TECHNISCHE
UNIVERSITÄT
DRESDEN

**Fakultät Mathematik** Institut für Stochastik, Professur für Statistik

# Mathematische Statistik

**Prof. Dr. Dietmar Ferger**

Wintersemester 2020/21

Mitschrift : Eric Kunze
E-Mail : `eric.kunze@mailbox.tu-dresden.de`

*Dieses Skript basiert auf den Mitschriften von Willi Sontopski und Felix Hilsky aus den vorherigen Semestern sowie den Bemerkungen und Korrekturen von Prof. Ferger im Wintersemester 2020/21.*

# Inhaltsverzeichnis

**1 Median** **3**

**2 Metrische Räume** **8**

**3 Zufallsvariablen in metrischen Räumen** **13**

3.1 Messbarkeit in metrischen Räumen . . . . . 13

3.2 Fast sichere Konvergenz . . . . . 15

3.3 Stochastische Konvergenz . . . . . 18

3.4 Konvergenz in Produkträumen . . . . . 19

**4 Verteilungskonvergenz von Zufallsvariablen in metrischen Räumen** **22**

4.1 Satz von Portmanteau . . . . . 22

**A Grundlagen** **34**

# Kapitel 1
# MEDIAN

Viele Schätzer in der Statistik sind definiert als Minimal- oder Maximalstelle von bestimmten **Kriteriumsfunktionen**, z. B. der **Maximum-Likelihood-Schätzer (MLS)** oder **Minimum-Quadrat-Schätzer (MQS, KQS)** oder **Bayes-Schätzer**. Allgemein nennt man solche Schätzer **M-Schätzer**.

**Beispiel 1.1 (Maximum-Likelyhood-Schätzer)** Gegeben seien $X_1, ..., X_n$ iid. $\sim f_\theta$. Dann ist $\hat{\theta}_n$ die Maximumstelle von der Funktion $l\colon \theta \mapsto \sum_{i=0}^n \log f_\theta(X_i)$ (sogenannte Log-Likelyhood-Funktion). Dabei kommen alle $\theta$ aus einer möglichen Menge $\Theta$ in Frage.

Ziel: Untersuchung des asymptotischen Verhaltens ($n \to \infty$) von M-Schätzern über einen funktionalen Ansatz. Als Beispiel betrachten wir nun den Median

Sei $X\colon (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}$ eine reelle Zufallsvariable mit Verteilungsfunktion $F_X\colon \mathbb{R} \to [0,1]$, d. h.

$$F_X(x) := \mathbb{P}(X \leq x) := \mathbb{P}(\{\omega \in \Omega : X(\omega) \leq x\}) = \mathbb{P}\left(X^{-1}((-\infty, x])\right),$$

also $X \sim F_X$. Definiere

$$\begin{aligned} Y(t) &:= \mathbb{E}(|X - t|) && (1.1)\\ &= \int_\Omega |X(\omega) - t| \; \mathbb{P}(\mathrm{d}\omega) \\ &\overset{\textit{Gleichung (A.2)}}{=} \int_\mathbb{R} |x - t| \; (\mathbb{P} \circ X^{-1})(\mathrm{d}x) \\ &= \int_\mathbb{R} |x - t| \; F_X(\mathrm{d}x) \qquad \forall t \in \mathbb{R} \end{aligned}$$

und

$$m \in \underset{t \in \mathbb{R}}{\arg\min}\, Y(t) \tag{1.2}$$

als (irgendeine) Minimalstelle der Funktion $Y$.

**Definition 1.2** Hierbei heißt $m \in \mathbb{R}$ **Median** von der Zufallsgröße $X$.

Wir haben den **Korrespondenzsatz** genutzt, der besagt, dass das Bildmaß $\mathbb{P} \circ X^{-1}$ eindeutig von der Verteilungsfunktion $F$ bestimmt ist. Das heißt also wir schreiben $F(\mathrm{d}x)$ und meinen $\mathbb{P} \circ X^{-1}(\mathrm{d}x)$

**Notation:** $F(m-) := F(m-0) := \lim_{t \uparrow m} F(t)$. Der Ausdruck ist wohldefiniert für Verteilungsfunktionen, denn diese sind rechtsseitig stetig und haben linksseitige Limiten.

In folgendem kleinen Lemma wollen wir die Menge aller Mediane charakterisieren.

**Lemma 1.3** Sei $X \sim F_X =: F$ integrierbar und $m \in \mathbb{R}$. Dann sind äquivalent:

(1) $F(m-) \le \frac{1}{2} \le F(m)$

(2) $\mathbb{E}\left[|X-t|\right] \ge \mathbb{E}|X-m| \qquad \forall t \in \mathbb{R}$

(3) $m$ ist Median

**Beispiel 1.4** Für $F$, $a < b \in \mathbb{R}$ mit

$$F(x) \begin{cases} < \frac{1}{2}, & \text{für } x < a \\ = \frac{1}{2}, & \text{für } a \le x < b \\ > \frac{1}{2}, & \text{für } x \ge b \end{cases}$$

ist die Menge der Mediane genau das Intervall $[a, b]$.

*Beweis.* **(1) ⇒ (2):** Sei $Q := \mathbb{P} \circ X^{-1}$ das zu $F$ gehörige Bildmaß. Setze

$$h(t) := \mathbb{E}\left[|X-t| - |X-m|\right] \overset{lin.}{=} Y(t) - Y(m)$$

Dann ist (2) äquivalent zu $h(t) \ge 0$ für alle $t \in \mathbb{R}$. Dies wollen wir nun zeigen.

Fall A: Sei $t < m$.

$$\begin{aligned} h(t) &\overset{\text{Trafo}}{=} \int_{\mathbb{R}} |x-t| - |x-m| \; F(\mathrm{d}x) \\ &= \int_{(-\infty,t]} \underbrace{|x-t| - |x-m|}_{=t-x-(m-x)=-(m-t)} F(\mathrm{d}x) + \int_{(t,m)} \underbrace{|x-t| - |x-m|}_{\underbrace{\underbrace{x-t}_{\ge 0} - \underbrace{(m-x)}_{\le m-t}}_{\ge -(m-t)}} F(\mathrm{d}x) \\ &\quad + \int_{[m,\infty)} \underbrace{|x-t| - |x-m|}_{x-t-(x-m)=m-t} F(\mathrm{d}x) \\ &\ge -(m-t) \cdot \underbrace{Q((-\infty,t])}_{F(t)} + (-(m-t) \cdot F(m-) - F(t)) + (m-t) \cdot \underbrace{Q([m,\infty))}_{1-\underbrace{Q((-\infty,m))}_{F(m-)}} \\ &= -\underbrace{(m-t)}_{\ge 0} \cdot \Big( \underbrace{1 - 2 \cdot F(m-)}_{\text{wegen 1. Ungl. in (1):} \ge 0} \Big) \\ &\ge 0 \end{aligned}$$

Fall B: Sei $t > m$. Wir verfahren ganz ähnlich:

$$\begin{aligned}
h(t) &= \int_{(-\infty,m]} |x-t| - |x-m|\, F(\mathrm{d}x) + \int_{(m,t]} |x-t| - |x-m|\, F(\mathrm{d}x) \\
&\quad + \int_{(t,\infty)} |x-t| - |x-m|\, F(\mathrm{d}x) \\
&= \int_{(-\infty,m]} t-x-(m-x)F(\mathrm{d}x) + \int_{(m,t]} t-x-(x-m)F(\mathrm{d}x) \\
&\quad + \int_{(t,\infty)} x-t-(x-m)F(\mathrm{d}x) \\
&\geq (t-m)F(m) - (t-m)(F(t)-F(m)) + (m-t)(1-F(t)) \\
&= (t-m)(F(m)-F(t)+F(m)-1+F(t)) \\
&= \underbrace{(t-m)}_{>0} \cdot (\underbrace{2\cdot F(m)-1}_{\text{wegen 2. Ungl. in (1)}\geq 0}) \\
&\geq 0
\end{aligned}$$

Fall C: Für $t = m$ ist die Aussage trivial.

**(2) $\Rightarrow$ (1):** Nach Annahme ist $h(t) \geq 0$ für alle $t \in \mathbb{R}$.

Fall A: Sei $t < m$. Die obige Rechnung im Fall 1 bei $\Rightarrow$ zeigt

$$\begin{aligned}
0 \leq h(t) &= -(m-t)\cdot F(t) + \int_t^m \underbrace{\underbrace{x}_{<m} - t - (m-x)}_{=2x-t-m\leq m-t} F(\mathrm{d}x) + (m-t)\cdot(1-F(m-)) \\
&\leq -(m-t)\cdot\Big(F(t) - 1 \underbrace{+F(m-) - F(m-)}_{=0} + F(t)\Big) \\
&= \underbrace{(m-t)}_{>0} \cdot (1 - 2\cdot F(t)) \\
&\Rightarrow \forall t < m:\ \ 0 \leq (m-t)\cdot(1-2\cdot F(t)) \\
&\Rightarrow \forall t < m:\ \ 0 \leq 1 - 2\cdot F(t) \\
&\Rightarrow \forall t < m:\ \ F(t) \leq \frac{1}{2} \\
&\overset{t\uparrow m}{\Rightarrow} F(m-) \leq \frac{1}{2} \qquad \text{(Def. linksseitiger Limes)}
\end{aligned}$$

Fall B: Sei $t > m$. Siehe 2. Fall, analog:

$$
\begin{aligned}
0 \leq h(t) &= (t-m) \cdot F(m) + \int_m^t \underbrace{t - x - (x-m)}_{=t+m-2x \leq t-m} F(\mathrm{d}x) + (m-t) \cdot (1 - F(t)) \\
&\leq (t-m) \cdot (F(m) + F(t) - F(m) - 1 + F(t)) \\
&= \underbrace{(t-m)}_{>0}(2F(t) - 1) \\
&\Rightarrow \forall t < m: \quad 0 \leq 2F(t) - 1 \\
&\Rightarrow \forall t < m: \quad F(t) \geq \frac{1}{2} \\
&\overset{t \downarrow m}{\Rightarrow} F(m) \geq \frac{1}{2} && \text{(Rechtsstetigkeit von } F\text{)}
\end{aligned}
$$

**(2) $\Rightarrow$ (3):** Die Aussage (2) ist offensichtlich äquivalent zur Definition des Medians. $\square$

### Bemerkung 1.5

(1) Lemma 1.3 besagt, dass $\{m \in \mathbb{R} : m \text{ erfüllt Punkt } (1)\}$ die Menge aller Mediane von $F$ ist. Der Median ist im Allgemeinen nicht eindeutig bestimmt.

(2) Im Allgemeinen gibt es mehrere Mediane. Üblicherweise *wählt* man $m := F^{-1}\left(\frac{1}{2}\right)$, wobei

$$F^{-1}(u) := \inf \{x \in \mathbb{R} : F(x) \geq u\} \quad \forall u \in (0,1)$$

die **Quantilfuntion** oder auch **verallgemeinerte Inverse** ist. Für weiterführende Literatur siehe [28, Seite 20]. Da

$$F\left(F^{-1}(u)-\right) \leq u \leq F\left(F^{-1}(u)\right) \qquad \forall u \in (0,1),$$

erfüllt $m = F^{-1}\left(\frac{1}{2}\right)$ die Bedingung (1) in Lemma 1.3 und ist somit ein Median, nämlich der kleinste.

(3) Die obige Funktion (1.1)

$$Y : \mathbb{R} \to \mathbb{R} \text{ mit } Y(t) = \int_{\mathbb{R}} |x - t| \ F(\mathrm{d}x) \qquad \forall t \in \mathbb{R}$$

ist stetig[1], aber im Allgemeinen nicht differenzierbar, z. B. falls $F \sim X$ eine diskrete Zufallsvariable ist. In diesem Fall ist somit die Minimierung über Differentiation nicht möglich.

(4) Sei $\mu = \mathbb{E}(X)$ der Erwartungswert von $X$. Dann gilt (Übung):

$$\mu = \arg\min_{t \in \mathbb{R}} \mathbb{E}\left[(X-t)^2\right] = \arg\min_{t \in \mathbb{R}} \mathbb{E}\left[(X-t)^2 - X^2\right].$$

(Das zweite $X^2$ wird abgezogen, da das $\arg\min$ nicht davon betroffen ist und so die Bedin-

[1] zur Stetigkeit von $Y$: nutze Folgenkriterium + dominierte Konvergenz mit Majorante $|X| + |t|$

gung $\mathbb{E}\left[X^2\right] < \infty$ entfällt.) Begründung:

$$\begin{aligned}\mathbb{E}\left[(X-t)^2) - X^2\right] &= \mathbb{E}\left[X^2 - 2tX + t^2 - X^2\right] \\ &= \mathbb{E}\left[X^2\right] - 2t\mu + t^2 - \mathbb{E}\left[X^2\right] = t^2 - 2t\mu\end{aligned}$$

Minimiere nun diese quadratische Funktion und erhalte die gewünschte Resultat.

In der Statistik identifizieren wir oft stillschweigend Zufallsgrößen mit ihren Realisationen, also $X \leftrightsquigarrow x$ was sich formal $X(\omega_0) = x$ schreiben lässt.

Zur Schätzung von $m$ seien $X_1, \dots, X_n \sim F$ iid Zufallsvariablen mit zugehöriger **empirischer Verteilungsfunktion**

$$F_n\colon \mathbb{R} \to [0,1] \quad \text{mit} \quad F_n(x) := \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbb{1}_{\{x_i \le x\}}(x) \qquad \forall x \in \mathbb{R}.$$

Tatsächlich ist $F_n$ die Verteilungsfunktion zum **empirischen Maß**

$$Q_n\colon \mathcal{B}(\mathbb{R}) \to [0,1] \quad \text{mit} \quad Q_n(B) := \frac{1}{n} \cdot \sum_{i=1}^{n} \delta_{x_i}(B)$$

wobei das **Dirac-Maß** in $t$ definiert ist als

$$\delta_t\colon \mathcal{B}(\mathbb{R}) \to [0,1] \quad \text{mit} \quad \delta_t(B) := \begin{cases} 0, & \text{falls } t \notin B \\ 1, & \text{falls } t \in B \end{cases} \qquad \forall B \in \mathcal{B}(\mathbb{R}) \ \forall t \in \mathbb{R}$$

Gemäß dem Satz von Gliwenko-Cantelli gilt:

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \stackrel{n\to\infty}{\longrightarrow} 0 \qquad \mathbb{P}\text{-fast sicher für alle Verteilungsfunktionen } F$$

Die empirische Verteilungsfunktion konvergiert also gegen die wahre Verteilungsfunktion.

**Erinnerung.** Für das Dirac-Maß $\delta_x\colon \mathcal{A} \to \mathbb{R}_+$ mit $\delta_x(A) := \mathbb{1}_A(x)$ gilt:

$$\int_{\mathbb{R}} f(t)\, \delta_x(\mathrm{d}t) = f(x) \qquad \forall x \in \mathbb{R} \tag{1.3}$$

**Notation.** Wir identifizieren eine Verteilung $\mathbb{P} \circ X^{-1}$ mit der zugehörigen Verteilungsfunktion $F_X$, also $\mathbb{P} \circ X^{-1} \leftrightarrow F_X$ und $F_X(\mathrm{d}x) := (\mathbb{P} \circ X^{-1}\, \mathrm{d}x)$.

**Erinnerung.** Das Lebesgue-Maß ist linear im Maß, d. h.:

$$\int_{\omega} f \mathrm{d}(a \cdot \mu + b \cdot \nu) = a \cdot \int_{\omega} f \mathrm{d}\mu + b \cdot \int_{\omega} f\, \mathrm{d}\nu \tag{1.4}$$

# Kapitel 2
# METRISCHE RÄUME

Sei $(\mathcal{S}, d)$ ein metrischer Raum.

