedingung.

Beweis: Sei o.B.d.A. $\mathbb{E}X_{n,i} = 0$ für alle $i = 1, \ldots, r_n, n \geqslant 1$. Sei $\delta > 0$ wie in der Ljapunov-Bedingung, die laut Annahme gilt. Für $\varepsilon > 0$ gilt

$$\begin{aligned}
\frac{1}{s_n^2} \sum_{i=1}^{r_n} \mathbb{E}\left[X_{n,i}^2 \mathbb{1}_{\{|X_{n,i}| > \varepsilon \cdot s_n\}}\right] &\overset{1 \leqslant \frac{|X_{n,i}|^\delta}{(\varepsilon \cdot s_n)^\delta}}{\leqslant} \frac{1}{s_n^2} \sum_{i=1}^{r_n} \mathbb{E}\left[X_{n,i}^2 \frac{|X_{n,i}|^\delta}{(\varepsilon \cdot s_n)^\delta} \mathbb{1}_{\{|X_{n,i}| > \varepsilon \cdot s_n\}}\right] \\
&\leqslant \frac{1}{\varepsilon^\delta} \frac{1}{s_n^{2+\delta}} \sum_{i=1}^{r_n} \mathbb{E}|X_{n_i}|^{2+\delta} \xrightarrow[n\to\infty]{} 0
\end{aligned}$$

□

12. Radon-Nikodym-Ableitungen

In diesem Kapitel sei $(\Omega, \mathcal{A})$ ein messbarer Raum und ν, μ generische Maße auf diesem Raum.

Signierte Maße

12.1. Definition: Ein signiertes Maß ist eine Abbildung $\varphi : \mathcal{A} \to \mathbb{R}$, die σ-additiv ist, i.e. für $A_n \in \mathcal{A}, n \geqslant 1$ disjunkt gilt

$$\varphi\left(\bigcup_{n\geqslant 1} A_n\right) = \sum_{n\geqslant 1} \varphi(A_n)$$

In unserem Kontext seien signierte Maße immer endlich und bilden $\emptyset$ auf 0 ab.

12.2. Beispiel: Einige Beispiele für signierte Maße sind

(i) $\phi(A) := \mu(A)$, wenn μ endlich ist.

(ii) $\varphi(A) := \mu(A) - \nu(A)$, wenn μ, ν endlich sind.

(iii) $\varphi(A) := \int_A f \, d\mu$, wenn $f \in L^1(\mu)$

12.3. Lemma: Seien $A_n \in \mathcal{A}, n \geqslant 1$, sodass entweder

(i) $A_1 \subseteq A_2 \subseteq \ldots \subseteq \bigcup_{n\geqslant 1} A_n := A$, oder

(ii) $A_1 \supseteq A_2 \supseteq \ldots \supseteq \bigcap_{n\geqslant 1} A_n := A$

Dann ist in beiden Fällen

$$\lim_{n\to\infty} \varphi(A_n) = \varphi(A)$$

i.e. signierte Maße sind von unten/oben stetig.

Beweis:

(i)

$$\begin{aligned}
\varphi(A) &= \varphi(A_1 \cup \bigcup_{n\geqslant 1} (A_{n+1} \setminus A_n)) \\
&= \varphi(A_1) + \sum_{n\geqslant 1} \varphi(A_{n+1} \setminus A_n) \\
&= \lim_{N\to\infty} \left[\varphi(A_1) + \sum_{n=1}^{N} \varphi(A_{n+1} \setminus A_n)\right] \\
&= \lim_{N\to\infty} \varphi(A_{N+1}) = \lim_{N\to\infty} \varphi(A_N)
\end{aligned}$$

(ii) Hier gilt $A_1^c \subseteq \ldots \subseteq \bigcup_{n\geqslant 1} A_n^c =: A^c$

$$\overset{\text{(i)}}{\Longrightarrow} \varphi(A^c) = \lim_{n\to\infty} \varphi(A_n^c)$$
$$\Longrightarrow \varphi(\Omega) - \varphi(A) = \lim_{n\to\infty} [\varphi(\Omega) - \varphi(A_n)] = \varphi(\Omega) - \lim_{n\to\infty} \varphi(A_n)$$
$$\Longrightarrow \varphi(A) = \lim_{n\to\infty} \varphi(A_n)$$

$\square$

4. Satz (Hahn/Jordan Decomposition Theorem): Sei φ ein signiertes Maß auf $(\Omega, \mathcal{A})$. Dann gibt es eine Partition (Hahn-Zerlegung) von Ω

$$\Omega = \Omega_+ \cup \Omega_-$$

mit $\Omega_+, \Omega_- \in \mathcal{A}$, sodass

$$\forall A \in \mathcal{A} : \varphi(A \cap \Omega_+) =: \varphi_+(A) \geqslant 0,\ \varphi(A \cap \Omega_-) =: \varphi_-(A) \leqslant 0$$

Daraus folgt sofort die Jordan-Zerlegung

$$\forall A \in \mathcal{A} : \varphi(A) = \varphi_+(A) + \varphi_-(A)$$

Beweis: Sei $\alpha := \sup\limits_{A\in\mathcal{A}} \varphi(A)$. Dann gibt es eine Menge $D \in \mathcal{A}$, sodass dieses Supremum angenommen wird, i.e. $\varphi(D) = \alpha$ (Details zur Konstruktion siehe unten). Setze nun $\Omega_+ := D$ und $\Omega_- := D^c$. Dann gilt

- Angenommen es gibt $A \in \mathcal{A}$ mit $\varphi(A \cap D) < 0$. Dann ist $D = (D \cap A) \cup (D \setminus A)$ und

  $$\alpha = \varphi(D) = \varphi(D \cap A) + \varphi(D \setminus A) < \varphi(D \setminus A)$$

  ein Widerspruch, da $D \setminus A \in \mathcal{A}$ und $\alpha = \sup\limits_{A\in\mathcal{A}} \varphi(A)$.

