

# **Stochastik II**

Prof. Dr. Bäuerle

Im Wintersemester 06/07

Das Team von <http://mitschriebwiki.nomeata.de/>

Dieses Dokument ist eine persönliche Vorlesungsmitschrift der  
Vorlesung Stochastik II im Wintersemester 2006/07 bei Prof. Dr. Bäuerle.

Das latexki-Team gibt keine Garantie für die  
Richtigkeit oder Vollständigkeit des Inhaltes und übernimmt keine  
Verantwortung für etwaige Fehler.  
Auch ist Frau Bäuerle nicht verantwortlich für den Inhalt dieses Skriptes.

# Inhaltsverzeichnis

<b>1</b>	<b>Maß-Integral und Erwartungswert</b>	<b>5</b>
<b>2</b>	<b>Eigenschaften des Maß-Integrals</b>	<b>15</b>
2.1	Konvergenzsätze . . . . .	15
2.2	Verhalten bei Transformationen . . . . .	16
2.3	Nullmengen und Maße mit Dichten . . . . .	18
2.4	Ungleichungen und Räume integrierbarer Funktionen . . . . .	21
<b>3</b>	<b>Produktmaße und Unabhängigkeit</b>	<b>25</b>
3.1	Der allgemeine Fall . . . . .	25
3.2	Reellwertige Abbildungen, Rechnen mit Verteilungen . . . . .	32
<b>4</b>	<b>Das starke Gesetz der großen Zahlen</b>	<b>37</b>
<b>5</b>	<b>Zentraler Grenzwertsatz von Lindeberg-Lévy</b>	<b>41</b>
5.1	Charakteristische Funktionen . . . . .	41
5.2	Umkehrsätze . . . . .	42
5.3	Verteilungskonvergenz . . . . .	44
<b>6</b>	<b>Zentraler Grenzwertsatz in <math>\mathbb{R}^n</math></b>	<b>57</b>
6.1	Mehrdimensionale Normalverteilung . . . . .	58
6.2	Zentraler Grenzwertsatz in $\mathbb{R}^d$ . . . . .	59
<b>7</b>	<b>Bedingte Erwartungswerte und Bedingte Verteilungen</b>	<b>61</b>
<b>8</b>	<b>Martingale und Stoppzeiten</b>	<b>71</b>
<b>9</b>	<b>Konvergenzsätze für Martingale</b>	<b>85</b>



# 1 Maß-Integral und Erwartungswert

Stochastik I: Ein Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$  bestehend aus:

- (i)  $\Omega \neq \emptyset$  bel. Menge, der Ergebnisraum
- (ii)  $\mathcal{A} \subset \mathcal{P}(\Omega)$  eine  $\sigma$ -Algebra, d.h.
  - $\Omega \in \mathcal{A}$
  - $A \in \mathcal{A} \implies A^c \in \mathcal{A}$
  - $A_1, A_2, \dots \in \mathcal{A} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$
- (iii)  $P : \mathcal{A} \rightarrow [0, 1]$  ein Wahrscheinlichkeitsmaß, d.h.
  - $P(\Omega) = 1$
  - $A_1, A_2, \dots \in \mathcal{A}$ , paarweise disjunkt  $\implies P(\sum_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$   
( $\sigma$ -Additivität)

Statt das Wahrscheinlichkeitsmaßes  $P$  betrachten wir jetzt eine allgemeine Funktion  $\mu : \mathcal{A} \rightarrow \mathbb{R}_+ \cup \{\infty\}$ , die beliebige positive Werte annehmen kann.

## Definition

Sei  $(\Omega, \mathcal{A})$  ein messbarer Raum. Eine Abbildung  $\mu : \mathcal{A} \rightarrow \mathbb{R}_+ \cup \{\infty\}$  heißt **Maß** auf  $(\Omega, \mathcal{A})$ , wenn  $\mu(\emptyset) = 0$  und  $\mu(\sum_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$  für alle paarweise disjunkten Ereignisse  $A_1, A_2, \dots$ ,  $(\Omega, \mathcal{A}, \mu)$  heißt **Maßraum**.

## Bemerkung

Da  $\mu(A) = \infty$  möglich, definieren wir:  $a + \infty = \infty \forall a \in \mathbb{R} \cup \{\infty\}$ .

## Definition

Sei  $\mu$  ein Maß auf  $(\Omega, \mathcal{A})$ .

1.  $\mu$  heißt **endlich**, falls  $\mu(\Omega) < \infty$ ,
2.  $\mu$  heißt  **$\sigma$ -endlich**, falls  $\exists$  eine Folge  $(A_i), i \in \mathbb{N}, A_i \in \mathcal{A}$  mit  $\bigcup_{i=1}^{\infty} A_i = \Omega$  und  $\mu(A_i) < \infty \forall i \in \mathbb{N}$ .

## Beispiel 1.1

- a) Sei  $(\Omega, \mathcal{A})$  ein messbarer Raum,  $\omega \in \Omega$  fest.

$$\delta_{\omega}(A) := \begin{cases} 1, & \omega \in A \\ 0, & \text{sonst} \end{cases}$$

für  $A \in \mathcal{A}$  definiert ein Maß.

$\delta_{\omega}$  heißt **Einpunktmaß** oder **Dirac-Maß** im Punkt  $\omega$ . Da  $\delta_{\omega}(\Omega) = 1$  ist  $\delta_{\omega}$  sogar ein Wahrscheinlichkeitsmaß.

- b)  $\mu := \sum_{\omega \in \Omega} \delta_\omega$  ist das **abzählende Maß** auf  $\Omega$ .  
 (Falls  $|A| < \infty$  :  $\mu(A) = |A|$  Anzahl der Elemente in  $A$ .)  
 $\mu$  ist endlich  $\Leftrightarrow \Omega$  ist endlich,  
 $\mu$  ist  $\sigma$ -endlich  $\Leftrightarrow \Omega$  ist abzählbar.
- c) Sei  $\Omega = \mathbb{R}$ ,  $\mathcal{A} = \mathfrak{B}(\mathbb{R})$  Borelsche  $\sigma$ -Algebra.

$$\mathfrak{B}(\mathbb{R}) = \sigma(\underbrace{\{(a, b], -\infty < a < b < \infty\}}_{=: \varepsilon \text{ Erzeuger}}) = \sigma(\varepsilon) := \bigcap_{\mathcal{A} \text{ } \sigma\text{-Algebra, } \varepsilon \subset \mathcal{A}} \mathcal{A}$$

Sei  $a, b \in \mathbb{R}$  mit  $a < b$ . Durch  $\lambda((a, b]) := b - a$  wird auf  $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$  ein Maß definiert, das sogenannte **Lebesgue-Maß**. Die Eindeutigkeit von  $\lambda$  folgt aus dem **Eindeutigkeitssatz für Maße**:

Sei  $\mathcal{A} = \sigma(\varepsilon)$  und  $\varepsilon$  durchschnittsstabil (d.h.:  $A, B \in \varepsilon \implies A \cap B \in \varepsilon$ ). Weiter seien  $\mu_1, \mu_2$  Maße auf  $\mathcal{A}$  mit  $\mu_1(A) = \mu_2(A) \forall A \in \varepsilon$ .  $\exists$  eine Folge  $(A_n)_{n \in \mathbb{N}} \subset \varepsilon$  mit  $A_n \uparrow \Omega$  und  $\mu_1(A_n) = \mu_2(A_n) < \infty \forall n$ , so gilt  $\mu_1 = \mu_2$ .

Eine nichttriviale Aufgabe ist es hier zu zeigen, dass  $\lambda$  auf ganz  $\mathfrak{B}(\mathbb{R})$  zu einem Maß fortgesetzt werden kann. (gezeigt von Carathéodory; s. z.B. Henze, Bauer)

Bei  $\Omega = \bar{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$ , ist  $\mathfrak{B}(\bar{\mathbb{R}}) := \{B \subset \bar{\mathbb{R}} \mid B \cap \mathbb{R} \in \mathfrak{B}(\mathbb{R})\} = \{B, B \cup \{\infty\}, B \cup \{-\infty\}, B \cup \{\infty, -\infty\} \mid B \in \mathfrak{B}(\mathbb{R})\}$  eine  $\sigma$ -Algebra (analog  $\mathfrak{B}((-\infty, \infty))$ ) und  $\bar{\lambda}(B) = \lambda(B) \forall B \in \mathfrak{B}(\mathbb{R})$  und  $\bar{\lambda}(\{\infty\}) = \bar{\lambda}(\{-\infty\}) = 0$

$\lambda$  ist nicht endlich, da  $\lambda((-\infty, a]) = \sum_{n=1}^{\infty} \underbrace{\lambda((a-n, a-n+1])}_{=1} = \infty$ , aber

$\sigma$ -endlich, da  $\bigcup_{n=1}^{\infty} (-n, n] = \mathbb{R}$ ,  $\lambda((-n, n]) < \infty \forall n \in \mathbb{N}$ .

- d) Seien  $\mu_n$  Maße,  $n \in \mathbb{N}$ , so ist

$$\mu := \sum_{n=1}^{\infty} b_n \mu_n$$

wieder ein Maß.

**Konvention:**  $a \cdot \infty = \infty \cdot a = \infty, a > 0, 0 \cdot \infty = 0$

Spezialfall:  $\mu_n = \delta_{\omega_n} (\omega_n \in \Omega), b \geq 0, \sum_{n=1}^{\infty} b_n = 1$

$$\mu = \sum_{n=1}^{\infty} b_n \delta_{\omega_n}$$

ist dann ein diskretes, auf  $\{\omega_1, \omega_2, \dots\}$  konzentriertes Wahrscheinlichkeitsmaß.

- e) Sei  $G : \mathbb{R} \rightarrow \mathbb{R}$  wachsend und rechtsseitig stetig (Eine Funktion mit diesen Eigenschaften heißt **maßdefinierende Funktion**. Gilt zusätzlich  $\lim_{x \rightarrow \infty} G(x) = 1, \lim_{x \rightarrow -\infty} G(x) = 0$ , dann ist  $G$  eine Verteilungsfunktion.)

$$\mu_G((a, b]) := G(b) - G(a)$$

für  $a, b \in \mathbb{R}, a < b$  definiert  $\mu_G$  ein Maß auf  $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ , das sogenannte **Lebesgue-Stieltjes-Maß** zu  $G$ . (Fortsetzungsproblem analog zu c) )

Ist  $G$  eine Verteilungsfunktion mit  $G(x) = \int_{-\infty}^x f(y)dy$  mit

$$f \geq 0 : \int_{-\infty}^{\infty} f(y)dy = 1,$$

so ist  $\mu_G((a, b]) = \int_a^b f(y)dy$  ein Wahrscheinlichkeitsmaß mit Dichte  $f$ .

### Bemerkung

Viele der in Stochastik I für Wahrscheinlichkeitsmaße besprochene Eigenschaften gelten auch für allgemeine Maße  $\mu$ , z.B.  $\mu$  ist stetig von unten, d.h.

$$\underbrace{A_n \uparrow}_{A_n \subset A_{n+1}} \text{ mit } \bigcup_{i=1}^{\infty} A_i = A \implies \mu(A) = \lim_{n \rightarrow \infty} \mu(A_n)$$

Bei der Stetigkeit von oben brauchen wir eine Zusatzbedingung:

$$\underbrace{A_n \downarrow}_{A_n \supset A_{n+1}} \text{ mit } \bigcap_{k=1}^{\infty} A_k = A, \underline{\mu(A_n) < \infty} \implies \mu(A) = \lim_{n \rightarrow \infty} \mu(A_n)$$

### Beispiel

Lebesgue-Maß:  $A_n = (-\infty, -n] \downarrow, \emptyset = \bigcap_{n=1}^{\infty} (-\infty, -n], \lim_{n \rightarrow \infty} \lambda((-\infty, -n]) = \infty \neq 0 = \lambda(\emptyset)$

### Definition

Seien  $(\Omega, \mathcal{A})$  und  $(\Omega', \mathcal{A}')$  zwei meßbare Räume. Eine Abbildung  $f : \Omega \rightarrow \Omega'$  heißt  $(\mathcal{A}, \mathcal{A}')$ -**messbar**, falls

$$f^{-1}(A') \in \mathcal{A}, \quad \forall A' \in \mathcal{A}'$$

$f$  mit dieser Eigenschaft heißt **Zufallsgröße**. Ist  $\Omega' = \mathbb{R}$ , dann **Zufallsvariable**.