### Beispiel 2.1 (Supremums-Metrik)

$$\mathcal{S} = C([0,1]) := \{f\colon [0,1] \to \mathbb{R} : f \text{ stetig}\}$$
$$d(f,g) := \sup_{t\in[0,1]} |f(t) - g(t)| \qquad \forall f,g \in C([0,1])$$

**Definition 2.2** (1) Für $x \in \mathcal{S}$ und $r > 0$ ist

$$B(x,r) := B_d(x,r) := \{y \in \mathcal{S} : d(x,y) < r\}$$

die **offene Kugel um Mittelpunkt $x$ und Radius $r$**.

(2) Sei $A \subseteq \mathcal{S}$. Dann ist:

- $A^\circ \dots$ das **Innere** von $A$
- $\overline{A} \dots$ die **abgeschlossene Hülle** von $A$
- $\partial A := \overline{A} \cap \overline{A^C} = \overline{A} \setminus A^\circ \dots$ der **Rand** von $A$
- $A^{\mathsf{C}} := \mathcal{S} \setminus A \dots$ das **Komplement** von $A$

(3) Die durch $d$ induzierte **Topologie** ist

$$\begin{aligned}\mathcal{G} := \mathcal{G}(\mathcal{S}) &:= \{G \subseteq \mathcal{S} : G \text{ ist offen bzgl. } d\}\\ &= \{G \subseteq \mathcal{S} : \forall x \in G : \exists r > 0 : B_d(x,r) \subseteq G\}\end{aligned}$$

und

$$\mathcal{F} := \mathcal{F}(\mathcal{S}) := \{F \subseteq \mathcal{S} : F \text{ ist abgeschlossen}\} = \left\{F : F = G^{\mathsf{C}}, G \in \mathcal{G}\right\}$$

ist die **Menge der abgeschlossenen Mengen**.

(4) Sei $\emptyset \neq A \subseteq \mathcal{S}$ und $x \in \mathcal{S}$. Dann ist $d(x,A) := \inf\{d(x,a) : a \in A\} \geq 0$ der **Abstand von $x$ zu $A$**.

(5) Es ist

$$C(\mathcal{S}) := \{f\colon S \to \mathbb{R} : f \text{ stetig}\}$$
$$C^b(\mathcal{S}) := \{f \in C(\mathcal{S}) : f \text{ beschränkt}\}$$
$$\|f\| := \|f\|_\infty := \sup_{x\in\mathcal{S}} |f(x)|$$

**Lemma 2.3** Sei $A \neq \emptyset$.

(1) Es ist $x \in \overline{A} \quad \Leftrightarrow \quad d(x, A) = 0$.

(2) Für alle $x, y \in \mathcal{S}$ gilt $|d(x, A) - d(y, A)| \leq d(x, y)$.

(3) Die Abbildung $d(\cdot, A)\colon \mathcal{S} \to [0, \infty)$, $x \mapsto d(x, A)$ ist gleichmäßig stetig.

*Beweis.* (zu 1) **(⇒)** Sei $x \in \overline{A}$. Dann existiert für alle $\varepsilon > 0$ ein $a \in A$ mit $d(x, a) < \varepsilon$. Somit

$$0 \leq d(x, A) \leq d(x, a) < \varepsilon \quad \forall \varepsilon > 0 \overset{\varepsilon \to 0}{\Rightarrow} d(x, A) = 0$$

**(⇐)** Sei $d(x, A) = 0$. Dann folgt aus der Infimumseigenschaft:

$$\forall \varepsilon > 0 \ \exists a \in A: \quad 0 \leq d(x, a) \leq 0 + \varepsilon = \varepsilon \quad \Rightarrow \quad x \in \overline{A}$$

(zu 2) Seien $x, y \in \mathcal{S}$. Dann gilt $d(x, a) \leq d(x, y) + d(y, a)$ mit der Dreiecksungleichung für alle $a \in A$ und somit

$$d(x, A) \leq d(x, y) + d(y, A) \quad \Rightarrow \quad d(x, A) - d(y, A) \leq d(x, y)$$

Vertauschen von $x$ und $y$ liefert $d(y, A) - d(x, A) \leq d(y, x) = d(x, y)$ und daraus folgt die Behauptung.

(zu 3) Folgt aus Punkt (2) da die Funktion $d(\cdot, A)$ Lipschitz-stetig und damit gleichmäßig stetig ist. □

**Satz 2.4 (Stetige Approximation von Indikatorfunktionen)** Zu $A \subseteq \mathcal{S}$ und $\varepsilon > 0$ existiert eine gleichmäßig stetige Funktion

$$f\colon \mathcal{S} \to [0, 1] \quad \text{mit der Eigenschaft} \quad f(x) = \begin{cases} 1 & \text{falls } x \in A \\ 0 & \text{falls } d(x, A) \geq \varepsilon \end{cases}$$

*Beweis.* Setze

$$\varphi\colon \mathbb{R} \to [0, 1] \quad \text{mit} \quad \varphi(t) := \begin{cases} 1 & \text{falls } t \leq 0 \\ 1 - t & \text{falls } 0 < t < 1 \\ 0 & \text{falls } t \geq 1 \end{cases}$$

Dann ist $\varphi$ gleichmäßig stetig auf $\mathbb{R}$. Sei

$$f(x) := \varphi\left(\frac{1}{\varepsilon} \cdot d(x, A)\right) \qquad \forall x \in \mathcal{S}$$

Dann hat dieses $f$ die gewünschte Eigenschaft wegen Lemma 2.3. □

**Definition 2.5** Ein metrischer Raum $(\mathcal{S}, d)$ heißt **separabel**

$$\begin{aligned} &:\Leftrightarrow \exists \text{ abzählbares } S_0 \subseteq \mathcal{S} : \mathcal{S} \subseteq \overline{S_0} \\ &\Leftrightarrow \exists \text{ abzählbares } S_0 \subseteq \mathcal{S} : \mathcal{S} = \overline{S_0} \\ &\Leftrightarrow \exists \text{ abzählbares } S_0 \subseteq \mathcal{S} : S_0 \text{ liegt dicht in } \mathcal{S} \end{aligned}$$

**Beispiel 2.6** Der metrische Raum $C([0,1])$ mit Supremums-Metrik ist separabel.

*Beweis.* Die Menge $S_0 := \{P : P \text{ ist Polynom mit rationalen Koeffizienten}\}$ ist abzählbar. Aus dem *Approximationssatz von Weierstraß* und der Dichtheit von $\mathbb{Q}$ folgt die Behauptung. □

**Definition 2.7** $\mathcal{G}_0 \subseteq \mathcal{G}$ heißt **Basis** von $\mathcal{G}$ genau dann, wenn sich jedes $G \in \mathcal{G}$ als Vereinigung von Mengen aus $\mathcal{G}_0$, so genannte **$\mathcal{G}_0$-Mengen**, darstellen lässt.

**Beispiel 2.8** Die Menge $\{B(x,r) : x \in \mathcal{S}, 0 < r \in \mathbb{Q}\}$ ist Basis von $\mathcal{G}$.

*Beweis.* Sei $G \in \mathcal{G}$. Dann gilt:

$$\forall x \in G \;\; \exists 0 < r_x \in \mathbb{Q} : B(x, r_x) \subseteq G \quad \Rightarrow \quad G = \bigcup_{x \in G} \{x\} \subseteq \bigcup_{x \in G} \underbrace{B(x, r_x)}_{\subseteq G} \subseteq G$$

Damit gilt also Gleichheit. □

**Satz 2.9** $\mathcal{S}$ separabel $\Leftrightarrow$ $\mathcal{G}$ hat eine abzählbare Basis

*Beweis.* **(⇒)** Sei $S_0 \subseteq \mathcal{S}$ abzählbar und dicht in $\mathcal{S}$. Wir zeigen, dass

$$\mathcal{G}_0 := \{B(x,r) : x \in S_0, 0 < r \in \mathbb{Q}\} \subseteq \mathcal{G}$$

eine Basis ist. Sei also $G$ offen. Dann folgt aus Beispiel Beispiel 2.8, dass

$$G = \bigcup_{x \in G} B(x, r_x), \qquad 0 < r_x \in \mathbb{Q} \;\; \forall x \in G \tag{$\star$}$$

Da $\overline{S_0} = \mathcal{S}$ gilt:

$$\begin{aligned} & \forall x \in G : \exists y_x \in S_0 : d(x, y_x) < \frac{r_x}{2} \\ \Rightarrow \quad & d(x,y) \overset{\triangle-\text{Ungl.}}{\leq} d(x, y_x) + d(y_x, x) < \frac{r_x}{2} + \frac{r_x}{2} = r_x \qquad \forall y \in B\left(y_x, \frac{r_x}{2}\right) \\ \Rightarrow \quad & B\left(y_x, \frac{r_x}{2}\right) \subseteq B(x, r_x) \qquad \forall x \in G \\ \Rightarrow \quad & G \overset{(\star)}{\supseteq} \bigcup_{x \in G} \underbrace{B\left(y_x, \frac{r_x}{2}\right)}_{\supseteq \{x\}} \supseteq \bigcup_{x \in G} \{x\} = G \end{aligned}$$

Damit gilt also Gleichheit und $\mathcal{G}_0$ ist eine Basis. Da $S_0$ abzählbar ist, ist $\mathcal{G}_0$ abzählbar.

($\Leftarrow$) Sei $\mathcal{G}_0$ abzählbare Basis von $\mathcal{G}$ und sei oBdA $\emptyset \notin \mathcal{G}_0$. Wähle für jedes $G \in \mathcal{G}_0$ ein $x_G \in G$ fest aus. Setze $S_0 := \{x_G : G \in \mathcal{G}_0\}$. Dann ist $S_0$ auch abzählbar. Es verbleibt die Dichtheit zu zeigen. Sei $x \in \mathcal{S}$ und $\varepsilon > 0$ beliebig. Da $B(x,\varepsilon)$ offen und $\mathcal{G}_0$ eine Basis ist, gilt:

$$\exists\, \mathcal{G}_{x,\varepsilon} \subseteq \mathcal{G}_0 \text{ mit } B(x,\varepsilon) = \bigcup_{G \in \mathcal{G}_{x,\varepsilon}} G \;\Rightarrow\; G \subseteq B(x,\varepsilon) \qquad \forall G \in \mathcal{G}_{x,\varepsilon}$$

Wähle ein $G$ von diesen aus. Dann gilt:

$$x_G \in G \subseteq B(x,\varepsilon) \;\Rightarrow\; x_G \in B(x,\varepsilon) \;\Rightarrow\; d(\underbrace{x_G}_{\in S_0}, x) < \varepsilon$$

□

**Satz 2.10** Seien $(\mathcal{S}, d)$ und $(\mathcal{S}', d')$ metrische Räume.

(1) Auf $\mathcal{S} \times \mathcal{S}'$ sind Metriken definiert durch

$$\begin{aligned}
d_1\left((x,x'),(y,y')\right) &:= \left((d(x,y))^2 + (d'(x',y'))^2\right)^{\frac{1}{2}} && \forall (x,x'),(y,y') \in \mathcal{S} \times \mathcal{S}' \\
d_2\left((x,x'),(y,y')\right) &:= \max\left\{d(x,y), d'(x',y')\right\} && \forall (x,x'),(y,y') \in \mathcal{S} \times \mathcal{S}' \\
d_3\left((x,x'),(y,y')\right) &:= d(x,y) + d'(x',y') && \forall (x,x'),(y,y') \in \mathcal{S} \times \mathcal{S}'
\end{aligned}$$

(2) Die Metriken $d_1$, $d_2$ und $d_3$ induzieren dieselbe Topologie $\mathcal{G}(\mathcal{S} \times \mathcal{S}')$ auf $\mathcal{S} \times \mathcal{S}'$, die sogenannte **Produkttopologie** von $\mathcal{G}(\mathcal{S})$ und $\mathcal{G}(\mathcal{S}')$.

(3) $\mathcal{G}(\mathcal{S} \times \mathcal{S}') = \left\{\bigcup_{\substack{G \in \mathcal{O} \\ G' \in \mathcal{O}'}} G \times G' : \mathcal{O} \subseteq \mathcal{G}(\mathcal{S}), \mathcal{O}' \subseteq \mathcal{G}(\mathcal{S}')\right\}$, d.h.

$$\{G \times G' : G \in \mathcal{G}(\mathcal{S}), G' \in \mathcal{G}(\mathcal{S}')\} \qquad \text{(Menge offener Rechtecke)}$$

bildet eine Basis von $\mathcal{G}(S \times \mathcal{S}')$.

*Beweis.* (zu 1) Überprüfung der Eigenschaften einer Metrik (zur Übung).

(zu 2) Punktweise gelten die Beziehungen:

$$d_2 \leq d_1 \leq \sqrt{2} \cdot d_2, \qquad \frac{1}{\sqrt{2}} \cdot d_3 \leq d_1 \leq d_3, \qquad d_2 \leq d_3 \leq 2 \cdot d_2$$

Beachte beim Nachweis, dass die $d_i$'s als Metriken größer oder gleich Null sind. Aus obigen Beziehungen folgt u. a.:

$$B_{d_2}\left(x, \frac{r}{\sqrt{2}}\right) \subseteq B_{d_1}(x,r)$$

denn

$$r > \sqrt{2} \cdot d_2(y,x) \geq d_1(y,x) \tag{2.1}$$

Damit folgt aus $G$ $d_1$-offen schon, dass $G$ $d_2$-offen ist, denn für $x \in G$ existiert $r > 0$ mit $B_{d_1}(x,r) \subseteq G$ und (2.1) liefert $B_{d_2}(x, \frac{r}{\sqrt{2}}) \subseteq G$. Damit ist $x$ auch ein ein $d_2$-innerer Punkt. Die anderen Relationen gelten analog.

(zu 3) **⊆:** Sei $G^* \in \mathcal{G}(\mathcal{S} \times \mathcal{S}')$ eine offene Menge in der Produkttopologie. Dann gilt

$$\forall x^* = (x, x') \in G^* : \exists r = r_{x^*} > 0 : \quad G^* = \bigcup_{x^* \in G^*} B\left(x^*, r_{x^*}\right).$$

Wegen Punkt (2) sei oBdA. $\mathcal{S}^* := \mathcal{S} \times \mathcal{S}'$ versehen mit der Metrik $d_2$. Dann gilt:

$$\begin{aligned} B_{d_2}\left(x^*, r_{x^*}\right) &= \left\{(y, y') \in \mathcal{S} \times \mathcal{S}' : \max\left\{d(x, y), d'(x', y')\right\} < r_{x^*}\right\} \\ &= \left\{(y, y') \in \mathcal{S} \times \mathcal{S}' : d(x, y) < r_{x^*} \wedge d'(x', y') < r_{x^*}\right\} \\ &= \underbrace{B_d\left(x, r_{x^*}\right)}_{\in \mathcal{G}(\mathcal{S})} \times \underbrace{B_{d'}\left(x', r_{x^*}\right)}_{\in \mathcal{G}(\mathcal{S}')} \end{aligned}$$

**⊇:** Sei zunächst $G \times G'$ $G, G'$ offen und $x^* = (x, x') \in G \times G'$. Also ist $x \in G$ und $x' \in G'$ und somit

$$\exists\, r, r' > 0 : B_d(x, r) \subseteq G \wedge B_{d'}(x', r') \subseteq G'$$

Setze $r^* := \min\{r, r'\} > 0$. Damit folgt

$$\begin{aligned} B_{d_2}\left(x^*, r^*\right) &\subseteq B_d(x, r) \times B_{d'}\left(x', r'\right) \\ &\subseteq G \times G' = G^* \\ &\Rightarrow G \times G' \in \mathcal{G}(\mathcal{S} \times \mathcal{S}') \\ &\Rightarrow \bigcup_{\substack{G \in \mathcal{O} \\ G' \in \mathcal{O}'}} G \times G' \subseteq \mathcal{G}(\mathcal{S} \times \mathcal{S}') \qquad \forall \mathcal{O} \subseteq \mathcal{G}(\mathcal{S}), \mathcal{O}' \subseteq \mathcal{G}(\mathcal{S}') \end{aligned}$$

da die Produkttopologie vereinigungsstabil ist. □

**Definition 2.11** Die Metriken $d_1, d_2$ und $d_3$ heißen **Produktmetriken**. Daher alternative Schreibweise $d \times d'$, also z. B. $d \times d' := \max\{d, d'\}$ usw.

