Dieses Dokument umfasst lediglich eine Auswahl des Stoffes der Vorlesung Stochastik I. Es wird keinerlei Haftung für Vollständigkeit, inhaltliche Korrektheit und sonstige Korrektheit übernommen. Hinweise auf Fehler jeder Art sind erwünscht. Die Nummerierung weicht von der Nummerierung im Skript ab, sollte in sich aber konsistent sein, die Verweise innerhalb des Dokuments beziehen sich also nicht auf die Nummerierung im Skript. Beweise sind mal als Skizzen, mal vollständig notiert.

1 Motivation

2 Inhalte, Prämaße, Elementarintegrale

Definition 2.1: Ein System $\mathscr{S}$ von Teilmengen von Ω heißt Präring über Ω, falls

(i) $\emptyset \in \mathscr{S}$

(ii) Für alle $A, B \in \mathscr{S}$ gibt es ein $n \in \mathbb{N}, c_1, ..., c_n \in \mathscr{S}$ disjunkt, so dass $A \backslash B = \bigcup_{i=1}^{m} c_i$

Definition 2.2: Sei $\mathscr{S}$ ein Präring über Ω. Eine Abbildung $\mu : \mathscr{S} \to [0, \infty]$ heißt Inhalt, falls

(i) $\mu(\emptyset) = 0$

(ii) Für alle $n \in \mathbb{N}, A_1, ..., A_n \in \mathscr{S}$ disjunkt mit $\bigcup_{i=1}^{n} A_i \in \mathscr{S}$ ist $\mu(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i)$ (endliche Additivität)

Gilt anstelle der endlichen Additivität die unendliche Additivität spricht man von einem Prämaß.

Definition 2.3: Eine Abbildung $F : \mathbb{R} \to \mathbb{R}$ heißt maßerzeugend, falls F monoton steigend und rechtsstetig ist. Gilt zudem, dass

$$\lim_{x \to -\infty} F(x) = 0 \text{ und } \lim_{x \to +\infty} F(x) = 1$$

so spricht man von einer Verteilungsfunktion.

Satz 2.4: *Sei F maßerzeugend, $\Omega := \mathbb{R}$ und $\mathscr{S} = \{(a, b] | a, b \in \mathbb{R}, a \leq b\}$. Dann definiert $\mu_F((a, b]) = F(b) - F(a)$ ein Prämaß auf $\mathscr{S}$.*

Sei $\mathscr{S}$ ein nichtleeres System von Teilmengen von Ω. Wir bezeichnen mit

$$\mathcal{R}(\mathscr{S}) = \{\bigcup_{i=1}^{n} A_i | A_i \cap A_j = \emptyset\}$$

das System aller endlichen disjunkten Vereinigungen von Elementen in $\mathscr{S}$.

Definition 2.5: Ein System $\mathcal{R} \subset 2^{\Omega}$ heißt Ring über Ω, falls

(i) $\emptyset \in \mathcal{R}$

(ii) $A, B \in \mathcal{R}$ disjunkt $\Rightarrow A \cup B \in \mathcal{R}$

(iii) $A, B \in \mathcal{R} \Rightarrow A \backslash B \in \mathcal{R}$.

Satz 2.6: *Sei $\mathscr{S} \subset 2^{\Omega}$ nichtleer. Dann sind äquivalent:*

(i) *$\mathscr{S}$ ist ein Präring*

(ii) *$\mathcal{R}(\mathscr{S})$ ist ein Ring, der von $\mathscr{S}$ erzeugte Ring.*

(iii) *$\mathscr{S}$ hat folgende Zerlegungseigenschaft: Zu beliebigen $A_1, ..., A_n \in \mathscr{S}$ existieren disjunkte $c_1, ..., c_l \in \mathscr{S}$, so dass für alle $i = 1, ..., n : A_i = \bigcup_{c_\lambda \subset A_i} c_\lambda$*

Definition 2.7: Sei $\mathscr{S}$ ein Präring über Ω.

(i) Wir bezeichnen mit

$$\varepsilon^+(\mathscr{S}) = \{f : \Omega \to [0, \infty) | f = \sum_{i=1}^{n} \alpha_i 1_{A_i}, n \in \mathbb{N}, A_i \in \mathscr{S}, \alpha_i \in [0, \infty)\}$$

die Menge der nicht-negativen elementaren Funktionen über $\mathscr{S}$.

(ii) Eine Abbildung $I : \varepsilon^+(\mathscr{S}) \to [0, \infty]$ nennen wir Elementarintegral über $\mathscr{S}$, falls

(1) $I(f + g) = I(f) + I(g)$

(2) $I(\alpha f) = \alpha I(f), \alpha \geq 0$

Satz 2.8: *Sei $\mathscr{S}$ ein Präring über Ω*

(i) *Ist μ ein Inhalt auf $\mathscr{S}$, so ist $I_\mu : \varepsilon^+(\mathscr{S}) \to [0, \infty]$, $\sum_{i=1}^{n} \alpha_i 1_{A_i} \mapsto \sum_{i=1}^{n} \alpha_i \mu(A_i)$ wohldefiniert und ein Elementarintegral.*

(ii) *Ist I ein Elementarintegral über $\mathscr{S}$ so definiert*

$$\nu : \mathcal{R}(\mathscr{S}) \to [0, \infty], A \mapsto I(1_A)$$

einen Inhalt auf $\mathcal{R}(\mathscr{S})$ und $I = I_\nu$.

Satz 2.9: *(Maßfortsetzungssatz) Sei $\mathscr{S}$ ein Präring über Ω und μ ein Inhalt auf $\mathscr{S}$. Dann existiert genau eine Fortsetzung von μ zu einem Inhalt auf $\mathcal{R}(\mathscr{S})$. Es gilt für $A \in \mathcal{R}(\mathscr{S})$:*

$$\mu(A) = I_\mu(1_A) = \sum_{i=1}^{n} \mu(A_i), A = \bigcup_{i=1}^{k} A_i, A_i \in \mathscr{S}$$

Ist μ ein Prämaß, so ist die kanonische Fortsetzung zu einem Inhalt auf $\mathcal{R}(\mathscr{S})$ ebenfalls ein Prämaß.

Satz 2.10: *(Charakterisierung von Prämaßen oder Beppo Levi für das Elementarintegral) Sei μ ein Inhalt auf einem Präring $\mathcal{R}$ über Ω und $I_\mu : \varepsilon^+(\mathcal{R}) \to [0, \infty]$ das zugehörige Elementarintegral. Dann sind äquivalent:*

(i) *μ ist ein Prämaß*

(ii) *μ ist stetig von unten, d.h. für alle $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{R}$ gilt: $A_n \subset A_{n+1}, n \in \mathbb{N}$ und $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{R} \Rightarrow \mu(\bigcup_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty} \mu(A_{n))}$*

(iii) *I_μ ist stetig von unten, das heißt für alle $(f_n)_{n\in\mathbb{N}}$ in $\varepsilon^+(\mathcal{R})$, $f \in \varepsilon^+(\mathcal{R})$ gilt: $f(\omega) = \sum_{n\in\mathbb{N}} f_n(\omega) \,\forall\, \omega \in \Omega \Rightarrow I_\mu(f) \leq \sum_{n\in\mathbb{N}} I_\mu(f_n)$*

(iv) *(Satz von Beppo Levi für Elementarintegral) Für alle $(f_n)_{n\in\mathbb{N}}$ in $\varepsilon^+(\mathcal{R})$ gilt:*

$$\sum_{n\in\mathbb{N}} f_n \in \varepsilon^+(\mathcal{R}) \Rightarrow I_\mu(\sum_{n\in\mathbb{N}} f_n) = \sum_{n\in\mathbb{N}} I_\mu(f_n)$$

Ist $\mu(A) < \infty$ für alle $A \in \mathcal{R}$, so ist (i) äquivalent zu:

(a) μ ist stetig von oben, d.h. für alle $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{R}$ gilt: $A_{n+1} \subset A_n, n \in \mathbb{N}$ und

$$\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{R} \Rightarrow \mu(\bigcap_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty} \mu(A_n)$$

(b) μ ist stetig von oben in $\emptyset$, d.h. für alle $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{R}$ mit $A_{n+1} \subset A_n$ und $\bigcap_{n\in\mathbb{N}} A_n = \emptyset$ gilt:

$$\lim_{n\to\infty} \mu(A_n) = 0$$

Satz 2.11: *(Produkt-Prämaße) Seien $\Omega_i \neq \emptyset$, $i = 1, 2$, $\mathscr{S}_i$ ein Präring über Ω_i und μ_i Prämaße auf $\mathscr{S}_i$. Sei μ_1 endlich (d.h. $\mu_1(A) < \infty$ für alle $A \in \mathscr{S}_1$). Dann:*

(i) *Das Mengensystem $\mathscr{S}_1 \times \mathscr{S}_2 = \{A_1 \times A_2 | A_1 \in \mathscr{S}_1$ und $A_2 \in \mathscr{S}_2\}$ ist ein Präring über $\Omega_1 \times \Omega_2$*

(ii) *Es gibt genau ein Prämaß auf $\mathcal{R}(\mathscr{S}_1 \times \mathscr{S}_2)$ so, dass für alle $A_i \in \mathscr{S}_i$ für $i = 1, 2$ $\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$. Dieses nennen wir das Produktprämaß von μ_1 und μ_2*

3 Maßräume und Wahrscheinlichkeitsräume

Definition 3.1: Sei $\Omega \neq \emptyset$. Ein Mengensystem $\mathcal{A} \subset 2^\Omega$ wird σ-Algebra genannt, falls

(i) $\Omega \in \mathcal{A}$

(ii) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$

(iii) $A_n \in \mathcal{A}, n \in \mathbb{N} \Rightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$

Das Paar $(\Omega, \mathcal{A})$ wird Messraum genannt. σ-Algebren sind unter Schnitt abgeschlossen.

Definition 3.2: (i) Prämaße auf σ-Algebren werden Maße genannt.

(ii) Ist $(\Omega, \mathcal{A})$ ein Messraum und μ ein Maß auf $\mathcal{A}$, so heißt das Tripel $(\Omega, \mathcal{A}, \mu)$ ein Maßraum.

Ist zudem $\mu(\Omega) = 1$, so spricht man von einem Wahrscheinlichkeitsmaß μ und einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mu)$.

Definition 3.3: Sei $\mathcal{M} \subset 2^\Omega$ ein System von Teilmengen von Ω. Dann heißt

$$\sigma(\mathcal{M}) = \bigcap \{\mathcal{A} | \mathcal{A} \ \sigma\text{-Algebra über } \Omega \text{ mit } \mathcal{M} \subset \mathcal{A}\}$$

die von $\mathcal{M}$ erzeugte σ-Algebra.

Satz 3.4: *(Maßfortsetzungssatz, Caratheodory) Sei μ_0 ein Prämaß auf einem Ring $\mathcal{R}$ über Ω. Dann existiert mindestens ein Maß μ auf $\sigma(\mathcal{R})$, sodass $\mu(A) = \mu_0(A)$ $\forall\, A \in \mathcal{R}$.*

Beweis: (Skizze). Für den Fall $\mu_0 < \infty$ für alle $A \in \mathcal{R}$: Man betrachte das äußere Maß

$$\mu^* : 2^\Omega \to [0, \infty], A \mapsto \inf\{\sum_{n=1}^{\infty} \mu_0(A_n) | (A_n)_{n\in\mathbb{N}} \in \mathcal{R}, A \subset \bigcup_{n\in\mathbb{N}} A_n\}$$

Dann gilt (Übung): $\mu^*(A) = \mu_0(A) \,\forall\, A \in \mathcal{R}$
WIr betrachten nun die induzierte Pseudometrik

$$\delta(A, B) = \mu^*(A \Delta B) A, B \in 2^\Omega$$

und den Abschluss von $\mathcal{R}$ bezüglich δ:
$\bar{R} = \{A \in 2^\Omega | \,\forall\, \varepsilon > 0 \,\exists\, A_\varepsilon \in \mathcal{R}$ sodass $\delta(A, A_\varepsilon < \varepsilon\}$. Dann ist $\bar{R}$ wieder ein Ring über Ω und für alle $B \in \bar{R}$ und $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{R}$ mit $\delta(B, A_n) \to 0$ gilt:

$$\mu_0(A_n) \to \mu^*(B)(n \to \infty)$$

Dann ist $\mu^*|_{\bar{R}}$ ein Inhalt und sogar ein Prämaß. Sei

$$\mathcal{A} = \{A \in 2^\Omega | A \cap B \in \bar{R} \text{ für alle } B \in \bar{R}\}$$

Dann ist $\mathcal{A}$ eine σ-Algebra und für $A \in \mathcal{A}$ gilt: $\mu^*(A) < \infty \Leftrightarrow A \in \bar{R}$. Damit kann man zeigen, dass $\mu^*|_{\mathcal{A}}$ ein Maß ist. Da $\sigma(\mathcal{R}) \subset \mathcal{A}$, ist $\mu^*|_{\sigma(\mathcal{R})}$ eine mögliche Fortsetzung. ∎

Definition 3.5: Ein Prämaß μ auf einem Präring $\mathscr{S}$ über Ω heißt

(i) endlich, falls $\mu(A) < \infty$ für alle $A \in \mathscr{S}$

(ii) σ-endlich, falls eine Folge $(A_n)_{n\in\mathbb{N}} \in \mathscr{S}$ existiert mit $\bigcup_{n\in\mathbb{N}} A_n = \Omega$ und $\mu(A_n) < \infty$ für alle $n \in \mathbb{N}$.