Im Folgenden sei  $(\Omega, \mathcal{A}, \mu)$  ein Maßraum. Ziel ist es, möglichst vielen Funktionen  $f : \Omega \rightarrow \mathbb{R}$  ein Integral bezüglich  $\mu$  zuzuordnen. Die Konstruktion erfolgt in drei Schritten:

- 1.) Sei  $\mathcal{E} := \{f : \Omega \rightarrow \mathbb{R} | f \geq 0, f \text{ ist } \mathcal{A}\text{-messbar}, f(\Omega) \text{ endlich}\}$  die Menge der **Elementarfunktionen** auf  $\Omega$ .

Ist  $f(\Omega) = \{\alpha_1, \dots, \alpha_n\}, \alpha_j \geq 0$ , so gilt:

$$f = \sum_{j=1}^n \alpha_j \mathbf{1}_{A_j}$$

mit  $A_j := f^{-1}(\{\alpha_j\})$  und  $\Omega = \sum_{j=1}^n A_j$ . Eine Darstellung von  $f$  mit dieser Eigenschaft heißt „Normaldarstellung“ von  $f$ .

Normaldarstellung ist nicht eindeutig.

### Definition

Ist  $f$  eine Elementarfunktion mit Normaldarstellung  $f = \sum_{j=1}^n \alpha_j \mathbf{1}_{A_j}$ , so heißt  $\int f d\mu := \sum_{j=1}^n \alpha_j \mu(A_j)$  das  **$\mu$ -Integral** von  $f$ . Schreibweise  $\int f d\mu = \mu(f)$ .

**Lemma 1.1 (Unabhängigkeit des Integrals von der Normaldarstellung)**

Für zwei Normaldarstellungen

$$f = \sum_{j=1}^n \alpha_j \mathbf{1}_{A_j} = \sum_{i=1}^m \beta_i \mathbf{1}_{B_i}$$

einer Funktion  $f \in \mathcal{E}$  gilt:

$$\sum_{j=1}^n \alpha_j \mu(A_j) = \sum_{i=1}^m \beta_i \mu(B_i)$$

**Beweis**

$$\text{Voraussetzung} \implies \Omega = \sum_{j=1}^n A_j = \sum_{i=1}^m B_i$$

$$\begin{aligned} \implies \mu(A_j) &\stackrel{\sigma\text{-Add.}}{=} \sum_{i=1}^m \mu(A_j \cap B_i) \\ \mu(B_i) &= \sum_{j=1}^n \mu(A_j \cap B_i) \end{aligned}$$

$$\mu(A_j \cap B_i) \neq 0 \implies A_j \cap B_i \neq \emptyset \implies \alpha_j = \beta_i$$

Insgesamt:

$$\begin{aligned} \sum_{j=1}^n \alpha_j \mu(A_j) &= \sum_{j=1}^n \sum_{i=1}^m \underbrace{\alpha_j}_{\beta_i} \mu(A_j \cap B_i) \\ &= \sum_{i=1}^m \beta_i \mu(B_i) \end{aligned} \quad \blacksquare$$

**Lemma 1.2 (Eigenschaften des  $\mu$ -Integrals)**

- a)  $\int \mathbf{1}_A d\mu = \mu(A)$  für  $A \in \mathcal{A}$
- b)  $\int (\alpha f) d\mu = \alpha \int f d\mu$  für  $f \in \mathcal{E}, \alpha \geq 0$
- c)  $\int (f + g) d\mu = \int f d\mu + \int g d\mu$  für  $f, g \in \mathcal{E}$
- d)  $f \leq g \implies \int f d\mu \leq \int g d\mu$  für  $f, g \in \mathcal{E}$

**Beweis**

a), b) klar



c) Sei  $f = \sum_{j=1}^n \alpha_j \mathbf{1}_{A_j}$ ,  $g = \sum_{i=1}^m \beta_i \mathbf{1}_{B_i}$

$$\begin{aligned}
 \Rightarrow f &= \sum_{j=1}^n \sum_{i=1}^m \alpha_j \mathbf{1}_{A_j \cap B_i} \\
 g &= \sum_{i=1}^m \sum_{j=1}^n \beta_i \mathbf{1}_{B_i \cap A_j} \\
 \text{also } f + g &= \sum_{j=1}^n \sum_{i=1}^m (\alpha_j + \beta_i) \mathbf{1}_{A_j \cap B_i} \\
 \Rightarrow \mu(f + g) &= \sum_{j=1}^n \sum_{i=1}^m (\alpha_j + \beta_i) \mu(A_j \cap B_i) \\
 &= \sum_{j=1}^n \alpha_j \sum_{i=1}^m \mu(A_j \cap B_i) + \sum_{i=1}^m \beta_i \underbrace{\sum_{j=1}^n \mu(A_j \cap B_i)}_{=\mu(B_i)} \\
 &= \mu(f) + \mu(g)
 \end{aligned}$$

d) folgt mit gleicher Darstellung wie in c) ■

### Bemerkung

- a) Ist  $f = \sum_{j=1}^n \alpha_j \mathbf{1}_{A_j} \in \mathcal{E}$ , aber nicht notwendig eine Normaldarstellung, so folgt aus Lemma 1.2 c)  $\int f d\mu = \sum_{j=1}^n \alpha_j \mu(A_j)$
- b) Ist  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum und  $X : \Omega \rightarrow \mathbb{R}_+$  eine Zufallsvariable mit endlich vielen Werten  $\{x_1, \dots, x_n\}$ , so gilt:

$$\begin{aligned}
 \int X dP &= \sum_{j=1}^n x_j P(X^{-1}(\{x_j\})) \\
 &= \sum_{j=1}^n x_j P^X(\{x_j\})
 \end{aligned}$$

$$\begin{aligned}
 (A_j &= X^{-1}(\{x_j\})) \\
 \text{Also: } \int X dP &= EX
 \end{aligned}$$

- 2.) Sei  $\mathcal{E}^+ := \{f : \Omega \rightarrow \bar{\mathbb{R}} \mid f \geq 0, f \text{ ist } \mathcal{A}\text{-messbar}\}$ . Wichtig: Elemente von  $\mathcal{E}^+$  kann man beliebig gut durch Elemente aus  $\mathcal{E}$  approximieren.

### Satz 1.1

Zu jedem  $f \in \mathcal{E}^+$  gibt es eine wachsende Folge  $(u_n)_{n \in \mathbb{N}}$  aus  $\mathcal{E}$  mit  $u_n \uparrow f$ , d.h.  $u_n \leq u_{n+1}$  und  $\lim_{n \rightarrow \infty} u_n = f$  (jeweils punktweise).

**Beweis**

Sei  $\alpha_n : \mathbb{R} \rightarrow [0, \infty]$  gegeben durch:

$$\alpha_n(x) := \begin{cases} 0, & \text{falls } x < 0 \\ \frac{j}{2^n}, & \text{falls } \frac{j}{2^n} \leq x < \frac{j+1}{2^n}, j = 0, 1, \dots, n2^n - 1 \\ n, & \text{falls } x \geq n \end{cases}$$

(Hier fehlt ein Bild)

$\alpha_n$  ist  $\mathfrak{B}$ -messbar.  $\alpha_n \uparrow$  und  $\lim_{n \rightarrow \infty} \alpha_n(x) = x$  für  $n \rightarrow \infty$ . Sei  $u_n := \alpha_n \circ f$ . Dann gilt  $u_n \in \mathcal{E}$  und  $u_n \uparrow f$ . ■

**Bemerkung**

Ist  $f$  beschränkt, so konvergiert die Folge  $(u_n)$  gleichmäßig gegen  $f$ , d.h.  $\lim_{n \rightarrow \infty} \sup_{\omega \in \Omega} |f(\omega) - u_n(\omega)| = 0$ .

**Definition**

Sei  $f \in \mathcal{E}^+$  und  $(u_n)$  eine wachsende Folge aus  $\mathcal{E}$  mit  $\lim_{n \rightarrow \infty} u_n = f$ . Dann heißt

$$\int f d\mu := \lim_{n \rightarrow \infty} \int u_n d\mu$$

das  $\mu$ -Integral von  $f$ . Wir zeigen, dass  $\int f d\mu$  wohldefiniert ist.

**Lemma 1.3**

Sind  $(u_n)$  und  $(v_n)$  wachsende Folgen aus  $\mathcal{E}$  mit  $\lim_{n \rightarrow \infty} u_n = \lim_{n \rightarrow \infty} v_n$ , so gilt:

$$\lim_{n \rightarrow \infty} \int u_n d\mu = \lim_{n \rightarrow \infty} \int v_n d\mu$$

**Beweis**

Wir zeigen zunächst:  $\lim_{n \rightarrow \infty} u_n \geq v$  mit  $v \in \mathcal{E} \implies \mu(v) \leq \lim_{n \rightarrow \infty} \mu(u_n)$

Denn: Sei  $v = \sum_{j=1}^m \alpha_j \mathbf{1}_{A_j}$  ( $\alpha_j \geq 0, A_j \in \mathcal{A}$ ) und  $0 < c < 1$  beliebig. Sei  $B_n := \{\omega | u_n(\omega) \geq cv(\omega)\} \in \mathcal{A}$ . Da  $u_n \geq cv \mathbf{1}_{B_n}$  folgt:

$$\mu(u_n) \geq c\mu(v \mathbf{1}_{B_n}) \quad (*)$$

Nach Voraussetzung:  $v \leq \lim_{n \rightarrow \infty} u_n, u_n \uparrow \implies B_n \uparrow \Omega, A_j \cap B_n \uparrow A_j$

$$\begin{aligned} \implies \mu(v) &= \sum_{j=1}^m \alpha_j \mu(A_j) = \lim_{n \rightarrow \infty} \sum_{j=1}^m \alpha_j \mu(A_j \cap B_n) \\ &= \lim_{n \rightarrow \infty} \mu(v \mathbf{1}_{B_n}) \end{aligned}$$

Nehme  $\lim_{n \rightarrow \infty}$  in  $(*)$ :  $\lim_{n \rightarrow \infty} \mu(u_n) \geq c\mu(v)$ . Da  $c < 1$  beliebig war, folgt die Behauptung.

Jetzt zur eigentlichen Aussage: Es gilt:  $v_k \leq \lim_{n \rightarrow \infty} u_n, u_k \leq \lim_{n \rightarrow \infty} v_n \xrightarrow{\text{Hilfsaussage}} \mu(v_k) \leq \lim_{n \rightarrow \infty} \mu(u_n), \mu(u_k) \leq \lim_{n \rightarrow \infty} \mu(v_n), \forall k \in \mathbb{N}$ .

$\lim_{k \rightarrow \infty}$  bei beiden Ungleichungen  $\implies$  Behauptung. ■

**Bemerkung**

- a) Die letzten beiden Definitionen sind verträglich
  - b) Die Eigenschaften von Lemma 1.2 gelten weiter.
- 3.)  $f : \Omega \rightarrow \bar{\mathbb{R}}$  ist  $\mathcal{A}$ -messbar (ohne Vorzeichenbeschränkung).  $f^+ := \max\{0, f\}$ ,  $f^- := -\min\{0, f\}$ ,  $f = f^+ - f^-$ ,  $|f| = f^+ + f^-$

**Definition**

Eine  $\mathcal{A}$ -messbare Funktion  $f : \Omega \rightarrow \bar{\mathbb{R}}$  heißt  $\mu$ -integrierbar, falls  $\int f^+ d\mu < \infty$ ,  $\int f^- d\mu < \infty$ . In diesem Fall heißt  $\int f d\mu = \mu(f) = \int f^+ d\mu - \int f^- d\mu$  das  **$\mu$ -Integral von  $f$** .