**Bemerkung 2.12** Analog zu obigen Definitonen lassen sich Produktmetriken für endlich viele metrische Räume $(\mathcal{S}_i, d_i)_{i \in \{1, \ldots, k\}}$ definieren, z. B.

$$d_1 \times \cdots \times d_k := \left(\sum_{i=1}^{k} d_i^2\right)^{\frac{1}{2}},$$

die wiederum dieselbe Produkttopologie induzieren. Die bisherigen Resultate gelten analog.

# Kapitel 3
# ZUFALLSVARIABLEN IN METRISCHEN RÄUMEN

## 3.1 Messbarkeit in metrischen Räumen

**Definition 3.1** Die durch den metrischen Raum $(\mathcal{S}, d)$ induzierte $\sigma$-Algebra

$$\mathcal{B} := \mathcal{B}(\mathcal{S}) := \mathcal{B}_d(\mathcal{S}) := \sigma\left((\mathcal{G}(\mathcal{S})\right) := \sigma(\mathcal{G}).$$

heißt **Borel-$\sigma$-Algebra**. Elemente $B \in \mathcal{B}(\mathcal{S})$ heißen **Borel-Mengen** in $\mathcal{S}$.

Beachte: $\mathcal{B}(\mathcal{S}) = \mathcal{B}_d(\mathcal{S})$ hängt im Allgemeinen von der Metrik $d$ ab.

**Lemma 3.2** Es gilt:

(1) $\mathcal{B}(\mathcal{S}) = \sigma\left(\mathcal{F}(\mathcal{S})\right)$

(2) Ist $f\colon (\mathcal{S}, d) \to (\mathcal{S}', d)$ stetig, so ist $f$ auch $\mathcal{B}_d(\mathcal{S})$-$\mathcal{B}_d(\mathcal{S}')$-messbar.

(3) Sei $\mathcal{G}_0$ abzählbare Basis von $\mathcal{G}(\mathcal{S})$. Dann gilt $\sigma(\mathcal{G}_0) = \mathcal{B}(\mathcal{S})$.

*Beweis.* (zu a) $\subseteq$: Sei $G \in \mathcal{G}(\mathcal{S})$. Dann gilt ist $G^{\complement} \in \mathcal{F}(\mathcal{S}) \subseteq \sigma\left(\mathcal{F}(\mathcal{S})\right)$. Da $\sigma\left(\mathcal{F}(\mathcal{S})\right)$ stabil unter Bildung von Komplementen ist, ist auch $G = \left(G^{\complement}\right)^{\complement} \in \sigma(\mathcal{F}(\mathcal{S}))$. Weiter folgt aus $\mathcal{G} \subseteq \sigma(\mathcal{F})$ schon $\sigma(\mathcal{G}) \subseteq \sigma(\mathcal{F})$.

$\supseteq$: analog

(zu b) Per Definition gilt $f^{-1}\left(\mathcal{B}_{d'}(\mathcal{S}')\right) = f^{-1}\left(\sigma\left(\mathcal{G}(\mathcal{S}')\right)\right)$. Ein Satz aus der Maßtheorie liefert uns weiter $f^{-1}\left(\sigma\left(\mathcal{G}(\mathcal{S}')\right)\right) = \sigma\left(f^{-1}\left(\mathcal{G}(\mathcal{S}')\right)\right)$. Bekannterweise sind für stetige Funktionen Urbilder offener Mengen wieder offen, d.h. es gilt $f^{-1}\left(\mathcal{G}(\mathcal{S}')\right) \subseteq \mathcal{G}(S)$. Somit ist $\sigma\left(f^{-1}\left(\mathcal{G}(\mathcal{S}')\right)\right) \subseteq \sigma\left(\mathcal{G}(\mathcal{S})\right) = \mathcal{B}(\mathcal{S})$.

(zu c) $\subseteq$: klar wegen $\mathcal{G}_0 \subseteq \mathcal{G}(\mathcal{S})$ und $\sigma$ monoton

$\supseteq$: Sei $G \in \mathcal{G}$. Dann existieren geeignete $G_i \in \mathcal{G}_0 \subseteq \sigma(\mathcal{G}_0)$, sodass $G = \bigcup_{i\in\mathbb{N}} G_i$. Damit ist $G \in \sigma(\mathcal{G}_0)$. Aus der Stabilität unter Vereinigungen folgt die Behauptung:

$$G = \bigcup_{i\in\mathbb{N}} G_i \subseteq \sigma\left(\mathcal{G}_0\right) = \mathcal{B}(\mathcal{S}) \in \sigma(\mathcal{G}_0),$$

da $\sigma$-Algebren *abzählbar* vereinigungsstabil sind. □

Vorlesung „Mathematische Statistik" bei Prof. Dr. Ferger im Wintersemester 2020/21

**Satz 3.3** Sei $(\mathcal{S}, d)$ separabler metrischer Raum. Dann gilt:

$$\mathcal{B}_{d\times d}(\mathcal{S}\times\mathcal{S}) = \mathcal{B}(S)\otimes\mathcal{B}(\mathcal{S})$$

Ohne Separabilität gilt nur „$\supseteq$“. Für $S_1$ und $S_2$ separabel gilt $\mathcal{B}_{d_1\times d_2}(\mathcal{S}_1\times\mathcal{S}_2) = \mathcal{B}(S_1)\otimes\mathcal{B}(\mathcal{S}_2)$ analog, auch erweiterbar auf endliche Produkte.

*Beweis.* Seien

$$\begin{aligned} \pi_1 &: \mathcal{S}\times\mathcal{S}\to\mathcal{S} \text{ mit } \pi_1(x,y) := x \qquad \forall (x,y)\in\mathcal{S}\times\mathcal{S} \\ \pi_2 &: \mathcal{S}\times\mathcal{S}\to\mathcal{S} \text{ mit } \pi_2(x,y) := y \qquad \forall (x,y)\in\mathcal{S}\times\mathcal{S} \end{aligned}$$

die **Projektionsabbildungen**. Dann gilt

$$\begin{aligned} \mathcal{B}(\mathcal{S})\otimes\mathcal{B}(\mathcal{S}) &= \sigma(\pi_1,\pi_2) = \sigma\left(\pi_1^{-1}\left(\sigma(\mathcal{G})\right)\cup\pi_2^{-1}\left(\sigma(\mathcal{G})\right)\right) && \text{(Definition)} \\ &= \sigma\left(\sigma\left(\pi_1^{-1}(\mathcal{G})\right)\cup\sigma\left(\pi_2^{-1}(\mathcal{G})\right)\right) \\ &\overset{(\star)}{=} \sigma\left(\pi_1^{-1}(\mathcal{G})\cup\pi_2^{-1}(\mathcal{G})\right) \\ &= \sigma\left(\{G\times S, S\times G' : G, G'\in\mathcal{G}\}\right) \\ &= \sigma\left(\left\{\overbrace{G\times G'}^{=(G\times S)\cap(S\times G')} : G, G'\in\mathcal{G}\right\}\right) \\ & && \text{(„}\subseteq\text{“, da } S\in\mathcal{G}\text{ ; „}\supseteq\text{“, da } \sigma\text{-Algebra } \cap\text{-stabil)} \\ &\overset{(\star\star)}{=} \sigma\left(\left\{\bigcup_{\substack{G\in\mathcal{O}\\ G'\in\mathcal{O}'}} G\times G' : \mathcal{O},\mathcal{O}'\subseteq\mathcal{G}\right\}\right) \\ &= \sigma\left(\mathcal{G}(\mathcal{S}\times\mathcal{S})\right) && \text{(Satz 2.10, Punkt (3))} \\ &\overset{\text{Def}}{=} \mathcal{B}(\mathcal{S}\times\mathcal{S}) \end{aligned}$$

Zum Nachweis von $(\star)$:

$\supseteq$: Setze $\mathcal{E} := \underbrace{\sigma\left(\pi_1^{-1}(\mathcal{G})\right)}_{\supseteq\pi_1^{-1}(\mathcal{G})}\cup\underbrace{\sigma\left(\pi_2^{-1}(\mathcal{G})\right)}_{\supseteq\pi_2^{-1}(\mathcal{G})} \supseteq \pi_1^{-1}(\mathcal{G})\cup\pi_2^{-1}(\mathcal{G}) =: \mathcal{H}$. Dann ist $\sigma(\mathcal{E})\supseteq\sigma(\mathcal{H})$.

$\subseteq$: Es ist $\pi_1^{-1}(\mathcal{G})\subseteq\left(\pi_1^{-1}(\mathcal{G})\cup\pi_2^{-1}(\mathcal{G})\right) = \mathcal{H}$. Also ist auch $\sigma\left(\pi_1^{-1}(\mathcal{G})\right)\subseteq\sigma(\mathcal{H})$. Analog erhalten wir auch $\sigma\left(\pi_2^{-1}(\mathcal{G})\right)\subseteq\sigma(\mathcal{H})$. Dann gilt

$$\mathcal{E} = \underbrace{\sigma\left(\pi_1^{-1}(\mathcal{G})\right)}_{\subseteq\sigma(\mathcal{H})}\cup\underbrace{\sigma\left(\pi_2^{-1}(\mathcal{G})\right)}_{\subseteq\sigma(\mathcal{H})}\subseteq\sigma(\mathcal{H}) \ ,$$

also auch $\sigma(\mathcal{E})\subseteq\sigma(\mathcal{H})$.

Bleibt Nachweis von $(\star\star)$:

$\subseteq$: ist klar mit $\mathcal{O}$ und $\mathcal{O}'$ einelementig (gilt auch ohne Separabilität)

$\supseteq$: Gemäß Satz 2.9 existiert abzählbare Basis $\mathcal{G}_0$ von $\mathcal{G}$. Seien $\mathcal{O}, \mathcal{O}\cdot \subseteq \mathcal{G}$. Sei

$$G^* = \bigcup_{\substack{G \in \mathcal{O} \\ G' \in \mathcal{O}'}} G \times G' = \bigcup_{\substack{G, G' \text{ offen} \\ G \times G' \subseteq G^*}} G \times G' \overset{(!)}{=} \bigcup_{\substack{G_0, G_0' \in \mathcal{G}_0 \\ G \times G_0' \subseteq G^*}} G_0 \times G_0'$$

eine abzählbare Vereinigung, da $\mathcal{G}_0$ Basis ist, also abzählbar. Somit gilt dann $G^* \in \sigma(\{G \times G' : G, G' \in \mathcal{G}\})$. □

**Definition 3.4** Sei $(\Omega, \mathcal{A})$ ein Messraum. Eine Abbildung $X\colon \Omega \to \mathcal{S}$, die $\mathcal{A}$-$\mathcal{B}(\mathcal{S})$-messbar ist, heißt **Zufallsvariable** (ZV) in dem metrischen Raum $(\mathcal{S}, d)$ über $(\Omega, \mathcal{A})$.

Sei $\mathbb{P}$ ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{A})$, also $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Das Bildmaß

$$\mathbb{P} \circ X^{-1} := \mathbb{P}_X := \mathcal{L}(X) := \mathcal{L}(X \mid \mathbb{P})$$
$$\left(\mathbb{P} \circ X^{-1}\right)(B) := \mathbb{P}\left(X^{-1}(B)\right) = \mathbb{P}\left(\{\omega \in \Omega : X(\omega) \in B\}\right) =: \mathbb{P}(X \in B) \qquad \forall B \in \mathcal{B}(\mathcal{S})$$

heißt **Verteilung** von $X$ unter $\mathbb{P}$.

**Satz 3.5** Sei $(\mathcal{S}, d)$ ein separabler metrischer Raum und seien $X, Y$ Zufallsvariablen in $(\mathcal{S}, d)$ über $(\Omega, \mathcal{A})$. Dann ist $d(X, Y)$ eine reelle Zufallsvariable.

*Beweis* (gute Prüfungsfrage). Die Abbildungen $X, Y\colon (\Omega, \mathcal{A}) \to (\mathcal{S}, \mathcal{B}(\mathcal{S}))$ sind messbar genau dann, wenn die Abbildung $(X, Y)\colon (\Omega, \mathcal{A}) \to (\mathcal{S} \times \mathcal{S}, \underbrace{\mathcal{B}(\mathcal{S}) \otimes \mathcal{B}(\mathcal{S})}_{=\mathcal{B}(\mathcal{S} \times \mathcal{S})})$ messbar ist. Jede Metrik ist bekanntlich stetig, also auch $d\colon (\mathcal{S} \times \mathcal{S}, \mathcal{G}(\mathcal{S} \times \mathcal{S})) \to \mathbb{R}$[1]. Dann folgt aus Lemma 3.2, dass $d\colon \mathcal{B}(\mathcal{S} \times \mathcal{S}) \to \mathcal{B}(\mathbb{R})$ messbar ist. Damit folgt die Behauptung, denn $d(X, Y) = d \circ (X, Y)$ ist messbar als Komposition von messbaren Abbildungen. □

## 3.2 Fast sichere Konvergenz

**Definition 3.6** Seien $X, X_n$ $(n \in \mathbb{N})$ Zufallsvariablen in einem separablen, metrischen Raum $(\mathcal{S}, d)$ über $(\Omega, \mathcal{A}, \mathbb{P})$.

$$X_n \overset{n\to\infty}{\longrightarrow} X \quad \mathbb{P}\text{-fast sicher} \quad :\Leftrightarrow \quad \mathbb{P}\left(\left\{\omega \in \Omega : d\left(X_n(\omega), X(\omega)\right) \overset{n\to\infty}{\longrightarrow} 0\right\}\right) = 1$$

Beachte: Die Definition von $M := \left\{\omega \in \Omega : d\left(X_n(\omega), X(\omega)\right) \overset{n\to\infty}{\longrightarrow} 0\right\}$ mengentheoretisch aufgeschrieben (Schnitt $\rightsquigarrow$ „für alle"; Vereinigung $\rightsquigarrow$ „es gibt") ergibt

$$M = \bigcap_{0 < \varepsilon \in \mathbb{Q}} \bigcup_{m \in \mathbb{N}} \bigcap_{n \geq m} \left\{d(X_n, X) < \varepsilon\right\} \overset{\text{Satz 3.5}}{\in} \mathcal{A}, \text{ denn } d(X_n, X)^{-1}\left((-\infty, \varepsilon)\right) \in \mathcal{A}\ .$$

[1] $\mathbb{R}$ sei mit der natürlichen Topologie versehen

Hierbei gilt $M \in \mathcal{A}$ aufgrund der Abzählbarkeit der beteiligten Schnitte und Vereinigungen. Man schreibt auch:

$$\mathbb{P}\left(\lim_{n\to\infty} X_n = X\right) = 1 \quad \text{oder} \quad d(X_n, X) \to 0 \ (n \to \infty) \ \ \mathbb{P}\text{-f. s.}$$

Die bekannten Regeln (Ergebnisse) für *reelle* Zufallsvariablen lassen sich mühelos verallgemeinern; so zum Beispiel der folgende Satz.