Satz 3.6: *(Maßeindeutigkeitssatz) Sei $\varepsilon \subset 2^\Omega$ und seinem μ_1, μ_2 Maße auf $\sigma(\varepsilon)$. Falls*

(i) *ε ist $\cap$-stabil, d.h. $E_1 \cap E_2 \in \varepsilon$ für alle $E_1, E_2 \in \varepsilon$*

(ii) *$\mu_1(E) = \mu_2(E)$ für alle $E \in \varepsilon$*

(iii) *Es gibt eine Folge $(E_n)_{n\in\mathbb{N}} \in \varepsilon$ mit $\bigcup_{n\in\mathbb{N}} E_n = \Omega$ und $E_n \subset E_{n+1}$, $\mu_1(E_n) < \infty$ für alle $n \in \mathbb{N}$.*

Dann ist $\mu_1(A) = \mu_2(A)$ für alle $A \in \sigma(\varepsilon)$. Somit kann jedes σ-endliche Prämaß auf einem Präring $\mathscr{S}$ auf genau eine Art und Weise zu einem Maß auf $\sigma(\mathcal{R}(\mathscr{S}))$ erweitert werden.

Was die Borelsche σ-Algebra über Ω im metrischen Raum (Ω, d) ist, ist klar.

Satz 3.7: *(Charakterisierung der W-Maße auf $\mathfrak{B}(\mathbb{R})$). Die Abbildung*

$$\{\text{Verteilungsfunktion auf } \mathbb{R}\} \to \{\text{W-Maße auf } \mathfrak{B}(\mathbb{R})\}$$
$$F \mapsto P_F$$

wobei $P_F((a,b] = F(b) - F(a)$, $a \leq b \in \mathbb{R}$ ist bijektiv.

Beweis: (Skizze)
-Die Wohldefiniertheit folgt aus dem Beispiel mit $\mathcal{F}_{(D)}$.
-Injetivität: Für $P_F = P_G$ gilt:

$$F(b) - F(a) = P_F((a,b]) = P_G((a,b]) = G(b) - G(a)$$

Da F, G Verteilungsfunktionen folgt mit $a \to -\infty$, dass $F(b) = G(b)$
-Surjektivität: Sei P ein W-Maß auf $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$. Sei $F(x) := P((-\infty, x])$. Dann klar.
Bleibt zu zeigen, dass F eine Verteilungsfunktion ist:
-Rechtstetigkeit
-Monotonie
-Normiertheit ∎

Definition 3.8: Dynkin System $\mathcal{D}$: Definiert wie σ-Algebra, nur ist in (iii) nur die disjunkte Vereinigung in $\mathcal{D}$.

Satz 3.9: *(Dynkin Lemma)*

(i) *Ein Dynkin-System $\mathcal{D}$ ist eine σ-Algebra genau dann, wenn $\mathcal{D}$ $\cap$-stabil ist.*

(ii) *Ist $\varepsilon \subset 2^{\Omega}$ $\cap$-stabil, so ist das Dynkin-erzeugte $=: \delta(\varepsilon) = \sigma(\varepsilon)$*

Beweis: (i) "$\Rightarrow$": klar.
"$\Leftarrow$": Sei $\mathcal{D}$ $\cap$-stabil, $(A_n)_{n\in\mathbb{N}} \subset \mathcal{D}$. Definiere

$$B_1 = A_1, B_i = A_i \cap (\bigcup_{j=1}^{i-1} B_j)^c \in \mathcal{D},\ \text{da} \bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} B_n \in \mathcal{D}.$$

(ii) "$\subset$": Jede σ-Algebra ist ein Dynkin-System.
"$\supset$": Wegen (i) zeigen wir nur dass $\delta(\varepsilon)\cap$-stabil ist.
Definiere für $B \in \delta(\varepsilon)$ das Mengensystem $\mathcal{D}_B := \{A \in \delta(\varepsilon) | A \cap B \in \delta(\varepsilon)\}$.
Zeige nun, dass $\delta(\varepsilon) \subset \mathcal{D}_B$ für alle $B \in \delta(\varepsilon)$.
-$\mathcal{D}_B$ ist ein Dynkin-System zeigen
-$\delta(\varepsilon) \subset D_B$ für $B \in \varepsilon \Rightarrow \delta(\varepsilon) \subset \mathcal{D}_A$ für alle $A \in \delta(\varepsilon)$ ■

Beweis: (von Satz 3.6) Sei $D_n := \{A \in \sigma(\varepsilon) | \mu_1(A \cap E_n) = \mu_2(A \cap E_n)\}$. Dies ist ein Dynkin System für alle $n \in \mathbb{N}$. Weiter ist wegen $\varepsilon\cap$-stabil, $A \cap E_n \in \varepsilon\, \forall\, A \in \varepsilon$. Es folgt, dass $\delta(\varepsilon) \subset D_n$ für alle $n \in \mathbb{N}$. Mit dem Dynkin-Lemma (3.9) folgt $\sigma(\varepsilon) = \delta(\varepsilon)$. Damit ist für alle $A \in \sigma(\varepsilon) \mu_1(A \cap E_n) = \mu_2(A \cap E_n)$.
Aus der Stetigkeit von unten des Maßes folgt mit $n \to \infty$

$$\mu_1(A) = \lim_{n\to\infty} \mu_1(A \cap E_n) = \lim_{n\to\infty} \mu_2(A \cap E_n) = \mu_2(A)$$ ■

4 Einige wichtige Verteilungen

Diskretes Zufallsexperiment: (Ω, p), wobei Ω höchstens abzählbar, $p : \Omega \to [0, \infty)$ und $\sum_{\omega \in \Omega} p(\omega) = 1$.

(a) **Diskrete Gleichverteilung und die hypergeometrische Verteilung**
Urnenmodelle: Mit/ohne Zurücklegen, mit/ohne Beachtung der Reihenfolge. Daraus konstruiert man die hypergeometrische Verteilung.

(b) **Binomialverteilung und Poissonverteilung**

Definition 4.1: Sei $\Omega = \{0, ..., n\}$, $B_{n,p}(k) = \binom{n}{k} p^k (1-p)^{n-k}$ für $k \in \Omega$ und festes $p \in [0,1]$. Dann ist $(\Omega, B_{n,p})$ ein diskretes Zufallsexperiment und das induzierte W-Maß auf 2^Ω wird Binomial(n,p)-Verteilung genannt. Im Fall $n = 1$ spricht man von der Bernoulli(p)-Verteilung. $B_{n,p}(k)$ gibt also die Wahrscheinlichkeit an, dass bei n unabhängigen Wiederholungen von Bernoulli(p)-Experimenten genau k Erfolge eintreten.

Satz 4.2: *Werden aus einer Urne mit r roten und s schwarzen Kugeln nacheinander mit Zurücklegen rein zufällig n Kugeln entnommen, so ist die Wahrscheinlichkeit dafür, dass darunter genau k rote Kugeln sind gegeben durch* $B_{n,\frac{r}{r+s}}(k)$ *für* $k = 0, ..., n$

Satz 4.3: *(Binomialapproximation der hypergeometrischen Verteilung) Hängt die Anzahl der roten Kugeln* r_N *und schwarzen Kugeln* s_N *so von* N *ab, dass* $s_N + r_N / \to \infty$ *und* $\frac{r_N}{s_N + r_N} \to p \in (0,1)$ *für* $N \to \infty$ *so gilt für alle* $n \in \mathbb{N}$ *und* $k = 0, ..., n$

$$\lim_{N \to \infty} H_{r_N, s_N, n}(k) = B_{n,p}(k)$$

Beweis: Hässliches Binomialherumgerechne. ∎

Satz 4.4: *(Poissonapproximation der Binomialverteilung) Hängt bei Binomialexperimenten die Erfolgswahrscheinlichkeit* p_N *so von der Anzahl der Versuche ab, dass* $np_n \to \lambda \in (0, \infty)$*, dann gilt für jedes* $k \in \mathbb{N}_0$*:*

$$\lim_{n \to \infty} B_{n,p_n}(k) = c^{-\lambda} \frac{\lambda^k}{k!}$$

Das durch die Zähldichte $\text{Poisson}_\lambda(k) := e^{-\lambda} \frac{\lambda^k}{k!}$ *induzierte W-Maß heißt Poissonverteilung.*

Beweis: Auch Herumrechnen. ∎

(c) **Wahrscheinlichkeitsmaße auf** $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$
Sei $f : \mathbb{R} \to \mathbb{R}$ eine Funktion, die über kompakten Intervallen Riemannintegrierbar ist mit $f(x) \geq 0$ und $\int\limits_{-\infty}^{+\infty} f(x) dx = 1$. Dann definiert $F : \mathbb{R} \to \mathbb{R}$,

$x \mapsto \int\limits_{-\infty}^{x} f(y)dy$ eine Verteilungsfunktion. Somit existiert ein eindeutig bestimmtes W-Maß P_F mit

$$P_F((a,b] = \int\limits_a^b f(x)dx.$$

f wir dann Dichte von P_F genannt.

$f(x)$	Name von P_f	Parameter
$\frac{1}{b-a}1_{[a,b]}(x)$	Gleichverteilung auf $[a,b]$	$a \leq b \in \mathbb{R}$
$\lambda e^{-\lambda x}1_{\{x\geq 0\}}$	Expontentialverteilung	$\lambda > 0$
$\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-(x-\mu)^2}{\sigma^2}}$	Normalverteilung (Standardnormalverteilung, falls $\mu = 0$, $\sigma = 1$)	$\mu \in \mathbb{R}, \sigma > 0$

5 Unabhängigkeit

(a) **Unabhängigkeit von Ereignissen und bedingte Wahrscheinlichkeiten**

Definition 5.1: Sei $(\Omega, \mathcal{A}, P)$ ein W-Raum.

(1) Ereignisse $A, B \in \mathcal{A}$ werden stochastisch unabhängig genannt, falls $P(A \cap B) = P(A)P(B)$

(2) Eine Familie von Ereignissen $(A_i)_{i \in I}, I \neq \emptyset$ beliebig heißt stochastisch unabhängig falls für alle endlichen Teilfamilien $(A_j)_{j \in J}, J \subset I$ endlich, gilt:

$$P(\bigcap_{j \in J} A_j) = \prod_{j \in J} P(A_j)$$

Die Familie $(A_i)_{i \in I}$ heißt paarweise unabhängig, falls für alle $i \neq j$: $P(A_i \cap A_j) = P(A_i)P(A_j)$

Definition 5.2: Sei $(\Omega, \mathcal{A}, P)$ ein W-Raum, $A, B \in \mathcal{A}$ mit $P(B) > 0$. Dann wird $P(A|B) = \frac{P(A \cap B)}{P(B)}$ bedingte Wahrscheinlichkeit von A gegeben B genannt.

Satz 5.3: *Sei $(\Omega, \mathcal{A}, P)$ ein W-Raum und $B \in \mathcal{A}$ mit $P(B) > 0$. Sei $\mathcal{A}_B = \{A \cap B | A \in \mathcal{A}\}$. Dann ist $\mathcal{A}_B$ eine σ-Algebra über B und $P_B : \mathcal{A} \to [0, \infty), A \mapsto P(A|B)$ ein W-Maß auf $\mathcal{A}_B$. Das Tripel $(B, \mathcal{A}_B, P_B)$ ist somit ein W-Raum.*

Satz 5.4: *Sei $(\Omega, \mathcal{A}, P)$ ein W-Raum $(B_i)_{i \in I}$ eine höchstens abzählbare Familie von disjunkten Ereignissen in $\mathcal{A}$ mit $\bigcup_{i \in I} B_i = \Omega$ und $P(B_i) > 0 \, \forall \, i \in I$. Sei $A \in \mathcal{A}$. Dann:*

(1) *Formel von der totalen Wahrscheinlichkeit*

$$P(A) = \sum_{i \in I} P(A|B_i)P(B_i)$$

(2) *Bayes-Formel: Falls $P(A) > 0$, gilt für jedes $j \in J$:*

$$P(B_j|A) = \frac{P(A|B_j P(B_j}{\sum_{i \in J} P(A|B_i)P(B_i)}$$

(b) **Unabhängigkeit von σ-Algebren**

Definition 5.5: Sei $(\Omega, \mathcal{A}, P)$ ein W-Raum und $(\mathcal{A}_i)_{i \in I}$ eine Familie von Unter-σ-Algebren von $\mathcal{A}$. $(\mathcal{A}_i)_{i \in I}$ wird unabhängig genannt $(I \neq \emptyset))$, falls die Familie von Ereignissen $(A_i)_{i \in I}$ unabhängig ist, wenn immer $A_i \in \mathcal{A}_i \, \forall \, i \in I$.

Satz 5.6: *Sei $(\Omega, \mathcal{A}, P)$ ein W-Raum und $(\mathcal{A}_i)_{i\in I}$ eine Familie von Unter-σ-Algebren von $\mathcal{A}$. Für jeden $i \in I$, sei ε_i ein $\cap$-stabiles Erzeugendensystem von $\mathcal{A}_i$ (d.h. $\sigma(\varepsilon_i) = \mathcal{A}_i$ und $A \cap B \in \varepsilon_i$ für alle $A, B \in \varepsilon_i$). Falls eine Familie von Ereignissen $(E_i)_{i\in I}$ unabhängig ist, wenn immer $E_i \in \varepsilon_i$ für alle $i \in I$, so ist die Familie $(\mathcal{A}_i)_{i\in I}$ unabhängig.*

(c) **Unabhängige Kopplung von Experimenten**
Produktraumkonstruktion.

6 Zufallsvariablen

Definition 6.1: Es seien $(\Omega, \mathcal{A}), (\Omega', \mathcal{A}'$ zwei Messräume. Wir betrachten Abbildungen, $X : \Omega \to \Omega'$ heißt $\mathcal{A}/\mathcal{A}'$-messbar, falls für alle $B \in \mathcal{A}'$
$X^{-1}(B) := \{\omega \in \Omega | X(\omega) \in B\} \in \mathcal{A}$. Messbare Abbildungen werden auch Zufallsvariablen genannt.