Schreibweise:  $\int f d\mu = \int f(\omega) \mu(d\omega) = \int_{\Omega} f d\mu$ ;  $\int_A f d\mu := \int f \cdot \mathbf{1}_A d\mu$

**Bemerkung** a) Die letzten beiden Definitionen sind verträglich

- b) Falls mindestens einer der Werte  $\int f^+ d\mu$ ,  $\int f^- d\mu$  endlich ist, so heißt  $f$  **quasi-integrierbar**.
- c) Ist  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum,  $X : \Omega \rightarrow \mathbb{R}$  eine Zufallsvariable, so gilt:  $EX$  existiert  $\iff X$  ist  $P$ -integrierbar. In diesem Fall:  $EX = \int X dP$
- d) Offenbar gilt:  $f$  ist integrierbar  $\iff |f|$  ist integrierbar

**Satz 1.2 (Eigenschaften des  $\mu$ -Integrals)**

Es seien  $f, g : \Omega \rightarrow \bar{\mathbb{R}}$   $\mu$ -integrierbar und  $c \in \mathbb{R}$ . Dann gilt:

- a)  $cf$  und  $f + g$  sind  $\mu$ -integrierbar und

$$\begin{aligned} \int cf d\mu &= c \int f d\mu \\ \int (f + g) d\mu &= \int f d\mu + \int g d\mu \end{aligned}$$

- b)  $f \leq g \implies \int f d\mu \leq \int g d\mu$

- c)  $|\int f d\mu| \leq \int |f| d\mu$

**Beweis** a)  $\alpha$  Sei  $c \geq 0$  (analog  $c \leq 0$ ):  $(cf)^+ = cf^+$ ,  $(cf)^- = cf^-$

Also ist  $cf$  integrierbar:  $\xrightarrow{\text{Satz 1.1}} \exists u_n^+ \uparrow f^+, u_n^+ \in \mathcal{E}$

$$\begin{aligned} \int cf^+ d\mu &= \lim_{n \rightarrow \infty} \int cu_n^+ d\mu \\ &= c \lim_{n \rightarrow \infty} \int u_n^+ d\mu \\ &= c \int f^+ d\mu \end{aligned}$$

Analog  $f^-$ .

β)  $|f + g| \leq |f| + |g| \implies f + g$   $\mu$ -integrierbar.

Sei zunächst  $f, g \in \mathcal{E}^+ \xrightarrow{\text{Satz 1.1}} \exists u_n \uparrow f, v_n \uparrow g, u_n, v_n \in \mathcal{E} \implies u_n + v_n \uparrow f + g, u_n + v_n \in \mathcal{E}$

Mit Lemma 1.2 folgt:

$$\begin{aligned} \int (f + g) d\mu &= \lim_{n \rightarrow \infty} \int (u_n + v_n) d\mu \\ &= \lim_{n \rightarrow \infty} \left( \int u_n d\mu + \int v_n d\mu \right) \\ &= \lim_{n \rightarrow \infty} \int u_n d\mu + \lim_{n \rightarrow \infty} \int v_n d\mu \\ &= \int f d\mu + \int g d\mu \end{aligned}$$

Sei jetzt  $f, g$  beliebig

$$\begin{aligned} (f + g)^+ - (f + g)^- &= f + g = f^+ - f^- + g^+ - g^- \implies (f + g)^+ + f^- + g^- = (f + g)^- + f^+ + g^+ \xrightarrow{\text{s.o.}} \int (f + g)^+ d\mu + \int f^- d\mu + \int g^- d\mu = \\ &= \int (f + g)^- d\mu + \int f^+ d\mu + \int g^+ d\mu \\ \implies \int (f + g) d\mu &= \int (f + g)^+ d\mu - \int (f + g)^- d\mu = \int f^+ d\mu - \int f^- d\mu + \int g^+ d\mu - \int g^- d\mu = \int f d\mu + \int g d\mu \end{aligned}$$

b) vergleiche Übung

c)  $f \leq |f|, -f \leq |f| \xrightarrow{\text{b) mit } g = |f|}$  Behauptung ■

**Bemerkung** Ist  $\mu = \lambda$  das Lebesgue-Maß, so heißt  $\int f d\mu = \int f d\lambda$  Lebesgue-Integral.

**Beispiel 1.2** a) Sei  $\delta_\omega$  das Dirac-Maß,  $f : \Omega \rightarrow \bar{\mathbb{R}}$  ist  $\delta_\omega$ -integrierbar falls  $f(\omega) < \infty$  und dann gilt

$$\int f d\delta_\omega = f(\omega)$$

Denn: Sei  $f \in \mathcal{E} \implies f = \sum_{j=1}^n \alpha_j \mathbf{1}_{A_j} \implies \int f d\delta_\omega = \sum_{j=1}^n \alpha_j \delta_\omega(A_j) = \alpha_k \cdot 1 = f(\omega)$

$f \in \mathcal{E}^+ : u_n \uparrow f, \int u_n d\delta_\omega = u_n(\omega) \uparrow f(\omega)$

$f$  allgemein  $\implies f = f^+ - f^-$

b) Sei  $(\mu_n)$  eine Folge von Maßen und  $\mu = \sum_{n=1}^\infty \mu_n$ . Für  $f : \Omega \rightarrow \bar{\mathbb{R}}$  gilt:

$$\begin{aligned} f \text{ ist } \mu\text{-integrierbar} &\iff \sum_{n=1}^\infty \int |f| d\mu_n < \infty \\ \int f d\mu &= \sum_{n=1}^\infty \int f d\mu_n \text{ (vergleiche Übung)} \end{aligned}$$

Spezialfall:  $(\Omega, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$ ,  $\mu = \sum_{n=1}^\infty \delta_n$  (Zählmaß auf  $\mathbb{N}$ )

$f$  ist  $\mu$ -integrierbar  $\iff \sum_{n=1}^\infty |f(n)| < \infty$ , dann  $\int f d\mu = \sum_{n=1}^\infty f(n)$ .

Summation ist ein Spezialfall von Integration. Sei  $\Omega = \{\omega_1, \omega_2, \dots\}$ ,  $\mathcal{A} =$

$\mathcal{P}(\Omega) \cdot \mu = P := \sum_{n=1}^{\infty} p_n \delta_{\omega_n}$  mit  $p_n \geq 0, \sum_{n=1}^{\infty} p_n = 1$  (Wahrscheinlichkeitsmaß).

Sei  $X : \Omega \rightarrow \mathbb{R}$  eine Zufallsvariable:

$$EX \text{ existiert} \iff \sum_{n=1}^{\infty} |X(\omega_n)| p_n < \infty \iff X \text{ ist } P\text{-integrierbar}$$

$$EX = \sum_{n=1}^{\infty} X(\omega_n) P_n = \sum_{n=1}^{\infty} X(\omega_n) P(\{\omega_n\}) = \int X dP$$

- c) Sei  $\Omega = [a, b]$  und  $\mathcal{A} = \mathfrak{B}_{[a,b]} = \{A \cap [a, b] \mid A \in \mathfrak{B}\}$  (Spur von  $\mathfrak{B}$  auf  $[a, b]$ )  
 $\mu(A) := \lambda(A) \forall A \in \mathcal{A}$ . Ist  $f : \Omega \rightarrow \mathbb{R}$  messbar und  $f$  Riemann-integrierbar, so ist  $f$  auch  $\mu$ -integrierbar und es gilt:

$$\int f d\mu = \int f(x) dx$$

(Hier fehlt ein Bild zur Veranschaulichung)

Das Lebesgue-Integral ist eine Erweiterung des Riemann-Integrals:

Sei  $f = \mathbf{1}_{\mathbb{Q} \cap [0,1]}$ .  $f$  ist nicht Riemann-integrierbar. Da  $f \in \mathcal{E}$  gilt:

$$\int f d\lambda = 0 \cdot \lambda(\mathbb{Q}^c \cap [0, 1]) + 1 \cdot \lambda(\mathbb{Q} \cap [0, 1]) = 0$$

Das letzte Gleichheitszeichen gilt wegen:

- (i)  $\lambda(\{a\}) = 0$ , da  $\{a\} = \cap_{n=1}^{\infty} [a, a + \frac{1}{n})$
- (ii)  $\lambda(\sum_{i=1}^{\infty} \{a_i\}) = \sum_{i=1}^{\infty} \lambda(\{a_i\}) = 0$

Vorsicht bei uneigentlichen Riemann-Integralen!  $\int_0^{\infty} \frac{\sin x}{x} dx$  ist Riemann-integrierbar, aber nicht Lebesgue-integrierbar.



## 2 Eigenschaften des Maß-Integrals

### 2.1 Konvergenzsätze

Im Folgenden sei  $(\Omega, \mathcal{A}, \mu)$  ein Maßraum und  $f, f_1, f_2, \dots : \Omega \rightarrow \bar{\mathbb{R}}$  messbare Funktionen.

**Satz 2.1 (Satz von Beppo Levi, Satz von der monotonen Konvergenz)**

Sind  $f, f_1, f_2, \dots \geq 0$  mit  $f_n \uparrow f$ , so gilt

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

**Beweis**  $\forall f_n \exists (u_{nm})_{m \in \mathbb{N}} \subset \mathcal{E}$  mit  $u_{nm} \uparrow f_n$  für  $m \rightarrow \infty$ . Sei  $h_m := \max\{u_{1m}, \dots, u_{mm}\} \implies h_m \uparrow$  und  $(h_m) \subset \mathcal{E}$ . Außerdem:  $u_{nm} \leq h_m$  für  $n \leq m$ .

Also:  $f_n = \sup_{m \in \mathbb{N}} u_{nm} = \sup_{m \geq n} u_{nm} \leq \sup_{m \in \mathbb{N}} h_m$  und  $h_m \leq f_m \leq f$ . Insgesamt:  $h_m \uparrow f$  und  $\lim_{m \rightarrow \infty} \int h_m d\mu = \int f d\mu$ . Mit  $\int h_m d\mu \leq \int f_m d\mu \leq \int f d\mu$  folgt die Behauptung. ■

Im Folgenden sei  $(\Omega, \mathcal{A}, \mu)$  ein Maßraum und  $f_1, f_2, f_3, \dots : \Omega \rightarrow \bar{\mathbb{R}}$  messbare Funktionen.