- Angenommen es gibt $A \in \mathcal{A}$ mit $\varphi(A \cap D^c) > 0$. Dann ist

  $$\varphi(D \cup (A \cap D^c)) = \varphi(D) + \varphi(A \cap D^c) > \varphi(D) = /\alpha$$

  ein Widerspruch, da $D \cup (A \cap D^c) \in \mathcal{A}$ und $\alpha = \sup\limits_{A\in\mathcal{A}} \varphi(A)$.

Zur Konstruktion von D:
Wähle $A_n, n \geqslant 1$, sodass $\lim\limits_{n\to\infty} \varphi(A_n) = \alpha$ und definiere für $n \geqslant 1$

$$\mathcal{B}_n := \left\{ B_I = \left(\bigcap_{I_i=1} A_i \right) \cup \left(\bigcap_{I_i=0} A_i \right) : I \in \{0,1\}^n \right\}$$

sukzessive feiner werdende Partitionen von Ω. Setze nun

$$C_n := \bigcup_{\substack{B_I \in \mathcal{B}_n \\ \varphi(B_I) > 0}} B_I$$

Nun ist

$$A_n = \bigcup_{\substack{B_I \in \mathcal{B}_n \\ I_n = 1}} B_I$$

sodass $\varphi(A_n) \leqslant \varphi(C_n)$ $(*)$. Außerdem ist

$$\varphi\left(\bigcup_{m=n}^{N} C_m\right) \leqslant \varphi\left(\bigcup_{m=n}^{N+1} C_m\right)$$

da $\left(\bigcup_{m=n}^{N+1} C_m\right) \setminus \left(\bigcup_{m=n}^{N} C_m\right)$ eine Vereinigung von Mengen B_I aus $\mathcal{B}_{N+1}$ mit $\varphi(B_I) > 0$ ist. $(**)$

Setze nun $D_i := \limsup\limits_{n\to\infty} C_n$ Dann folgt

$$\begin{aligned}
\alpha &= \lim_{n\to\infty} \varphi(A_n) \\
&\overset{(*)}{\leqslant} \liminf_{n\to\infty} \varphi(C_n) \\
&\overset{(**)}{\leqslant} \liminf_{n\to\infty} \limsup_{N\to\infty} \varphi\left(\bigcup_{m=n}^{N} C_m\right) \\
&\overset{\text{S.v.U.}}{=} \liminf_{n\to\infty} \varphi\left(\bigcup_{m=n}^{\infty} C_m\right) \\
&\overset{\text{S.v.O.}}{=} \varphi\left(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} C_m\right) \\
&= \varphi\left(\limsup_{n\to\infty} C_n\right) = \varphi(D) \leqslant \alpha
\end{aligned}$$

□

12.5. Definition:

(i) μ und ν sind gegenseitig singular (kurz $\mu \perp \nu$), wenn

$$\exists M \in \mathcal{A} : \mu(M) = 0, \nu(M^c) = 0$$

Die Relation $\perp$ ist symmetrisch, i.e. $\mu \perp \nu \iff \nu \perp \mu$.

(ii) ν ist absolut stetig bezüglich μ, (kurz $\nu \ll \mu$) wenn

$$\forall A \in \mathcal{A} : \mu(A) = 0 \implies \nu(A) = 0$$

Die Relation $\ll$ ist transitiv, i.e. $\nu \ll \mu, \mu \ll \lambda \implies \nu \ll \lambda$.

12.6. Lemma: Seien μ und ν zwei endliche Maße, die NICHT gegenseitig singular sind. Dann gibt es ein $\varepsilon > 0$ und $N \in \mathcal{A}$ mit $\mu(B) > 0$, sodass

$$\forall A \in \mathcal{A} : \varepsilon \cdot \mu(A \cap B) \leqslant \nu(A \cap B)$$

Beweis: Für $n \geqslant 1$ definiere ein signiertes Maß $\varphi_n := \nu - \mu/n$. Seien nun $\Omega_+^{(n)}$ und $\Omega_-^{(n)}$ die Hahn-Zerlegung von φ_n. Setze $M := \bigcup_{n\geqslant 1} \Omega_+^{(n)}$ und $M^c = \bigcap_{n\geqslant 1} \Omega_-^{(n)}$. Da $\varphi_1 \leqslant \ldots \leqslant \lim_{n\to\infty} \varphi_n = \nu$, gilt $\Omega_+^{(1)} \supseteq \ldots \supseteq M^c$ und mit der S.v.O. für signierte Maße $\forall n \geqslant 1 : \varphi_n(M^c) \leqslant 0$.
Damit folgtper Definition für alle $n \geqslant 1$

$$\nu(M^c) - \frac{1}{n}\mu(M^c) \leqslant 0 \iff \nu(M^c) \leqslant \frac{1}{n}\mu(M^c)$$

und damit $\nu(M^c) = 0$. Laut Annahme git aber $\nu \not\perp \mu$ und damit $\mu(M) > 0$. Also folgt

$$0 < \mu(M) \leqslant \sum_{n\geqslant 1} \mu(\Omega_+^{(n)})$$

und damit gibt es ein $m \geqslant 1$, sodass $\mu(\Omega_+^{(m)}) > 0$ $(*)$. Mit der Definition von φ_m gilt aber

$$\forall A \in \mathcal{A} : \nu(A \cap \Omega_+^{(m)}) - \frac{1}{m}\mu(A \cap \Omega_+^{(m)}) \geqslant 0 \iff \nu(A \cap \Omega_+^{(m)}) \geqslant \frac{1}{m}\mu(A \cap \Omega_+^{(m)})$$

Setze also $B := \Omega_+^{(m)}$ mit $\varepsilon := m^{-1}$. Dann gilt wegen $(*)$ auch $\mu(B) = \mu(\Omega_+^{(m)}) > 0$. □

12.7. Satz (Radon–Nikodym Theorem für σ-endliche Maße): Seien μ und ν zwei σ-endliche Maße, sodass $\nu \ll \mu$. Dann hat ν eine Dichte bezüglich μ, i.e. es gibt ein $f : \Omega \to \mathbb{R}$ messbar, sodass

$$\forall A \in \mathcal{A} : \nu(A) = \int_A f \; d\mu \tag{7}$$

Wir schreiben $f = \dfrac{d\nu}{d\mu}$. Falls insbesondere $g : \Omega \to \mathbb{R}$ eine weitere messbare Abbildung ist, die (7) erfüllt, dann gilt $f = g$ μ-a.e.