Ist $(\Omega', \mathcal{A}') = (\mathbb{R}^D, \mathscr{B}(\mathbb{R}^D))$, so spricht man von einer $\mathbb{R}^D$-wertigen Zufallsvariable. Ist $(\Omega', \mathcal{A}') = (\bar{\mathbb{R}}, \bar{\mathscr{B}})$, wobei $\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$ und $B \in \bar{\mathscr{B}}$ genau dann, wenn $B \cap \mathbb{R} \in \mathscr{B}(\mathbb{R})$ so spricht man von einer numerischen Zufallsvariable.

Satz 6.2: *Seien $(\Omega, \mathcal{A}), (\Omega', \mathcal{A}'$ zwei Messräume. Dann ist $X : \Omega \to \Omega'$ $\mathcal{A}/\mathcal{A}'$-messbar, falls ein Erzeugendensystem $\varepsilon \subset 2^{\Omega}$ von $\mathcal{A}'$ existiert, so dass $X^{-1}(A) \in \mathcal{A}$ für alle $A \in \varepsilon$.*

Satz 6.3: *Seien $(\Omega_i, \mathcal{A}_i), i = 1, 2, 3$ Messräume. Seien $X_1 : \Omega_1 \to \Omega_2$ $\mathcal{A}_1/\mathcal{A}_2$-messbar und $X_2 : \Omega_2 \to \Omega_2$ $\mathcal{A}_2/\mathcal{A}_3$-messbar. Dann ist*

$$X_2 \circ X_1 : \Omega_1 \to \Omega_3 \ \mathcal{A}_1/\mathcal{A}_3\text{-messbar.}$$

Endliche Produkte und endliche Summen von Zufallsvariablen $X_1, ..., X_D : \Omega \to \mathbb{R}$ sind $\mathbb{R}^D$-wertige Zufallsvariablen.

Satz 6.4: *Sei $(X_n)_{n \in \mathbb{N}}$ eine Folge von numerischen Zufallsvariablen. Dann:*

(i) $\sup\limits_{n \in \mathbb{N}} X_n, \inf\limits_{n \in \mathbb{N}} X_n, \limsup\limits_{n \to \infty} X_n, \liminf\limits_{n \to \mathbb{N}} X_n$ *sind numerische Zufallsvariablen.*

(ii) $A := \{\omega | \lim\limits_{n \to \infty} X_n(\omega)$ *existiert in* $[-\infty, +\infty]\} \in \mathcal{A}$ *und*

$$X = \begin{cases} \lim\limits_{n \to \infty} X_n(\omega), & \omega \in \mathcal{A} \\ 0, & \text{sonst} \end{cases}$$

Definition 6.5: Sei $(\Omega, \mathcal{A}, P)$ ein W-Raum und $(\Omega', \mathcal{A}')$ ein Messraum. Ist $X : \Omega \to \Omega'$ eine Zufallsvariable, so heißt das auf $(\Omega', \mathcal{A}')$ definierte W-Maß P_X, wobei $P_X(B) = P(X^) - 1(B)$, $B \in \mathcal{A}'$, die Verteilung von X. Man überprüft leicht, dass P_X ein W-Maß auf $(\Omega', \mathcal{A}')$ ist. Ist X $\mathbb{R}^D$-wertig, so wird

$$F_X(x) = P_X((-\infty, x]), x \in \mathbb{R}^D$$

Es handelt sich dabei um die Verteilungsfunktion von P_X.

Konvention 6.6: (i) $\{X \in B\} := \{\omega | X(\omega) \in B\}$

(ii) Wir sagen eine Eigenschaft gilt μ-fast sicher, falls die Menge der ω, für die die Eigenschaft nicht gilt Maß 0 unter μ hat. Zum Beispiel: $X \leq Y$ μ-fast sicher bedeutet: $\mu(\{X > Y\}) = \mu(\{\omega | X(\omega) > Y(\omega)\}) = 0$

(iii) Zwei Zufallsvariablen X, Y heißen gleich is auf Modifikation, falls $X = Y$ μ-fast sicher.

(iv) $X : (\Omega, \mathcal{A}) \to (\Omega', \mathcal{A}')$ bedeutet $X : \Omega \to \Omega'$ ist $\mathcal{A}/\mathcal{A}'$-messbar.

Satz 6.7: *Sei $X : (\Omega, \mathcal{A}) \to (\Omega', \mathcal{A}')$ eine Zufallsvariable. Dann ist*

$$\sigma(X) := \{X^{-1}(A) | A \in \mathcal{A}'\} \subset \mathcal{A}$$

eine Unter-σ-Algebra von $\mathcal{A}$, die von X erzeugt σ-Algebra.

Definition 6.8: Sei $(\Omega, \mathcal{A}, P)$ ein W-Raum. Eine Familie von Zufallsvariablen $(X_i)_{i \in I}, I \neq \emptyset$ mit Werten in $(\Omega_i, \mathcal{A}_i)$ (d.h. jedes X_i ist $\mathcal{A}/\mathcal{A}_i$- messbar) heißt unabhängig, falls die Familie von σ-Algebren $(\sigma(X_i)_{i \in I}$ unabhängig ist, also genau dann, wenn jede endliche Teilfamilie es ist.

Satz 6.9: *Sei $(\Omega, \mathcal{A}, P)$ ein W-Raum und $(X_i)_{i=1,\dots,n}$ eine endliche Familie von Zva mit Werten in $(\Omega_i, \mathcal{A}_i)$. Zu jedem $\mathcal{A}_i$ sei ε_i ein $\cap$-stabiler Erzeuger mit $\Omega_i \in \varepsilon_i$. Dann sind äquivalent:*

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

(ii) $P(\bigcap\limits_{i=1}^{n} \{X_i \in E_i\}) = \prod\limits_{i=1}^{n} P(\{X_i \in E_i\}) \, \forall \, E_i \in \varepsilon_i, i = 1, .., n$

(iii) $P_X = P_{X_1} \otimes \dots \otimes P_{X_n}$, *wobei* $X = (X_1, \dots, X_n)$

Beispiel 6.10: Sei X eine standardnormalverteilte Zufallsvariable auf $(\Omega, \mathcal{A}, P)$. Seien $Y = |X|$ und $Z := \operatorname{sgn}(X)$. Dann:

$$F_Y(y) = P(\{|X| \leq y\}) = \begin{cases} 0 & y < 0 \\ \sqrt{\frac{2}{\pi}} \int\limits_0^y e^{\frac{-u^2}{2}} & y \geq 0 \end{cases}$$

$$F(Z) = P(\{\operatorname{sgn}(x) \leq z\}) \begin{cases} 0 & z < -1 \\ \frac{1}{2} & -1 \leq z < 1 \\ 1 & z \geq 1 \end{cases}$$

$$F_{(y,z)}(y, z) = P(\{|X| \leq y\} \cap \{\operatorname{sgn}(X) \leq z\}) = \begin{cases} 0 & y < 0 \text{ oder } z < -1 \\ \frac{1}{\sqrt{2\pi}} \int\limits_0^y e^{\frac{-u^2}{2}} du & y \geq 0 \text{ und } -1 \leq z < 1 \\ \sqrt{\frac{2}{\pi}} \int\limits_0^y e^{-\frac{u^2}{2}} du, & y \geq 0 \text{ und } z \geq 1 \end{cases}$$

Satz 6.11: *Sei $(X_i)_{i=1,..,n+1}$ eine Familie von $(\Omega_i, \mathcal{A}_i)$-wertigen Zva. Dann: $(X_1, \dots, X_{n+1})$ ist unabhängig genau dann, wenn $(X_1, \dots, X_n)$ unabhängig ist und X_{n+1} unabhängig von $(X_1, \dots, X_n)$ ist.*

Satz 6.12: *Sei $(X_i)_{i \in I}$ eine Familie von unabhängigen $(\Omega_i, \mathcal{A}_i)$-wertigen Zva. Seien $Y_i : (\Omega_i, \mathcal{A}_i) \to (\Omega'_i, \mathcal{A}'_i)$, $i \in I$. Dann ist die Familie $(Y_i \circ X_i)_{i \in I}$ ebenfalls unabhängig.*

7 Maßintegrale und Erwartungswerte

(a) **Konstruktion des Maßintegrals**

Wir wollen Erwartungswerte für allgemeine Zufallsvariablen einführen. Es bietet sich an Integrale der Form $\int xp_X(dx)$ zu nehmen. Sei in diesem Abschnitt $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Für Elemente in $\varepsilon^+(\mathcal{A})$ haben wir bereits das Elementarintegral (Satz 2.8) eingeführt. Bezeichne nun

$$\bar{\varepsilon}(\mathcal{A}) := \{f : \Omega \to [0, \infty] | f \mathcal{A}/\bar{B}\text{-messbar}\}$$

die Menge der nichtnegativen numerischen, messbaren Funktionen. Wir nennen eine Folge $(f_n)_{n\in\mathbb{N}}$ in $\varepsilon^+(\mathcal{A})$ approximierend für eine Funktion f, falls

(1) $f_n(\omega) \leq f_{n+1}(\omega)$ für alle $n \in \mathbb{N}, \omega \in \Omega$.
(2) $\sup\limits_{n\in\mathbb{N}} f_n(\omega) = f(\omega)$ für alle $\omega \in \Omega$.

Es gilt: $f \in \bar{\varepsilon}(\mathcal{A}) \Leftrightarrow \exists\,(f_n)_{n\in\mathbb{N}}$ in $\varepsilon^+(\mathcal{A})$ approximierend für f.

Definition 7.1: Sei $f \in \bar{\varepsilon}(\mathcal{A})$ und $(f_n)_{n\in\mathbb{N}}$ eine approximierende Folge in $\varepsilon^+(\mathcal{A})$. Dann heißt $\int\limits_\Omega f(\omega)\mu(d\omega) := \int f d\mu := \sup\limits_{n\in\mathbb{N}} I_\mu(f_n)$ Maßintegral von f bezüglich μ. Es ist wohldefiniert. Die explizite Darstellung ist:

$$\int f d\mu = \sup\{I_\mu(g) | g \leq f, g \in \varepsilon^+(\mathcal{A})\}$$

Definition 7.2: Sei $f : \Omega \to \bar{\mathbb{R}}$ $\mathcal{A}/\bar{B}$-messbar. Sei $f_+ = \max\{f, 0\}, f_- = \min\{f, 0\}$

(1) f heißt μ-quasi-integrierbar falls $\int f_+ d\mu < \infty$ oder $\int f_- d\mu < \infty$. In diesem Fall definiert man das μ-integral von f durch

$$\int f d\mu = \int f_+ d\mu - \int f_- d\mu$$

. Es kann die Werte $+\infty$ und $-\infty$ annehmen.

(2) f heißt μ-quasi-integrierbar falls $\int f_+ d\mu < \infty$ und $\int f_- d\mu < \infty$. In diesem Fall ist $\int f d\mu \in \mathbb{R}$.

(3) Sei $(\Omega, \mathcal{A}, P)$ ein W-Raum und X eine numerische Zufallsvariable, die P-quasi-integrierbar ist. Dann wird

$$E[X] := \int X dP$$

Erwartungswert von X genannt.

Satz 7.3: *Für eine messbare numerische Funktion f sind äquivalent:*

(1) *f ist μ-integrierbar*
(2) *Es gibt μ-integrierbare Funktionen $f_1, f_2 \in \bar{\varepsilon}(\mathcal{A})$ mit $f = f_1 - f_2$*
(3) *Es gibt ein μ-integrierbares $g \in \bar{\varepsilon}(\mathcal{A})$ mit $|f| \leq g$*
(4) *$|f|$ ist μ-integrierbar*

Falls (2) gilt so folgt $\int f d\mu = \int f_1 d\mu - \int f_2 d\mu$

Satz 7.4: *Seien f, g μ-integrierbar und $\alpha \in \mathbb{R}$. Dann sind $\alpha f, f + g$ (falls überall definiert),* $\max f, g, \min f, g$ *μ-integrierbar und es gelten:*

$$\int \alpha f d\mu = \alpha \int f d\mu, \quad \int f + g d\mu = \int f d\mu + \int g d\mu.$$

Ist $f \leq g$ so folgt $\int f d\mu \leq \int g d\mu$.

(b) **Konvergenzsätze und L^p-Räume**

Satz 7.5: *(Integration und Nullmengen)*

(1) *Für $f \in \bar{\varepsilon}(\mathcal{A})$ gilt:*

$$\int f d\mu \Leftrightarrow f = 0 \text{ } \mu\text{-fast sicher } (\mu(\{\omega | f(\omega) \neq 0\}) = 0)$$

(2) *Seien f, g numerische messbare Funktionen mit $f = g$ μ-fast sicher. Dann: Ist f μ-integrierbar, so auch g und $\int f d\mu = \int g d\mu$. Dies gilt auch falls $f, g \in \bar{\varepsilon}(\mathcal{A})$*
(3) *Ist f μ-integrierbar, so $|f| < \infty$ μ-fast sicher.*

Satz 7.6: *(Beppo-Levi) Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge in $\bar{\varepsilon}(\mathcal{A})$ mit $f_n \leq f_{n+1}$ für alle $n \in \mathbb{N}$. Sei f eine numerische messbare Funktion mit*

$$\lim_{n \to \infty} f_n = f \text{ } \mu\text{-f.s.}$$

. Dann ist f μ-quasi-integrierbar und $\int f d\mu = \lim_{n \to \infty} \int f_n d\mu$ ($\lim_{n \to \infty} f_n = f \mu$-fast sicher heißt: $\exists\, N \in \mathcal{A} : \mu(N) = 0$ und $\lim_{n \to \infty} f_n(\omega) = f(\omega) \,\forall\, \omega \in N^c$)

Beweis: Folgt aus dem Satz von Beppo-Levi für das Elementarintegral und der Konstruktion des μ-Integrals. Zunächst für $f = \sup_{n \in \mathbb{N}} f_n$ und dann durch Abänderung auf einer Nullmenge für den allgemeinen Fall: siehe Analysis III. ■

Lemma 7.7: *Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge $\bar{\varepsilon}(\mathcal{A})$. Dann:*

$$\int \liminf_{n \in \mathbb{N}} f_n d\mu \leq \liminf_{n \to \infty} \int f_n d\mu$$

Beweis: Folgt direkt aus Beppo Levi, wähle die monoton steigende Folge $g_n := \inf_{m \geq n} f_m$ ■

Satz 7.8: *(Lebesgue) Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von messbaren numerischen Funktionen.Sei $g \in \bar{\varepsilon}(\mathcal{A})$ mit $\int g d\mu < \infty$ und $|f_n| \leq g$. Falls für eine messbare numerische Funktionen f gilt, dass $f = \lim_{n\to\infty} f_n$ μ-fast sicher, so ist f μ-integrierbar und*

$$\int f d\mu = \lim_{n\to\infty} \int f_n d\mu$$

.