**Satz 3.7** Der fast-sichere Grenzwert ist $\mathbb{P}$-fast sicher eindeutig:

$$X_n \overset{n\to\infty}{\longrightarrow} X \quad \mathbb{P}\text{-fast sicher} \ \text{ und } X_n \overset{n\to\infty}{\longrightarrow} X' \quad \mathbb{P}\text{-fast sicher} \quad \Rightarrow \quad X = X' \quad \mathbb{P}\text{-fast sicher}$$

*Beweis.* Es ist $\{X \neq X'\} \subseteq \{X_n \not\to X\} \cup \{X_n \not\to X'\}$ und $\mathbb{P}(X_n \not\to X) + \mathbb{P}(X_n \not\to X') = 0 + 0$. Also ist auch $\mathbb{P}(X \neq X') = 0$ und mit dem Gegenereignis gilt $\mathbb{P}(X = X') = 1 - \mathbb{P}(X \neq X') = 1$. □

**Satz 3.8** Seien $(X_n)_{n\in\mathbb{N}}$ und $X$ Zufallsvariablen im separablen, metrischen Raum $(\mathcal{S}, d)$ und sei $f\colon (\mathcal{S}, d) \to (\mathcal{S}', d')$ $\mathcal{B}_d(S)$ - $\mathcal{B}_{d'}(S')$-messbar und stetig in $X$ $\mathbb{P}$-fast sicher. [a] Dann gilt:

$$X_n \overset{n\to\infty}{\longrightarrow} X \quad \mathbb{P}\text{-fast sicher} \quad \Rightarrow \quad f(X_n) \overset{n\to\infty}{\longrightarrow} f(X) \quad \mathbb{P}\text{-fast sicher}$$

[a] „stetig in $X$ $\mathbb{P}$-fast sicher“ bedeutet, dass $\mathbb{P}(\{\omega \in \Omega : f \text{ in } X(\omega) \text{ stetig}\}) = 1$. Insbesondere ist $\{\omega \in \Omega : f \text{ in } X(\omega) \text{ stetig}\}$ messbar, was keine Selbstverständlichkeit ist.

*Beweis.* Wegen der Folgenstetigkeit gilt $\left\{X_n \overset{n\to\infty}{\longrightarrow} X\right\} \cap \{f \text{ stetig in } X\} \subseteq \left\{f(X_n) \overset{n\to\infty}{\longrightarrow} f(X)\right\}$. Da abzählbare Schnitte von Einsmengen stets Einsmengen sind, d. h.

$$\forall i \in \mathbb{N} : \mathbb{P}(E_i) = 1 \quad \Rightarrow \quad \mathbb{P}\left(\bigcap_{i\in\mathbb{N}} E_i\right) = 1 \qquad \forall E_i \subseteq \Omega \text{ mit } \mathbb{P}(E_i) = 1$$

folgt

$$1 = \mathbb{P}\left(\left\{X_n \overset{n\to\infty}{\longrightarrow} X \text{ und } f \text{ stetig in } X\right\}\right) \leq \mathbb{P}\left(\left\{f(X_n) \overset{n\to\infty}{\longrightarrow} f(X)\right\}\right) \leq 1$$

und somit schon $\mathbb{P}\left(\left\{f(X_n) \overset{n\to\infty}{\longrightarrow} f(X)\right\}\right) = 1$. □

**Satz 3.9 (Konvergenz-Kriterium)**

$$X_n \overset{n\to\infty}{\longrightarrow} X \quad \mathbb{P}\text{-fast sicher} \quad \Leftrightarrow \quad \forall \varepsilon > 0 : \lim_{n\to\infty} \mathbb{P}\left(\sup_{m\geq n} d(X_m, X) > \varepsilon\right) = 0$$

*Beweis.* Man ersetze im Beweis für den Fall reeller Zufallsvariablen $|X_n - X|$ durch $d(X_n, X)$.

Und beachte, dass alle Schlussfolgerungen bestehen bleiben. Hier ausführlich:

$$\begin{aligned}
X_n \overset{n\to\infty}{\longrightarrow} X \ \mathbb{P}\text{-fast sicher} \quad &\Leftrightarrow \quad 1 = \mathbb{P}\left(\bigcup_{0<\varepsilon}\bigcap_{m\in\mathbb{N}}\bigcup_{n\geq m} d(X_n,X) < \varepsilon\right) \\
&\Leftrightarrow \quad 0 = \mathbb{P}\left(\bigcap_{0<\varepsilon}\bigcup_{m\in\mathbb{N}}\bigcap_{n\geq m} d(X_n,X) \geq \varepsilon\right) \\
&\Leftrightarrow \quad 0 = \mathbb{P}\left(\bigcup_{m\in\mathbb{N}}\bigcap_{n\geq m} d(X_n,X) \geq \varepsilon\right) \qquad \forall \varepsilon > 0
\end{aligned}$$

Da für $m \to \infty$ die Mengen $\left\{\bigcap_{n\geq m} d(X_n,X) \geq \varepsilon\right\}$ kleiner werden, ist dies äquivalent zu

$$\begin{aligned}
\forall \varepsilon > 0 \ \forall \delta > 0 \ \exists m \in \mathbb{N}: \quad &\mathbb{P}\left(\bigcap_{n\geq m} d(X_n,X) \geq \varepsilon\right) < \delta \\
\Leftrightarrow \quad \forall \varepsilon > 0: \quad &\lim_{m\to\infty} \mathbb{P}\left(\bigcap_{n\geq m} d(X_n,X) \geq \varepsilon\right) = 0 \\
\Leftrightarrow \quad \forall \varepsilon > 0: \quad &\lim_{m\to\infty} \mathbb{P}\left(\sup_{n\geq m} d(X_n,X) \geq \varepsilon\right) = 0
\end{aligned}$$

□

Ein sehr nützliches Kriterium ist Folgendes:

**Satz 3.10**

$$\forall \varepsilon > 0: \quad \sum_{n\in\mathbb{N}_{\geq 1}} \mathbb{P}\left(d(X_n,X) > \varepsilon\right) < \infty \quad \Longrightarrow \quad X_n \overset{n\to\infty}{\longrightarrow} X \quad \mathbb{P}\text{-fast sicher}$$

*Beweis.* Setze $A_n(\varepsilon) := \{d(X_n,X) > \varepsilon\} \in \mathcal{A}$ wegen Satz 3.5. Dann folgt aus dem *ersten Borel-Cantelli-Lemma* $\mathbb{P}\left(\limsup_{n\to\infty} A_n(\varepsilon)\right) = 0$ für alle $\varepsilon > 0$. Mit

$$\begin{aligned}
\liminf_{n\to\infty}\left(A_n(\varepsilon)^{\complement}\right) \overset{\text{Def}}{=} \bigcup_{m\in\mathbb{N}}\bigcap_{n\geq m}(A_n(\varepsilon))^{\complement} &= \left(\bigcap_{m\in\mathbb{N}}\bigcup_{n\geq m}\{d(X_n,X) > \varepsilon\}\right)^{\complement} \\
&\overset{\text{Def}}{=} \left(\limsup_{n\to\infty} A_n(\varepsilon)\right)^{\complement}
\end{aligned}$$

folgt dann

$$1 = \mathbb{P}\left(\left(\limsup_{n\to\infty} A_n(\varepsilon)\right)^{\complement}\right) = \mathbb{P}\left(\liminf_{n\to\infty}\left(A_n(\varepsilon)^{\complement}\right)\right) \qquad \forall \varepsilon > 0$$

Da abzählbare Durchschnitte von Eins-Mengen (also Mengen mit $\mathbb{P}$-Maß 1) wieder Eins-Mengen sind, folgt schließlich:

$$\mathbb{P}\left(\underbrace{\bigcap_{0<\varepsilon\in\mathbb{Q}}\bigcup_{m\in\mathbb{N}}\bigcap_{n\geq m}\{d(X_n,X) \leq \varepsilon\}}_{\{X_n\to X\}=\{d(X_n,X)\to 0\}}\right) = 1$$

□

**Bemerkung.** Die $\mathbb{P}(d(X_n, X) > \varepsilon)$ werden in der Statistik **Fehlerwahrscheinlichkeit** oder **tail probability** genannt. Um Fehlerwahrscheinlichkeiten abzuschätzen, also die Voraussetzung für diesen Satz für einen speziellen Fall zu zeigen, nutzt man häufig sogenannte Maximalungleichungen wie die Markov-Ungleichung und die Tschebychew-Ungleichung.

Weitere Eigenschaften der fast sicheren Konvergenz von Zufallsvariablen in metrischen Räumen finden sich z. B. in [13, Kapitel 8.2].

## 3.3 Stochastische Konvergenz

**Definition 3.11** $X_n$ **konvergiert stochastisch** bzw. **in Wahrscheinlichkeit** gegen $X$, in Symbolen:

$$X_n \xrightarrow{\mathbb{P}} X \quad :\Leftrightarrow \quad \forall \varepsilon > 0 : \mathbb{P}\left(\{d(X_n, X) > \varepsilon\}\right) \xrightarrow{n\to\infty} 0$$

d. h. $d(X_n, X) \xrightarrow{\mathbb{P}} 0$.

**Satz 3.12** $X_n \xrightarrow{n\to\infty} X \quad \mathbb{P}\text{-fast sicher} \quad \Rightarrow \quad X_n \xrightarrow{\mathbb{P}} X$

*Beweis.* Gemäß Satz 3.9 gilt

$$\forall \varepsilon > 0 : \quad 0 \leq \mathbb{P}\left(d(X_n, X) > \varepsilon\right) \leq \mathbb{P}\left(\sup_{m\geq n} d(X_m, X) > \varepsilon\right) \xrightarrow{n\to\infty} 0 \qquad \square$$

Die Umkehrung von Satz 3.12 gilt im Allgemeinen *nicht*, aber es gilt das folgende Teilfolgenkriterium.

**Satz 3.13 (Teilfolgenkriterium für stochastische Konvergenz)** Folgende Aussagen sind äquivalent:

(1) $X_n \xrightarrow{\mathbb{P}} X$

(2) Zu jeder Teilfolge (TF) $(X_{n'})$ von $(X_n)_{n\in\mathbb{N}}$ existiert eine Teilfolge $(X_{n''})$ von $(X_{n'})$ derart, dass $X_{n''} \xrightarrow{n''\to\infty} X$ $\mathbb{P}$-fast sicher.

Die Notation $(X_{n''})$ stammt aus [2].

*Beweis.* Der Beweis verläuft wie in $\mathbb{R}$.

($\Rightarrow$) Sei $\varepsilon > 0$ beliebig und $(X_{n_k})_k$ eine Teilfolge von $(X_n)_n$. Da $\mathbb{P}\left(\{d(X_n, X) > \varepsilon\}\right) \to 0$, können wir eine Teilteilfolge $(X_{n_{k_i}})_i := (Y_i)_i$ finden mit $\mathbb{P}\left(\{d(Y_i, X) > \varepsilon\}\right) < i^{-2}$. Dann ist für alle $n \in \mathbb{N}$ und mit $n \to \infty$

$$\mathbb{P}\left(\sup_{i\geq n} d(Y_i, X) < \varepsilon\right) = \mathbb{P}\left(\bigcup_{i\geq n} \{d(Y_i, X) < \varepsilon\}\right) \leq \sum_{i=n}^{\infty} \mathbb{P}\left(\{d(Y_i, X) < \varepsilon\}\right) \leq \sum_{i=n}^{\infty} i^{-2} \to 0$$

Laut dem Konvergenz-Kriterium in Satz 3.9 bedeutet das $Y_i \to X$ $\mathbb{P}$-fast sicher.

($\Leftarrow$) Wir nehmen an, es gäbe ein $\varepsilon > 0$ und eine Teilfolge $(X_{n_k})_k$ der Folge $(X_n)_n$, sodass

$\mathbb{P}(d(X_{n_k}, X) > \varepsilon) > \delta > 0$ für ein passendes $\delta$. Dann konvergiert nach Punkt (2) eine Teilteilfolge $(X_{n_{k_i}})_i$ $\mathbb{P}$-fast sicher gegen $X$, also laut Satz 3.12 auch stochastisch. Das widerspricht aber der obigen Annahme über $(X_{n_k})_k$, die sich auch auf $(X_{n_{k_i}})_i$ überträgt.

□

Mit dem Teilfolgenkriterium lassen sich Rechenregeln für fast sichere Konvergenz auf stochastische Konvergenz übertragen.

**Korollar 3.14** (a) $X_n \xrightarrow{\mathbb{P}} X$ und $X_n \xrightarrow{\mathbb{P}} X' \quad\Rightarrow\quad X = X' \quad \mathbb{P}$-fast sicher

(b) Gelte $X_n \xrightarrow{\mathbb{P}} X$ in $(\mathcal{S}, d)$ und $f\colon (\mathcal{S}, d) \to (\mathcal{S}', d')$ messbar mit $f$ stetig in $X$ $\mathbb{P}$-fast sicher. Dann gilt auch $f(X_n) \xrightarrow{\mathbb{P}} f(X)$.

*Beweis.* (zu a) Wegen $X_n \xrightarrow{\mathbb{P}} X$ existiert nach Satz 3.13 eine Teilfolge $(X_{n'}) \subseteq (X_n)_{n\in\mathbb{N}}$ mit $X_{n'} \xrightarrow{n'\to\infty} X$ $\mathbb{P}$-fast sicher. Zu $(X_{n'})$ existiert (wegen $X_n \xrightarrow{\mathbb{P}} X'$ und Satz 3.13) eine Teilfolge $(X_{n''})$ von $(X_{n'})$, sodass aus $X_{n''} \xrightarrow{n\to\infty} X'$ $\mathbb{P}$-fast sicher mit Satz 3.7 schon $X = X'$ fast sicher folgt.

(zu b) Zu einer beliebigen Teilfolge $(X_{n'})$ von $(X_n)$ existiert eine Teilfolge $(X_{n''})$ von $(X_{n'})$ mit $X_{n''} \to X$ $\mathbb{P}$-fast sicher. Mit Satz 3.8 folgt $f(X_{n''}) \to f(X)$ $\mathbb{P}$-fast sicher und mit Satz 3.13 angewendet auf $(f(X_n))_n$ gilt $f(X_n) \xrightarrow{\mathbb{P}} f(X)$. □

## 3.4 Konvergenz in Produkträumen

Seien $(\mathcal{S}, d)$ und $(\mathcal{S}', d')$ separable metrische Räume. Dann ist auch $(\mathcal{S}\times\mathcal{S}', d\times d')$ ein separabler metrischer Raum. Dies folgt z. B. aus dem *Satz von der koordinatenweise Konvergenz*:

$$(a_n, a_n') \xrightarrow{d\times d'} (a, a') \quad\Leftrightarrow\quad (a_n) \xrightarrow{d} a \text{ und } (a_n') \xrightarrow{d'} a' \tag{3.1}$$

Es gibt auch eine „stochastische Versionen" dieses Satzes.

**Satz 3.15 (Koordinatenweise Konvergenz vom Zufallsgrößen)**

(1) $(X_n, X_n') \xrightarrow{d\times d'} (X, X')$ $\mathbb{P}$-f. s. $\Leftrightarrow$ $X_n \xrightarrow{d} X$ $\mathbb{P}$-f. s. und $X_n' \xrightarrow{d'} X'$ $\mathbb{P}$-f. s.

(2) $(X_n, X_n') \xrightarrow{\mathbb{P}} (X, X')$ $\Leftrightarrow$ $X_n \xrightarrow{\mathbb{P}} X$ und $X_n' \xrightarrow{\mathbb{P}} X'$

*Beweis.* (zu 1) Nach Gleichung (3.1) folgt aus $(X_n, X_n') \xrightarrow{d\times d'} (X, X')$ $\mathbb{P}$-fast sicher auch $X_n \to X$ und $X_n' \to X'$ $\mathbb{P}$-fast sicher. Da der Schnitt von Eins-Mengen wieder eine Eins-Menge ist, ist dies wiederum äquivalent zu $X_n \to X$ $\mathbb{P}$-fast sicher und $X_n' \to X'$ $\mathbb{P}$-fast sicher. *Achtung* – die „Stellung" der Fast-Sicherheit spielt eine Rolle!