Beweis: Unterscheide hier die Fälle, wo μ, ν beide endlich sind, und den allgemeinen Fall (μ, ν beide σ-endlich).

1. Fall (μ, ν beide endlich): Betrachte die Menge

$$\mathcal{G} := \left\{ g \geqslant 0 : g \text{ messbar und } \forall A \in \mathcal{A} : \int_A g \; d\mu \leqslant \nu(A) \right\}$$

Dann gilt $0 \in \mathcal{G}$ und für $g, h \in \mathcal{G}$ ist $\max(g, h) \in \mathcal{G}$, denn für $A \in \mathcal{A}$ gilt

$$\int_A \max(g,h) \; d\mu = \int\limits_{A\cap\{g\geqslant h\}} g \; d\mu + \int\limits_{A\cap\{g<h\}} h \; d\mu \overset{g,h,\in\mathcal{G}}{\leqslant} \nu(A \cap \{g \geqslant h\}) + \nu(A \cap \{g < h\}) = \nu(A).$$

Setze nun $m := \sup_{g\in\mathcal{G}} \int g \; d\mu < \nu(A)$ und wähle eine Folge $g_n \in \mathcal{G}, n \geqslant 1$, sodass

$$\lim_{n\to\infty} \int g_n \; d\mu = m.$$

Setze $f_n := \max_{i\leqslant n} g_i$ und $f := \sup_{n\geqslant 1} g_n$, sodass

- $f_n \in \mathcal{G}$ für alle $n \geqslant 1$,

- $f_n(\omega) \xrightarrow[n\to\infty]{} f(\omega)$ für alle $\omega \in \Omega$, monoton nichtfallend in n,
- $m \geqslant \int f_n \; d\mu \geqslant \int g_n \; d\mu \xrightarrow[n\to\infty]{} m$,

und daher

$$\int f \; d\mu \overset{\text{MONK}}{=} \lim_{n\to\infty} \int f_n \; d\mu = m.$$

Setze

$$\nu_s(A) = \nu(A) - \int_A f \; d\mu, \;\; A \in \mathcal{A}.$$

Dann ist ν_s ein endliches Maß (einfache Überlegung) und $\nu_s \perp \mu$. Angenommen nicht: Mit Lemma 12.6 gibt es $\varepsilon > 0, b \in \mathcal{A}$ mit $\mu(B) > 0$, sodass $\forall A \in \mathcal{A} : \varepsilon \cdot \mu(A \cap B) \leqslant \nu_s(A \cap B)$. Damit folgt für $A \in \mathcal{A}$

$$\begin{aligned}\int_A (f + \varepsilon \cdot \mathbb{1}_B) \; d\mu &= \int_A f \; d\mu + \varepsilon \cdot \mu(A \cap B) \\ &\leqslant \int_A f \; d\mu + \nu_s(A \cap B) \\ &\leqslant \int_A f \; d\mu + \nu_s(A) = \nu(A)\end{aligned}$$

Damit folgt $f + \varepsilon \cdot \mathbb{1}_B \in \mathcal{G}$ und

$$\sup_{g\in\mathcal{G}} \int g \; d\mu \geqslant \int (f + \varepsilon \cdot \mathbb{1}_B) \; d\mu = \int f \; d\mu \varepsilon \cdot \mu(B) > \int f \; d\mu = \sup_{g\in\mathcal{G}} \int g \; d\mu$$

ein Widerspruch.

Es gilt per Definition

$$\exists M \in \mathcal{A} : \mu(M) = 0, \;\; \nu_s(M^c) = 0.$$

Da aber $\nu \ll \mu$ ist $\nu(M) = 0$ und per Konstruktion

$$0 = \nu(M) 0 \geqslant \nu_s(M) \geqslant 0$$

und damit

$$\nu_s(\Omega) = \nu_s(M) + \nu(M^c) = 0$$

Also ist ν_s das Nullmaß und $\nu(A) = \int_A f \; d\mu$, womit die erste Aussage folgt.
Sei nun g eine weitere Dichte von ν bzgl. μ (also eine messbare Abbildung, die (7) erfüllt). Setze $A := \{f > g\}$. Dann ist

$$\begin{aligned}0 &= \nu(A) - \nu(A) \\ &= \int_A f \; d\mu - \int_A g \; d\mu \\ &= \int_A f - g \; d\mu \\ &= \int \mathbb{1}_{\{f-g>0\}}(f - g) \; d\mu\end{aligned}$$

und da im Integral $f - g > 0$, gilt $\mathbb{1}_A \overset{a.e.}{=} 0$ und $\mu(A) = 0$ Ein ähnliches Argument folgt für $B := \{f < g\}$ und die Aussage folgt schließlich mit der σ-Subadditivität.

2. Fall: (μ, ν beide σ-endlich): Wähle $A_n \in \mathcal{A}, n \geqslant 1$ mit $A_1 \subseteq \ldots \subseteq \bigcup_{n\geqslant 1} A_n = \Omega$, sodass $\mu(A_n) < \infty$ für alle $n \geqslant 1$. Wähle $B_n \in \mathcal{A}$ mit denselben Eigenschaften für ν. Setze nun $C_n := A_n \cap B_n$ für $n \geqslant 1$, sodass $C_n \in \mathcal{A}$ und $C_1 \subseteq \ldots \subseteq \bigcup_{n\geqslant 1} C_n = \Omega$ und $\mu(C_n), \nu(C_n) < \infty$ für alle $n \geqslant 1$. Disjunktisiere nun C_n durch $M_1 := C_1$ und $M_{n+1} := C_{n+1} \setminus C_n$ für alle $n \geqslant 1$. Dann gilt $\Omega = \bigsqcup_{n\geqslant 1} M_n$ und $\mu(M_n), \nu(M_n) < \infty$ für alle $n \geqslant 1$. Definiere nun endlich Maße ν_n und μ_n wie folgt:

$$\nu_n(A) := \nu(A \cap M_n)$$
$$\mu_n(A) := \mu(A \cap M_n)$$

für alle $n \geqslant 1$ und $A \in \mathcal{A}$. Laut Voraussetzung ist $\nu \ll \mu$ und daher auch $\nu_n \ll \mu_n$ für alle $n \geqslant 1$. Mit dem 1. Fall hat also für jedes $n \geqslant 1$ ν_n eine Dichte f_n bzgl. μ_n. Setze nun

$$f := \sum_{n\geqslant 1} f_n \cdot \mathbb{1}_{M_n}$$

Dann ist $f \geqslant 0$ und messbar und für $A \in \mathcal{A}$ gilt

$$\begin{aligned}
\int_A f \, d\mu &= \int \sum_{n\geqslant 1} f_n \cdot \mathbb{1}_{M_n} \cdot \mathbb{1}_A \, d\mu \\
&\overset{\text{MONK}}{=} \sum_{n\geqslant 1} \int f_n \cdot \mathbb{1}_{M_n} \cdot \mathbb{1}_A \, d\mu \\
&= \sum_{n\geqslant 1} \int_A f_n \cdot \mathbb{1}_{M_n} \, d\mu \\
&= \sum_{n\geqslant 1} \int_A f_n \, d\mu_n \\
&= \sum_{n\geqslant 1} \nu_n(A) \\
&= \sum_{n\geqslant 1} \nu(A \cap M_n) = \nu(A)
\end{aligned}$$

Die Eindeutigkeit folgt wie im endlichen Fall. □

12.8. Satz (Radon–Nikodym Theorem für signierte Maße) folgt. Sei φ ein signiertes Maß und μ ein σ-endliches Maß, sodass $\varphi \ll \mu$.Dann hat φ eine Dichte bezüglich μ, i.e. es gibt ein $f : \Omega \to \mathbb{R}$ messbar, sodass

$$\forall A \in \mathcal{A} : \varphi(A) = \int_A f \, d\mu$$

Wir schreiben $f = \dfrac{d\varphi}{d\mu}$. Falls insbesondere $g : \Omega \to \mathbb{R}$ eine weitere messbare Abbildung ist, die (7) erfüllt, dann gilt $f = g$ μ-a.e.

Beweis: Betrachte die Hahn-Zerlegung $\Omega = \Omega_+ \cup \Omega_-$ und die Jordan-Zerlegung $\varphi = \varphi_+ - \varphi_-$. Dann sind φ_+ und φ_- beide endliche Maße und $\varphi_+ \ll \mu$, $\varphi_- \ll \mu$. Setze also

$$f_{(+)} := \frac{d\varphi_+}{d\mu}, \; f_{(-)} := \frac{d\varphi_-}{d\mu}, \; f := f_{(+)} - f_{(-)}$$

Die Messbarkeit von f folgt sofort aus der Messbarkeit von $f_{(+)}, f_{(-)}$ laut Satz 12.7. Weiters gilt $f_{(+)}, f_{(-)} \in L^1(\mu)$ und damit auch $f \in L^1(\mu)$. Dass f eine Dichte von φ bezüglich μ ist folgt mit

$$\varphi(A) = \varphi_+(A) - \varphi_-(A) = \int_A f_{(+)}\ d\mu - \int f_{(-)}\ d\mu = \int_A f\ d\mu$$

Ist g eine weitere Dichte von φ bezüglich μ, dann ist $g \cdot \mathbb{1}_{\Omega_+}$ eine Dichte von φ_+ bezüglich μ und mit Satz 12.7 gilt $g \cdot \mathbb{1}_{\Omega_+} \overset{a.e.}{=} f_{(+)}$. Dasselbe folgt für $g \cdot \mathbb{1}_{\Omega_-}$ und $f_{(-)}$. Da die Vereinigung zweier Nullmengen wieder eine Nullmenge ist folgt

$$g \overset{a.e.}{=} f$$

$\square$

13. Bedingte Erwartungswerte und Wahrscheinlichkeiten

13.1. Definition: Betrachte einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ mit einer sub-σ-Algebra $\mathcal{G} \subseteq \mathcal{A}$ und eine Zufallsvariable $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$, i.e. $\mathcal{A}$-messbar und integrierbar. Für $G \in \mathcal{G}$ sei $\nu(G) := \int_G X \, d\mathbb{P} = \mathbb{E}\left[\mathbb{1}_G X\right]$ und $\mathbb{P}_\mathcal{G}(G) := \mathbb{P}(G)$. Dann ist ν ein signiertes Maß auf $(\Omega, \mathcal{G})$ und $\mathbb{P}_\mathcal{G}$ ist ein Wahrscheinlichkeitsmaß (und damit endlich) auf $(\Omega, \mathcal{G})$, sodass $\nu \ll \mathbb{P}_\mathcal{G}$.
Der bedingte Erwartungswert von X bezüglich $\mathcal{G}$ ist dann definiert als

$$\mathbb{E}[X \parallel \mathcal{G}] := \frac{d\nu}{d\mathbb{P}_\mathcal{G}}$$

Für eine weitere Zufallsvariable Y auf $(\Omega, \mathcal{A}, \mathbb{P})$, setzt man

$$\mathbb{E}[X \parallel Y] := \mathbb{E}[X \parallel \sigma(Y)]$$

Für $A \in \mathcal{A}$ setzt man

$$\mathbb{P}(A \parallel \mathcal{G}) := \mathbb{E}[\mathbb{1}_A \parallel \mathcal{G}]$$

Bemerkung: $\mathbb{E}[X \parallel \mathcal{G}]$ ist fast sicher eindeutig, d.h. falls Y eine $\mathcal{G}$-messbare Zufallsvariable ist und $\forall G \in \mathcal{G} : \mathbb{E}[Y \cdot \mathbb{1}_G] = \mathbb{E}[X \cdot \mathbb{1}_G]$, dann gilt $Y \overset{a.s.}{=} \mathbb{E}[X \parallel \mathcal{G}]$. Mit Satz kann man den bedingten Erwartungswert $\mathbb{E}[X \parallel \mathcal{G}]$ auch für quasiintegrierbare X definieren (Details Übung).