**Satz 2.2 (Lemma von Fatou)**

Gilt  $f_n \geq 0, n \in \mathbb{N}$ , so folgt

$$\int \liminf_{n \rightarrow \infty} f_n d\mu \leq \liminf_{n \rightarrow \infty} \int f_n d\mu$$

**Beweis** Sei  $g_n := \inf_{m \geq n} f_m, f := \liminf_{n \rightarrow \infty} f_n$ , so gilt  $g_n \uparrow f$  und mit Satz 2.1  $\int \liminf_{n \rightarrow \infty} f_n d\mu = \lim_{n \rightarrow \infty} \int g_n d\mu = \liminf_{n \rightarrow \infty} \int g_n d\mu \leq \liminf_{n \rightarrow \infty} \int f_n d\mu$  ■

**Satz 2.3 (Satz von Lebesgue oder Satz von der majorisierten Konvergenz)**

Es gelte  $\lim_{n \rightarrow \infty} f_n(\omega) = f(\omega) \forall \omega \in \Omega$ . Existiert eine  $\mu$ -integrierbare Funktion  $g : \Omega \rightarrow \mathbb{R}$  mit der Eigenschaft  $|f_n(\omega)| \leq g(\omega) \forall \omega \in \Omega, \forall n \in \mathbb{N}$ , so folgt:

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

**Beweis** Sei  $g_n := |f_n - f|, h := |f| + g$ . Wegen  $|h| \leq 2g$  ist  $h$   $\mu$ -integrierbar. Außerdem gilt

$$\begin{aligned} h - g_n &= |f| + g - |f_n - f| \geq |f| + g - |f_n| - |f| \\ &= g - |f_n| \geq 0 \end{aligned}$$

wegen  $g_n \rightarrow 0$  gilt  $h - g_n \rightarrow h$ , also folgt mit Satz 2.2

$$\begin{aligned} \int h d\mu &= \int \liminf_{n \rightarrow \infty} (h - g_n) d\mu \\ &\leq \liminf_{n \rightarrow \infty} \int (h - g_n) d\mu \\ &= \underbrace{\int h d\mu}_{< \infty} - \limsup_{n \rightarrow \infty} \int g_n d\mu \end{aligned} \quad \blacksquare$$

$\Rightarrow \limsup_{n \rightarrow \infty} \int g_n d\mu \leq 0$  Wegen  $g_n \geq 0$  bedeutet dies:

$$\lim_{n \rightarrow \infty} \int |f_n - f| d\mu = \lim_{n \rightarrow \infty} \int g_n d\mu = 0$$

und damit

$$\left| \int f_n d\mu - \int f d\mu \right| = \left| \int (f_n - f) d\mu \right| \leq \int |f_n - f| d\mu \rightarrow 0$$

**Bemerkung 2.1** Für Wahrscheinlichkeitsmaße lautet Satz 2.3:

Ist  $(X_n)_{n \in \mathbb{N}}$  eine Folge von Zufallsvariablen, so dass  $X_n \xrightarrow{f.s.} X$  ( $X$  ist dann automatisch wieder eine Zufallsvariable) und es gibt eine Zufallsvariable  $Y$  mit  $|X_n| \leq Y \forall n \in \mathbb{N}$  und  $EY < \infty$ , so gilt  $\lim_{n \rightarrow \infty} EX_n = EX$ .

Oft kommt man mit einer Majorante der Form  $Y \equiv c, c \in \mathbb{R}$  zum Ziel.

## 2.2 Verhalten bei Transformationen

Es sei  $(\Omega, \mathcal{A}, \mu)$  ein Maßraum und  $(\Omega', \mathcal{A}')$  ein messbarer Raum und  $T : \Omega \rightarrow \Omega'$  eine  $(\mathcal{A}, \mathcal{A}')$ -messbare Abbildung. Aus Stochastik 1 ist bekannt (vgl. §5.2, Verteilung), dass durch

$$\mu^T : \mathcal{A}' \rightarrow [0, \infty], \mu^T(A') := \underbrace{\mu(T^{-1}(A'))}_{\in \mathcal{A}} = \mu(\{\omega \in \Omega \mid T(\omega) \in A'\})$$

ein Maß auf  $(\Omega', \mathcal{A}')$  definiert wird (Maßtransport).  $\mu^T$  heißt **Bildmaß** von  $\mu$  unter der Transformation  $T$ .

Ist  $X = T$  eine Zufallsgröße auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$  mit Werten in  $(\Omega', \mathcal{A}')$ , so nennt man  $\mu^T = P^X$  die Verteilung von  $X$ . Sei nun weiter  $f : \Omega' \rightarrow \mathbb{R}$  messbar.

$$\begin{array}{ccc} \text{Skizze: } (\Omega, \mathcal{A}) & \xrightarrow{T} & (\Omega', \mathcal{A}') \\ & \searrow f \circ T & \downarrow f \\ & & (\mathbb{R}, \mathfrak{B}) \end{array}$$



**Satz 2.4 (Integration bezüglich des Bildmaßes, Transformationssatz)**

Mit den obigen Bezeichnungen und Voraussetzungen gilt:  $f$  ist genau dann  $\mu^T$ -integrierbar, wenn  $f \circ T$   $\mu$ -integrierbar ist.

Dann gilt:

$$\int f d\mu^T = \int (f \circ T) d\mu$$

**Beweis**

(i) Falls  $f = \mathbf{1}_A, (A \in \mathcal{A})$  gilt

$$\begin{aligned} \int f d\mu^T &= \mu^T(A) \\ &= \mu(T^{-1}(A)) \\ &= \int \mathbf{1}_{T^{-1}(A)} d\mu \\ &= \int \mathbf{1}_A \circ T d\mu \\ &= \int f \circ T d\mu \end{aligned}$$

wegen Satz 1.2(a) folgt damit die Aussage für  $f \in \mathcal{E}$

(ii) Sei jetzt  $f \geq 0 \implies \exists (u_n)_{n \in \mathbb{N}} \subset \mathcal{E}$  mit  $u_n \uparrow f$  und  $\int f d\mu^T = \lim_{n \rightarrow \infty} \int u_n d\mu^T$ .  
Offenbar gilt  $u_n \circ T \in \mathcal{E}, (u_n \circ T) \uparrow (f \circ T)$

Also folgt:

$$\begin{aligned} \int f d\mu^T &= \lim_{n \rightarrow \infty} \int u_n d\mu^T \\ &\stackrel{(i)}{=} \lim_{n \rightarrow \infty} \int (u_n \circ T) d\mu \\ &= \int (f \circ T) d\mu \end{aligned}$$

(iii) Ist  $f : \Omega' \rightarrow \mathbb{R}$  eine beliebige  $(\mathcal{A}', \mathfrak{B})$ -messbare Abbildung so gilt

$$\begin{aligned} \int f^+ d\mu^T < \infty &\iff \int f^+ \circ T d\mu < \infty \\ \int f^- d\mu^T < \infty &\iff \int f^- \circ T d\mu < \infty \end{aligned}$$

Da  $(f \circ T)^+ = f^+ \circ T, (f \circ T)^- = f^- \circ T$ , folgt  $f$   $\mu^T$ -integrierbar  $\iff f \circ T$

$\mu$ -integrierbar

$$\begin{aligned}
 \int f d\mu^T &= \int f^+ d\mu^T - \int f^- d\mu^T \\
 &\stackrel{(ii)}{=} \int f^+ \circ T d\mu - \int f^- \circ T d\mu \\
 &= \int (f \circ T)^+ d\mu - \int (f \circ T)^- d\mu \\
 &= \int f \circ T d\mu. \quad \blacksquare
 \end{aligned}$$

**Bemerkung 2.2** Das Beweisverfahren (zuerst für  $f \in \mathcal{E}$  (bzw.  $f = \mathbf{1}_A$ ), dann für  $f \in \mathcal{E}^+$ , dann für  $f$  beliebig) heißt **algebraische Induktion** und wird häufig verwendet.

## 2.3 Nullmengen und Maße mit Dichten

Im Folgenden sei  $(\Omega, \mathcal{A}, \mu)$  ein Maßraum.

**Definition 2.1**  $N \in \mathcal{A}$  heißt  $\mu$ -Nullmenge, falls  $\mu(N) = 0$ .

**Definition 2.2** Ist  $(A)$  eine Aussage, die von  $\omega \in \Omega$  abhängt, so sagen wir, dass  $(A)$   $\mu$ -**fast überall** ( $\mu$ -f.ü.) gilt, wenn  $(A)$  wahr ist  $\forall \omega$  außerhalb einer  $\mu$ -Nullmenge. Ist  $\mu = P$  ein Wahrscheinlichkeitsmaß, so sagt man  $P$ -fast-überall oder  $P$ -fast sicher ( $P$ -f.s.)

### Satz 2.5

$f, g : \Omega \rightarrow \mathbb{R}$  seien  $(\mathcal{A}, \mathfrak{B})$  messbar.

- a) Sei  $f \geq 0$ . Dann gilt:  $\int f d\mu = 0 \iff f = 0, \mu$ -f.ü.
- b) Ist  $f$   $\mu$ -integrierbar und gilt  $f = g$   $\mu$ -f.ü., so ist auch  $g$   $\mu$ -integrierbar mit  $\int f d\mu = \int g d\mu$ .

### Beweis

- a) Sei  $N := \{\omega \in \Omega \mid f(\omega) \neq 0\}$ .  $N \in \mathcal{A}$ , da  $f$  messbar.
  - (i) Annahme:  $\int f d\mu = 0$ .  
 Sei  $A_n := \{\omega \in \Omega \mid f(\omega) \geq \frac{1}{n}\} \implies A_n \uparrow N$  und  $\mu(N) = \lim_{n \rightarrow \infty} (\mu(A_n))$ .  
 Außerdem gilt  $0 = \int f d\mu \geq \int \frac{1}{n} \cdot \mathbf{1}_{A_n} d\mu = \frac{1}{n} \cdot \mu(A_n) \geq 0$   
 $\implies \mu(A_n) = 0 \forall n \in \mathbb{N} \implies \mu(N) = 0$ , also  $f = 0$   $\mu$ -f.ü.
  - (ii) Annahme:  $N$  ist  $\mu$ -Nullmenge.  
 Sei  $g \in \mathcal{E}$ ,  $g(\Omega) = \{\alpha_1, \dots, \alpha_n\}$ ,  $g \leq f$ .  
 $\implies g = \sum_{j=1}^n \alpha_j \circ \mathbf{1}_{A_j}$ .  
 Falls  $\alpha_j > 0 \implies A_j \subset N \implies \int g d\mu = 0 \xrightarrow{\text{L.1.3}} \int f d\mu = 0$ .

b) Seien zunächst  $f, g \geq 0$ ,  $N := \{f \neq g\} \xrightarrow{a)}$

$$\begin{aligned} \int f d\mu &= \int_N f d\mu + \int_{N^c} f d\mu \\ &= 0 + \int_{N^c} g d\mu \\ &= \int_N g d\mu + \int_{N^c} g d\mu \\ &= \int g d\mu \end{aligned}$$

Insbesondere:  $\int f d\mu < \infty \iff \int g d\mu < \infty$ .

Seien nun  $f, g$  beliebig. Wegen  $\{f^+ = g^+\} \supset \{f = g\} \subset \{f^- = g^-\}$  gilt auch  $f^+ = g^+$  und  $f^- = g^-$   $\mu$ -f.ü. und mit dem vorigen Teil folgt die Behauptung.

■

**Bemerkung 2.3** Im Folgenden sei  $L^1(\Omega, \mathcal{A}, \mu) := \{f : \Omega \rightarrow \mathbb{R} \mid f \text{ ist messbar und } \mu\text{-integrierbar}\}$  (ist ein Vektorraum) und wir definieren

$f \sim_\mu g : \iff f = g$   $\mu$ -f.ü. und  $\sim_\mu$  ist Äquivalenzrelation auf  $\{f : \Omega \rightarrow \mathbb{R} \mid f \text{ ist messbar}\}$ . Sei  $f^{[\mu]}$  die Äquivalenzklasse zu  $f$ .

Mit Satz 2.5: Entweder alle oder keines der Elemente in  $f^{[\mu]}$  ist  $\mu$ -integrierbar und die Integrale sind ggfs. gleich. Außerdem gilt:

$$f_1 \in f^{[\mu]}, g_1 \in g^{[\mu]} \implies f_1 + g_1 \in (f + g)^{[\mu]}.$$

$\implies$  Man kann zum Raum der Äquivalenzklassen übergehen:  $L^1(\Omega, \mathcal{A}, \mu) / \sim_\mu$

Mit  $\|f^{[\mu]}\|_1 := \int |f| d\mu$  ist eine Norm definiert; sie ist wohldefiniert, da  $\int f_1 d\mu = \int f_2 d\mu \forall f_1, f_2 \in f^{[\mu]}$ .

Wichtig:  $f \mapsto \int |f| d\mu =: \|f\|$  ist auf  $L^1(\Omega, \mathcal{A}, \mu)$  keine Norm, da  $\|f\| = 0 \implies f \equiv 0$  im Allgemeinen falsch ist!