(zu 2) Die linke Seite ist wegen (3.1) und Satz 3.13 äquivalent dazu, dass für alle Teilfolgen $(X_{n'}, X_{n'}') \subseteq (X_n, X_n')$ eine Teilteilfolge $(X_{n''}, X_{n''}')$ existiert mit $(X_{n''}, X_{n''}') \xrightarrow{n\to\infty} 0$ fast sicher. Also gilt wegen Punkt (1)

$$X_{n''} \to X \text{ f. s. und } X'_{n''} \to X' \text{ f. s.}$$

Somit existiert für alle Teilfolgen $(X_{n'}) \subseteq (X_n)$ eine Teilteilfolge $(X_{n''}) \subseteq (X_{n'})$ mit $X_{n''} \to X$ fast sicher. Dies ist wegen (3.1) äquivalent zu $X_n \xrightarrow{\mathbb{P}} X$. Analog gilt auch $X'_n \to X'$. □

### 3.4.1 Gleichheit in Verteilung

**Definition 3.16** Seien $X_i$ $(i = 1, 2)$ Zufallsvariablen in $(\mathcal{S}, d)$ über $(\Omega_i, \mathcal{A}_i, \mathbb{P}_i)$. Sie heißen **gleich in Verteilung**, in Zeichen

$$X_1 \overset{\mathrm{d}}{=} X_2 \quad :\Leftrightarrow \quad \mathbb{P}_1 \circ X_1^{-1} = \mathbb{P}_2 \circ X_2^{-1}$$

Wir wollen Verteilungsgleichheit in folgendem Satz charakterisieren.

**Satz 3.17** Es gilt:

(a) Seien $P$ und $Q$ Wahrscheinlichkeitsmaße auf $\mathcal{B}(\mathcal{S})$. Dann gilt:

$$P \equiv Q \quad \Leftrightarrow \quad \int f \ \mathrm{d}P = \int f \ \mathrm{d}Q \qquad \forall f \in C^b(\mathcal{S}) \text{ glm. stetig}$$

(b) Es gilt

$$X \overset{\mathrm{d}}{=} Y \quad \Leftrightarrow \quad \mathbb{E}\left[f(X)\right] = \int_{\Omega_1} f(X_1) \ \mathrm{d}\mathbb{P}_1 = \int_{\Omega_2} f(X_2) \ \mathrm{d}\mathbb{P}_2 = \mathbb{E}\left[f(Y)\right]$$

für alle $f \in C^b(\mathcal{S})$ gleichmäßig stetig.

*Beweis.* (zu a) **(⇒)** Klar.

**(⇐)** Aus Lemma 3.2, Punkt (1) wissen wir, dass $\mathcal{B}(\mathcal{S}) = \sigma(\mathcal{F}(\mathcal{S}))$ und $\mathcal{F}$ ist durchschnittsstabil. Wegen des *Maßeindeutigkeitssatz* reicht es zu zeigen, dass $\mathbb{P}(F) = \mathbb{Q}(F)$ für alle $F \in \mathcal{F}(S)$. Sei nun $F \subseteq \mathcal{S}$ abgeschlossen. Setze $f_k(x) := \varphi(k \cdot d(x, F))$ (vgl. Satz 2.4, $\varphi$ wie dort). Aus Lemma 2.3 folgt, dass die $f_k$ beschränkt und gleichmäßig stetig sind mit $f_k \searrow \mathbb{1}_F$ für $k \to \infty$. Also gilt:

$$\mathbb{P}(F) = \int \mathbb{1}_F \ \mathrm{d}\mathbb{P} = \int \lim_{k\to\infty} f_k \ \mathrm{d}\mathbb{P}$$

und mit monotoner Konvergenz

$$= \lim_{k\to\infty} \int f_k \ \mathrm{d}\mathbb{P} = \lim_{k\to\infty} \int f_k \ \mathrm{d}\mathbb{Q}$$

Wiederum mit monotoner Konvergenz

$$= \int \lim_{k\to\infty} f_k \, \mathrm{d}\mathbb{Q} = \int \mathbb{1}_F \, \mathrm{d}\mathbb{Q} = \mathbb{Q}(F)$$

Da $F$ beliebig war, folgt die Behauptung.

(zu b) Dies folgt aus Punkt (a) mit dem Transformationssatz (A.1):

$$\begin{aligned} X \overset{\mathrm{d}}{=} Y &\overset{3.16}{\Leftrightarrow} \mathbb{P}\circ X^{-1} = \mathbb{P}\circ Y^{-1} \\ &\overset{(a)}{\Leftrightarrow} \int_{\mathcal{S}} f \, \mathrm{d}(\mathbb{P}\circ X^{-1}) = \int_{\mathcal{S}} f \, \mathrm{d}(\mathbb{P}\circ Y^{-1}) \qquad \forall f \in C^b(\mathcal{S}) \text{ glm. stetig} \\ &\Leftrightarrow \mathbb{E}\left[f(X)\right] = \mathbb{E}\left[f(Y)\right] \qquad \forall f \in C^b(\mathcal{S}) \text{ glm. stetig} \end{aligned}$$

Verwende dafür im letzten Schritt den Transformationssatz, d.h.

$$\int_{\mathcal{S}} f \, \mathrm{d}(\mathbb{P}\circ X^{-1}) \overset{\text{Trafo}}{=} \int_{\Omega} f\circ X \, \mathrm{d}\mathbb{P} \overset{\text{Def}}{=} \mathbb{E}\left[f\circ X\right] = \mathbb{E}\left[f(X)\right] \qquad \square$$

# Kapitel 4
# VERTEILUNGSKONVERGENZ VON ZUFALLSVARIABLEN IN METRISCHEN RÄUMEN

## 4.1 Satz von Portmanteau

Seien $X$ und $(X_n)_{n\in\mathbb{N}}$ Zufallsvariablen in $(\mathcal{S}, d)$ über $(\Omega, \mathcal{A}, \mathbb{P})$. Dann sind

$$P := \mathbb{P} \circ X^{-1} \qquad \text{und} \qquad P_n := \mathbb{P} \circ X_n^{-1} \qquad (n \in \mathbb{N})$$

Wahrscheinlichkeitsmaße auf $\mathcal{B}(\mathcal{S})$.

**Definition 4.1** (1) Seien $P$ und $(P_n)_{n\in\mathbb{N}}$ Wahrscheinlichkeitsmaße auf $\mathcal{B}(S)$. Dann **konvergiert** $P_n$ **schwach** gegen $P$, in Zeichen

$$P_n \xrightarrow{\text{w}} P \;:\Leftrightarrow\; \int f \;\mathrm{d}P_n \overset{n\to\infty}{\longrightarrow} \int f \;\mathrm{d}P \qquad \forall f \in C^b(\mathcal{S})$$

(2) $X_n$ **konvergiert in Verteilung** gegen $X$ in Raum $(\mathcal{S}, d)$, in Zeichen

$$X_n \xrightarrow{\text{d}} X \text{ in } (\mathcal{S}, d) \;:\Leftrightarrow\; \mathbb{P} \circ X_n^{-1} \xrightarrow{\text{w}} \mathbb{P} \circ X^{-1}$$

Das d steht für „distribution“. Alternative Schreibweise: $X_n \xrightarrow{\mathcal{L}} X$.

Der folgende Satz gibt eine äquivalente Charakterisierung von $\xrightarrow{\text{w}}$ bzw. $\xrightarrow{\text{d}}$.

**Theorem 4.2 (Portmanteau-Theorem)** Folgende Aussagen sind äquivalent:

(1) $P_n \xrightarrow{\text{w}} P$

(2) $\int f \;\mathrm{d}P_n \longrightarrow \int f \;\mathrm{d}P \qquad \forall f \in C^b(\mathcal{S})$ glm. stetig

(3) $\limsup_{n\to\infty} P_n(F) \leq P(F) \qquad \forall F \in \mathcal{F}(\mathcal{S})$

(4) $\liminf_{n\to\infty} P_n(G) \geq P(G) \qquad \forall G \in \mathcal{G}(\mathcal{S})$

(5) $\lim_{n\to\infty} P_n(B) = P(B) \qquad \forall B \in \mathcal{B}(\mathcal{S})$ mit $P(\underbrace{\partial B}_{\in\mathcal{F}(\mathcal{S})}) = 0$

Die Mengen $B \in \mathcal{B}(\mathcal{S})$ mit $P(\partial B) = 0$ heißen $P$**-randlos**.

*Beweis.* **(1) ⇒ (2):** Folgt aus Definition 4.1.

**(2) ⇒ (3):** Sei $F \in \mathcal{F}(\mathcal{S})$. Der Beweis von Satz 3.17 zeigt, dass es eine Folge $(f_k)_{k\in\mathbb{N}}$ von gleichmäßig stetigen und beschränkten Funktionen auf $\mathcal{S}$ gibt mit $f_k \downarrow \mathbb{1}_F$. Dann gilt:

$$\limsup_{n\to\infty} P_n(F) = \limsup_{n\to\infty} \int \underbrace{\mathbb{1}_F}_{\le f_k} \ \mathrm{d}P_n \overset{\text{Mon.}}{\le} \limsup_{n\to\infty} \int f_k \ \mathrm{d}P_n \overset{\text{Vor.}}{=} \int f_k \ \mathrm{d}P \quad \forall k \in \mathbb{N} \ .$$

Mit monotoner Konvergenz folgt daraus $\int f_k \ \mathrm{d}P \longrightarrow \int \mathbb{1}_F \ \mathrm{d}P = P(F)$ und somit für $k \to \infty$ auch (3).

**(3) $\Leftrightarrow$ (4):** Nutze den Übergang zum Komplement sowie die Rechenregeln für lim inf und lim sup. Sei $G$ offen, das heißt $G \in \mathcal{G}(\mathcal{S})$ und damit $G^{\mathsf{C}} \in \mathcal{F}(\mathcal{S})$. Dann gilt

$$\liminf_{n\to\infty} P_n(G) = \liminf_{n\to\infty} \Big(1 - \mathbb{P}_n(G^{\mathsf{C}})\Big) = 1 - \underbrace{\limsup_{n\to\infty} P_n(G^{\mathsf{C}})}_{\le P(G^{\mathsf{C}})} \ge 1 - P(G^{\mathsf{C}}) = P(G) \ .$$

**(3) $\Rightarrow$ (1):** Sei $f \in C^b(\mathcal{S})$ beliebig. Wir zeigen zunächst, dass

$$\limsup_{n\to\infty} \int f \ \mathrm{d}P_n \le \int f \ \mathrm{d}P \tag{4.1}$$

**Schritt 1:** Sei $0 \le f < 1$. Setze $F_i := \left\{f \ge \frac{i}{k}\right\} = \left\{x \in \mathcal{S} : f(x) \ge \frac{i}{k}\right\} = f^{-1}([\frac{i}{k}, \infty))$ für alle $0 \le i \le k$ und $k \in \mathbb{N}$. Dann gilt $F_i \in \mathcal{F}$ für alle zulässigen $i$, da die Urbilder abgeschlossener Mengen und der stetigen Funktion $f$ wieder abgeschlossen sind. Wegen

$$\int_{\mathcal{S}} f \ \mathrm{d}P \overset{\text{Lin.}}{=} \sum_{i=1}^{k} \int \mathbb{1}_{\left\{\frac{i-1}{k} \le f < \frac{i}{k}\right\}} \cdot f \ \mathrm{d}P$$

folgt mit der Monotonie des Integrals und der Abschätzung des Integranten nach oben und unten mittels der Indikatorfunktion

$$\sum_{i=1}^{k} \underbrace{\frac{i-1}{k}}_{=\frac{i}{k}-\frac{1}{k}} \cdot P\left(\frac{i-1}{k} \le f < \frac{i}{k}\right) \le \int f \ \mathrm{d}P \le \sum_{i=1}^{k} \frac{1}{k} \cdot P\Bigg(\underbrace{\frac{i-1}{k} \le f < \frac{i}{k}}_{F_{i-1}\setminus F_i}\Bigg) \ . \tag{4.2}$$

Die rechte Summe in (4.2) ist gleich

$$\begin{aligned}
&\frac{1}{k} \cdot \sum_{i=1}^{k} i \cdot (P(F_{i-1}) - P(F_i)) \quad \text{da } F_i \subseteq F_{i-1} \\
&= \frac{1}{k} \cdot \Big(P(F_0) - P(F_1) + P(F_1) - 2 \cdot P(F_2) + 3 \cdot P(F_2) - 3 \cdot P(F_3) + \\
&\qquad + \cdots + (k-1) \cdot P(F_{k-2}) - (k-1) \cdot P(F_{k-1}) + k \cdot P(F_{k-1}) - k \cdot P(F_k)\Big) \\
&= \frac{1}{k} \cdot \Big(\underbrace{P(F_0)}_{=1} + P(F_1) + P(F_2) + \cdots + P(F_{k-1}) - k \cdot \underbrace{P(F_k)}_{=0}\Big) \quad (0 \le f < 1) \\
&= \frac{1}{k} + \frac{1}{k} \cdot \sum_{i=1}^{k-1} P(F_i) \ .
\end{aligned}$$

Da die linke Summe in (4.2) gleich der rechten Summe in (4.2) minus $\frac{1}{k}$ ist, folgt

$$\frac{1}{k}\sum_{i=1}^{k-1} P(F_i) \leq \int f \ \mathrm{d}P \leq \frac{1}{k} + \frac{1}{k}\sum_{i=1}^{k-1} P(F_i) \ . \tag{4.3}$$

Man beachte, dass (4.3) für *jedes* Wahrscheinlichkeitsmaß $P$ gilt, also insbesondere auch für $P_n$. Damit folgt

$$\begin{aligned}
\limsup_{n\to\infty} \int f \ \mathrm{d}P_n &\overset{(4.3)}{\leq} \frac{1}{k} + \frac{1}{k}\sum_{i=1}^{k-1} \underbrace{\limsup_{n\to\infty} P_n(F_i)}_{\overset{(3)}{\leq} P(F_i) \ \forall i} \\
&\overset{(3)}{\leq} \frac{1}{k} + \underbrace{\frac{1}{k}\sum_{i=1}^{k-1} P(F_i)}_{\overset{(4.3)}{\leq} \int f \ \mathrm{d}P} \\
&\overset{(4.3)}{\leq} \frac{1}{k} + \int f \ \mathrm{d}P
\end{aligned}$$

für alle $k \in \mathbb{N}$. Der Grenzübergang $k \to \infty$ liefert schließlich (4.1).

**Schritt 2:** Da $f \in C^b(\mathcal{S})$ beliebig, gilt wegen Beschränktheit von $f$, dass $a, b \in \mathcal{S}$ existieren, sodass $a \leq f < b$ gilt. Damit ist die Funktion

$$g(x) := \frac{f(x) - a}{b - a}$$

stetig und es gilt $0 \leq g < 1$. Mit dem Ergebnis von Schritt 1 erhalten wir

$$\begin{aligned}
\limsup_{n\to\infty} \int f \ \mathrm{d}P_n &= \limsup_{n\to\infty} \int (b-a) \cdot g + a \ \mathrm{d}P_n \\
&= \limsup_{n\to\infty} \left( (b-a) \cdot \int g \ \mathrm{d}P_n + a \right) \\
&\leq (b-a) \cdot \underbrace{\limsup_{n\to\infty} \int g \ \mathrm{d}P_n}_{\leq \int g \ \mathrm{d}P} + a \\
&\leq (b-a) \cdot \int g \ \mathrm{d}P + a \\
&\overset{\text{Lin.}}{=} \int f \ \mathrm{d}P
\end{aligned}$$

Damit ist (4.1) gezeigt. Der Übergang zu $-f$ in (4.1) liefert

$$\begin{aligned}
\liminf_{n\to\infty} \int f \ \mathrm{d}P_n = \liminf_{n\to\infty} - \int -f \ \mathrm{d}P_n &= - \limsup_{n\to\infty} \int \underbrace{-f}_{\in C^b(\mathcal{S})} \ \mathrm{d}P \\
&\overset{(4.1)}{\geq} - \int -f \ \mathrm{d}P \\
&\overset{\text{Lin.}}{=} \int f \ \mathrm{d}P \qquad \forall f \in C^b(\mathcal{S}) \ .
\end{aligned}$$

Daraus folgt nun

$$\int f \ \mathrm{d}P \leq \liminf_{n\to\infty} \int f \ \mathrm{d}P_n \leq \limsup_{n\to\infty} \int f \ \mathrm{d}P_n \overset{(4.1)}{\leq} \int f \ \mathrm{d}P \ ,$$

was via $\lim_{n\to\infty} \int f \ \mathrm{d}P_n = \int f \ \mathrm{d}P$ für alle $f \in C^b(S)$ nun (1) impliziert.