13.2. Beispiel: Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum mit einer sub-σ-Algebra $\mathcal{G} \subseteq \mathcal{A}$. Für $X = \mathbb{1}_A$ und $Y = \mathbb{1}_B$ mit $A, B \in \mathcal{A}$ mit $0 < \mathbb{P}(B) < 1$ ist

$$\mathbb{P}(A \parallel \mathbb{1}_B)(\omega) = \begin{cases} \mathbb{P}(A \parallel B)(\omega) & \text{wenn } \omega \in B \\ \mathbb{P}(A \parallel B^c)(\omega) & \text{wenn } \omega \notin B \end{cases}$$

Weiters ist $\Omega \in \sigma(B)$, sodass für $\mathbb{1}_\Omega = 1$ gilt:

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{E}\mathbb{1}_A = \mathbb{E}[\mathbb{1}_A \cdot \mathbb{1}_\Omega] \\ &\overset{\text{Def.}}{=} \int_\Omega \mathbb{E}[\mathbb{1}_A \parallel \sigma(B)] \, d\mathbb{P} \\ &= \int_\Omega \mathbb{P}(A \parallel \mathbb{1}_B) \, d\mathbb{P} \\ &\overset{\text{s.o.}}{=} \mathbb{P}(A \parallel B) \cdot \mathbb{P}(B) + \mathbb{P}(A \parallel B^c) \end{aligned}$$

Eigenschaften bedingter Erwartungswerte

13.3. Proposition: Sei $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ und $\mathcal{G} \subseteq \mathcal{A}$ eine sub-σ-Algebra.

(i) Sind $\sigma(X)$ und $\mathcal{G}$ unabhängig, dann gilt $\mathbb{E}[X \parallel \mathcal{G}] \overset{a.s.}{=} \mathbb{E}[X]$.

(ii) Ist X $\mathcal{G}$–$\mathcal{B}(\mathbb{R})$-messbar, dann ist $\mathbb{E}[X \parallel \mathcal{G}] \overset{a.s.}{=} X$

Beweis:

(i) Die konstante Zufallsvariable $\mathbb{E}[X]$ ist natürlich $\mathcal{G}$-messbar und für $G \in \mathcal{G}$ sind $\mathbb{1}_G$ und X unabhängig (per Definition), sodass

$$\mathbb{E}[\mathbb{E}[X] \cdot \mathbb{1}_G] = \mathbb{E}[X] \cdot \mathbb{E}[\mathbb{1}_G] \overset{\text{u.a.}}{=} \mathbb{E}[X \cdot \mathbb{1}_G]$$

(ii) Wenn X $\mathcal{G}$–$\mathcal{B}(\overline{\mathbb{R}})$-messbar ist, ist X eine Dichte gemäß Definition 13.1, da trivial

$$\mathbb{E}[X \cdot \mathbb{1}_G] = \mathbb{E}[X \cdot \mathbb{1}_G]$$

$\square$

Bemerkung: Damit gilt für $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ sofort

- $\mathbb{E}[X \parallel \{\emptyset, \Omega\}] \overset{a.s.}{=} \mathbb{E}[X]$
- $\mathbb{E}[X \parallel \mathcal{A}] \overset{a.s.}{=} X$

13.4. Proposition: Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum mit einer sub-σ-Algebra $\mathcal{G} \subseteq \mathcal{A}$, $X_n, n \geqslant 1$ eine Folge von $(\Omega, \mathcal{A}, \mathbb{P})$-integrierbaren Zufallsvariablen und X, Y $(\Omega, \mathcal{A}, \mathbb{P})$-integrierbare Zufallsvariablen. Seien außerdem $a, b \in \mathbb{R}$. Dann gilt

(i) Ist $X \overset{a.s.}{=} a$, dann ist $\mathbb{E}[X \parallel \mathcal{G}] \overset{a.s.}{=} a$

(ii) $\mathbb{E}[aX + bY \parallel \mathcal{G}] \overset{a.s.}{=} a \cdot \mathbb{E}[X \parallel G] + b \cdot \mathbb{E}[Y \parallel \mathcal{G}]$ **(Linearität)**

(iii) Ist $X \overset{a.s.}{\leqslant} Y$, dann ist $\mathbb{E}[X \parallel \mathcal{G}] \leqslant \mathbb{E}[Y \parallel \mathcal{G}]$ **(Monotonie)**

(iv) $|\mathbb{E}[X \parallel \mathcal{G}]| \overset{a.s.}{\leqslant} \mathbb{E}[|X| \parallel \mathcal{G}]$ **(Dreiecksungleichung)**

(v) Für $X_1 \leqslant X_2 \leqslant \ldots \leqslant \lim_{n\to\infty} X_n$ fast sicher, ist **(MONK)**

$$\lim_{n\to\infty} \mathbb{E}[X_n \parallel \mathcal{G}] \overset{a.s.}{=} \mathbb{E}\left[\lim_{n\to\infty} X_n \parallel \mathcal{G}\right]$$

(vi) Falls $\forall n \geqslant 1 : Y \overset{a.s.}{\leqslant} X_n$ und $\liminf_{n\to\infty} X_n \in L^1$, dann ist **(Fatou's Lemma)**

$$\mathbb{E}[\liminf_{n\to\infty} X_n \parallel \mathcal{G}] \overset{a.s.}{\leqslant} \liminf_{n\to\infty} \mathbb{E}[X_n \parallel \mathcal{G}]$$

(vii) Falls $X_n \xrightarrow[n\to\infty]{a.s.} X$ und $\forall n \geqslant 1 : |X_n| \overset{a.s.}{\leqslant} Y$, dann ist $X \in L^1$ und **(DOMK)**

$$\lim_{n\to\infty} \mathbb{E}[X_n \parallel \mathcal{G}] \overset{a.s.}{=} \mathbb{E}[X \parallel \mathcal{G}]$$

Beweis:

(i) folgt sofort aus Proposition 13.3 und der Tatsache, dass X u.a. von jedem $\mathcal{G}$ ist.