**Satz 2.6**  $(L^1(\Omega, \mathcal{A}, \mu) / \sim_\mu, \|\cdot\|_1)$  ist ein Banachraum.

**Definition 2.3** Es seien  $\mu, \nu$  Maße auf dem messbaren Raum  $(\Omega, \mathcal{A})$ . Gilt dann  $\mu(A) = 0 \implies \nu(A) = 0 \forall A \in \mathcal{A}$ , so heißt  $\nu$   **$\mu$ -stetig**, in Zeichen  $\nu \ll \mu$ . Man sagt auch, dass  $\mu$  das Maß  $\nu$  dominiert.

**Satz 2.7** und Definition

Sei  $(\Omega, \mathcal{A}, \mu)$  ein Maßraum und  $f : \Omega \rightarrow \mathbb{R}_+$   $(\mathcal{A}, \mathfrak{B})$ -messbar. Dann wird durch  $\nu : \mathcal{A} \rightarrow \mathbb{R}_+ \cup \{\infty\}$ ,  $\nu(A) := \int_A f d\mu$  ein Maß auf  $(\Omega, \mathcal{A})$  definiert. Man nennt  $\nu$  das **Maß mit der Dichte**  $f$  bzgl.  $\mu$  und  $f$  eine  **$\mu$ -Dichte** von  $\nu$ . Schreibweise:  $f = \frac{d\nu}{d\mu}$

**Beweis** Wir weisen nach, dass  $\nu$  ein Maß ist:

$\nu \geq 0$  ist klar, da  $f$  nach  $\mathbb{R}_+$  abbildet;

$$(i) \quad \mu(\emptyset) = \int f \cdot \mathbf{1}_\emptyset d\mu = 0.$$

- (ii) Seien  $A_1, A_2, \dots$  paarweise disjunkt und  $A = \sum_{n=1}^{\infty} A_n$ .  
 Wegen  $f \cdot \mathbf{1}_{\sum_{k=1}^n A_k} \uparrow f \cdot \mathbf{1}_A$  folgt mit Satz 2.1:

$$\begin{aligned}
 \nu\left(\sum_{n=1}^{\infty} A_n\right) &= \int f \cdot \mathbf{1}_A d\mu \\
 &= \lim_{n \rightarrow \infty} \left( \int f \cdot \underbrace{\mathbf{1}_{\sum_{k=1}^n A_k}}_{=\sum_{k=1}^n \mathbf{1}_{A_k}} d\mu \right) \\
 &= \lim_{n \rightarrow \infty} \left( \int \sum_{k=1}^n f \cdot \mathbf{1}_{A_k} d\mu \right) \\
 &= \lim_{n \rightarrow \infty} \left( \sum_{k=1}^n \underbrace{\left( \int f \cdot \mathbf{1}_{A_k} d\mu \right)}_{=\nu(A_k)} \right) \\
 &= \sum_{k=1}^{\infty} \nu(A_k)
 \end{aligned}$$

**Satz 2.8 (Satz von Radon-Nikodym)**

Seien  $\mu, \nu$  Maße auf dem messbaren Raum  $(\Omega, \mathcal{A})$ ,  $\mu$  sei  $\sigma$ -endlich. Dann gilt:  
 $\nu$  ist genau dann  $\mu$ -stetig, wenn  $\nu$  eine Dichte bzgl.  $\mu$  hat.

**Beweis**  $\nu$  hat Dichte bzgl.  $\mu \implies \nu(A) = \int_A f d\mu = \int f \cdot \mathbf{1}_A d\mu \xrightarrow{\text{S.2.5a}} \nu \ll \mu$ .  
 Die andere Richtung siehe z.B. Henze, Stochastik II. ■

**Satz 2.9** Seien  $\mu$  und  $\nu$  Maße auf  $(\Omega, \mathcal{A})$ ,  $\nu$  habe  $\mu$ -Dichte  $f$ . Dann gilt für alle  $(\mathcal{A}, \mathfrak{B})$ -messbaren Abbildungen  $g : \Omega \rightarrow \mathbb{R}$ :

$g$  ist genau dann  $\nu$ -integrierbar, wenn  $g \cdot f$   $\mu$ -integrierbar ist und in diesem Fall ist  $\int g d\nu = \int g \cdot f d\mu$ .

**Beweis** Übung. ■

**Bemerkung 2.4** Merkgel:  $\int g d\nu = \int g \cdot \frac{d\nu}{d\mu} d\mu$ .

**Beispiel 2.1** Sei  $\mu = \lambda$  das Lebesgue-Maß und  $\nu = P^X$  die Verteilung einer Zufallsvariablen  $X$ . Ist  $X$  absolutstetig, so gilt (Stochastik I):

$$P^X(B) = \int_B f_X(x) dx$$

mit  $f_X : \mathbb{R} \rightarrow \mathbb{R}_+ \cup \{\infty\}$  und

$$EX = \int_{\Omega} X dP = \int_{\mathbb{R}} x P^X(dx) = \int_{\mathbb{R}} x \cdot f_X(x) dx.$$

mit den Sätzen 2.4 und 2.9.

## 2.4 Ungleichungen und Räume integrierbarer Funktionen

Hier stellen wir einige Hilfsmittel für später zusammen. Der folgende Satz behandelt den Spezialfall von Wahrscheinlichkeitsmaßen.

**Satz 2.10** *Es sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum,  $X : \Omega \rightarrow \mathbb{R}$  eine Zufallsvariable und  $\gamma > 0$ . Dann gilt:*

$$P(|X| \geq a) \leq \frac{1}{a^\gamma} \cdot E|X|^\gamma \quad \forall a > 0.$$

Existiert die Varianz von  $X$ , so gilt:

$$P(|X - EX| \geq a) \leq \frac{1}{a^2} \cdot \text{Var}(X) \quad \forall a > 0.$$

(Ungleichung von Tschebyschef, siehe Abschnitt 7.6, Stochastik I)

### Beweis

Sei  $Y : \Omega \rightarrow \mathbb{R}$  definiert durch:

$$Y(\omega) = \begin{cases} a, & \text{falls } |X(\omega)| \geq a \\ 0, & \text{sonst} \end{cases}$$

$$\implies |Y| \leq |X|$$

$$\implies |Y|^\gamma \leq |X|^\gamma \quad \forall \gamma > 0$$

$$\implies a^\gamma P(|X| \geq a) = a^\gamma P(|Y| \geq a) = E|Y|^\gamma \leq E|X|^\gamma$$

Für Teil 2 setze  $\tilde{X} := X - EX$  und  $\gamma = 2$ . ■

Sei  $I \subset \mathbb{R}$  ein offenes Intervall und  $\Phi : I \rightarrow \mathbb{R}$  eine konvexe Funktion, d.h.

$$\Phi(\alpha x + (1 - \alpha)y) \leq \alpha\Phi(x) + (1 - \alpha)\Phi(y) \quad \forall x, y \in I, \forall \alpha \in [0, 1]$$

Außerdem gilt  $\forall y \in I, \exists m \in \mathbb{R}$ , mit

$$\Phi(x) \geq \Phi(y) + m(x - y)$$

### Satz 2.11 (Jensensche Ungleichung)

*Es seien  $I \subset \mathbb{R}$  ein offenes Intervall,  $\Phi : I \rightarrow \mathbb{R}$  konvex und  $X$  eine Zufallsvariable mit  $E|X| < \infty, E|\Phi(X)| < \infty$  und  $P(X \in I) = 1$ . Dann gilt:*

$$EX \in I \text{ und } \Phi(EX) \leq E\Phi(X)$$

### Beweis

Falls  $I = (-\infty, \infty)$  ist automatisch  $EX \in I$ . Ist  $X < a$  P-f.s. so gilt:  $EX \leq Ea = a$ .

Falls  $E(a - X) = 0$  folgt, da  $a - X \geq 0 \xrightarrow{\text{Satz 2.5}} X = a$  P-f.s. Widerspruch!

D.h., falls  $I = (\cdot, a) \subset (-\infty, a) \implies EX < a$ . Analog untere Schranke  $\implies EX \in I$ .

Mit der Vorüberlegung folgt ( $y = EX, x = X(\omega)$ )

$$\Phi(X) \geq \Phi(EX) + m(X - EX) \quad \text{P-f.s.}$$

für ein  $m \in \mathbb{R}$ . Erwartungswert auf beiden Seiten führt zur Behauptung (Nullmengen können wir vernachlässigen). ■

**Beispiel 2.2**

Für  $\Phi(x) = |x|$ ,  $\Phi(x) = x^2$  folgt:  $|EX| \leq E|X|$ ,  $(EX)^2 \leq EX^2$ . ( $\implies EX^2 - (EX)^2 = \text{Var } X \geq 0$ )

Im Folgenden sei  $(\Omega, \mathcal{A}, \mu)$  wieder ein Maßraum.

**Definition**

Eine messbare Funktion  $f : \Omega \rightarrow \mathbb{R}$  heißt  **$p$ -fach  $\mu$ -integrierbar**, wenn  $\int |f|^p d\mu < \infty$  mit  $p > 0$ .

$$L^p(\Omega, \mathcal{A}, \mu) := \{f : \Omega \rightarrow \mathbb{R} \mid \int |f|^p d\mu < \infty\}$$

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

Wie im vorigen Abschnitt ist  $L^p$  bzw.  $L^p(\Omega, \mathcal{A}, \mu) / \sim_\mu$  ein Vektorraum über  $\mathbb{R}$  und  $\|f\|_p$  auf den Äquivalenzklassen eine Norm.

**Satz 2.12**

- a) (Höldersche Ungleichung) Es seien  $p > 1$ ,  $f \in L^p(\Omega, \mathcal{A}, \mu)$ ,  $g \in L^q(\Omega, \mathcal{A}, \mu)$ , wobei  $\frac{1}{p} + \frac{1}{q} = 1$ . Dann folgt:  $f \cdot g \in L^1(\Omega, \mathcal{A}, \mu)$  und es gilt:

$$\|f \cdot g\|_1 \leq \|f\|_p \cdot \|g\|_q$$

- b) (Minkowskische Ungleichung) Es seien  $p \geq 1$  und  $f, g \in L^p(\Omega, \mathcal{A}, \mu)$ . Dann folgt  $f + g \in L^p(\Omega, \mathcal{A}, \mu)$  und es gilt:

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

**Beweis**

- a) Falls  $\int |f|^p d\mu = 0 \xrightarrow{\text{Satz 2.5}} f = 0 \text{ } \mu\text{-f.s.}$  und die Ungleichung ist richtig. Sei also  $\|f\|_p > 0$  und  $\|g\|_q > 0$  (gleiches Argument).  $x \mapsto \log x$  ist konkav, d.h. es gilt:  $\alpha \log(a) + (1 - \alpha) \log(b) \leq \log(\alpha a + (1 - \alpha)b) \quad \forall a, b > 0, 0 < \alpha < 1$ .  $\exp(\cdot)$  auf beiden Seiten:

$$a^\alpha b^{1-\alpha} \leq \alpha a + (1 - \alpha)b \quad \forall a, b \geq 0, 0 < \alpha < 1$$

Setze  $a := \frac{|f(\omega)|^p}{\|f\|_p^p}$ ,  $b := \frac{|g(\omega)|^q}{\|g\|_q^q}$ ,  $\alpha = \frac{1}{p}$  ( $\omega$  beliebig)

$$\implies \frac{|f(\omega)| \cdot |g(\omega)|}{\|f\|_p \cdot \|g\|_q} \leq \frac{1}{p} \frac{|f(\omega)|^p}{\|f\|_p^p} + \frac{1}{q} \frac{|g(\omega)|^q}{\|g\|_q^q}$$

$$\implies |f(\omega)| \cdot |g(\omega)| \leq \frac{1}{p} |f(\omega)|^p \|f\|_p^{1-p} \|g\|_q + \frac{1}{q} |g(\omega)|^q \|g\|_q^{1-q} \|f\|_p$$

$$\begin{aligned} \xrightarrow{\text{Int. über } \omega} \|f \cdot g\|_1 &\leq \frac{1}{p} \|f\|_p^p \|f\|_p^{1-p} \|g\|_q + \frac{1}{q} \|g\|_q^q \|g\|_q^{1-q} \|f\|_p \\ &= \frac{1}{p} \|f\|_p \|g\|_q + \frac{1}{q} \|g\|_q \|f\|_p \end{aligned}$$