Definition 7.9: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $p \in [1, \infty)$. Eine Messbare reellwertige Funktion f heißt p-fach μ-integrierbar, falls $|f|^p \mu$-integrierbar ist. Wir schreiben $\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ für die Menge der p-fach integrierbaren Funktionen. Für $f \in \mathcal{L}^p(\mu)$ sei

$$||f||_p := (\int |f|^p d\mu)^{\frac{1}{p}}.$$

Satz 7.10: *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{L}^p(\mu)$ so, dass $|f_n| \leq g$ für ein μ-integrierbares $g \in \bar{\varepsilon}(\mathcal{A})$ mit $\int |g|^p d\mu < \infty$.*
Sei f eine reellwertige messbare Funktion mit $\lim_{n\in\mathbb{N}} f_n = f$ μ-fast sicher. Dann konvergiert (f_n) gegen f in $\mathcal{L}^p(\mu)$, d.h.

$$||f_n - f||_p \to 0 \ (n \to \infty$$

Satz 7.11: *Hölder-Ungleichung und Minkowski-Ungleichung*

(c) **Integration bezüglich Produktmaßen**

Satz 7.12: *(Fubini) Seien $(\Omega_i, \mathcal{A}_i, \mu_i), i = 1, 2, \sigma$-endliche Maßräume und $f : \Omega_1 \times \Omega_2 \to \bar{\mathbb{R}}$ $\mathcal{A}_1 \otimes \mathcal{A}_2/\bar{\mathscr{B}}$-messbar. Dann :*

(1) *Für alle $\omega_1 \in \Omega_1$ ist f_{ω_1} $\mathcal{A}_2/\bar{\mathscr{B}}$ und für alle $\omega_2 \in \Omega_2$ ist $f_{\omega_2} \mathcal{A}_1/\bar{\mathscr{B}}$-messbar.*

(2) *Ist f nicht negativ, so ist $\omega_1 \mapsto \int f_{\omega_1} d\mu_2$ $\mathcal{A}_1/\bar{\mathscr{B}}$-messbar, $\omega_2 \mapsto \int f_{\omega_2} d\mu$ $\mathcal{A}_2/\bar{\mathscr{B}}$-messbar und*

$$\int f d(\mu_1 \otimes \mu_2) = \int (\int (f_{\omega_1 d\mu_2}) \mu_1(d\omega_1) = \int (\int f_{\omega_2} d\mu_1) \mu_2(d\omega_2) (*)$$

(3) *Ist f $\mu_1 \otimes \mu_2$- integrierbar, so ist f_{ω_1} für μ_1-fast alle ω_1 μ_2-integrierbar, f_{ω_2} für μ_2-fast alle $\omega_2 \in \Omega_2$ μ_1-integrierbar. Die fast überall definierten Funktionen $\omega_1 \mapsto \int f_{\omega_1} d\mu_2$ (bzw. $\omega_2 \mapsto \int f_{\omega_2} d\mu_1$) sind μ_1-integrierbar (bzw. μ_2-integrierbar) und $(*)$ gilt.*

8 Methoden zur Berechnung von Maßintegralen u. Erwartungswerten

(a) **Transformationssatz und partielle Integration**
Seien $(\Omega, \mathcal{A}, \mu)$ ein Maßraum $(\Omega', \mathcal{A}')$ ein Messraum und $T : \Omega \to \Omega'$ $\mathcal{A}/\mathcal{A}'$-messbar. Das *Bildmaß* von μ unter T auf $(\Omega', \mathcal{A}')$ ist definiert durch $\mu_T(A') = \mu(T^{-1}(A')), A' \in \mathcal{A}'$.

Satz 8.1: *Ist in obiger Situation $f : (\Omega', \mathcal{A}') \to (\bar{\mathbb{R}}, \bar{\mathscr{B}})$, so ist f μ_T-integrierbar genau dann wenn $f \circ T$ μ-integrierbar ist. In diesem Fall gilt:*

$$\int\limits_{\Omega'} f d\mu_T = \int\limits_{\Omega} f \circ T d\mu$$

Satz 8.2: *(Partielle Integration) Seien F, G maßerzeugend und G stetig. Dann gilt für $a \leq b$:*

$$F(b)G(b) = F(a)G(a) + \int\limits_{(a,b]} G(s)F(ds) + \int\limits_{(a,b]} F(s)G(ds)$$

Beweis: Verwende Fubini und die Stetigkeit von G. Dann rechnen. ∎

Satz 8.3: *Sei X eine reellwertige (oder num.) Zufallsvariable auf $(\Omega, \mathcal{A}, P)$. Dann:*

$$X \text{ ist } P\text{-integrierbar} \Leftrightarrow \int\limits_0^\infty ((1 - F(x)) + F(-x))dx < \infty,$$

wobei F die Verteilungsfunktion von X ist. In diesem Fall gilt:

(1) $E[|X|] = \int\limits_0^\infty ((1 - F(x)) + F(-x))dx$

(2) $E[X] = \int\limits_0^\infty ((1 - F(x)) - F(-x))dx$

(Integrale können als uneigentliche Riemann-Integrale interpretiert werden)

(b) **Maße mit Dichten**

Satz 8.4: (1) *Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, $f \in \bar{\varepsilon}(\mathcal{A})$. Dann wird durch*

$$\nu(A) := \int 1_A f d\mu, \ A \in \mathcal{A}$$

ein Maß auf $(\Omega, \mathcal{A})$ definiert. f wird Dichte von ν bezüglich μ genannt. ν ist ein W-Maß genau dann, wenn $\int f d\mu = 1$.

(2) *Eine numerische messbare Funktion g ist genau dann ν-integrierbar, wenn $g * f$ μ-integrierbar ist. In diesem Fall gilt:*

$$\int g d\nu = \int f g d\mu$$

9 Varianz, Co-Varianz und Faltung

Definition 9.1: Seien X, Y numerische Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$

(i) Ist X P-integrierbar, so heißt

$$\mathrm{Var}(X) = E[(X - E[X])^2]$$

die Varianz von X und $\sqrt{\mathrm{Var}(X)}$ die Standardabweichung von X.

(ii) Sind X, Y und XY P-integrierbar, so wird

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

die Kovarianz von X und Y genannt. X und Y heißen unkorreliert, falls $\mathrm{Cov}(X, Y) = 0$

(iii) Falls $X, Y \in \mathcal{L}^2(\Omega, \mathcal{A}, P)$ und $\mathrm{Var}(X) \neq 0 \neq \mathrm{Var}(Y)$, so wird

$$\mathrm{Corr}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$$

die Korrelation von X und Y genannt.

(iv) Ist $X \in \mathcal{L}^n(\Omega, \mathcal{A}, P)$ für ein $n \in \mathbb{N}$, so heißt $E[X^2]$ das n-te Moment von X und $E[|X|^n]$ das n-te Absolutmoment von X.

Satz 9.2: *Seien X, Y numerische Zufallsvariablen. Dann:*

(i) $X \in \mathcal{L}^2(\Omega, \mathcal{A}, P) \Leftrightarrow X$ *P-integrierbar und Var$(X) < \infty$. In diesem Fall gilt:*
   (1) $Var(X) = E[X^2] - E[X]^2$
   (2) $Var(aX + b) = a^2 \; Var(X)$
   (3) $Var(X) = \min\limits_{a \in \mathbb{R}} E[(X - a)^2]$

(ii) *Sind X und Y in $L^2(\Omega, \mathcal{A}, P)$ mit $Var(X) \neq 0 \neq Var(Y)$, so ist*

$$|Corr(X, Y)| \leq 1 \textit{ und}$$

$$|Corr(X, Y)| = 1 \Leftrightarrow \exists\, (a, b) \in \mathbb{R}^2, a \neq 0 : Y = aX + b \textit{ P-fast sicher}$$

Satz 9.3: *Es seien $X_1, ..., X_n$ unabhängige rellwertige Zufallsvariablen.*

(i) *Sind $X_1, ..., X_n$ nicht negativ so ist*

$$E[\prod_{i=1}^{n} X_i] = \prod_{i=1}^{n} E[X_i] (*)$$

(ii) *Sind* $X_1, ..., X_n$ P*-integrierbar, so gilt* $\prod\limits_{i=1}^{n} X_i$ P*-integrierbar und* $(*)$ *gilt. Insbesondere sind* $X_1, ..., X_n$ *dann paarweise unkorreliert.*

Satz 9.4: *(Bienaymé) Sind* $X_1, ..., X_n$ *paarweise unkorreliert, so ist*
$Var(\sum\limits_{i=1}^{n} X_i) = \sum\limits_{i=1}^{n} Var(X_i)$.

Beweis: $\mathrm{Var}(\sum\limits_{i=1}^{n} X_i) = E[(\sum\limits_{i=1}^{n} X_i - E[\sum\limits_{i=1}^{n} X_i])^2] = E[(\sum\limits_{i=1}^{n} X_i - E[X_i])^2]$
$= \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} E[(X_i - E[X_i])(X_j - E[X_j])] = \sum\limits_{i=1}^{n} \mathrm{Var}(X_i)$ ■

Definition 9.5: Seien X und Y zwei unabhängige $\mathbb{R}^D$-wertige Zufallsvariablen mit Verteilungen P_X und P_Y. Dann heißt die Verteilung $X + Y$ die Faltung der beiden Verteilung P_X und P_Y und wird mit $P_X * P_Y$ notiert.

Satz 9.6: *(Faltungsformel) Es seien* X *und* Y *unabhängige* $\mathbb{R}^D$*-wertige Zufallsvariablen. Dann gilt für alle* $A \subset \mathscr{B}_D$:

$$P_X * P_Y = \int\limits_{\mathbb{R}^D} P_X(-y + A) P_Y(dy) = \int\limits_{\mathbb{R}^D} P_Y(-x + A) P_X(dx)$$

. Haben P_X *und* P_Y *Dichten* f_X *und* f_Y *bezüglich des Lebesquemaßes* $\lambda^{\otimes D}$ *so ist*

$$f_{X+Y}(z) := \int\limits_{\mathbb{R}^D} f_X(z - y) f_Y(y) \lambda^{\otimes D}(dy) = \int\limits_{\mathbb{R}^D} f_Y(z - x) f_X(x) \lambda^{\otimes D}(dx)$$

die Dichte von $P_X * P_Y$ *bezüglich* $\lambda^{\otimes D}$.

10 Konvergenz von Zufallsvariablen

LAUT MATTHIAS WICHTIG FÜR DIE PRÜFUNG

Definition 10.1: Seien $(X_n)_{n\in\mathbb{N}}$ eine FOlge von reellwertigen Zufallsvariablen und X eine $\mathbb{R}$-wertige Zufallsvariable auf einem W-Raum $(\Omega, \mathcal{A}, P)$. Wir sagen:

(i) $(X_n)_{n\in\mathbb{N}}$ konvergiert gegen X in Wahrscheinlichkeit ($X_n \xrightarrow{P} X$) falls $\forall \varepsilon > 0 \; P(\{|X_n - X| > \varepsilon\}) \to 0 (n \to \infty)$

(ii) $(X_n)_{n\in\mathbb{N}}$ konvergiert gegen X P-fast sicher ($X_n \xrightarrow{f.s.} X$), falls $P(\{\lim\limits_{n\to\infty} X_n = X\}) = 1$

(iii) $(X_n)_{n\in\mathbb{N}}$ konvergiert gegen X im p-ten Mittel ($X_n \xrightarrow{L^p} X$), $p \in [1, \infty)$ falls $E[|X_n - X|^p] \to 0 (n \to \infty)$.