(ii) $a \cdot \mathbb{E}[X \parallel \mathcal{G}] + b \cdot \mathbb{E}[Y \parallel \mathcal{G}]$ ist als Linearkombination $\mathcal{G}$-messbarer Funktionen wieder $\mathcal{G}$-messbar und für $G \in \mathcal{G}$ gilt:

$$\begin{aligned}\int_G a \cdot \mathbb{E}[X \parallel \mathcal{G}] + b \cdot \mathbb{E}[Y \parallel \mathcal{G}] \; d\mathbb{P} &= a \int_G \mathbb{E}[X \parallel \mathcal{G}] \; d\mathbb{P} + b \int_G \mathbb{E}[Y \parallel \mathcal{G}] \; d\mathbb{P} \\ &= a \cdot \mathbb{E}[X \cdot \mathbb{1}_G] + b \cdot \mathbb{E}[Y \cdot \mathbb{1}_G] \\ &= \mathbb{E}\left[(aX + bY) \cdot \mathbb{1}_G\right]\end{aligned}$$

Die Aussage folgt mit Definition 13.1 und Satz 12.8.

(iii) Hier ist insbesondere $X_+ \overset{a.s.}{\leqslant} Y_+$, sodass $0 \overset{a.s.}{\leqslant} Y_+ - X_+$. Sei $G := \{\mathbb{E}[X_+ \parallel \mathcal{G}] > \mathbb{E}[Y_+ \parallel \mathcal{G}]\}$. Dann ist G messbar und

$$\int_G \mathbb{E}[X_+ \parallel \mathcal{G}] - \mathbb{E}[Y_+ \parallel \mathcal{G}] \; d\mathbb{P} \overset{\text{(ii)}}{=} \int_G \mathbb{E}[X_+ - Y_+ \parallel \mathcal{G}] \; d\mathbb{P} = \mathbb{E}[(X_+ - Y_+) \cdot \mathbb{1}_G]$$

wobei das linke Integral nicht-negativ und das rechte Integral nicht-positiv ist, sodass $\mathbb{P}(G) = 0$ folgt. Ein ähnliches Argument gilt für X_- und Y_-, sodass folgt

$$\mathbb{E}[X \parallel \mathcal{G}] = \mathbb{E}[X_+ \parallel \mathcal{G}] - \mathbb{E}[X_- \parallel \mathcal{G}] \leqslant \mathbb{E}[Y_+ \parallel \mathcal{G}] - \mathbb{E}[Y_- \parallel \mathcal{G}] = \mathbb{E}[Y \parallel \mathcal{G}]$$

wobei alle Relationen fast sicher gelten.

(iv) Aus $X \leqslant |X|$ folgt mit (iii), dass

$$\mathbb{E}[X \parallel \mathcal{G}] \overset{a.s.}{\leqslant} \mathbb{E}[|X| \parallel \mathcal{G}]$$

Aus $-X \leqslant |X|$ folgt mit (ii) und (iii), dass

$$-\mathbb{E}[X \parallel \mathcal{G}] \overset{a.s.}{=} \mathbb{E}[-X \parallel \mathcal{G}] \overset{a.s.}{\leqslant} \mathbb{E}[|X| \parallel \mathcal{G}]$$

womit die Aussage aus der Kombination beider Fälle folgt.

(v) Mit (iii) gilt $\mathbb{E}[X_1 \parallel \mathcal{G}] \leqslant \mathbb{E}[X_2 \parallel \mathcal{G}] \leqslant \ldots \leqslant \lim_{n\to\infty} \mathbb{E}[X_n \parallel \mathcal{G}]$. Damit ist $\lim_{n\to\infty} \mathbb{E}[X_n \parallel \mathcal{G}]$ als Grenzwert $\mathcal{G}$-messbarer Funktionen messbar. Mit (iv) ist $|\mathbb{E}[X_1 \parallel \mathcal{G}]| \overset{a.s.}{\leqslant} \mathbb{E}[|X_1| \parallel \mathcal{G}]$, und damit $\mathbb{E}[X_1 \parallel \mathcal{G}] \in L^1$. Mit MONK gilt für $G \in \mathcal{G}$

$$\begin{aligned}\int_G \lim_{n\to\infty} \mathbb{E}[X_n \parallel \mathcal{G}] \; d\mathbb{P} &= \lim_{n\to\infty} \int_G \mathbb{E}[X_n \parallel \mathcal{G}] \; d\mathbb{P} \\ &= \lim_{n\to\infty} \mathbb{E}[X_n \cdot \mathbb{1}_G] \\ &\overset{\text{MONK}}{=} \mathbb{E}[X \cdot \mathbb{1}_G]\end{aligned}$$

womit per Definition 13.1 die Aussage folgt.

(vi) Für $Z_n := \inf_{k \geqslant n} X_k$ gilt $Z_1 \leqslant \ldots \leqslant \lim_{n\to\infty} Z_n = \liminf_{n\to\infty} X_n \in L^1$. Es gilt $Z_n \overset{a.s.}{\leqslant} X_k$ für alle $k \geqslant n$, sodass mit (iii)

$$\mathbb{E}[Z_n \parallel \mathcal{G}] \overset{a.s.}{\leqslant} \inf_{k \geqslant n} \mathbb{E}[X_k \parallel \mathcal{G}]$$

und damit

$$\mathbb{E}[\liminf_{n\to\infty} X_n \parallel \mathcal{G}] \overset{\text{(v)}}{=} \lim_{n\to\infty} \mathbb{E}[Z_n \parallel \mathcal{G}] \leqslant \lim_{n\to\infty} \inf_{k \geqslant n} \mathbb{E}[X_k \parallel \mathcal{G}] = \liminf_{n\to\infty} \mathbb{E}[X_n \parallel \mathcal{G}]$$

wobei in der letzten Ungleichung die ersten beiden Relationen als fast sicher zu verstehen sind.