$\implies$  Behauptung

- b) Wegen  $|f + g| \leq |f| + |g|$  gilt  $\|f + g\|_p \leq \|f\|_p + \|g\|_p$ . Also genügt es die Ungleichung für  $f + g \geq 0$  zu beweisen. Falls  $p = 1$  folgt  $\|f + g\|_1 = \int (f + g) d\mu = \int f d\mu + \int g d\mu = \|f\|_1 + \|g\|_1$ . Sei also  $p > 1$ . Mit  $(f + g)^p \leq (2 \cdot \max\{f, g\})^p \leq 2^p(|f|^p + |g|^p) \implies (f + g) \in L^p$ , also  $\|f + g\|_p < \infty$ . Sei  $q := \frac{1}{1-\frac{1}{p}}$ . Anwendung von Teil a) liefert:

$$\begin{aligned} \|f + g\|_p^p &= \int f(f + g)^{p-1} d\mu + \int g(f + g)^{p-1} d\mu \\ &\stackrel{\text{a)}}{\leq} (\|f\|_p + \|g\|_p) \|(f + g)^{p-1}\|_q \quad (*) \end{aligned}$$

Wegen  $(p - 1)q = p$  gilt:

$$\|(f + g)^{p-1}\|_q = \left( \int (f + g)^{(p-1)q} d\mu \right)^{\frac{1}{q}} = \|f + g\|_p^{\frac{p}{q}} = \|f + g\|_p^{p-1}$$

Falls  $\|f + g\|_p = 0$  ist die Ungleichung richtig. Sei also  $\|f + g\|_p > 0$ . Nehme (\*) und teile durch  $\|f + g\|_p^{p-1}$  auf beiden Seiten  $\implies$  Behauptung. ■

### Bemerkung

Falls  $p = q = 2, \Omega = \{1, \dots, n\}, \mathcal{A} = \mathcal{P}(\Omega), \mu = \sum_{k=1}^n \delta_k, f(i) = a_i, g(i) = b_i$ , bekommt man:

$$\sum_{i=1}^n a_i b_i \leq \left( \sum_{i=1}^n a_i^2 \right)^{\frac{1}{2}} \cdot \left( \sum_{i=1}^n b_i^2 \right)^{\frac{1}{2}}$$

In diesem Fall ist Satz 2.12 a) die Cauchy-Schwarz-Ungleichung.

Lineare Algebra:  $|\langle a, b \rangle| \leq \|a\| \cdot \|b\| \quad \forall a, b \in \mathbb{R}^n$ . Das motiviert

### Satz 2.13

Es sei  $(\Omega, \mathcal{A}, \mu)$  ein Maßraum und  $L^2(\Omega, \mathcal{A}, \mu) / \sim_\mu$  der Raum der  $\sim_\mu$ -Äquivalenzklassen quadratisch  $\mu$ -integrierbarer Funktion  $f : \Omega \rightarrow \mathbb{R}$ .

Dann ist  $\langle f, g \rangle := \int f \cdot g d\mu$  hierauf ein Skalarprodukt, durch den  $L^2(\Omega, \mathcal{A}, \mu) / \sim_\mu$  zu einem Hilbertraum wird.

**Beweis** siehe Henze, Stochastik II ■

### Bemerkung

- a)  $(L^p(\Omega, \mathcal{A}, \mu) / \sim_\mu, \|\cdot\|_p)$  ist ein Banachraum für  $p \geq 1$ .
- b) Ist  $\Phi : L^p(\Omega, \mathcal{A}, \mu) \rightarrow \mathbb{R}$  stetig und linear, so existiert ein  $g \in L^q(\Omega, \mathcal{A}, \mu)$  mit  $\Phi(f) = \int f \cdot g d\mu \quad \forall f \in L^p(\Omega, \mathcal{A}, \mu)$ .





## 3 Produktmaße und Unabhängigkeit

### 3.1 Der allgemeine Fall

Im Folgenden sei  $I \neq \emptyset$  eine beliebige Indexmenge.  $\forall i \in I$  sei  $(\Omega_i, \mathcal{A}_i)$  ein messbarer Raum. Weiter sei  $\Omega := \times_{i \in I} \Omega_i$  ein neuer Ergebnisraum. Wir definieren die **Projektion** auf die  $i$ -te Koordinate  $\Pi_i : \Omega \rightarrow \Omega_i$  durch  $\Pi_i(\omega) = \omega_i$ .

**Definition** Die **Produkt- $\sigma$ -Algebra**  $\mathcal{A} := \bigotimes_{i \in I} \mathcal{A}_i$  ist die kleinste  $\sigma$ -Algebra mit der Eigenschaft, dass für alle  $i \in I$  die Abbildung  $\Pi_i$   $(\mathcal{A}, \mathcal{A}_i)$ -messbar ist. Genauer:

$$\mathcal{A} := \sigma \left( \bigcup_{i \in I} \{ \Pi_i^{-1}(A_i) \mid A_i \in \mathcal{A}_i \} \right)$$

**Bemerkung** Sei  $J \subset I$ ,  $\Pi_J : \Omega \rightarrow \times_{i \in J} \Omega_i$ ,  $\Pi_J(\omega)(j) = \omega_j$  ( $j \in J$ ) die Projektion auf die  $J$ -Koordinaten, so bildet

$$\left\{ \Pi_J^{-1}(A_J) \mid A_J \in \bigotimes_{i \in J} \mathcal{A}_i, J \subset I, J \text{ endlich} \right\}$$

ein durchschnittstabiles Erzeugendensystem von  $\mathcal{A}$ . Man nennt diese Mengen auch **Zylindermengen** mit endlicher Basis.

$$\left( A_J = A_{i_1} \times \cdots \times A_{i_{|J|}}, \Pi_J^{-1}(A_J) = \bigcap_{k=1}^{|J|} \Pi_{i_k}^{-1}(A_{i_k}) \right)$$

**Beispiel 3.1** Ist  $I = \{1, \dots, n\}$  endlich, so ist (vgl. Stochastik I, §8):

$$\mathcal{A} = \bigotimes_{i=1}^n \mathcal{A}_i = \sigma(\{A_1 \times \cdots \times A_n \mid A_i \in \mathcal{A}_i, i \in \{1, \dots, n\}\})$$

Wir betrachten zunächst den Fall  $|I| = 2$ . Gegeben seien zwei Maßräume  $(\Omega_1, \mathcal{A}_1, \mu_1)$  und  $(\Omega_2, \mathcal{A}_2, \mu_2)$ . Weiter sei  $\Omega = \Omega_1 \times \Omega_2$ ,  $\mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2$ . Wir müssen nun ein Produktmaß konstruieren.

**Lemma 3.1** Für alle  $A \in \mathcal{A}$ ,  $\omega_1 \in \Omega_1$ ,  $\omega_2 \in \Omega_2$  gilt:

$$\begin{aligned} A_{\omega_1} &:= \{ \omega_2 \in \Omega_2 \mid (\omega_1, \omega_2) \in A \} \in \mathcal{A}_2 \text{ und} \\ A_{\omega_2} &:= \{ \omega_1 \in \Omega_1 \mid (\omega_1, \omega_2) \in A \} \in \mathcal{A}_1. \end{aligned}$$

$A_{\omega_i}$  heißt  $\omega_i$ -Schnitt von  $A$  für  $i = 1, 2$ .

- hier fehlt eine Skizze -

**Beweis** Sei  $\omega_1 \in \Omega_1$ . Dann ist  $\mathcal{A}' := \{A \in \mathcal{A} \mid A_{\omega_1} \in \mathcal{A}_2\} \subset \mathcal{A}$ , also die Menge der Mengen, für die das Lemma gilt, eine  $\sigma$ -Algebra, denn:

(i)

$$\Omega_{\omega_1} = \Omega_2 \in \mathcal{A}_2 \implies \Omega \in \mathcal{A}'$$

(ii)

$$\begin{aligned} (\Omega \setminus A)_{\omega_1} &= \{\omega_2 \mid (\omega_1, \omega_2) \notin A\} \\ &= \{\omega_2 \mid (\omega_1, \omega_2) \in A\}^C \\ &= \Omega_2 \setminus \underbrace{A_{\omega_1}}_{\in \mathcal{A}_2} \in \mathcal{A}_2 \end{aligned}$$

$$\implies (\Omega \setminus A)_{\omega_1} \in \mathcal{A}'.$$

(iii)

$$\left( \bigcup_{n=1}^{\infty} A_n \right)_{\omega_1} = \bigcup_{n=1}^{\infty} (A_n)_{\omega_1} \in \mathcal{A}_2 \implies \left( \bigcup_{n=1}^{\infty} A_n \right)_{\omega_1} \in \mathcal{A}'$$

$$\text{Wegen } (A_1 \times A_2)_{\omega_1} = \begin{cases} A_2 & , \omega_1 \in A_1 \\ \emptyset & , \omega_1 \notin A_1 \end{cases} \in \mathcal{A}_2 \text{ gilt:}$$

$$\sigma(\{A_1 \times A_2 \mid A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\}) \subset \mathcal{A}', \text{ also gilt } \mathcal{A} = \mathcal{A}'$$

mit der Voraussetzung von oben. Aus Symmetriegründen gilt die entsprechende Aussage auch für  $A_{\omega_2}$ ,  $\omega_2 \in \Omega_2$ . ■

**Lemma 3.2** Die Maße  $\mu_1, \mu_2$  seien  $\sigma$ -endlich. Dann gilt für alle  $A \in \mathcal{A}$ :

$$\begin{aligned} \omega_1 &\mapsto \mu_2(A_{\omega_1}) \text{ ist } (\mathcal{A}_1, \mathfrak{B}_{(-\infty, \infty]})\text{-messbar,} \\ \omega_2 &\mapsto \mu_1(A_{\omega_2}) \text{ ist } (\mathcal{A}_2, \mathfrak{B}_{(-\infty, \infty]})\text{-messbar.} \end{aligned}$$

**Beweis**  $\mu_2$   $\sigma$ -endlich  $\implies \exists (B_n)_{n \in \mathbb{N}} \subset \mathcal{A}_2$  mit  $B_n \uparrow \Omega_2$  und  $\mu_2(B_n) < \infty \quad \forall n \in \mathbb{N}$ .

Setze  $f_A(\omega_1) := \mu_2(A_{\omega_1})$ ,  $f_{A,n}(\omega_1) := \mu_2(A_{\omega_1} \cap B_n)$ . Sei  $\mathcal{D} := \{D \in \mathcal{A} \mid f_{D,n} \text{ ist } (\mathcal{A}_1, \mathfrak{B})\text{-messbar}\}$  für ein festes  $n$ . Dann gilt:

- (i)  $f_{\Omega,n} = \mu_2(\Omega_2 \cap B_n) = \mu_2(B_n)$
- (ii)  $f_{D^C,n} = \mu_2(B_n) - f_{D,n}$ , also  $D \in \mathcal{D} \implies D^C \in \mathcal{D}$
- (iii)  $f_{\sum_{i=1}^{\infty} D_i,n} = \sum_{i=1}^{\infty} f_{D_i,n}$ , also  $D_i \in \mathcal{D} \implies \sum_{i=1}^{\infty} D_i \in \mathcal{D}$

Damit ist  $\mathcal{D}$  ein Dynkin-System (vgl. Stochastik 1).

Wegen  $f_{A_1 \times A_2,n}(\omega_1) = \mu_2(A_2 \cap B_n) \cdot \mathbf{1}_{A_1}(\omega_1)$  ist  $f_{A_1 \times A_2,n}$  für  $A_1 \in \mathcal{A}_1$ ,  $A_2 \in \mathcal{A}_2$  messbar und daher  $A_1 \times A_2 \in \mathcal{D}$ .