**(3) $\Rightarrow$ (5):** Sei $B \in \mathcal{B}(\mathcal{S})$ mit $P(\partial B) = 0$. Zunächst gilt

$$0 = P(\overbrace{\partial B}^{\overline{B}\setminus B^\circ}) = P(\overline{B}) - P(B^\circ) \Rightarrow P(B^\circ) \leq P(B) \leq P(\overline{B}) = P(B^\circ) \ .$$

Damit folgt dann durch

$$\begin{aligned} P(\overline{B}) &\overset{(3)}{\geq} \limsup_{n\to\infty} \underbrace{P_n(\overline{B})}_{\geq P_n(B)} \geq \limsup_{n\to\infty} P_n(B) \geq \liminf_{n\to\infty} \underbrace{P_n(B)}_{\geq P_n(B^\circ)} \\ &\overset{(3)\Leftrightarrow(4)}{\geq} P\left(B^\circ\right) = P(\overline{B}) = P(B) \end{aligned}$$

die Gleichheit von Limes inferior und Limes superior und somit existiert der Grenzwert $\lim_{n\to\infty} P_n(B) = P(B)$.

**(5) $\Rightarrow$ (3):** Sei $F \in \mathcal{F}$ beliebig. Dann gilt für alle $\varepsilon > 0$

$$\partial \left\{x \in \mathcal{S} : d(x,F) \leq \varepsilon\right\} \subseteq \left\{x \in \mathcal{S} : d(x,F) = \varepsilon\right\} \tag{4.4}$$

denn: Sei $x \in \partial \left\{x \in \mathcal{S} : d(x,F) \leq \varepsilon\right\}$. Dann gilt mit der Charakterisierung des Abschlusses als Grenzwerte von Folgen einer Menge:

$$\begin{aligned} &\exists (x_n)_{n\in\mathbb{N}} : \forall n \in \mathbb{N} : d(x_n, F) \leq \varepsilon \text{ und } \lim_{n\to\infty} x_n = x \ , \\ &\exists (\zeta_n)_{n\in\mathbb{N}} : \forall n \in \mathbb{N} : d(\zeta_n, F) > \varepsilon \text{ und } \lim_{n\to\infty} \zeta_n = x \ . \end{aligned}$$

Da $d(\cdot, F)$ stetig ist gemäß Lemma 2.3 (3), folgt

$$\varepsilon \leq \lim_{n\to\infty} d(\zeta_n, F) = d(x,F) = \lim_{n\to\infty} d(x_n, F) \leq \varepsilon \ .$$

Wegen (4.4) sind

$$A_\varepsilon := \partial \left\{x \in \mathcal{S} : d(x,F) \leq \varepsilon\right\} \qquad \forall \varepsilon > 0$$

paarweise disjunkt, da bereits die Obermengen paarweise disjunkt sind. Dann folgt

$$E := \{\varepsilon > 0 : P(A_\varepsilon) > 0\} \text{ ist höchstens abzählbar,} \tag{4.5}$$

denn:

$$E = \bigcup_{m\in\mathbb{N}} \underbrace{\left\{\varepsilon > 0 : P(A_\varepsilon) \geq \frac{1}{m}\right\}}_{=:E_m} .$$

Es gilt $|E_m| \leq m$ — Angenommen es existieren $0 < \varepsilon_1 < \dots < \varepsilon_{m+1}$ mit

$$P(A_{\varepsilon_i}) \geq \frac{1}{m} \quad (1 \leq i \leq m+1) \;\Rightarrow\; 1 \geq P\left(\dot{\bigcup_{i=1}^{m+1}} A_{\varepsilon_i}\right) = \sum_{i=1}^{m+1} \underbrace{P\left(A_{\varepsilon_i}\right)}_{\geq \frac{1}{m}} \geq (m+1) \cdot \frac{1}{m} > 1$$

Das ist ein Widerspruch. Damit ist $E$ abzählbare Vereinigung endlicher Mengen, also höchstens abzählbar unendlich. Damit liegt das Komplement

$$E^{\mathsf{C}} = \{\varepsilon > 0 : P(A_\varepsilon) = 0\}$$

dicht in $[0,\infty)$. (Dies kann man durch Widerspruch zeigen: angenommen es existiert ein $x \geq 0$ mit $x \notin \overline{E^{\mathsf{C}}}$. Gemäß einer Übungsaufgabe existiert dann ein $r > 0$ mit $B(x,r) \cap E^{\mathsf{C}} = \emptyset$, also mit $B(x,r) \subseteq \left(E^{\mathsf{C}}\right)^{\mathsf{C}} = E$. Da $B(x,r)$ überabzählbar ist, müsste auch $E$ überabzählbar sein, was eben nicht der Fall ist). Daraus folgt insbesondere

$$\exists (\varepsilon_k)_{k\in\mathbb{N}} \subseteq \mathbb{R} \text{ mit } \varepsilon_k \downarrow 0 : \forall k \in \mathbb{N} : F_k := \{x \in \mathcal{S} : d(x,F) \leq \varepsilon_k\} \text{ ist } P\text{-randlos.}$$

Beachte $A_{\varepsilon_k} = \partial F_k$. Wähle also $\varepsilon_k \in E^{\mathsf{C}}$ für alle $k \in \mathbb{N}$. Da $F \subseteq F_k$ für alle $k \in \mathbb{N}$ ist, gilt:

$$\limsup_{n\to\infty} P_n(F) \leq \limsup_{n\to\infty} \underbrace{P_n(F_k)}_{\text{konv.}} \overset{(5)}{=} P(F_k) \qquad \forall k \in \mathbb{N}$$

Für $k \to \infty$ folgt daraus $\limsup\limits_{n\to\infty} P_n(F) \leq \lim_{k\to\infty} P(F_k) = P(F)$. Die letzte Gleichheit gilt, weil $P$ $\sigma$-stetig von oben ist und $F_k \downarrow F$. Dies gilt, da $(\varepsilon_k)$ monton fallend ist und somit $F_1 \supseteq F_2 \supseteq \dots$ gilt; sowie $\bigcap\limits_{k\in\mathbb{N}} F_k = F$, denn:

$$x \in \bigcap_{k\in\mathbb{N}} F_k \qquad \Leftrightarrow \qquad x \in F_k \quad \forall k \in \mathbb{N} \qquad \Leftrightarrow \qquad d(x,F) \leq \varepsilon_k \quad \forall k \in \mathbb{N}$$

Dies impliziert dann $d(x,F) = 0$, was wegen Lemma 2.3 äquivalent zu $x \in \overline{F} = F$ ist. □

Falls $\mathcal{U} = \{B \in \mathcal{B}(S) : P(\partial B) = 0\}$, so besagt Theorem 4.2 in Punkt (5), dass $P_n(B) \to P(B)$ für alle $B \in \mathcal{U}$, also punktweise Konvergenz auf der Klasse $\mathcal{U}$, die schwache Konvergenz $P_n \xrightarrow{\text{w}} P$ nach sich zieht. Es gibt weitere solcher Klassen.

**Theorem 4.3** Sei $\mathcal{U} \subseteq \mathcal{B}(\mathcal{S})$ mit

(i) $\mathcal{U}$ ist endlich schnittstabil, d.h. mit $A, B \in \mathcal{U}$ ist auch $A \cap B \in \mathcal{U}$

(ii) Jedes offene $G \subseteq \mathcal{S}$ ist abzählbare Vereinigung von Mengen aus $\mathcal{U}$.

Dann gilt: $P_n(A) \overset{n\to\infty}{\longrightarrow} P(A) \;\forall A \in \mathcal{U} \;\Rightarrow\; P_n \xrightarrow{\text{w}} P$.

*Beweis.* Seien $A_1, \dots, A_m \in \mathcal{U}$. Dann gilt:

$$\begin{aligned}
P_n\left(\bigcup_{i=1}^m A_i\right) &= \sum_{k=1}^m (-1)^{k-1} \cdot \sum_{1\le i_1<\dots<i_k\le m} \underbrace{P_n(\overbrace{A_{i_1}\cap\dots\cap A_{i_k}}^{\in\mathcal{U}})}_{\stackrel{n\to\infty}{\longrightarrow} P(A_{i_1}\cap\dots\cap A_{i_k})} \\
&\stackrel{n\to\infty}{\longrightarrow} \sum_{k=1}^m \sum_{1\le i_1<\dots<i_k\le m} P\left(A_{i_1}\cap\dots\cap A_{i_k}\right) \\
&\stackrel{\text{allg. Add.}}{=} P\left(\bigcup_{i=1}^m A_i\right)
\end{aligned}$$

Sei $G \in \mathcal{G}$, also offen. Wegen Voraussetzung (ii) ist $G$ darstellbar als $G = \bigcup_{i\in\mathbb{N}} A_i$ mit $(A_i)i \in \mathbb{N} \subseteq \mathcal{U}$. Es gilt $G_m := \bigcup_{i=1}^m A_i \uparrow G$ für $m \to \infty$. Mit der $\sigma$-Stetigkeit von $P$ gilt

$$\begin{aligned}
&\forall\varepsilon>0\ \exists m_0\in\mathbb{N}\ \forall m>m_0:\quad P(G)-P(G_{m_0})\le\varepsilon \\
\Rightarrow\ &\forall\varepsilon>0\ \exists m_0\in\mathbb{N}:\quad P(G)-\varepsilon\le P(G_{m_0}) \stackrel{\text{s.o.}}{=} \lim_{n\to\infty} P_n(\underbrace{G_{m_0}}_{\subseteq G}) \le \liminf_{n\to\infty} P_n(G) \quad .
\end{aligned}$$

Mit $\varepsilon \to 0$ folgt daraus $\liminf_{n\to\infty} P_n(G) \ge P(G)$. Da $G \in \mathcal{G}$ beliebig war, folgt nun die Behauptung aus Theorem 4.2 (4). □

**Korollar 4.4** Sei $\mathcal{U} \subseteq \mathcal{B}(S)$ endlich durchschnittsstabil mit

$$\forall x\in S\ \forall\varepsilon>0\ \exists A\in\mathcal{U}:\quad x\in A^\circ\subseteq A\subseteq B(x,\varepsilon) \quad . \tag{4.6}$$

Ist $(\mathcal{S}, d)$ separabel, so gilt

$$\Big(\forall A\in\mathcal{U}: P_n(A)\stackrel{n\to\infty}{\longrightarrow} P(A)\Big) \quad\Rightarrow\quad P_n \stackrel{\text{w}}{\longrightarrow} P \quad .$$

*Beweis.* Gemäß Satz 2.9 hat $\mathcal{G}$ eine abzählbare Basis. Nach dem *Satz von Lindelőf* (vergleiche [24]):

Für jede offene Überdeckung einer beliebigen Teilmenge von $\mathcal{S}$ existiert eine *abzählbare* Teilüberdeckung.

Sei nun $G \in \mathcal{G}$ beliebig. Für alle $x \in G$ existiert ein $\varepsilon_x > 0$ mit $B(x, \varepsilon_x) \subseteq G$. Gemäß (4.6) findet man ein $A_x \in \mathcal{U}$ mit $x \in A_x^\circ \subseteq A_x \subseteq B(x, \varepsilon_x) \subseteq G$. Also folgt

$$G = \bigcup_{x\in G}\{x\} \subseteq \bigcup_{x\in G} A_x^\circ \subseteq G$$

Somit ist $\{A_x^\circ : x \in G\}$ eine offene Teilüberdeckung von $G$. Aus dem Satz von Lindelőf existieren nun $A_{x_i} \in \mathcal{U}$ $(i \in \mathbb{N})$ mit

$$G \subseteq \bigcup_{i\in\mathbb{N}} A_{x_i}^\circ \subseteq \bigcup_{i\in\mathbb{N}} A_{x_i} \subseteq G \quad\Rightarrow\quad G = \bigcup_{i\in\mathbb{N}} A_{x_i}$$

Also erfüllt $\mathcal{U}$ die Voraussetzung (i) und (ii) in Theorem 4.3 und es folgt die Behauptung. □

Als Anwendung/ Beschreibung der schwachen Konvergenz im $\mathcal{S} = \mathbb{R}$. Wir erinnern uns an die **schwache Konvergenz von Verteilungsfunktionen** $(F_n)_{n\in\mathbb{N}}$ gegen $F$, in Zeichen

$$F_n \rightharpoonup F \quad :\Leftrightarrow \quad F_n(x) \overset{n\to\infty}{\longrightarrow} F(x) \qquad \forall x \in C_F := \{x \in \mathbb{R} : F \text{ ist stetig in } x\}$$

**Korollar 4.5** Seien $P$ und $(P_n)_{n\in\mathbb{N}}$ Wahrscheinlichkeitsmaße auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mit zugehörigen Verteilungsfunktionen $F$ und $(F_n)_{n\in\mathbb{N}}$ (d. h. $F(x) = P((-\infty, x]))$ für alle $x \in \mathbb{R}$, $n \in \mathbb{N}$). Dann gilt:

$$P_n \xrightarrow{\text{w}} P \quad \Leftrightarrow \quad F_n \rightharpoonup F$$

*Beweis.* **($\Rightarrow$)** Sei $x \in \mathbb{R}$ und $B := (-\infty, x]$. Dann ist $\partial B = \{x\}$ und somit

$$P(\partial B) = 0 \quad \Leftrightarrow \quad P(\{x\}) = F(x) - \underbrace{F(x-0)}_{\text{Grenzwert}} = 0 \qquad \forall x \in C_F \ . \tag{4.7}$$

Damit folgt für $x \in C_F$

$$F_n(x) \overset{\text{Def}}{=} P_n\left(\underbrace{(-\infty, x]}_{=B}\right) = P_n(B) \overset{n\to\infty}{\longrightarrow} P(B) \overset{\text{Def}}{=} F(x) \qquad \forall x \in C_F$$

gemäß Theorem 4.2 (5), da $B$ $P$-randlos ist für alle $x \in C_F$ gemäß (4.7).