Satz 10.2: *In der Situation von Def.10.1 betrachten wir die Eigenschaften*

(i) $X_n \xrightarrow{f.s} X$

(ii) $\forall \varepsilon > 0 \lim\limits_{N\to\infty} P(\{\sup\limits_{n\geq N} |X_n - X| > \varepsilon\}) = 0$

(iii) $\forall \varepsilon > 0 \sum\limits_{n=1}^{\infty} P(\{|X_n - X| > \varepsilon\}) < \infty$

(iv) $X_n \xrightarrow{P} X$

Dann gilt: (iii) $\Rightarrow$ (i) $\Leftrightarrow$ (ii) $\Rightarrow$ (iv)

Lemma 10.3: *(Borel-Cantelli) Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum, $(A_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{A}$. Dann:*

(i) $\sum\limits_{n=1}^{\infty} P(A_n) < \infty \Rightarrow P(\bigcap\limits_{n=1}^{\infty} \bigcup\limits_{m=n}^{\infty} A_m) = 0$

(ii) $\sum\limits_{n=1}^{\infty} P(A_n) = +\infty$ *und* $(A_n)_{n\in\mathbb{N}}$ *unabhängig* $\Rightarrow P(\bigcap\limits_{n=1}^{\infty} \bigcup\limits_{m=n}^{\infty} A_m) = 1$

Beweis: Stetigkeit des W-Maßes von oben impliziert:

$$p := P(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m) = \lim_{n\to\infty} (P(\bigcup_{m=n}^{\infty} A_m))(*)$$

(i) Es folgt aus der σ-Subadditivität: $p \leq \lim\limits_{n\to\infty} \sum\limits_{m=n}^{\infty} P(A_m) = 0$

(ii) Mit $(A_n)_{n\in\mathbb{N}}$ sind auch deren Komplemente unabhängig. Inbesondere gilt $\forall\, n,k \in \mathbb{N}$

$$P(\bigcap_{m=n}^{n+k} A_m^c) = \prod_{m=n}^{n+k}(1-P(A_n)).$$

Aus $(*)$ folgt

$$\begin{aligned} p &= \lim_{n\to\infty}\lim_{k\to\infty} P(\bigcup_{m=n}^{n+k} A_m) = 1 - \lim_{n\to\infty}\lim_{k\to\infty} P(\bigcap_{m=n}^{n+k} A_m^c) \\ &= 1 - \lim_{n\to\infty}\lim_{k\to\infty} \prod_{m=n}^{n+k}(1-P(A_m)) \end{aligned}$$

Verwende nun die Ungleichung $\ln(1-x) \le -x$ für $0 \le x \le 1$ und erhalte

$$\log(\prod_{m=n}^{n+k}(1-P(A_m)) = \sum_{m=n}^{n+k}\log(1-P(A_m)) \le -\sum_{m=n}^{n+k}P(A_m) \to -\infty (k\to\infty)$$

Damit folgt die Aussage. ■

Beweis: (von Satz 10.2)
(iii) $\Rightarrow$ (i): Folgt aus dem Borel-Cantelli-Lemma:

$$\begin{aligned} X_n \xrightarrow{f.s.} X &\Leftrightarrow \forall\, k\in\mathbb{N} P(\bigcup_{N_0=1}^{\infty}\bigcap_{n=N_0}^{\infty}\{|X_n - X| \le \frac{1}{k}\}) = 1 \\ &\Leftrightarrow \forall\,\varepsilon > 0 P(\bigcup_{N_0=1}^{\infty}\bigcap_{n=N_0}^{\infty}\{|X_n - X| \le \varepsilon\}) = 1 \\ &\Leftrightarrow \forall\,\varepsilon > 0 P(\bigcap_{N_0=1}^{\infty}\bigcup_{n=N_0}^{\infty}\{|X_n - X| > \varepsilon\}) = 0 \end{aligned}$$

(i) $\Rightarrow$ (ii) Aus der Stetigkeit von P von oben gilt:

$$\begin{aligned} X_n \xrightarrow{f.s.} X &\Leftrightarrow \forall\,\varepsilon > 0 \lim_{N_0\to\infty} P(\bigcup_{n=N_0}^{\infty}\{|X_n - X| > \varepsilon\}) = 0 \\ &\Leftrightarrow \forall\,\varepsilon > 0 \lim_{N_0\to\infty} P(\{\sup_{n\ge N_0}|X_n - X| > \varepsilon\}) = 0 \end{aligned}$$

(ii) $\Rightarrow$ (iv): Klar, da $P(\{|X_n - X|\}) \le P(\{\sup_{k\ge n}|X_k - X| > \varepsilon\}) \xrightarrow{n\to\infty} 0$ ■

Satz 10.4: *Sei $(X_n)_{n\in\mathbb{N}}$ eine FOlge von $\mathbb{R}$-wertigen Zufallsvariablen auf $(\Omega, \mathcal{A}, P)$. Dann sind äquivalent:*

(a) $(X_n)_{n\in\mathbb{N}}$ *konvergiert fast sicher*

(b) *Es gibt ein* $N \in \mathcal{A}$ *mit* $P(N) = 0$ *so, dass* $\forall \omega \in N^c (X_n(\omega))_{n\in\mathbb{N}}$ *Cauchy-Folge in* $\mathbb{R}$ *ist.*

(c) $\forall \varepsilon > 0 \lim\limits_{N\to\infty} P(\{\sup\limits_{k\geq N} |X_k - X_N| > \varepsilon\}) = 0$

Satz 10.5: (i) *In der Situation von Def. 10.1 sind äquivalent:*

(a) $X_n \xrightarrow{P} X$

(b) $E[|X_n - X| \wedge 1] \to 0 (n \to \infty)$

(c) *Jede Teilfolge von* $(X_n)_{n\in\mathbb{N}}$ *hat eine weitere Teilfolge, die gegen* X *P-fast sicher konvergiert.*

(ii) $(X_n)_{n\in\mathbb{N}}$ *konvergiert in Wahrscheinlichkeit* $\Leftrightarrow$ *das folgende Cauchy-Krit. ist erfüllt* $\forall \varepsilon > 0, \delta > 0 \, \exists N_0 \in \mathbb{N} \, \forall m, n \geq N_0 P(\{|X_n - X_m| > \varepsilon\}) \leq \delta$

Satz 10.6: *Sei* X *eine numerische Zufallsvariable auf* $(\Omega, \mathcal{A}, P)$

(i) *Markov-Ungleichung: Für alle* $p \geq 1$ *und* $\varepsilon > 0$ *gilt:*

$$P(\{|X| > \varepsilon\}) \leq \frac{E[|X|]^p}{\varepsilon^p}$$

(ii) *Tschebyschev-Ungleichung: Ist* X *P-integrierbar, so gilt:*

$$P(\{|X - E[X]| \geq \varepsilon\}) \leq \frac{Var(X)}{\varepsilon}$$

Lemma 10.7: *Falls* $X_n \xrightarrow{P} X$*, so gilt:* $X_n \xrightarrow{P} \bar{X} \Leftrightarrow X = \bar{X}$ *P-fast sicher.*

Beweis: (von Satz 10.5) Skizze:

(i) $(a) \Rightarrow (b)$: Jede Teilfolge von $(X_n)_{n\in\mathbb{N}}$ konvergiert gegen X in Wahrscheinlichkeit und erfüllt nach (ii) das Cauchy-Kriterium. Verwende Satz 10.2 und Lemma 10.7 und erhalte die Aussage.
$(c) \Rightarrow (b)$: Auf $[0,1]$ beschränkte Folge $m_n := E[|X_n - X| \wedge 1]$. Dann Beweis durch Widerspruch.
$(b) \Rightarrow (a)$: Folgt aus der Markovungleichung mit $p = 1$.

(ii) Schritt 1: Zeige, dass das Cauchy-Kriterium gilt (nachprüfen).
Schritt 2: Zeige mit Hilfe des Cauchy-Kriterium und des
Borel-Cantelli-Lemmas, dass eine konvergente Teilfolge $X_{n_k} \xrightarrow{f.s.} X$.
Schritt 3: Zeige hier, dass $X_n \xrightarrow{P} X$. ■

Satz 10.8: *In der Situation von Def 10.1 gilt:*

$$\begin{array}{ccccc} & & \exists \text{ Teilfolge } X_{n_k} \xrightarrow{f.s.} X & & \\ & & \Uparrow (2) & & \\ X_n \xrightarrow{f.s.} X & \overset{(1)}{\Longrightarrow} & X_n \xrightarrow{P} X & \overset{(5)}{\Longleftarrow} & X_n \xrightarrow{L^p} X \text{ falls } |X_n| \leq k \\ & & \Uparrow (3) & & P\text{-f.s. für ein } k \geq 0 \\ & & X_n \xrightarrow{L^p} X \;\; p \geq 1 & & \\ & & \Uparrow (4) & & \\ & & X_n \xrightarrow{L^q} X \;\; q \geq p & & \end{array}$$

Beweis: (1) Satz 10.2
(2) Satz 10.5
(3) Beweis von Folgerung 7.15 (4) Nach (3) gilt: $X_n \xrightarrow{L^1} X$. Somit $E[|X_n - X| \wedge 1] \leq E[|X_n - X|] \xrightarrow{n\to\infty} 0$
(5) Sei $\varepsilon > 0$. Dann:

$$P(\{|\frac{X_n - X}{2k}|^p \geq \varepsilon\}) = P(\{|X_n - X| \geq 2k\varepsilon^{\frac{1}{p}}\}) \xrightarrow{n\to\infty} 0$$

Also

$$|\frac{X_n - X}{2k}|^p \to 0$$

in Wahrscheinlichkeit. Aus Satz 10.5 folgt:

$$E[|X_n - X|^p] = (2k)^p E[|\frac{X_n - X}{2k}|^p \wedge 1] \xrightarrow{n\to\infty} 0 \qquad \blacksquare$$

Beispiel 10.9: Gegenbeispiel, dass die Umkehrung von (1) nicht gilt: Sei X_n Bernoulli (p_n)-verteilt für $n \in \mathbb{N}$ und $(X_n)_{n\in\mathbb{N}}$ unabhängig. Dann gilt für alle $\varepsilon \in (0,1)$

$$P(\{|X_n| > \varepsilon\}) = P(\{X_n = 1\}) = p_n \text{ und } E[|X_n|^p] = p_n$$

Also folgt aus Satz 10.2 mit $X := 0$

- $X_n \xrightarrow{f.s.} 0 \Leftrightarrow \sum_{n=1}^{\infty} p_n < \infty$
- $p_n \to 0 \Leftrightarrow X_n \xrightarrow{P} X$ und $X_n \xrightarrow{L^p} X \Leftrightarrow p_n \to 0$

In diesem Spezialfall gilt $X_n \xrightarrow{P} X \Leftrightarrow X_n \xrightarrow{L^p} X$. Aus Konvergenz in Wahrscheinlichkeit und aus Konvergenz im p-ten Mittel folgt also i.A. nicht die fast sichere Konvergenz.
Gegenbeispiel, dass die Umkehrung von (4) nicht gilt: Sei $(\Omega, \mathcal{A}, P) = (\mathbb{R}, \mathscr{B}, \lambda_{[0,1]})$, $\lambda_{[0,1]}$ ist die Gleichverteilung auf dem Einheitsintervall. Sei

$$X_n(\omega) = \begin{cases} e^n, & \omega \leq \frac{1}{n} \\ 0, & \omega > \frac{1}{n} \end{cases}$$

Dann gilt: $X_n(\omega) \to 0$ für alle $\omega \in (0,1]$ also $X_n \to 0$ P-fast sicher und somit auch in Wahrscheinlichkeit. $(X_n)_{n\in\mathbb{N}}$ konvergiert aber nicht im p-ten Mittel für $p \geq 1$. Denn nach Satz 10.8 (4) kommt nur 0 als Grenzwert in Frage aber

$$E[|X_n|^p] = \frac{e^n}{n} \xrightarrow{n\to\infty} \infty$$

Konvergenz von Reihen unabhängiger Zufallsvariablen

Satz 10.10: *(Lévy) Sei $(X_n)_{n\in\mathbb{N}}$ eine unabhängige Folge von reellwertigen Zufallsvariablen auf $(\Omega, \mathcal{A}, P)$. Dann:*

$$\sum_{n=1}^{\infty} X_n \text{ konvergiert in Wahrscheinlichkeit } \Leftrightarrow \sum_{n=1}^{\infty} X_n \text{ konvergiert } P\text{-f.s.}$$

Lemma 10.11: *(Ungleichung von Ottaviani) Es seien $X_1, ..., X_n$ unabhängige reellwertige Zufallsvariablen auf $(\Omega, \mathcal{A}, P)$. Sei*

$$S_k = \sum_{i=1}^{k} X_i, k = 1, ..., n.$$

Dann gilt für alle $\eta, \delta > 0$ und $m = 1, ..., n$

$$P(\{\max_{m\leq k\leq n} |S_k| > \eta + \delta\}) \min_{m\leq k\leq n} P(\{|S_n - S_k| \leq \eta\}) \leq P(\{|S_n| > \delta\})$$

Beweis: (von Satz 10.10, Lévy)
"$\Rightarrow$": Sei $S_n = \sum\limits_{k=1}^{n} X_k, n \in \mathbb{N}$. Verwende dann die Ungleichung von Ottaviani, später 10.5 (ii) und später Satz 10.4, um die fast-sichere Konvergenz von $(S_n)_{n\in\mathbb{N}}$ zu zeigen.
"$\Leftarrow$": Satz 10.2 für $(S_n)_{n\in\mathbb{N}}$. ■

Satz 10.12: *(Zwei-Reihen-Kriterium, Kolmogorov) Sei $(X_n)_{n\in\mathbb{N}}$ eine unabhängige Folge von $\mathbb{R}$-wertigen Zufallsvariablen in $L^2(\Omega, \mathcal{A}, P)$. Falls $\sum\limits_{n=1}^{\infty} E[X_n]$ und $\sum\limits_{n=1}^{\infty} Var(X_n)$ konvergent sind, so konvergiert $\sum\limits_{n=1}^{\infty} X_n$ P-f.s.*

Beweis: Seien $k \geq m$. Dann ist mit Bienaymée

$$E[(\sum_{n=m}^{k} X_n)^2] = \mathrm{Var}(\sum_{n=m}^{k} X_n) + E[\sum_{n=m}^{k}]^2 = \sum_{n=m}^{k} \mathrm{Var}(X_n) + (\sum_{n=m}^{k} E[X_n])^2$$

Also ist $(\sum\limits_{k=1}^{n} X_k)_{n\in\mathbb{N}}$ Cauchy Folge in $\mathcal{L}^2(\Omega, \mathcal{A}, P)$ und die Aussage folgt aus Satz 10.9 (iv) und Satz 10.13. ■

11 Die Gesetze der großen Zahlen

Definition 11.1: Es sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen in $\mathcal{L}^1(\Omega, \mathcal{A}, P)$. Man sagt $(X_n)_{n\in\mathbb{N}}$ erfüllt das starke (schwache) Gesetz der großen Zahlen, falls

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (X_i - E[X_i]) = 0 \ P\text{-f.s. (in Wahrscheinlichkeit)}$$

Satz 11.2: *(Schwaches GGZ) Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von paarweise unkorrelierten Zufallsvariablen in $\mathcal{L}^2(\Omega, \mathcal{A}, P)$ mit*

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{i=1}^{n} Var(X_i) = 0(*)$$

Dann erfüllt $(X_n)_{n\in\mathbb{N}}$ das schwache Gesetzt der großen Zahlen. Hinreichende Bedingungen sind

(i) *$(X_n)_{n\in\mathbb{N}}$ sind identisch verteilt, d.h. $P_{X_n} = P_{X_1} \,\forall\, n \in \mathbb{N}$*

(ii) *$\exists\, c > 0 \,\forall\, n \in \mathbb{N} \ \ Var(X_n) \leq c$*

(iii) *$\sum\limits_{n=1}^{\infty} \frac{Var(X_n)}{n^2} < \infty$*

Beweis: $E[|\frac{1}{n}\sum(X_i E[X_i])|]^2 = \frac{1}{n^2}\sum\limits_{i=1}^{n} \text{Var}(X_i) \to 0$ für $n \to \infty$. Aber die L^2-Konvergenz zieht die Konvergenz in Wahrscheinlichkeit nach sich. Klar: $(i) \Rightarrow (ii) \Rightarrow (*)$ und $(iii) \Rightarrow (i)$: Kronecker (nächstes Lemma). ∎

Lemma 11.3: *(Kronecker) Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen, $(b_n)_{n\in\mathbb{N}}$ monoton steigende Folge in $\mathbb{R}$ mit $b_n \xrightarrow{n\to\infty} +\infty$. Falls $\sum\limits_{n\in\mathbb{N}} x_n$ konvergiert, so folgt*

$$\lim_{n\to\infty} \frac{1}{b_n} \sum_{i=1}^{n} b_i x_i = 0$$

.