$\mathcal{D}$  enthält also das durchschnittstabile Erzeugendensystem von  $\mathcal{A}$ .

$\xrightarrow{\text{St.1, S.4.3}} \mathcal{D} = \mathcal{A} \implies f_{A,n}$  ist  $(\mathcal{A}_1, \mathfrak{B})$ -messbar  $\forall A \in \mathcal{A}$ ,  $n \in \mathbb{N}$ .

Wegen  $f_A = \sup_{n \in \mathbb{N}} \{f_{A,n}\}$  folgt die Behauptung. ■

**Definition 3.1 und Satz:**

Sind  $\mu_1, \mu_2$   $\sigma$ -endlich, so existiert genau ein Maß  $\mu$  auf  $\mathcal{A}_1 \otimes \mathcal{A}_2$  mit  $\mu(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2) \forall A_1 \in \mathcal{A}_1, \forall A_2 \in \mathcal{A}_2$ .  $\mu$  heißt **Produktmaß** von  $\mu_1$  und  $\mu_2$ , Schreibweise:  $\mu = \mu_1 \otimes \mu_2$ . Für  $\mu$  gilt<sup>1</sup>:

$$\mu(A) = \int \mu_2(A_{\omega_1}) \mu_1(d\omega_1) = \int \mu_1(A_{\omega_2}) \mu_2(d\omega_2) \quad \forall A \in \mathcal{A}$$

Schließlich ist  $\mu$  auch  $\sigma$ -endlich.

**Beweis** Es seien wieder  $f_A(\omega_1) = \mu_2(A_{\omega_1})$ . Seien  $A_n \in \mathcal{A}, n \in \mathbb{N}, A_n$  paarweise disjunkt und  $\sum_{n=1}^{\infty} A_n = A$ . Es folgt:

$$\begin{aligned} \int f_A d\mu_1 & \stackrel{\text{stetig von unten}}{=} \int \lim_{n \rightarrow \infty} \left( f_{\sum_{i=1}^n A_i} \right) d\mu_1 \\ & \stackrel{\text{monotone Konvergenz}}{=} \lim_{n \rightarrow \infty} \left( \int f_{\sum_{i=1}^n A_i} d\mu_1 \right) \\ & = \lim_{n \rightarrow \infty} \left( \sum_{i=1}^n \left( \int f_{A_i} d\mu_1 \right) \right) \\ & = \sum_{i=1}^{\infty} \left( \int f_{A_i} d\mu_1 \right) \end{aligned}$$

Außerdem ist  $\int f_{\emptyset} d\mu_1 = \int 0 d\mu_1 = 0$ .

Also ist  $\Pi : \mathcal{A} \rightarrow [0, \infty]$ ,  $\Pi(A) := \int f_A d\mu_1$  ein Maß auf  $\mathcal{A}$ . Nach Konstruktion gilt:

$$\Pi(A_1 \times A_2) = \int \mu_2(A_2) \cdot \mathbf{1}_{A_1} d\mu_1 = \mu_2(A_2) \cdot \mu_1(A_1).$$

Analog ist  $\Pi'(A) := \int \mu_1(A_{\omega_2}) \cdot \mu_2(d\omega_2)$  ein Maß mit  $\Pi'(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2)$ , d.h.  $\Pi$  und  $\Pi'$  stimmen auf dem durchschnittstabilen Erzeuger  $\{A_1 \times A_2 | A_i \in \mathcal{A}_i\}$  überein. Der Eindeigkeitsatz für Maße (vgl. Übung) liefert  $\Pi = \Pi' =: \mu$  auf ganz  $\mathcal{A}$ .  $\sigma$ -Endlichkeit ist klar. ■

---

<sup>1</sup>Anmerkung:  $\int_{\Omega} f d\mu = \int f(\omega) \mu(d\omega)$

Wie integriert man bzgl.  $\mu_1 \otimes \mu_2$ ?

Ist  $f : \Omega \rightarrow \mathbb{R}$  eine Abbildung, so sei

$$\begin{aligned} f_{\omega_1} : \Omega_2 &\rightarrow \mathbb{R}, & f_{\omega_1}(\omega_2) &:= f(\omega_1, \omega_2), \\ f_{\omega_2} : \Omega_1 &\rightarrow \mathbb{R}, & f_{\omega_2}(\omega_1) &:= f(\omega_1, \omega_2). \end{aligned}$$

**Lemma 3.3** *Ist  $f(\mathcal{A}, \mathfrak{B})$ -messbar, so ist  $f_{\omega_1}(\mathcal{A}_2, \mathfrak{B})$ -messbar  $\forall \omega_1 \in \Omega_1$  und  $f_{\omega_2}$  ist  $(\mathcal{A}_1, \mathfrak{B})$ -messbar  $\forall \omega_2 \in \Omega_2$ .*

**Beweis**

$$\begin{aligned} f_{\omega_1}^{-1}(B) &= \{\omega_2 \in \Omega_2 \mid f(\omega_1, \omega_2) \in B\} \\ &= (\{\omega \in \Omega \mid f(\omega) \in B\})_{\omega_1} \\ &= \left( \underbrace{f^{-1}(B)}_{\in \mathcal{A}} \right)_{\omega_1} \in \mathcal{A}_2 \quad \forall B \in \mathfrak{B}. \quad \blacksquare \end{aligned}$$

**Satz 3.1 (Satz von Fubini, Teil I, auch: Satz von Tonelli)**

*Es seien  $\mu_1$  und  $\mu_2$   $\sigma$ -endlich sowie  $f : \Omega \rightarrow \mathbb{R}_+$   $(\mathcal{A}, \mathfrak{B})$ -messbar<sup>2</sup>. Dann ist*

$$\begin{aligned} \omega_1 &\mapsto \int f_{\omega_1} d\mu_2 \quad (\mathcal{A}_1, \mathfrak{B}_{(-\infty, \infty]})\text{-messbar und} \\ \omega_2 &\mapsto \int f_{\omega_2} d\mu_1 \quad (\mathcal{A}_2, \mathfrak{B}_{(-\infty, \infty]})\text{-messbar und es gilt:} \end{aligned}$$

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

**Beweis** mit algebraischer Induktion.

(1) Falls  $f = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}$  erhält man mit  $(\mathbf{1}_A)_{\omega_2}(\omega_1) = \mathbf{1}_{A_{\omega_2}}(\omega_1)$  die Beziehung

$$\int f_{\omega_2} d\mu_1 \stackrel{\text{lin.}}{=} \sum_{i=1}^n \alpha_i \int \mathbf{1}_{(A_i)_{\omega_2}} d\mu_1 = \sum_{i=1}^n \alpha_i \mu_1((A_i)_{\omega_2})$$

$$\stackrel{\text{L.3.2}}{\implies} \omega_2 \mapsto \int f_{\omega_2} d\mu_1 \text{ ist messbar.}$$

$$\implies \int \left( \int f_{\omega_2} d\mu_1 \right) \mu_2(d\omega_2) = \sum_{i=1}^n \alpha_i \int \mu_1((A_i)_{\omega_2}) \mu_2(d\omega_2)$$

$$\stackrel{\text{D.3.1}}{=} \sum_{i=1}^n \alpha_i \cdot \mu_1 \otimes \mu_2(A_i) = \int f d(\mu_1 \otimes \mu_2).$$

<sup>2</sup>Dass hier  $f \geq 0$  gilt, ist wesentlich für Fubini I; den allgemeinen Fall behandelt Fubini II.

(2)  $f \geq 0$ ,  $f$   $(\mathcal{A}, \mathfrak{B})$ -messbar.

$\implies \exists (u_n)_{n \in \mathbb{N}} \subset \mathcal{E}$  mit  $u_n \uparrow f$  und  $\int f d\mu = \lim_{n \rightarrow \infty} (\int u_n d\mu)$ .

Wegen  $(u_n)_{\omega_2} \uparrow f_{\omega_2}$  und  $g_n(\omega_2) := \int (u_n)_{\omega_2} d\mu_1 \uparrow \int f_{\omega_2} d\mu_1 \forall \omega_2 \in \Omega_2$  ist nach Schritt 1  $\int g_n(\omega_2) \mu_2(d\omega_2) = \int u_n d(\mu_1 \otimes \mu_2)$ . Mit dem Satz von der monotonen Konvergenz folgt:

$$\begin{aligned} \int \left( \int f_{\omega_2} d\mu_1 \right) \mu_2(d\omega_2) &= \lim_{n \rightarrow \infty} \left( \int g_n d\mu_2 \right) \\ &= \lim_{n \rightarrow \infty} \left( \int u_n d(\mu_1 \otimes \mu_2) \right) \\ &= \int f d(\mu_1 \otimes \mu_2). \end{aligned}$$

Wiederhole die Schritte mit  $\omega_2$  statt mit  $\omega_1$  und erhalte den Rest der Behauptung. ■

Bevor wir den Satz von Fubini für allgemeine  $f$  beweisen, benötigen wir folgende Überlegung:

**Bemerkung 3.1** Ist  $(\Omega, \mathcal{A}, \mu)$  Maßraum,  $A \in \mathcal{A}$  mit  $\mu(A^C) = 0$ ,  $f : A \rightarrow \mathbb{R}$ , so nennen wir  $f$   $(\mathcal{A}, \mathfrak{B})$ -messbar,  $\mu$ -integrierbar, etc., wenn dies auf die folgende Fortsetzung  $\bar{f}$  von  $f$  zutrifft:

$$\bar{f} : \Omega \rightarrow \mathbb{R}, \quad \bar{f}(\omega) := \begin{cases} f(\omega) & \omega \in A \\ 0 & \text{sonst} \end{cases} \quad \text{und schreiben dann } \int f d\mu \text{ statt } \int \bar{f} d\mu.$$

**Satz 3.2 (Satz von Fubini, Teil II)**

Es seien  $\mu_1$  und  $\mu_2$   $\sigma$ -endlich,  $f : \Omega \rightarrow \mathbb{R}$   $(\mu_1 \otimes \mu_2)$ -integrierbar.

Dann sind  $\mu_1$ -fast alle  $f_{\omega_1}$   $\mu_2$ -integrierbar und  $\mu_2$ -fast alle  $f_{\omega_2}$   $\mu_1$ -integrierbar.

Weiter sind die Integrale

$$\omega_1 \mapsto \int f_{\omega_1} d\mu_2$$

und

$$\omega_2 \mapsto \int f_{\omega_2} d\mu_1$$

als Funktionen von  $\omega_1$  bzw.  $\omega_2$  im obigen Sinne  $\mu_1$ - bzw.  $\mu_2$ -integrierbar und es gilt:

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

**Beweis**

Es gilt  $|f|_{\omega_1} = |f_{\omega_1}|$ ,  $f_{\omega_1}^+ = (f_{\omega_1})^+$  und  $f_{\omega_1}^- = (f_{\omega_1})^-$ .