**($\Leftarrow$)** Das Mengensystem $\mathcal{U} := \{(a, b] : a, b \in C_F \text{ mit } a \leq b\}$ ist endlich durchschnittsstabil. Ferner ist die Menge

$$D_F := C_F^{\complement} = \{x \in \mathbb{R} : F \textit{ nicht} \text{ stetig in } x\} = \{x \in \mathbb{R} : P(\{x\}) > 0\}$$

höchstens abzählbar (vgl. (4.3) im Beweis von Theorem 4.2). Damit ist $D_F^{\complement} = C_F$ dicht in $\mathbb{R}$. Also folgt

$$\begin{aligned}&\forall x \in \mathbb{R} : \forall \varepsilon > 0 : \exists A = (a, b] \in \mathcal{U} : \\ &\qquad x \in (a, b) = A^\circ \subseteq A = (a, b] \subseteq B(x, \varepsilon) = (x - \varepsilon, x + \varepsilon)\end{aligned}$$

denn mit der Dichtheit von $C_F$ gilt: In $(x - \varepsilon, x)$ muss ein $a \in C_F$ existieren und analog findet man ein $b \in (x, x + \varepsilon)$. Somit erfüllt $\mathcal{U}$ die Voraussetzungen von Korollar 4.4. Klar: $\mathcal{S} = \mathbb{R}$ ist separabel, da $\mathbb{Q}$ abzählbar und dicht in $\mathbb{R}$ ist. Schließlich gilt:

$$P_n((a, b]) = F_n(b) - F_n(a) \overset{a,b\in C_F}{\longrightarrow} F(b) - F(a) = P((a, b]) \qquad \forall a, b \in C_F$$

und mit Korollar 4.4 folgt daraus $P_n \xrightarrow{\text{w}} P$. □

**Bemerkung 4.6** (1) Seien $X$ und $(X_n)_{n\in\mathbb{N}}$ reelle Zufallsvariablen über $(\Omega, \mathcal{A}, \mathbb{P})$ mit Verteilungsfunktionen $F_n$ bzw. $F$. Dann gilt:

$$\begin{aligned} X_n \xrightarrow{\mathrm{d}} X &\overset{\text{Def}}{\Leftrightarrow} P_n = \mathbb{P} \circ X_n^{-1} \xrightarrow{\mathrm{w}} \mathbb{P} \circ X^{-1} = P \\ &\overset{\text{Korollar 4.5}}{\Leftrightarrow} \mathbb{P}(X_n \leq x) \overset{n\to\infty}{\longrightarrow} \mathbb{P}(X \leq x) \qquad \forall x \in C_F \\ &\Leftrightarrow F_n \rightharpoonup F \end{aligned} \tag{4.8}$$

Damit ist das klassische Konzept der Verteilungskonvergenz reeller Zufallsvariablen nur ein Spezialfall für $\mathcal{S} = \mathbb{R}$.

(2) Es gibt Verallgemeinerung von Korollar 4.5 bzw (4.8) auf $\mathcal{S} = \mathbb{R}^k$: Seien

$$X = \left(X^{(1)}, \ldots, X^{(k)}\right), X_n = \left(X_n^{(1)}, \ldots, X_n^{(k)}\right), \qquad (\Omega, \mathcal{A}) \to \left(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k)\right)$$

Zufallsvariablen in $\mathbb{R}^k$. Dann gilt:

$$X_n \xrightarrow{\mathrm{d}} X \text{ in } \mathbb{R}^k \quad \Leftrightarrow \quad \mathbb{P}(X_n \leq x) \overset{n\to\infty}{\longrightarrow} \mathbb{P}(X \leq x) \qquad \forall x = (x_1, \ldots, x_k) \in \mathbb{R}^k$$

wobei $x_i$ Stetigkeitsstelle der Verteilungsfunktion von $X^{(i)}$ ist für alle $i \in \{1, \ldots, k\}$. Beweis ist analog zu Korollar 4.5, vergleiche [29, Satz 5.58].

Der schwache Limes einer Folge $(P_n)_{n\in\mathbb{N}}$ ist eindeutig.

**Lemma 4.7** $P_n \xrightarrow{\mathrm{w}} P$ und $P_n \xrightarrow{\mathrm{w}} Q \quad \Rightarrow \quad P = Q$

*Beweis.* Gemäß Definition gilt:

$$\int f \, \mathrm{d}P_n \to \int f \, \mathrm{d}P \quad \text{und} \quad \int f \, \mathrm{d}P_n \to \int f \, \mathrm{d}Q \qquad \forall f \in C^b(\mathcal{S})$$

Der Grenzwert von reellen Zahlenfolgen eindeutig ist, folgt $\int f \, \mathrm{d}P = \int f \, \mathrm{d}Q$ für alle $f \in C^b(\mathcal{S})$ und mit Satz 3.17 folgt daraus schließlich $P = Q$. □

In Bemerkung 4.6 wurde Korollar 4.5 über Definition 4.1 übertragen in eine Version für Verteilungskonvergenz von reellen Zufallsvariablen. Analog lassen sich auch Theorem 4.2, Theorem 4.3 und Korollar 4.4 sowie Lemma 4.7 auf Zufallsvariablen übertragen.

**Satz 4.8 (Portmanteau-Theorem)** Folgende Aussagen sind äquivalent:

(1) $X_n \xrightarrow{\mathrm{d}} X$ in $(\mathcal{S}, d)$

(2) $\mathbb{E}\left[f(X_n)\right] \overset{n\to\infty}{\longrightarrow} \mathbb{E}\left[f(X)\right] \qquad \forall f \in C^b(\mathcal{S})$ gleichmäßig stetig

(3) $\limsup_{n\to\infty} \mathbb{P}(X_n \in F) \leq \mathbb{P}(X \in F) \qquad \forall F \in \mathcal{F}$

(4) $\liminf_{n\to\infty} \mathbb{P}(X_n \in G) \geq \mathbb{P}(X \in G) \qquad \forall G \in \mathcal{G}$

(5) $\mathbb{P}(X_n \in B) \overset{n\to\infty}{\longrightarrow} \mathbb{P}(X \in B) \qquad \forall B \in \mathcal{B}(\mathcal{S})$ mit $\mathbb{P}(X \in \partial B) = 0$

*Beweis.* Wende Theorem 4.2 an auf $P_n := \mathbb{P} \circ X_n^{-1}$ und $P := \mathbb{P} \circ X^{-1}$. Beachte z. B.

$$P_n(F) = \mathbb{P} \circ X_n^{-1}(F) \overset{\text{Def}}{=} \mathbb{P}\left(X_n^{-1}(F)\right) = \mathbb{P}\left((\{\Omega \in \Omega : X_n(\Omega) \in F\}\right) = \mathbb{P}(X_n \in F)$$

und

$$\int f \;\mathrm{d}P_n = \int_{\mathcal{S}} f \;\mathrm{d}(\mathbb{P} \circ X_n^{-1}) \overset{\text{(A.1)}}{=} \int_{\Omega} f(X_n) \;\mathrm{d}\mathbb{P} = \mathbb{E}f(X_n) \;. \qquad \square$$

Ferner erhält man unter den jeweiligen Voraussetzungen in Theorem 4.3 bzw. Korollar 4.4:

$$\mathbb{P}(X_n \in A) \overset{n\to\infty}{\longrightarrow} \mathbb{P}(X \in A) \qquad \forall A \in \mathcal{U} \qquad \Longrightarrow \qquad X_n \overset{\mathrm{d}}{\longrightarrow} X \text{ in } (\mathcal{S}, d)$$

und aus Lemma 4.7:

$$X_n \overset{\mathrm{d}}{\longrightarrow} X, \quad X_n \overset{\mathrm{d}}{\longrightarrow} X' \quad \Longrightarrow \quad X \overset{\mathrm{d}}{=} X' \;.$$

### 4.1.1 Das Continuous Mapping Theorem (CMT)

Sei $h\colon (\mathcal{S}, d) \to (\mathcal{S}', d')$ messbar. Wir wollen nun Bedingung an $h$ finden, sodass die beiden folgenden Implikationen gelten:

(1) $P_n \overset{\mathrm{w}}{\longrightarrow} P \;\Rightarrow\; P_n \circ h^{-1} \overset{\mathrm{w}}{\longrightarrow} P \circ h^{-1}$

(2) $X_n \overset{\mathrm{d}}{\longrightarrow} X$ in $(\mathcal{S}, d) \;\Rightarrow\; h(X_n) \overset{\mathrm{d}}{\longrightarrow} h(X)$ in $(\mathcal{S}', d')$

Man beachte dabei, dass

$$\begin{aligned} P \circ (h(X_n))^{-1} &= P \circ (h \circ X_n)^{-1} = \left(P \circ X_n^{-1}\right) \circ h^{-1} \\ P \circ (h(X))^{-1} &= \left(P \circ X^{-1}\right) \circ h^{-1} \end{aligned}$$

gilt. Also folgt (2) aus (1). Ist $h$ stetig auf $\mathcal{S}$, dann gilt (1) schon. Um das zu zeigen, sei $f \in C^b(\mathcal{S}')$ beliebig — es gilt:

$$\begin{aligned} \int_{\mathcal{S}'} f \;\mathrm{d}\left(P_n \circ h^{-1}\right) \overset{\text{(A.1)}}{=} \int_{\mathcal{S}} \underbrace{f \circ h}_{\in C^b(\mathcal{S})} \;\mathrm{d}P_n \overset{\text{Def.}}{\longrightarrow} \int f \circ h \;\mathrm{d}P \overset{\text{(A.1)}}{=} \int f \;\mathrm{d}(P \circ h^{-1}) \\ \overset{\text{Def.}}{\Rightarrow} P_n \circ h^{-1} \overset{\mathrm{w}}{\longrightarrow} P \circ h^{-1} \;. \end{aligned}$$

Auf Stetigkeit von $h$ kann im Allgemeinen nicht verzichtet werden, denn es gilt das nachfolgende Gegenbeispiel.

**Beispiel 4.9** Sei $\mathcal{S} = \mathcal{S}' = [0,1]$ und $h\colon [0,1] \to [0,1]$ mit

$$h(x) := \begin{cases} 1, & \text{falls } x \in \{0\} \cup \left\{\frac{1}{2n} : n \in \mathbb{N}\right\} \\ 0, & \text{sonst} \;. \end{cases}$$

Sei $P_n := \delta_{\frac{1}{n}}$ $(n \in \mathbb{N})$ und $P := \delta_0$, wobei $\delta_x$ das Dirac-Maß bezeichnet. Dann gilt $P_n \overset{\mathrm{w}}{\longrightarrow} P$,

denn
$$\int f \ \mathrm{d}P_n = f\left(\frac{1}{n}\right) \overset{n\to\infty}{\longrightarrow} f(0) = \int f \ \mathrm{d}P \qquad \forall f \in C^b\left([0,1]\right) \ .$$
Aber wegen $\delta_x \circ h^{-1} = \delta_{h(x)}$ gilt für
$$Q_n := P_n \circ h^{-1} = \delta_{\frac{1}{n}} \circ h^{-1} = \delta_{h\left(\frac{1}{n}\right)} \ , \qquad Q := P \circ h^{-1} = \delta_1$$
das Folgende:
$$\int f \ \mathrm{d}Q_n = f\left(h\left(\frac{1}{n}\right)\right) = \begin{cases} f(1), & \text{falls } n \text{ gerade} \\ 0, & \text{falls } n \text{ ungerade} \end{cases} .$$
Sei $f \in C^b([0,1])$ mit $f(0) \neq f(1)$ (z. B. $f = \mathrm{id}$). Folglich ist die Folge $(\int f \ \mathrm{d}Q_n)_{n\in\mathbb{N}}$ divergent. Somit ist $\int f \ \mathrm{d}Q_n \to \int f \ \mathrm{d}Q$, was $Q_n \overset{\mathrm{w}}{\longrightarrow} Q$ zeigt.

Die Forderung der Stetigkeit lässt sich aber abschwächen so, dass Punkt (1) noch gilt. Dazu definieren wir die **Menge der Unstetigkeitsstellen**
$$D_h := \{x \in \mathcal{S} : h \text{ nicht stetig in } x\} \ .$$

**Lemma 4.10** Für alle $h\colon \mathcal{S} \to \mathcal{S}'$ gilt $D_h \in \mathcal{B}(\mathcal{S})$. $h$ muss insbesondere nicht einmal messbar sein.

*Beweis* (nach [1, Appendix]). Mit der Dreiecksungleichung überlegt man sich leicht:
$$\begin{aligned} h \text{ stetig in } x \quad &\Leftrightarrow \ \forall 0 < \varepsilon \in \mathbb{Q} : \exists 0 < \delta \in \mathbb{Q} : \forall y,z \in \mathcal{S} : \\ &\qquad d(x,y) < \delta \text{ und } d(x,z) < \delta \ \Rightarrow \ d'\left(h(y), h(z)\right) < \varepsilon \ . \end{aligned}$$
Damit folgt
$$D_h = \bigcup_{0<\varepsilon\in\mathbb{Q}} \bigcap_{0<\delta\in\mathbb{Q}} \underbrace{\{x \in \mathcal{S} : \exists y,z \in \mathcal{S} \text{ mit } d(x,y) < \delta \text{ und } d(x,y) < \delta \text{ und } d'\left(h(y),h(z)\right) \geq \varepsilon\}}_{=:A_{\varepsilon,\delta}} \ . \tag{4.9}$$
Zeige nun
$$A_{\varepsilon,\delta} \in \mathcal{G}(\mathcal{S}) \qquad \forall \varepsilon,\delta > 0 \ . \tag{4.10}$$
Dazu sei $x_0 \in A_{\varepsilon,\delta}$. Dann existieren $y,z \in \mathcal{S}$ mit $d(x_0,y) < \delta$ und $d(z_0,z) < \delta$, aber $d'\left(h(y),h(z)\right) \geq \varepsilon$. Wähle
$$\delta_0 := \min\left\{\delta - d(x_0,y), \delta - d(x_0,z)\right\} > 0 \ .$$
Ferner gilt
$$B(x_0,\delta_0) \subseteq A_{\varepsilon,\delta}, \tag{4.11}$$
denn: Sei $x \in B(x_0,\delta_0)$. Dann gilt $d(x,y) \leq d(x,x_0) + d(x_0,y) < \delta_0 + d(x_0,y) < \delta$ und analog $d(x,z) < \delta$ Also folgt $x \in A_{\varepsilon,\delta}$, denn $d'\left((h(y),h(z)\right) \geq \varepsilon$. Wegen (4.11) ist $x_0$ innerer Punkt von $A_{\varepsilon,\delta}$. Also folgt (4.10) und mit (4.9) dann die Behauptung. □

**Satz 4.11 (Continuous Mapping Theorem – CMT)** Sei $h\colon (\mathcal{S}, d) \to (\mathcal{S}', d')$ $\mathcal{B}_d(\mathcal{S})$-$\mathcal{B}_{d'}(\mathcal{S}')$-messbar. Dann gilt:

(1) $P_n \xrightarrow{\text{w}} P$ und $P(D_h) = 0 \;\Rightarrow\; P_n \circ h^{-1} \xrightarrow{\text{w}} P \circ h^{-1}$

(2) $X_n \xrightarrow{\text{d}} X$ in $(\mathcal{S}, d)$ und $\mathbb{P}(X \in D_h) = 0 \;\Rightarrow\; h(X_n) \xrightarrow{\text{d}} h(X)$ in $(\mathcal{S}', d')$

(3) Sei $h$ stetig. Dann gilt: $X_n \xrightarrow{\text{d}} X$ in $(\mathcal{S}, d) \;\Rightarrow\; h(X_n) \xrightarrow{\text{d}} h(X)$ in $(\mathcal{S}', d')$.

*Beweis. Der Beweis ist eine beliebte Prüfungsfrage laut Prof. Ferger.*
Wegen des Portmanteau-Theorems (Theorem 4.2) zeigen wir die äquivalente Aussage über den Limes superior.

(zu 1) Sei $F \in \mathcal{F}(\mathcal{S}')$ (d. h. $F \subseteq \mathcal{S}'$ abgeschlossen) beliebig. Dann gilt:

$$\begin{aligned}
\limsup_{n\to\infty} P_n \circ h^{-1}(F) &= \limsup_{n\to\infty} P_n \left(\underbrace{h^{-1}(F)}_{\subseteq \overline{h^{-1}(F)}}\right) && (4.12)\\
&\le \limsup_{n\to\infty} P_n \left(\underbrace{\overline{h^{-1}(F)}}_{\in \mathcal{F}(\mathcal{S})}\right)\\
&\overset{4.2}{\le} P\left(\overline{h^{-1}(F)}\right) && (4.13)
\end{aligned}$$

Wir zeigen nun, dass

$$\overline{h^{-1}(F)} \subseteq h^{-1}(F) \cup D_h \tag{4.14}$$

gilt. Sei dazu $x \in \overline{h^{-1}(F)}$.