Satz 11.4: *(Starkes GGZ, Kolmogorov) Sei $(X_n)_{n\in\mathbb{N}}$ eine unabhängige Folge von Zva in $\mathcal{L}^1(\Omega, \mathcal{A}, P)$. Dann erfüllt $(X_n)_{n\in\mathbb{N}}$ das starke GGZ, falls*

(i) *$(X_n)_{n\in\mathbb{N}}$ identisch verteilt; oder*

(ii) *$\sum\limits_{i=1}^{\infty} \frac{Var(X_n)}{n^2} < \infty$*

Beweis: (i) Ohne Einschränkung $E[X_i] = 0$. Sei

$$\bar{X}_n(\omega) = \begin{cases} X_n(\omega), & |X_n(\omega)| < n \\ 0, & |X_n(\omega)| \geq n \end{cases}$$

Zeige:

(a) $(X_n)_{n\in\mathbb{N}}$ erfüllt das starke GGZ $\Leftrightarrow$ $(\bar{X}_n)_{n\in\mathbb{N}}$ erfüllt das starke GGZ

(b) $\bar{X}_{n n\in\mathbb{N}}$ erfüllt das starke GGZ

(a) Aus Lesbesque und dem Töplitz Lemma folgt

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to 0 \ P\text{-f.s.} \Leftrightarrow \frac{1}{n}\sum_{i=1}^{n}(\bar{X}_i - E[\bar{X}_i]) \to 0 \ P\text{-f.s.}$$

(b) $(\bar{X}_n)_{n\in\mathbb{N}}$ ist unabhängig. Dann Teil (ii) nachrechnen.

(ii) Sei $S_n = \sum\limits_{i=1}^{n}(X_i - E[X_i])$. Dann $\frac{1}{n}S_n = \frac{1}{n}\sum\limits_{k=1}^{n} k \underbrace{\frac{X_k - E[X_k]}{k}}_{=:Y_k}$. Nach dem Kronecker Lemma ist $\frac{1}{n}S_n \to 0$ P-f.s. falls $\sum\limits_{k=1}^{\infty} Y_k$ P-f.s. konvergiert. Letzteres folgt aus dem Zwei-Reihen-Satz, da $E[Y_k] = 0$ und
$\sum\limits_{k=1}^{\infty} \operatorname{Var}(Y_n) = \sum\limits_{n=1}^{\infty} \frac{1}{n^2} \operatorname{Var}(X_n) < \infty$ ∎

Beispiel 11.5: (i) Weierstraßscher Approximationssatz

(ii) Monte-Carlo-Integration

12 Konvergenz von Wahrscheinlichkeitsmaßen

In diesem Abschnitt: (S, d) metrischer Raum und $\mathcal{S} = \mathscr{B}(S)$ die Borelsche σ-Algebra über S.

Definition 12.1: Sei $(P_n)_{n\in\mathbb{N}}$ eine Folge von W-Maßen auf $\mathcal{S}$ und P ein W-Maß auf $\mathcal{S}$

(i) Man sagt $(P_n)_{n\in\mathbb{N}}$ konvergiert schwach gegen P ($P_n \xrightarrow{w} P$), falls für alle stetigen und beschränkten Funktionen $f : S \to \mathbb{R}$

$$\int f dP_n \xrightarrow{n\to\infty} \int f dP$$

(ii) $(P_n)_{n\in\mathbb{N}}$ heißt wesentlich konvergent gegen P, falls für alle $A \in \mathcal{S}$ mit $P(\partial A) = 0$ gilt: $P_n(A) \xrightarrow{n\to\infty} P(A)$

Definition 12.2: Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von S-wertigen Zva mit Verteilung P_{X_n}, $n \in \mathbb{N}$. Sei X eine weitere S-wertige Zva. mit Verteilung P_X. Man sagt $(X_n)_{n\in\mathbb{N}}$ konvergiert in Verteilung gegen X, falls $P_{X_n} \xrightarrow{w} P_X$ (d.h. $E[f(X_n)] \to E[f(X)]$ für alle stetigen, beschränkten f). Hier können die Zva X_n von verschiedenen W-Räumen aus abbilden.

Satz 12.3: *(Portmanteau-Theorem) Die folgenden Aussagen sind im Kontext von Def.12.1 äquivalent:*

(i) $P_n \xrightarrow{w} P$

(ii) *Für alle abgeschlossenen Mengen $A \in S$ gilt:* $\limsup\limits_{n\to\infty} P_n(A) \leq P(A)$

(iii) *Für alle offenen Mengen $U \in S$ gilt:* $\liminf\limits_{n\to\infty} P_n(U) \geq P(U)$

(iv) $P_n \Rightarrow P$

Beweis: (i) $\Rightarrow$ (ii): Sei $d(x, A) = \inf\limits_{y\in A} d(x, y)$. Für $n \in \mathbb{N}$ sei $g_n(x) = 1 - (nd(x, A) \wedge 1)$. Dann $g_n(x) \geq 1_A, g_n \to 1_A(x)$ punktweise, g_n stetig und beschränkt. Aus (i) folgt:

$$\limsup_{n\to\infty} P_n(A) \leq \limsup_{n\to\infty} \int g_k dP_n = \int g_k dP \ \ \forall\, k \in \mathbb{N}$$

Da $|g_k| \leq 1$ folgt aus dem Satz von Lebesque

$$\limsup_{n\to\infty} P_n(A) \leq \int 1_A dP = P(A)$$

(ii) $\Rightarrow$ (iii) : Sei U offen. Dann ist $A := U^c$ abgeschlossen. Also folgt aus (ii): $\liminf\limits_{n\to\infty} P_n(U) = 1 - \limsup\limits_{n\to\infty} P_n(A) \geq 1 - P(A) = P(U)$
(iii) $\Rightarrow$ (ii) : analog.
(iii) $\Rightarrow$ (iv) : Sei $U = A^\circ$ und $K = \bar{A}$. Dann folgt aus (iii) $\Leftrightarrow$ (ii), da $P(\partial A) = 0$

$$\limsup_{n\to\infty} P_n(A) \leq \limsup_{n\to\infty} P_n(\bar{A}) \leq P(\bar{A}) \leq P(A) + P(\partial A) = P(A)$$
$$\liminf_{n\to\infty} P_n(A) \geq \liminf_{n\to\infty} P_n(A^\circ) \geq P(A^\circ) = P(A^\circ) + P(\partial A) \geq P(A)$$

(iv) $\Rightarrow$ (i): Sei $f : S \to \mathbb{R}$ stetig und beschränkt durch $M \geq 0$.
Sei $D = \{t \in \mathbb{R} | P(\{f = t\}) \neq 0\}$. Da Mengen der Form $f^{-1}(\{t\}), t \in D$ disjunkt sind und P ein endliches Maß ist, ist D höchstens abzählbar. Sei ohne Einschränkung $M \notin D$, $\varepsilon > 0$ fest. Wähle eine Partition $\pi = (t_0, ..., t_k)$ von $[-M, M]$, so dass $t_i \notin D$ für $i = 1, .., k-1$ und $|\pi| = \max\limits_{i=1,...,k} (t_i - t_{i-1}) \leq \varepsilon$. Sei:

$$B_i = \{t_i \leq f < t_{i+1}\}, i = 0, ..., k-2 \text{ und } B_{k-1} := \{t_{k-1} \leq f \leq t_k\}.$$

Da f stetig ist, ist $f^{-1}((t_i, t_{i+1}))$ offen. Also ist $\partial B_i \subset f^{-1}(\{t_i\}) \cup f^{-1}(\{t_{i+1}\})$. Somit ist

$$P(\partial B_i) \leq P(\{f = t_i\}) + P(\{f = t_{i+1}\}) = 0.$$

Also liefert (iv)

$$\sum_{i=0}^{k-1} t_i P_n(B_i) \xrightarrow{n\to\infty} \sum_{i=0}^{k-1} t_i P(B_i)$$

Somit gilt für hinreichend großes $n \geq N_0$:

$$\begin{aligned} |\int f dP_n - \int f dP| &= |\sum_{i=0}^{k-1} \int 1_{B_i}(f - t_i) dP_n| + |\sum_{i=0}^{k-1} t_i P_n(B_i) - \sum_{i=0}^{k-1} t_i P(B_i)| \\ &+ |\sum_{i=0}^{k-1} \int 1_{B_i}(t_i - f) dP| \leq 3\varepsilon \end{aligned}$$

■

Satz 12.4: *(Helly-Bray) Im Fall $S = \mathbb{R}$ gilt:*

$$P_n \xrightarrow{w} P \Leftrightarrow F_n(x) \to F(x) \text{ für alle Stetigkeitsstellen } x \text{ von } F$$

Hierbei sind F_n, F die Verteilungsfunktion zu P_n, P.

Beweis: "$\Rightarrow$": $x \in \mathbb{R} \Rightarrow \partial(-\infty, x] = \{x\}$. Also

$$P(\partial(-\infty, x]) = 0 \Leftrightarrow P(\{x\}) = 0 \Leftrightarrow x \text{ ist Stetigkeitsstelle von } .F$$

Aus dem Portmanteautheorem ((i) $\Rightarrow$ (iv)) folgt für alle Stetigkeitsstellen x von F:

$$F_n(x) = P_n((-\infty, x]) \to P((-\infty, x]) = F(x)$$

"$\Leftarrow$": Sei $U \subset \mathbb{R}$ offen. Dann existiert eine abzählbare Familie von offenen Intervallen $(a_k, b_k)_{k\in\mathbb{N}}$, paarweise disjunkt mit $U = \bigcup\limits_{k\in\mathbb{N}} (a_k, b_k)$.

Sei $\varepsilon > 0$ fest. Ist $I_k := (a_k, b_k) \neq \emptyset$, wähle ein Teilintervall $I_k \supset I'_k = (a'_k, b'_k]$ so, dass $P(I'_k) \geq P(I_k) - \varepsilon 2^{-k}$ und dass a'_k, b'_k Stetigkeitsstellen von F sind. Also (die erste Ungleichung gilt mit Fatou)

$$\begin{aligned}\liminf_{n\to\infty} P_n(U) &= \liminf_{n\to\infty} \sum_{k=1}^{\infty} P_n(I_k) \geq \sum_{k=1}^{\infty} \liminf_{n\to\infty} P_n(I_k) \geq \sum_{k=1}^{\infty} \liminf_{n\to\infty} \underbrace{P_n(I'_k)}_{F_n(b'_k)-F_n(a'_k)} \\ &= \sum_{k=1}^{\infty} F(b'_k) - F(a'_k) = \sum_{k=1}^{\infty} P(I'_k) \geq \sum_{k=1}^{\infty} P(I_k) - \varepsilon = P(U) - \varepsilon\end{aligned}$$

Mit $\varepsilon \to 0$ folgt (iii) des Portmanteautheorems , somit schwache Konvergenz. ■

Satz 12.5: *Sei S separabel (d.h. es gibt eine höchstens abzählbare Teilmenge von S, die dicht in S liegt). Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von S-wertigen Zva auf $(\Omega, \mathcal{A}, P)$ und X eine weitere S-wertige Zva auf $(\Omega, \mathcal{A}, P)$. Dann gilt*

$$X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$$

d.h. dann impliziert Konvergenz in Wahrscheinlichkeit die Konvergenz in Verteilung.

Satz 12.6: *Sei S separabel. Seien $X_n, Y_n, X, n \in \mathbb{N}$ S-wertige Zva auf $(\Omega, \mathcal{A}, P)$. Dann:*

$$X_n \xrightarrow{d} X, d(X_n, Y_n) \xrightarrow{P} 0 \Rightarrow Y_n \xrightarrow{d} X$$

Satz 12.7: *(Slutsky) Sei $S = \mathbb{R}$. Seien X_n, Y_n, X reellwertige Zva auf $(\Omega, \mathcal{A}, P)$. Sei $a \in \mathbb{R}$. Falls $X_n \xrightarrow{d} X$ und $Y_n \xrightarrow{P} a$, so gilt:*

- $X_n + Y_n \xrightarrow{d} X + a$
- $Y_n X_n \xrightarrow{d} aX$
- $\frac{X_n}{Y_n} \xrightarrow{d} \frac{X}{a}$, *falls* $a \neq 0$.