(vii) Mit DOMK ist $X \in L^1$ und $\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X]$. Mit (iii) und (iv) gilt

$$|\mathbb{E}[X_n \parallel \mathcal{G}]| \overset{a.s.}{\leqslant} \mathbb{E}[|X_n| \parallel \mathcal{G}] \overset{a.s.}{\leqslant} \mathbb{E}[Y_n \parallel \mathcal{G}]$$

Mit (vi) gilt

$$\mathbb{E}[X \parallel \mathcal{G}] \overset{a.s.}{=} \mathbb{E}[\liminf_{n\to\infty} X_n \parallel \mathcal{G}] \overset{a.s.}{\leqslant} \liminf_{n\to\infty} \mathbb{E}[X_n \parallel \mathcal{G}]$$

wobei die letze Ungleichung folgt, da laut Annahme $\mathbb{E}[-Y \parallel \mathcal{G}] \leqslant \mathbb{E}[X_n \parallel \mathcal{G}]$ für alle $n \geqslant 1$. Ebenfalls gilt

$$\begin{aligned} -\mathbb{E}[X \parallel \mathcal{G}] &\overset{a.s.}{=} \mathbb{E}[-X \parallel \mathcal{G}] \\ &\overset{a.s.}{=} \mathbb{E}[\liminf_{n\to\infty} (-X_n) \parallel \mathcal{G}] \\ &\overset{a.s.}{\leqslant} \liminf_{n\to\infty} \mathbb{E}[-X_n \parallel \mathcal{G}] \\ &\overset{a.s.}{=} -\limsup_{n\to\infty} \mathbb{E}[X_n \parallel \mathcal{G}] \end{aligned}$$

Damit gilt

$$\limsup_{n\to\infty} \mathbb{E}[X_n \parallel \mathcal{G}] \overset{a.s.}{\leqslant} \mathbb{E}[X \parallel \mathcal{G}] \overset{a.s.}{\leqslant} \liminf_{n\to\infty} \mathbb{E}[X_n \parallel \mathcal{G}]$$

und damit folgt die Aussage. □

13.5. Proposition: Sei $\mathcal{G} \subseteq \mathcal{A}$ eine sub-σ-Algebra und seien X, U Zufallsvariablen, sodass $X, UX \in L^1$. Falls U $\mathcal{G} - \mathcal{B}(\mathbb{R})$-messbar ist, dann gilt

$$\mathbb{E}[UX \parallel \mathcal{G}] \overset{a.s.}{=} U \cdot \mathbb{E}[X \parallel \mathcal{G}]$$

Beweis: $U \cdot \mathbb{E}[X \parallel \mathcal{G}]$ ist $\mathcal{G}$-messbar.

I. $U = \mathbb{1}_H, H \in \mathcal{G}$

$$\forall G \in \mathcal{G} : \int_G U \cdot \mathbb{E}[X \parallel \mathcal{G}] \; d\mathbb{P} = \int_{G \cap G} \mathbb{E}[X \parallel \mathcal{G}] \; d\mathbb{P} = \mathbb{E}[X \cdot \mathbb{1}_{G \cap H}] = \mathbb{E}[UX \cdot \mathbb{1}_G]$$

II. U einfach
folgt aus I. und der Linearität des bedingten Erwartungswertes.

III. $U \geqslant 0$
Wähle $U_n, n \geqslant 1$ einfach und $\mathcal{G}$-messbar mit $0 \leqslant U_n \nearrow U$. Dann gilt $U_n X \to UX$ und $|U_n X| = |U_n||X| \leqslant |U||X|$. Es folgt

$$\mathbb{E}[UX \parallel \mathcal{G}] \overset{\text{DOMK}}{=} \lim_{n\to\infty} \mathbb{E}[U_n X \parallel \mathcal{G}] \overset{\text{II.}}{=} \lim_{n\to\infty} U_n \cdot \mathbb{E}[X \parallel \mathcal{G}] = U \cdot \mathbb{E}[X \parallel \mathcal{G}]$$

IV. U allgemein

Schreibe $U = U_+ - U-$. Dann gilt

$$\begin{aligned}
\mathbb{E}[UX \parallel \mathcal{G}] &= \mathbb{E}[(U_+ - U_-)X \parallel \mathcal{G}] \\
&\overset{a.s.}{=} \mathbb{E}[U_+X \parallel \mathcal{G}] - \mathbb{E}[U_-X \parallel \mathcal{G}] \\
&\overset{a.s.}{=} (U_+) \cdot \mathbb{E}[X \parallel \mathcal{G}] - (U_-) \cdot \mathbb{E}[X \parallel \mathcal{G}] \\
&\overset{a.s.}{=} (U_+ - U_-) \cdot \mathbb{E}[X \parallel \mathcal{G}] = U \cdot \mathbb{E}[X \parallel \mathcal{G}]
\end{aligned}$$

□

13.6. Proposition: Seien $\mathcal{G}$ und $\mathcal{H}$ sub-σ-Algebren, sodass $\mathcal{H} \subseteq \mathcal{G} \subseteq \mathcal{A}$. Sei $X \in L^1$. Dann gilt

$$\mathbb{E}[X \parallel \mathcal{H}] \overset{\text{(i)}}{=} \mathbb{E}[\mathbb{E}[X \parallel \mathcal{H}] \parallel \mathcal{G}] \overset{\text{(ii)}}{=} \mathbb{E}[\mathbb{E}[X \parallel \mathcal{G}] \parallel \mathcal{H}]$$

fast sicher.