**Fall 1:** Sei $x \in D_h$. Dann folgt daraus schon, dass $x \in h^{-1}(F) \cup D_h$.

**Fall 2:** Sein nun $x \notin D_h$, d.h. $h$ ist stetig in $x$. Per Definition des Abschlusses existiert eine Folge $(x_n)_{n\in\mathbb{N}} \subseteq h^{-1}(F) \subseteq \mathcal{S}$ mit $x_n \overset{n\to\infty}{\longrightarrow} x$. Für jede solche Folge gilt

$$\underbrace{h(x_n)}_{\in F} \overset{n\to\infty}{\longrightarrow} h(x) \;\Rightarrow\; h(x) \in \overline{F} = F \;\Rightarrow\; x \in h^{-1}(F) \;\Rightarrow\; x \in h^{-1}(F) \cup D_h\ .$$

Mit (4.14) setzt sich die Ungleichungskette (4.12) fort und wir erhalten

$$P\left(\overline{h^{-1}(F)}\right) \overset{(4.14)}{\le} P\left(h^{-1}(F) \cup D_h\right) \le P\left(h^{-1}(F)\right) + \underbrace{P(D_h)}_{=0} = P \circ h^{-1}(F)\ .$$

Schließlich liefern (4.12) und Theorem 4.2(3) das gewünschte Resultat $P_n \circ h^{-1} \xrightarrow{\text{w}} P \circ h^{-1}$.

(zu 2) Dies folgt direkt aus (1) nach Definition Definition 4.1, denn es ist

$$\begin{aligned} P\circ(h(X_n))^{-1} = P\circ(h\circ X_n)^{-1} = (P\circ X_n^{-1})\circ h^{-1} &\xrightarrow{\text{w}} P\circ X^{-1} \\ &\xrightarrow{\text{w}} (P\circ X^{-1})\circ h^{-1} = P\circ(h(X))^{-1} \\ \overset{\text{Def}}{\Rightarrow}\ h(X_n) &\xrightarrow{\text{d}} h(X)\ . \end{aligned}$$

(zu 3) Dies folgt wiederum direkt aus (2), da für eine stetige Funktion $h$ die Menge $D_h$ stets leer ist. □

**Bemerkung 4.12 (Choque-Kapazität)** Es gibt eine Verallgemeinerung des Wahrscheinlichkeitsbegriffs: die so genannte **Choque-Kapazität**. Auf der Menge aller Choque-Kapazitätäten lässt sich eine Topologie $\tau_{\text{weak}}$ einführen (die so genannte *schwache Topologie*) und ein entsprechendes Portmanteau-Theorems beweisen, vgl. [12] und „Dietmar Ferger (2019): *unveröffentlicht*".



■

# Kapitel A
# GRUNDLAGEN

- $f : \mathbb{R} \to \mathbb{R}$ heißt **Borel-messbar**, wenn $f^{-1}(M) \in \mathcal{B}(\mathbb{R})$ für alle $M \in \mathcal{B}(\mathbb{R})$.
- Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $(S, \mathcal{B})$ ein Messraum. Eine **Zufallsvariable (ZV)** $X$ ist eine $(\mathcal{A}, \mathcal{B})$-messbare Abbildung $X : \Omega \to S$. Wenn $S = \mathbb{R}$, dann heißt $X$ **Zufallsgröße (ZG)**.
- Die **Verteilung** einer Zufallsgröße $X$ ist Wahrscheinlichkeitsmaß $\mu$ auf $\mathcal{B}$ definiert durch

$$\mu_X : \mathcal{B} \to [0,1], \qquad \mu_X(B) := \mathbb{P}(X \in B) := \mathbb{P}(X^{-1}(B)) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in B\}) \ .$$

- Die **Dichte** ist

$$p \colon \mathbb{R}^n \to [0, \infty] \text{ mit } \mu_X(A) = \int_A p(x) \ \mathrm{d}\mathbb{P}(x) \qquad \forall A \in \mathcal{B}(\mathbb{R}^n)$$

- **Verteilungsfunktion** von $X$ ist

$$F_X : \mathbb{R} \to [0,1], \qquad F_X(x) = \mathbb{P}(X < x) = \mu_X((-\infty, x)) = \int_{-\infty}^{x} p(t) \ \mathrm{d}t$$

  mit der Eigenschaft $F'(x) = p(x)$ für alle $x \in \mathbb{R}$ fast überall.
- Zwei Zufallsvariablen $X_1 : \Omega \to S_1$ und $X_2 : \Omega \to S_2$ heißen **unabhängig**, wenn

$$\mathbb{P}(X_1 \in B_1) \cdot \mathbb{P}(X_2 \in B_2) = \mathbb{P}(X_1 \in B_1, X_2 \in B_2) = \mathbb{P}(\{\omega \in \Omega : X_1(\omega) \in B_1 \text{ und } X_2(\omega) \in B_2\})$$

  für alle $B_1 \in \mathcal{B}_1$ und alle $B_2 \in \mathcal{B}_2$.
- Sei $X$ eine $\mathbb{P}$-integrierbare oder nichtnegative Zufallsgröße. Dann ist der **Erwartungswert von** $X$ definiert als

$$\mathbb{E}[X] := \int_\Omega X(\omega) \ \mathrm{d}\mathbb{P}(\omega) \overset{\text{A.1}}{=} \int_\mathbb{R} x \ \mathrm{d}\mu_X(x) = \int_\mathbb{R} x \cdot p(x) \ \mathrm{d}x$$

  und hat folgende Eigenschaften ($X, Y$ seien Zufallsgrößen):

  (1) Linearität: $\mathbb{E}[aX + bY] = a \cdot \mathbb{E}[X] + b \cdot \mathbb{E}[Y]$

  (2) $X = c \in \mathbb{R}$ fast sicher konstant $\Rightarrow$ $\mathbb{E}[X] = c$

  (3) $a \leq X \leq b$ fast sicher konstant $\Rightarrow$ $a \leq \mathbb{E}(X) \leq b$

  (4) $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$

  (5) $X \geq 0$ fast sicher und $\mathbb{E}(X) = 0$ $\Rightarrow$ $X = 0$ fast sicher

(6) $X, Y$ unabhängig $\Rightarrow$ $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$

- Zwei ZG heißen **unkorreliert**, falls $\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$.
- Für $X \in L^2(\mathbb{P})$ ist die **Varianz**

$$\mathbb{V}\mathrm{ar}(X)(X) := \mathbb{E}[(X - \mathbb{E}[X])^2] = \int_\Omega (X - \mathbb{E}[X])^2 \, \mathrm{d}\mathbb{P} = E[X^2] - \mathbb{E}[X]^2$$

mit den Eigenschaften

$$\mathbb{V}\mathrm{ar}(X)(a \cdot X + b) = a^2 \cdot \mathbb{V}\mathrm{ar}(X)(X) \ ,$$
$$\mathbb{V}\mathrm{ar}(X)(X) = 0 \ \Leftrightarrow \ X \text{ ist konstant fast sicher} \ ,$$
$$\mathbb{V}\mathrm{ar}(X)(X+Y) = \mathbb{V}\mathrm{ar}(X)(X) + \mathbb{V}\mathrm{ar}(X)(Y) + \underbrace{2\mathbb{E}[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])]}_{=0, \text{ falls } X, Y \text{ unkorreliert sind}} \ .$$

**Satz A.1 (Korrespondenzsatz)** Jede Verteilungsfunktion $F$ ist Verteilungsfunktion eines eindeutigen Wahrscheinlichkeitsmaßes $\mathbb{P}$. Dieses Maß $\mathbb{P}$ ist durch

$$\mathbb{P}_F((-\infty, x]) := F(x)$$

eindeutig bestimmt. Umgekehrt bestimmt jedes Wahrscheinlichkeitsmaß eine eindeutige Verteilungsfunktion über

$$F_\mathbb{P}(x) := \mathbb{P}((-\infty, x]) \ .$$

Somit ist die Zuordnung der Verteilungsfunktionen zu den Wahrscheinlichkeitsverteilungen bijektiv.

**Bemerkung A.2 (Notation)** $\mathbb{P}(\mathrm{d}x) = \mathrm{d}\mathbb{P}(x)$. Außerdem bedeutet $F(\mathrm{d}x)$ oftmals auch, dass man bzgl. jenem Maß $Q$ integriert, was durch $F$ eindeutig festgelegt ist.

**Satz A.3 (Transformationssatz)** Seien $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $(S, \mathcal{F})$ ein Messraum. Sei $g\colon S \to \mathbb{R}$ eine messbare Funktion und $X\colon \Omega \to S$ eine Zufallsvariable. Dann gilt:

$$\int_\Omega g(X(\omega)) \, \mathbb{P}(\mathrm{d}\omega) = \int_S g(s) \, \mathbb{P}_X(\mathrm{d}s) \tag{A.1}$$

Hierbei ist $\mathbb{P}_X = \mu_X = \mathbb{P} \circ X^{-1}$ die Verteilung von $X$. Für den Standardfall reeller Zahlen ergibt sich mit $f_X$ als Dichte von $X$:

$$\int_\Omega g(X(\omega)) \, \mathbb{P}(\mathrm{d}\omega) = \int_\mathbb{R} g(x) \, \mathbb{P}_X(\dot{x}) = \int_\mathbb{R} g(x) \cdot f_X(x)\mathrm{d}x \ . \tag{A.2}$$

**Satz A.4 (majorisierte Konvergenz)** Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Maßraum und sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von $\mathbb{P}$-mesbaren Funktionen (z. B. ZV) $f_n\colon \Omega \to \mathbb{R} \cup \{\infty\}$. Die Folge $(f_n)_{n\in\mathbb{N}}$ konvergiere $\mathbb{P}$-fast überall gegen eine $\mathbb{P}$-messbare Funktion $f$ und es gilt $|f_n| \leq g$ $\mathbb{P}$-fast überall für alle $n \in \mathbb{N}$ für eine $\mathbb{P}$-integrierbare Funktion $g\colon \Omega \to \mathbb{R}$. Dann sind $f_n$ und $f$ $\mathbb{P}$-integrierbar und es gilt

$$\int_\Omega f(\omega)\ \mathrm{d}\mathbb{P}(\omega) = \int_\Omega \lim_{n\to\infty} f_n(\omega)\ \mathrm{d}\mathbb{P}(\omega) = \lim_{n\to\infty} \int_\Omega f_n(\omega)\ \mathrm{d}\mathbb{P}(\omega)$$

**Satz A.5 (Monotone Konvergenz)** Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Maßraum. Ist $(f_n)_{n\in\mathbb{N}}$ eine Folge nicht-negativer, messbarer Funktionen $f_n\colon \Omega \to [0,\infty]$, die $\mathbb{P}$-fast-überall monoton wachsend gegen eine messbare Funktion $f\colon \Omega \to [0,\infty]$ konvergiert, so gilt:

$$\int_\Omega f(\omega)\ \mathrm{d}\mathbb{P}(\omega) = \int_\Omega \lim_{n\to\infty} f_n(\omega)\ \mathrm{d}\mathbb{P}(\omega) = \lim_{n\to\infty} \int_\Omega f_n(\omega)\ \mathrm{d}\mathbb{P}(\omega)$$

# LITERATURVERZEICHNIS


[1] BILLINGSLEY, P. : *Convergence of probability measures.* John Wiley & Sons, 1968

[2] BILLINGSLEY, P. : *Probability and measure.* John Wiley & Sons, 1995

[3] BORODIN, A. N. ; SALMINEN, P. : *Handbook of Brownian motion-facts and formulae.* Birkhäuser, 2002

[4] CHOW, Y. S. ; TEICHER, H. : *Probability theory: independence, interchangeability, martingales.* Springer Science & Business Media, 1997

[5] CSÖRGŐ, M. ; HORVÁTH, L. : *Limit theorems in change-point analysis.* John Wiley & Sons Chichester, 1997

[6] CZADO, C. ; SCHMIDT, T. : *Mathematische Statistik.* Springer, 2011

[7] DUDLEY, R. M.: *Real Analysis and Probability.* Chapman and Hall/CRC, 1999

[8] FERGER, D. : Asymptotic distribution theory of change-point estimators and confidence intervals based on bootstrap approximation. In: *Mathematical Methods of Statistics* 3 (1994), Nr. 4, S. 362

[9] FERGER, D. : Moment equalities for sums of random variables via integer partitions and Faà di Bruno's formula. In: *Turkish Journal of Mathematics* 38 (2014), Nr. 3, S. 558–575

[10] FERGER, D. : Optimal constants in the Marcinkiewicz–Zygmund inequalities. In: *Statistics & Probability Letters* 84 (2014), S. 96–101

[11] FERGER, D. : On the supremum of a Brownian bridge standardized by its maximizing point with applications to statistics. In: *Statistics & Probability Letters* 134 (2018), S. 63–69

[12] FERGER, D. : Weak convergence of probability measures to Choquet capacity functionals. In: *Turkish Journal of Mathematics* 42 (2018), S. 1747–1764

[13] GÄNSSLER, P. ; STUTE, W. : *Wahrscheinlichkeitstheorie.* Berlin , Heidelberg [u.a.] : Springer, 1977. – ISBN 9780387084183

[14] HAND, D. J.: Statistical Decision Theory: Estimation, Testing, and Selection by Friedrich Liese, Klaus-J. Miescke. In: *International Statistical Review* 76 (2008), Nr. 3, S. 450–450

[15] HEUSER, H. : *Funktionalanalysis.* 2006

[16] HJORT, N. L. ; POLLARD, D. : Asymptotics for minimisers of convex processes. In: *arXiv preprint arXiv:1107.3806* (2011)

[17] JACOD, J. ; PROTTER, P. : *Probability essentials.* Springer Science & Business Media, 2000

[18] KALLENBERG, O. : *Foundations of modern probability.* Springer, 1997

[19] Klenke, A. : *Wahrscheinlichkeitstheorie*. Bd. 1. Springer, 2008

[20] Prokhorov, Y. V.: Convergence of random processes and limit theorems in probability theory. In: *Theory of Probability & Its Applications* 1 (1956), Nr. 2, S. 157–214

[21] Rockafellar, R. : Convex analysis. In: *Princeton Univ., Princeton, NJ* (1972)

[22] Royden, H. L. ; Fitzpatrick, P. : *Real analysis*. Bd. 32. Macmillan New York, 1988

[23] Schmidt, K. D.: *Maß und Wahrscheinlichkeit*. Springer-Verlag, 2011

[24] Schubert, H. : *Topologie eine Einführung*. 4. Aufl. Stuttgart : Teubner, 1975. – ISBN 3519122006

[25] Shorack, G. R. ; Wellner, J. A.: *Empirical processes with applications to statistics*. Bd. 59. Siam, 1986

[26] Smirnov, N. V.: Limit distributions for the terms of a variational series. In: *Trudy Matematicheskogo Instituta imeni VA Steklova* 25 (1949), S. 3–60

[27] Whitt, W. : Weak convergence of probability measures on the function space C([0,unendlich)). In: *The Annals of Mathematical Statistics* 41 (1970), Nr. 3, S. 939–944

[28] Witting, H. ; Müller-Funk, U. : *Mathematische Statistik 1 Parametrische Verfahren bei festem Stichprobenumfang*. Stuttgart : Teubner, 1985. – ISBN 3519020262

[29] Witting, H. ; Müller-Funk, U. : *Mathematische Statistik II: Asymptotische Statistik: Parametrische Statistik und nichtparametrische Modelle*. Stuttgart : Teubner, 1995