Beweis: siehe Übung. ■

Definition 12.8: Eine Menge $\mathcal{P}$ von W-Maßen auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ wird relativ kompakt genannt, falls jede Folge in $\mathcal{P}$ eine gegen ein Wahrscheinlichkeitsmaß schwach konvergente Teilfolge hat.

Definition 12.9: Eine Menge $\mathcal{P}$ von W-Maßen auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ heißt straff, falls für jedes $\varepsilon > 0$ eine kompakte Menge $K \subset \mathbb{R}$ existiert, so dass

$$\sup_{P\in\mathcal{P}} P(\mathbb{R}\backslash K) \leq \varepsilon.$$

Lemma 12.10: *Jede relativ kompakte Menge von W-Maßen P auf $\mathscr{B}(\mathbb{R})$ ist straff.*

Satz 12.11: *(Prohorov) Eine Menge $\mathcal{P}$ von W-Maßen auf $\mathscr{B}(\mathbb{R})$ ist relativ kompakt genau dann wenn sie straff ist.*

Es bezeichne J die Menge von verallgemeinerten Verteilungsfunktion, also von Funktionen $G : \mathbb{R} \to \mathbb{R}$ mit den folgenden Eigenschaften:

- G ist monoton steigend
- G ist rechtsstetig
- $0 \geq \lim\limits_{x \to \infty} G(x) \leq \lim\limits_{x \to +\infty} G(x) \leq 1$

Satz 12.12: *(Helly): Die Menge J (w.o.) ist schwach folgenkompakt, d.h. zu jeder Folge (G_n) in J existiert ein $G \in J$ und eine Teilfolge (n_k) mit $G_{n_k}(x) \xrightarrow{k \to \infty} G(x)$ für alle Stetigkeitsstellen von G.*

Beweis: Sei $T = \{x_1, x_2, ...\} \subset \mathbb{R}$ abzählbar und dicht. Nach Konstruktion gilt dann: $G_{n_j^{(j)}}(x) \to G_T(x_m)(*)$ Wir zeigen:

(i) $G \in J$, wobei $G(x) = \inf\{G_T(y) | y > x, y \in T\}$, also G ist monoton steigend(klar), rechtsstetig (durch Widerspruch) und $0 \leq G(-\infty) \leq G(+\infty) \leq 1$(auch klar).

(ii) $G_{n_j^{(j)}}(x) \to G(x)$ für alle Stetigkeitsstellen x von G.

zu (ii): Sei z_0 eine Stetigkeitsstelle von G. Für $z_0 < y \in T$ gilt nach $(*)$:

$$\limsup_{j \to \infty} G_{n_j^{(j)}}(z_0) \leq \limsup_{j \to \infty} G_{n_j^{(j)}}(y) = G_T(y)$$

Also $\limsup\limits_{j \to \infty} G_{n_j^{(j)}}(z_0) \leq G(z_0)$. Sei andererseits $z_1 \in \mathbb{R}$ und $\bar{y} \in T$ mit $z_1 < \bar{y} < z_0$. Dann:

$$G(z_1) \leq G_T(\bar{y}) = \liminf_{j \to \infty} G_{n_j^{(j)}}(\bar{y}) \leq \liminf_{j \to \infty} G_{n_j^{(j)}}(z_0).$$

Da G stetig in z_0 ist, folgt mit $z_1 \to z_0$ so, dass

$$\liminf_{j \to \infty} G_{n_j^{(j)}}(z_0) \geq G(z_0).$$

Zusammen $G(z_0) = \lim\limits_{j \to \infty} G_{n_j^{(j)}}(z_0)$ ■

Beweis: (von Prohorov) "$\Rightarrow$": Sei $\mathcal{P}$ eine straffe Menge von W-Maßen auf $\mathscr{B}(\mathbb{R})$, $(P_n)_{n \in \mathbb{N}}$ eine Folge in $\mathcal{P}$ und $(F_n)_{n \in \mathbb{N}}$ die zugehörige Folge von Verteilungsfunktionen. Nach Helly existiert eine verallgemeinerte Verteilungsfkt G und eine Teilfolge (n_k) mit $F_{n_k}(x) \to G(x)$ an allen Stetigkeitsstellen x von G. Bleibt z.Z. dass G Verteilungsfunktion ist. Denn dann konvergiert (P_{n_k}) schwach gegen das durch G

bestimmte W-Maß nach dem Satz von Helly Bray. Dazu: Sei $\varepsilon > 0$. Aufgrund der Straffheit eine kompakte Menge $K \subset \mathbb{R}$ mit $\forall\, k\ P_{n_k}(K) > 1 - \varepsilon$.
Wähle $a < b$ Stetigkeitsstellen von G mit $K \subset [a, b]$. Dann:

$$1 - \varepsilon < P_{n_k}(K) \leq P_{n_k}((a, b]) = F_{n_k}(b) - F_{n_k}(a) \to G(b) - G(a).$$

Also $G(b) - G(a) \geq 1 - \varepsilon$ und somit $G(+\infty) - G(-\infty) \geq 1 - \varepsilon$ und also $G(+\infty) - G(-\infty) \geq 1$ mit $\varepsilon \to 0$. Da $G(+\infty) \leq 1$ und $G(-\infty) \geq 0$, folgt $G(+\infty) = 1$ und $G(-\infty) = 0$. Also ist G Verteilungsfunktion und somit $\mathcal{P}$ relativ kompakt.
"$\Leftarrow$": Lemma 12.10 ■

Satz 12.13: *Sei $(P_n)_{n\in\mathbb{N}}$ eine Folge in einer straffen Menge $\mathcal{P}$ von W-Maßen auf $\mathscr{B}(\mathbb{R})$. Falls jede schwach konvergente Teilfolge (P_{n_k}) gegen dasselbe W-Maß P konvergiert, so konvergiert (P_n) schwach gegen P.*

13 Charakteristische Funktionen

Definition 13.1: (i) Sei P ein W-Maß auf $(\mathbb{R}^D, \mathscr{B}(\mathbb{R}^D))$. Dann heißt

$$\varphi(t) = \int_{\mathbb{R}^D} e^{i\langle t,x\rangle} P(dx),\ t \in \mathbb{R}^D$$

Die charakteristische Funktion von P. Dabei ist $\langle \cdot, \cdot \rangle$ das Euklidische Skalarprodukt im $\mathbb{R}^D$, i ist die imaginäre Einheit und

$$\begin{aligned}\int_{\mathbb{R}^D} e^{i\langle t,x\rangle} P(dx) &:= \int_{\mathbb{R}^D} \operatorname{Re}(e^{i\langle t,x\rangle}) P(dx) + i \cdot \int_{\mathbb{R}^D} \operatorname{Im}(e^{i\langle t,x\rangle}) P(dx) \\ &= \int_{\mathbb{R}^D} \cos(\langle t,x\rangle) P(dx) + i \cdot \int_{\mathbb{R}^D} \sin(\langle t,x\rangle) P(dx)\end{aligned}$$

(ii) Für eine $\mathbb{R}^D$-wertige Zva X ist $\varphi_X(t) = E[e^{i\langle t,X\rangle}] = \int_{\mathbb{R}^D} e^{i\langle t,x\rangle} P_X(dx)$ die charakteristische Funktion von X.

Beispiel 13.2: (i) Sei X Bernoulli(λ)-verteilt. Dann $\varphi_X(t) = e^{it}p + (1-p)$.

(ii) Ist X standardnormalverteilt, so wissen wir aus den Übungen, dass

$$E[X^{2n}] = \frac{1}{2^n}\frac{(2n)!}{n!},\ E[X^{2n+1}] = 0\ (*)$$

Also (3. Gleichheit: Fubini)

$$\begin{aligned}\varphi_X(t) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{itx} e^{-\frac{x^2}{2}}\, dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \sum_{k=0}^{\infty} \frac{(itx)^k}{k!} e^{-\frac{x^2}{2}}\, dx \\ &= \sum_{k=0}^{\infty} (it)^k \cdot \frac{1}{k!} \underbrace{\int_{\mathbb{R}} x^k \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx}_{E[X^k]} = \sum_{m=0}^{\infty} \frac{(it)^{2m}}{(2m)!} \frac{1}{2^m} \frac{(2m)!}{m!} = e^{-\frac{t^2}{2}}\end{aligned}$$

Ist X $\mathcal{N}(\mu, \sigma)$-verteilt, also normalverteilt mit Erwartungswert μ und Varianz σ^2, so ist $Y = \frac{X-\mu}{\sigma}$ $\mathcal{N}(0,1)$-verteilt. Also:

$$\varphi_X(t) = E[e^{itx}] = E[e^{it(\mu+\sigma Y)}] = e^{it\mu}\varphi_Y(t\sigma) = e^{it\mu - \frac{1}{2}\sigma^2 t^2}$$

Satz 13.3: *Sei X eine $\mathbb{R}$-wertige Zva mit $E[|X|^n] < \infty$ für ein $n \in \mathbb{N}$. Dann existiert für alle $r \leq n$ die r-te Ableitung der char. Funktion φ_X von X; ist stetig und es gilt:*

$$\varphi_X^{(r)}(t) = \int (ix)^r e^{itx} P(dx).\ \textit{Inbes. } E[X^r] = \frac{\varphi^{(r)}(0)}{i^r}$$

Satz 13.4: *Sei X eine $\mathbb{R}$-wertige Zva mit $E[|X|^n] < \infty$ für ein $n \in \mathbb{N}$. Dann:*

$$\varphi_X(t) = \sum_{r=0}^{n} \frac{(it)^r}{r!} E[X^r] + \frac{t^n}{n!}\theta_n(t),$$

wobei θ_n stetig ist mit $\theta_n(0) = 0$.

Satz 13.5: *(Umkehrformel) Sei P ein W-Maß auf $\mathscr{B}(\mathbb{R})$ mit Verteilungsfkt F und char. Fkt. φ. Dann:*

(i) *Für alle Stetigkeitsstellen $a < b$ von F gilt:*

$$F(b) - F(a) = \lim_{c\to+\infty} \int\limits_{[-c,c]} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t)\lambda(dt)$$

(ii) *Falls $\int\limits_{-\infty}^{\infty} |\varphi(t)|\lambda(dt) < \infty$, so hat P eine Dichte f bzgl λ und*

$$f(x) = \frac{1}{2\pi} \int\limits_{\mathbb{R}} e^{-itx} \varphi(x)\lambda(dx)$$

Satz 13.6: *(Eindeutigkeitssatz) Seien P, Q W-Maße auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ die dieselbe char. Fkt. φ haben. Dann $P(A) = Q(A)$ für alle $A \in \mathscr{B}(\mathbb{R})$.*

Beweis: Folgt aus dem Maßeindeutigkeitssatz. ■

Satz 13.7: *Ein Vektor $(X_1, ..., X_D)$ von $\mathbb{R}$-wertigen Zva ist unabhängig genau dann, wenn*

$$\varphi_X(t_1, ..., t_D) = \prod_{d=1}^{D} \varphi_{X_d}(t_d) \text{ für alle } t = (t_1, ..., t_D) \in \mathbb{R}^D.$$

Satz 13.8: *Seien X, Y unabhängige reellwertige Zva mit char. Fkt. φ_X, φ_Y. Dann ist die char. Fkt φ_Z der Faltung $Z = X + Y$ gegeben durch:*

$$\varphi_Z(t) = \varphi_X(t)\varphi_Y(t) \,\forall\, t \in \mathbb{R}.$$

Satz 13.9: *(Stetigkeitssatz; Levy) Sei $(P_n)_{n\in\mathbb{N}}$ eine Folge von W-Maßen auf $\mathscr{B}(\mathbb{R})$ und $(\varphi_n)_{n\in\mathbb{N}}$ die zugehörige Folge von char. Funktionen. Dann:*

(i) *Konvergiert $(P_n)_{n\in\mathbb{N}}$ schwach gegen ein W-Maß P auf $\mathscr{B}(\mathbb{R})$, so gilt: $\lim\limits_{n\to\infty} \varphi_n(t) = \varphi(t)$ für alle $t \in \mathbb{N}$, wobei φ char. Funktion von P ist.*

(ii) *Falls für $t \in \mathbb{R}$ gilt: $\varphi_n(t) \to \varphi(t)$ für eine in $t = 0$ stetige Funktionen φ, so ist φ chark. Fkt. eines W-Maßes P auf $\mathscr{B}(\mathbb{R})$ mit $P_n \xrightarrow{w} P$.*

Beweis: (i) klar, da $x \mapsto \cos(tx), x \mapsto \sin(tx)$ sind stetig und beschränkt.