Beweis:

(i) folgt sofort, da $\mathbb{E}[X \parallel \mathcal{H}]$ $\mathcal{H}$-messbar und damit auch $\mathcal{G}$-messbar ist. Mit Proposition 13.3 gilt

$$\mathbb{E}[X \parallel \mathcal{H}] \overset{a.s.}{=} \mathbb{E}[\mathbb{E}[X \parallel \mathcal{H}] \parallel \mathcal{G}]$$

(ii) $\mathbb{E}[\mathbb{E}[X \parallel \mathcal{G}] \parallel \mathcal{H}]$ ist $\mathcal{H}$-messbar und damit auch $\mathcal{G}$-messbar. Es gilt

$$\begin{aligned}
\forall H \in \mathcal{H} : \int_H \mathbb{E}[\mathbb{E}[X \parallel \mathcal{G}] \parallel \mathcal{H}] \; d\mathbb{P} &= \int_H \mathbb{E}[X \parallel \mathcal{G}] \; d\mathbb{P} \\
&= \int_H X \; d\mathbb{P}
\end{aligned}$$

da $H \in \mathcal{H} \implies H \in \mathcal{G}$. Mit Definition 13.1 und Satz 12.8 folgt die Aussage. □

Bedingte Verteilungen

13.7. Satz: Sei $\mathcal{G} \subseteq \mathcal{A}$ eine sub-σ-Algebra und sei X eine reelwertige Zufallsvariable. Dann existiert eine Funktion $\mu : \mathcal{B}(\mathbb{R}) \times \Omega \to [0,1]$ mit den folgenden Eigenschaften:

(i) Für $\omega \in \Omega$ fest ist $\mu(\cdot\,, \omega) : \mathcal{B}(\mathbb{R}) \to [0,1]$ ein Wahrscheinlichkeitsmaß.

(ii) Für $B \in \mathcal{B}(\mathbb{R})$ fest, gilt für $\mu(B, \cdot\,) : \Omega \to [0,1]$, dass $\mu(B, \cdot\,) = \mathbb{P}(X \in B \parallel \mathcal{G})$ fast sicher.

Man nennt $\mu(\cdot\,, \omega)$ die bedingte Verteilung von X gegeben $\mathcal{G}$.

Beweis: Für $q \in \mathbb{Q}$ sei $F(q, \omega) := \mathbb{P}(X \leqslant q \parallel \mathcal{G})(\omega)$. Für $q, r \in \mathbb{Q}$ mit $q \leqslant r$ gilt damit (Monotonie des bedingten Erwartungswertes, Proposition 13.4)

$$F(q, \omega) \leqslant F(r, \omega) \tag{8}$$

für alle w außerhalb einer $\mathbb{P}$-Nullmenge. Mit DOMK für bedingte Erwartungswerte (Proposition 13.4) gilt

$$\forall q \in \mathbb{Q} : F(q, \omega) = \lim_{n \to \infty} F\left(q + \frac{1}{n}, \omega\right) \tag{9}$$

für ω außerhalb einer $\mathbb{P}$-Nullmenge und (DOMK)

$$\lim_{n\to\infty} F(-n,\omega) = 0 \text{ und } \lim_{n\to\infty} F(n,\omega) = 1 \tag{10}$$

außerhalb einer $\mathbb{P}$-Nullmenge. Sei N die abzählbare Vereinigung der in (10), (11) und (12) auftretenden Nullmengen Dann ist N wieder eine $\mathbb{P}$-Nullmenge und (10), (11), (12) gelten für jedes $\omega \in N^c$.

Sei nun $t \in \mathbb{R}$. Für $\omega \in N^c$ setze

$$F(t,\omega) := \inf\{F(q,\omega) : t \leqslant q, q \in \mathbb{Q}\}$$

und für $\omega \in N$ setze $F(t,\omega) := F(t)$ für eine beliebige fest cdf F. Mit (10), (11) und (12) ist $F(\cdot,\omega)$ eine cdf für jedes $\omega \in \Omega$. Für $\omega \in \Omega$ sei nun $\mu(\cdot\,\omega)$ das durch $F(\cdot\,.\omega)$ bestimmte Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Damit gilt (i).

Die Mengenfamilie

$$\mathcal{L} := \{B \in \mathcal{B}(\mathbb{R}) : \mu(B,\cdot\,) \text{ ist } \mathcal{G}\text{-messbar}\}$$

ist ein λ-System (leicht zu prüfen) und enthält das π-System $\mathcal{M} := \{(-\infty, q] : q \in \mathcal{Q}\}$. Mit dem λ–π-Theorem gilt

$$\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{M}) = \lambda(\mathcal{M}) \subseteq \mathcal{L} \subseteq \mathcal{B}(\mathbb{R})$$

Also gilt $\mathcal{L} = \mathcal{B}(\mathbb{R})$ und $\mu(B,\cdot\,)$ ist für jedes $B \in \mathcal{B}(\mathbb{R})$ $\mathcal{G}$-messbar.

Für $B = (-\infty, q]$ mit $q \in \mathbb{Q}$ ist $\mu(B,\cdot\,) = \mathbb{P}(X \leqslant q \parallel \mathcal{G})$ laut Konstruktion, sodass für $G \in \mathcal{G}$ gilt

$$\mathbb{E}[\mathbb{1}_G \cdot \mu(B,\cdot\,)] = \mathbb{P}(G \cap \{X \in B\}) \tag{11}$$

Für $G \in \mathcal{G}$ fest gilt (13) für jede Menge $B \in \mathcal{M}$. Sei nun

$$\mathcal{Z} := \{B \in \mathcal{B}(\mathbb{R}) : \mathbb{E}[\mathbb{1}_G \cdot \mu(B,\cdot\,)] = \mathbb{P}(G \cap \{X \in B\})\}$$

Nun sind in (14) beide Seiten der Gleichung endliche Maße, die auf dem π-System $\mathcal{M}$ übereinstimmen. Mit Korollar 2.9 folgt, dass die beiden Maße auch auf $\sigma(\mathcal{M}) = \mathcal{B}(\mathbb{R})$ übereinstimmen. Also gilt (13) für alle $B \in \mathcal{B}(\mathbb{R})$. Mit Definition 13.1 folgt Aussage (ii). $\square$