(ii) Schritt 1: $\{P_n | n \in \mathbb{N}\}$ ist straff.
Schritt 2: Nach dem Satz von Prohorov ist $\{P_n | n \in \mathbb{N}\}$ relativ kompakt, es ex. also eine schwach konvergente Teilfolge $(P_{n_k})_{k\in\mathbb{N}}$ gegen ein W-Maß P. Nach (i) entlang der Teilfolge ist φ die charakteristische Funktion von P. Sei $\bar{P}$ schwacher Grenzwert einer weiteren konvergenten Teilfolge. Dann ist φ auch char. Funktion von $\bar{P}$. Aus dem Eindeutigkeitssatz folgt $P = \bar{P}$. Aus Satz 12.13 folgt $P_n \xrightarrow{w} P$. ■

14 Der zentrale Grenzwertsatz von Lindeberg-Levy

Sei $(X_n)_{n\in\mathbb{N}}$ Folge von unabh. identisch verteilten Zva in $\mathcal{L}^2(\Omega, \mathcal{A}, P)$ mit $\mu := E[X_1]$ und $\sigma^2 := \text{Var}X_1 > 0$. Dann gilt für $0 \le \alpha < \frac{1}{2}$:

$$\sum_{n=1}^{\infty} \frac{\text{Var}(X_n)}{n^{2-2\alpha}} = \sigma^2 \sum_{n=1}^{\infty} \frac{1}{n^{2-2\alpha}} < \infty.$$

Nach Satz 10.12 konvergiert $\sum\limits_{n=1}^{\infty} \frac{X_n - E[X_n]}{n^{1-\alpha}}$ P-f.s., sodass nach dem Lemma von Kronecker (Lemma 11.3) gilt

$$\frac{1}{n^{1-\alpha}} \sum_{k=1}^{n} (X_k - E[X_k]) \to 0 \ P\text{-f.s.}$$

Satz 14.1: *(Zentraler Grenzwertsatz von Lindeberg-Levy) Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von unabh. identisch verteilten Zva in $\mathcal{L}^2(P)$ mit $\mu := E[X_1]$, $\sigma^2 := Var(X_1) > 0$. Sei*

$$\Phi(y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

die Verteilungsfunktion der Standardnormalverteilung. Dann

$$\sup_{y\in\mathbb{R}} |P(\{\frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{n} (X_k - E[X_k]) \le y\}) - \Phi(y)| \to 0$$

Beweis: Sei $S_n = \frac{1}{\sigma\sqrt{n}} \sum\limits_{k=1}^{n} (X_k - E[X_k])$. Dann:

$$\begin{aligned}\varphi_{S_n}(t) &= E[\exp\{i\frac{t}{\sqrt{n}} \sum_{k=1}^{n} \frac{X_k - E[X_k]}{\sigma}\}] = \prod_{k=1}^{n} E[\exp\{i\frac{t}{\sqrt{n}}(\frac{X_k - E[X_k]}{\sigma})\}] \\ &= (\varphi_{\frac{X_1 - E[X_1]}{\sigma}}(\frac{t}{\sqrt{n}}))^n\end{aligned}$$

wobei die letzte Gleichheit gilt, da identisch verteilt. Sei $X = \frac{X_1 - E[X_1]}{\sigma}$. Folgerung 13.5 mit $n = 2$ liefert:

$$\varphi_X(\frac{t}{\sqrt{n}}) = 1 + \frac{1}{2}\frac{t^2}{n} + \frac{1}{2}\frac{t^2}{n}\nu(\frac{t}{\sqrt{n}})$$

Für eine stetige Funktion ν mit $\nu(0) = 0$. Also:

$$\varphi_{S_n}(t) = (1 + \frac{1}{2}\frac{t^2}{n} - \frac{1}{2}\frac{t^2}{n}\nu(\frac{t}{\sqrt{n}}))^n \xrightarrow{k\to\infty} e^{-\frac{1}{2}t^2}$$

Aufgrund des Stetigkeitssatzes 13.12 konvergiert S_n in Verteilung gegen eine standardnormalverteilte Zva. Der Satz von Helly-Bray (Satz 12.4) zieht die punktweise Konvergenz der Verteilungsfunktion nach sich. Gleichmäßige Konvergenz: Übung. ■

Beispiel 14.2: 100 unabh. Würfe einer fairen Münze. Wir hoch ist die Wahrscheinlichkeit, dass mind. 60 Mal Kopf fällt?

Beispiel 14.3: Monte-Carlo-Integration, Fortsetzung.

15 Multivariate Normalverteilung

Definition 15.1: Sei $\mu \in \mathbb{R}^D$ und Σ eine symmetrische, positiv semidefinite $D \times D$-Matrix. Das W-Maß auf $(\mathbb{R}^D, \mathscr{B}(\mathbb{R}^D))$ mit charakteristischer Funktion

$$\varphi(t) = e^{i\langle t,y\rangle - \frac{1}{2}t^T\Sigma t}, t \in \mathbb{R}^D$$

heißt multivariate Normalverteilung mit Paramtern (μ, Σ) kurz: $\mathcal{N}(\mu, \Sigma)$ -Verteilung. Eine $\mathbb{R}^D$-wertige Zva X heißt multivariate Normalverteiltung falls ihr Verteilung P_X eine multivariate Normalverteilung ist.

Satz 15.2: *Sei Σ eine symm. pos.semidefinite $D \times D$-Matrix und $\mu \in \mathbb{R}^D$. Dann ex. ein W-Maß auf $(\mathbb{R}^D, \mathscr{B}(\mathbb{R}^D))$, sodass die charakteristische Funktion von P gegeben ist durch*

$$\varphi(t) = e^{i\langle t,\mu\rangle - \frac{1}{2}t^T\Sigma t},$$

d.h. die $\mathcal{N}(\mu, \Sigma)$-Verteilung existiert.

Satz 15.3: *Eine $\mathbb{R}^D$-wertige Zufallsvariable X ist genau dann multivariat normalverteilt, wenn für alle $a \in \mathbb{R}^D$ $\langle a, X\rangle$ (1-dim.) normalverteilt.*

Satz 15.4: *Sei $X = (X_1, ..., X_D)^T$ multivariat normalverteilt. Dann sind äquivalent:*

(i) *$X_1, ..., X_D$ sind paarweise disjunkt.*

(ii) *$X_1, ..., X_D$ sind unabhängig.*

Nächstes Ziel: Multivariate Version des zentralen Grenzwertsatzes

Satz 15.5: *(Cramér-Wold Device) Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von $\mathbb{R}^D$-wertigen Zva und X eine $\mathbb{R}^D$-wertige Zva. Dann sind äquivalent:*

(i) *$X_n \to X$ in Verteilung*

(ii) *$\forall\, a \in \mathbb{R}^D : \langle a, X_n\rangle \to \langle a, X\rangle$ in Verteilung.*

Beweis: (i) $\Rightarrow$ (ii): Sei $f : \mathbb{R} \to \mathbb{R}$ stetig und beschränkt. Sei $a \in \mathbb{R}^D$. Dann ist $f_a(x) := f(\langle a, X_n\rangle)$ stetig und beschränkt. Also

$$E[f(\langle a, X_n\rangle)] = \int \underbrace{f(\langle a, x\rangle)}_{f_a(x)} P_{X_n}(dx) \to \int f(\langle a, x\rangle) P_X(dx) = E[f(\langle a, X\rangle)]$$

Also: $\langle a, X_n\rangle \xrightarrow{d} \langle a, X\rangle$.
(ii) $\Rightarrow$ (i): Da $x \mapsto \mathrm{Re}(e^{ix})$ und $x \mapsto \mathrm{Im}(e^{ix})$ stetig und beschränkt sind folgt aus (ii) für alle $t \in \mathbb{R}^D$:

$$\varphi_{X_n}(t) = E[e^{i\langle t,X_n\rangle}] \to E[e^{i\langle t,X\rangle}] = \varphi_X(t)$$

Die multivariate Version des Stetigkeitssatzes impliziert, dass $X_n \xrightarrow{d} X$. ■

Satz 15.6: *(Multivariater zentraler Grenzwertsatz)*
Sei $(X_n)_{n\in\mathbb{N}} = ((X_{n,1}, ..., X_{n,D}))_{n\in\mathbb{N}}$ *eine Folge von* $\mathbb{R}^D$*-wertigen Zvaauf* $(\Omega, \mathcal{A}, P)$*, die unabh. und identisch verteilt sind. Es gelte* $X_{1,d} \in \mathcal{L}^2(P)$ *für alle* $d = 1, ..., D$*. Sei*

$$\mu = (E[X_{1,1}, ..., E[X_{1,D}]])^T \text{ und } \Sigma = (Cov(X_{1,i}, X_{1,j}))_{i,j=1,...,D}.$$

Dann gilt:

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k - \mu \xrightarrow{d} Z,$$

wobei Z $\mathcal{N}(0, \Sigma)$*-verteilt ist.*

Beweis:

- $\underline{D = 1,\, \Sigma > 0}$: $(\frac{1}{\sqrt{n\Sigma}} \sum_{k=1}^{n} X_k - \mu)_{n\in\mathbb{N}}$ konvergiert in Verteilung gegen eine $\mathcal{N}(0,1)$-verteilte Zva nach dem zentralen Frenzwertsatz von Lindeberg-Lévy. Multiplikation mit $\sqrt{\Sigma}$ liefert die Behauptung.
- $\underline{D = 1,\, \Sigma = 0}$: Dann ist $X_k = \mu$ P-f.s. und also $0 = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k - \mu \xrightarrow{d} 0$ und 0 ist $\mathcal{N}(0,0)$-verteilt.
- $\underline{D \geq 2}$: Ohne Einschränkung $\mu = 0$. Sei Z $\mathcal{N}(0, \Sigma)$-verteilt. Sei $a \in \mathbb{R}^D$. Setze $X_k^{(a)} = \langle a, X_k \rangle, S_n^{(a)} = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k^{(a)}$. Dann:

  $$E[X_k^{(a)}] = 0, \mathrm{Var}(X_k^{(a)}) = a^T \Sigma a.$$

  Daher liefert der Fall $D = 1$:

  $$S_n^{(a)} \xrightarrow{d} \langle a, Z \rangle,$$

  da $\langle a, Z \rangle$ $\mathcal{N}(0, a^T \Sigma a)$-verteilt ist. Da $S_n^{(a)} = \langle a, \frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k \rangle$, folgt die Aussage mit dem Cramér-Word Device. ∎

16 Der Kolmogorovsche Erweiterungssatz

Sei $\mathbb{R}^{\mathbb{N}} := \{(x_n)_{n\in\mathbb{N}} | x_n \in \mathbb{R} \,\forall n \in \mathbb{N}\}$ die Menge aller reellwertigen Folgen. Ist $X = (X_n)_{n\in\mathbb{N}}$ eine Folge von reellwertigen Zva auf $(\Omega, \mathcal{A}, P)$, so kann man X als Abbildung von $\Omega \to \mathbb{R}^{\mathbb{N}}$ auffassen. Diese Abbildung wird messbar, wenn wir sie mit einer geeigneten σ-Algebra ausstatten.

Definition 16.1: Die Borelsche σ-Algebra auf $\mathbb{R}^{\mathbb{N}}$ ist die von Mengen der Form $\{x \in \mathbb{R}^{\mathbb{N}} | (x_1, ..., x_n) \in B\}, n \in \mathbb{N}, B \in \mathscr{B}(\mathbb{R}^n)$ erzeugte σ-Algebra. Wir notieren sie mit $\mathscr{B}(\mathbb{R}^N)$.

Lemma 16.2: *Sei $X : \Omega \to \mathbb{R}^{\mathbb{N}}$. Dann sind äquivalent:*

(i) *X ist $\mathcal{A}/\mathscr{B}(\mathbb{R}^N)$-messbar.*

(ii) *Für alle $n \in \mathbb{N}$ ist $(X_1, ..., X_n)^T : \Omega \to \mathbb{R}^N$ $\mathcal{A}/\mathscr{B}(\mathbb{R}^n)$-messbar.*

(iii) *Für alle $n \in \mathbb{N}$ ist $X_n : \Omega \to \mathbb{R}$ $\mathcal{A}/\mathscr{B}(\mathbb{R})$-messbar.*

Satz 16.3: *(Kolmogorov) Seien $P_n, n \in \mathbb{N}$ W-Maße auf $\mathscr{B}(\mathbb{R}^n)$, die die Konsistenzbedingung (K) erfüllen. Dann existiert genau ein W-Maß P auf $\mathscr{B}(\mathbb{R}^{\mathbb{N}})$ mit*

$$P(\{x \in \mathbb{R}^{\mathbb{N}} | (x_1, ..., x_n) \in B\}) = P_n(B) \text{ für alle } n \in \mathbb{N}, B \in \mathscr{B}(\mathbb{R}^n)$$

Im Beweis benötigen wir:

Lemma 16.4: *Seien P W-Maß auf $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$. Dann existiert für alle $B \in \mathscr{B}(\mathbb{R}^n)$ und $\varepsilon > 0$ eine kompakte Menge $K \subset B$ mit $P(B \backslash K) \leq \varepsilon$.*

Beweis: (von Satz 16.3)

(i) Für $n \in \mathbb{N}$ und $B \in \mathscr{B}(\mathbb{R}^n)$ sei $T_n(B) := \{x \in \mathbb{R}^{\mathbb{N}} | (x_1, ..., x_n) \in B\}$. Sei $K^{\mathbb{N}} := \{T_n(B) | n \in \mathbb{N}, B \in \mathscr{B}(\mathbb{R}^n)\}$. Dann prüfen: $K^{\mathbb{N}}$ ist ein Ring.

(ii) Sei $\bar{P} : K^{\mathbb{N}} \to [0, 1], T_n(B_n) \mapsto P_n(B_n)$. Dann ist $\bar{P}$ wohldefiniert und ein Inhalt (nachrechnen).

(iii) $\bar{P}$ ist ein Prämaß. Nach Satz 2.10 müssen wir zeigen, dass $\bar{P}$ stetig von oben in $\emptyset$ ist. Durch Widerspruch.

(iv) Nach dem Satz von Caratheodory und dem Maßeindeutigkeitssatz ex. genau eine Fortsetzung P von $\bar{P}$ zu einem Maß auf $(\mathbb{R}^{\mathbb{N}}, \mathscr{B}(\mathbb{R}^{\mathbb{N}}))$, sodass

$$P(T_n(B_n)) = \bar{P}(T_n(B_n)) - P_n(B_n).$$

Dies ist ein W-Maß, da $P(\mathbb{R}^{\mathbb{N}}) = P(T_1(\mathbb{R})) = P_1(\mathbb{R}) = 1$. ■

Korollar 16.5: *Sei* $(P_n)_{n\in\mathbb{N}}$ *eine Folge von W-Maßen auf* $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))_{n\in\mathbb{N}}$, *die die Konsistenzbedingung (K) erfüllt. Dann ex. ein W-Raum* $(\Omega, \mathcal{A}, P)$ *und eine Folge* $(X_n)_{n\in\mathbb{N}}$ *von reeelwertigen Zva auf* $(\Omega, \mathcal{A}, P)$, *die* $(P_n)_{n\in\mathbb{N}}$ *als endliche Randverteilung hat.*