Also folgt aus Satz 3.1.:

$$\begin{aligned} \int |f| d\mu &= \int \left( \int |f_{\omega_1}| d\mu_2 \right) \mu_1(d\omega_1) < \infty \quad (\text{das ist die Voraussetzung}) \\ \implies \mu_1 \left( \left\{ \omega_1 \mid \int |f_{\omega_1}| d\mu_2 = \infty \right\} \right) &= 0 \\ \implies f_{\omega_1} &\text{ ist } \mu_1\text{-f.ü. } \mu_2\text{-integrierbar.} \end{aligned}$$

Satz 3.1. angewandt auf  $f_{\omega_1}^+$  und  $f_{\omega_1}^-$  ergibt, dass

$$\omega_1 \mapsto \int f_{\omega_1} d\omega_2 = \left( \int f_{\omega_1}^+ d\mu_2 - \int f_{\omega_1}^- d\mu_2 \right)$$

$(\mathcal{A}, \mathfrak{B})$ -messbar ist (auf einer  $\mu_1$ -Nullmenge könnte “ $\infty - \infty$ ” stehen und die Funktion wäre dort nicht definiert, siehe hierzu aber die vorstehende Bemerkung) und

$$\begin{aligned} \int \left( \int f_{\omega_1} d\mu_2 \right) \mu_1(d\omega_1) &= \int \left( \int f_{\omega_1}^+ d\mu_2 - \int f_{\omega_1}^- d\mu_2 \right) \mu_1(d\omega_1) \\ &= \int f^+ d\mu - \int f^- d\mu \\ &= \int f d\mu. \end{aligned}$$

Der Rest folgt mit dem Symmetrieargument. ■

### Bemerkung 3.2

a) Der Satz von Fubini läßt sich wie folgt schreiben:

$$\begin{aligned} \int f d(\mu_1 \otimes \mu_2) &= \int \int f(\omega_1, \omega_2) \mu_1(d\omega_1) \mu_2(d\omega_2) \\ &= \int \int f(\omega_1, \omega_2) \mu_2(d\omega_2) \mu_1(d\omega_1) \end{aligned}$$

Die Integrationsreihenfolge spielt also keine Rolle.

b) Sind messbare Räume  $(\Omega_i, \mathcal{A}_i)$  ( $i \in I$ ) gegeben mit  $|I|$  endlich und  $|I| > 2$ , so erhält man ein Maß  $\mu := \bigotimes_{i \in I} \mu_i$  auf der Produkt- $\sigma$ -Algebra durch schrittweises Ausführen von Produkten mit 2 Faktoren. Insbesondere gilt auf Rechteckmengen  $A_1 \times \cdots \times A_n$  mit  $A_i \in \mathcal{A}_i$  ( $i = 1, \dots, n$ ):

$$\mu(A_1 \times \cdots \times A_n) = \prod_{i=1}^n \mu_i(A_i).$$

Da die Rechteckmengen ein durchschnittstabiler Erzeuger von  $\mathcal{A}$  sind, folgt wegen der Eindeutigkeit von  $\mu$ :

$$(\mu_1 \otimes \mu_2) \otimes \mu_3 = \mu_1 \otimes (\mu_2 \otimes \mu_3) \quad (\text{Assoziativität des Maßprodukts})$$

### Satz 3.3

Auf  $(\Omega, \mathcal{A})$  existiert genau ein Wahrscheinlichkeitsmaß  $P := \bigotimes_{i \in I} P_i$  mit

$$P^{\Pi_J} = \bigotimes_{i \in J} P_i \quad \forall J \subset I, J \text{ endlich.}$$

**Beweis** Siehe z.B. Bauer, Henze, Stochastik II S.8.13. ■

$$\begin{array}{ccc}
\mu & & \mu^T \\
(\Omega, \mathcal{A}) & \xrightarrow{T} & (\Omega', \mathcal{A}') \\
P & & P^{\Pi_J} \\
(\Omega, \mathcal{A}) & \xrightarrow{\Pi_J} & (\times_{i \in J} \Omega_i, \otimes_{i \in J} \mathcal{A}_i)
\end{array}$$

z.B.  $P((\times_{i \in J} \mathcal{A}_i) \times (\times_{j \notin J} \Omega_j)) = \prod_{i \in J} P_i(\mathcal{A}_i)$ ,  $A = \times_{i \in J} A_i$

**Definition** Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum und  $(\Omega'_i, \mathcal{A}'_i)$  ein messbarer Raum  $\forall i \in I$ .  $X_i : \Omega \rightarrow \Omega'_i$  seien Zufallsgrößen. Die Familie  $(X_i)_{i \in I}$  heißt **stochastisch unabhängig** genau dann, wenn  $\forall J \subset I, J$  endlich und  $\forall A'_j \in \mathcal{A}'_j, j \in J$

$$\frac{P(\cap_{j \in J} \{X_j \in A'_j\})}{P^X(\times_{j \in J} A'_j \times \times_{i \notin J} \Omega_i)} = \prod_{j \in J} \frac{P(X_j \in A'_j)}{P^{X_j}(A'_j)}$$

**Bemerkung** Bei der Überprüfung der Bedingung kann man sich auf  $A_j \in \mathcal{E}_j$  beschränken, wobei  $\mathcal{E}_j$  ein durchschnittsstabiler Erzeuger von  $\mathcal{A}_j$  ist.

In der Situation der vorigen Definition gilt für  $\Omega' := \times_{i \in I} \Omega_i, \mathcal{A}' := \otimes_{i \in I} \mathcal{A}_i$ :

$$X : \Omega \rightarrow \Omega', (X(\omega))(i) := X_i(\omega), \forall i \in I, \omega \in \Omega$$

ist  $(\mathcal{A}, \mathcal{A}')$ -messbar (vgl. Ü 2.1), d.h.  $X$  transportiert  $P$  zu einem Wahrscheinlichkeitsmaß  $P^X$  auf  $(\Omega', \mathcal{A}')$ .  $P^X$  nennt man auch **gemeinsame Verteilung** der Zufallsgrößen  $X_i, i \in I$ .

### Satz 3.4

Die Familie  $X = (X_i)_{i \in I}$  ist genau dann unabhängig, wenn

$$P^X = \otimes_{i \in I} P^{X_i}$$

**Beweis** Folgt aus der Definition und S.3.3. ■

### Bemerkung

- (i) Unabhängigkeit der  $(X_i)_{i \in I}$  ist äquivalent dazu, dass jede endliche Teilfamilie  $(X_i)_{i \in J}, J \subset I, (J \text{ endlich})$ , unabhängig ist.
- (ii) Sei  $\Omega'_i = \mathbb{R}, X = (X_1, \dots, X_d)$  ein Zufallsvektor und  $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ .  $F_X(x_1, \dots, x_d) = P^X((-\infty, x_1] \times \dots \times (-\infty, x_d]) = P(X_1 \leq x_1, \dots, X_d \leq x_d)$  ist die gemeinsame Verteilungsfunktion. Da  $\mathcal{E} = \{(-\infty, x] : x \in \mathbb{R}^d\}$  durchschnittsstabiler Erzeuger von  $\mathfrak{B}^d$  ist, sind  
 $X_1, \dots, X_d$  unabhängig  $\iff F_X(x_1, \dots, x_d) = F_{X_1}(x_1) \cdots F_{X_d}(x_d) \forall x \in \mathbb{R}^d$ .  
 Falls Dichten existieren:  
 $X_1, \dots, X_d$  unabhängig  $\iff f_X(x_1, \dots, x_d) = f_{X_1}(x_1) \cdots f_{X_d}(x_d) \forall x \in \mathbb{R}^d$

- (iii) Als Wahrscheinlichkeitsraum für das Experiment “ $\infty$ -oft Münze werfen” kann man z.B.  $\Omega = \{0, 1\}^{\mathbb{N}}$ ,  $\mathcal{A} = \otimes_{i \in \mathbb{N}} \mathcal{P}(\{0, 1\})$ ,  $P = \otimes_{i \in \mathbb{N}} (\frac{1}{2}(\delta_0 + \delta_1))$  wählen. S.3.3 impliziert, dass es zu jedem vorgegebenen Wahrscheinlichkeitsmaß eine Folge von unabhängigen und identisch verteilten Zufallsvektoren gibt. Man kann beim Münzexperiment auch  $([0, 1], \mathfrak{B}_{[0,1]}, \lambda_{[0,1]})$ ,  $X_n(\omega) = \lfloor 2^n \cdot \omega \rfloor \bmod 2$  wählen. (vgl. Bsp 13.2 St I)

## 3.2 Reellwertige Abbildungen, Rechnen mit Verteilungen

Wir betrachten den Spezialfall  $(\Omega_i, \mathcal{A}_i, \mu_i) = (\mathbb{R}, \mathfrak{B}, \lambda)$  für  $i = 1, \dots, d$ . Hier folgt:  $\Omega = \mathbb{R}^d$ ,  $\mathcal{A} = \otimes_{i=1}^d \mathcal{A}_i = \sigma(\{(a_1, b_1] \times \dots \times (a_d, b_d] : a_i \leq b_i, a_i, b_i \in \mathbb{R}, i = 1, \dots, d\}) = \mathfrak{B}^d$ .

$P = \lambda^d, \lambda^d((a_1, b_1] \times \dots \times (a_d, b_d]) = \prod_{i=1}^d (b_i - a_i) \hat{=} \text{Volumen}$ . Was passiert, wenn  $(a, b]$  mit einer Abbildung  $\Psi$  transformiert wird?

### Satz 3.5 (Transformationssatz für das $d$ -dimensionale Lebesgue-Maß)

Es seien  $U, V \subset \mathbb{R}^d$  offen und  $\Psi : U \rightarrow V$  eine bijektive, stetig differenzierbare Abbildung. Gilt dann  $\det(\Psi'(x)) \neq 0 \forall x \in U$ , so hat das Bildmaß der Einschränkung von  $\lambda^d$  auf  $U$  unter  $\Psi$  bzgl. der Einschränkung von  $\lambda^d$  auf  $V$  die Dichte

$$\frac{d(\lambda_U^d)^\Psi}{d\lambda_V^d}(y) = \frac{1}{|\det \Psi'(\Psi^{-1}(y))|} \quad \forall y \in V.$$

**Beweis** Henze, Stochastik II. ■

### Bemerkung

- (a) Unter den Voraussetzungen von S.3.5 ist auch  $\Psi^{-1}$  stetig differenzierbar und die Kettenregel liefert:

$$\det(\Psi'(\Psi^{-1}(y))) \cdot \det((\Psi^{-1})'(y)) = 1.$$

Es gilt also

$$\frac{d(\lambda_U^d)^\Psi}{d\lambda_V^d}(y) = |\det(\Psi^{-1})'(y)| \quad \forall y \in V.$$

- (b) Mit S.2.4 gilt:

$$\int_U f(\Psi(x)) dx \stackrel{\text{S.2.4}}{=} \int_V f(y) d(\lambda_U^d)^\Psi = \int_V f(y) |\det(\Psi^{-1})'(y)| dy$$

bzw.

$$\int_U g(x) dx = \int_V g(\Psi^{-1}(y)) |\det(\Psi^{-1})'(y)| dy$$



**Beispiel 3.2** Transformation auf Polarkoordinaten

Hier:  $d=2$ .  $U = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 > 0 \text{ oder } x_2 \neq 0\}$ ,  $V = (0, \infty) \times (-\pi, \pi)$ ,  $\Psi : U \rightarrow V, (x_1, x_2) \xrightarrow{\Psi} (r, \Phi)$  bijektiv.  $(\Psi^{-1})_1(r, \Phi) = r \cos \Phi$ ,  $(\Psi^{-1})_2(r, \Phi) = r \sin \Phi$ .

$$\implies (\Psi^{-1})'(r, \Phi) = \begin{pmatrix} \cos \Phi & -r \sin \Phi \\ \sin \Phi & r \cos \Phi \end{pmatrix}$$

$$\implies \frac{d(\lambda_U^d)^\Psi}{d(\lambda_V^d)} = r \cos^2 \Phi + r \sin^2 \Phi = r \quad \forall (r, \Phi) \in V.$$

Wir bekommen:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x_1, x_2) dx_1 dx_2 = \int_{-\pi}^{\pi} \int_0^{\infty} r \cdot g(r \cos \Phi, r \sin \Phi) dr d\Phi$$

Im Folgenden sei  $X = (X_1, \dots, X_d) : \Omega \rightarrow \mathbb{R}^d$  ein Zufallsvektor.

**Satz 3.6 (Transformationssatz für Wahrscheinlichkeitsdichten)**

Es seien  $U$  und  $V$  offene Teilmengen von  $\mathbb{R}^d$  und  $\Psi : U \rightarrow V$  eine bijektive, stetige und differenzierbare Abbildung mit der Eigenschaft

$$\det \Psi'(x) \neq 0 \quad \forall x \in U.$$

Ist dann  $X$  ein Zufallsvektor auf  $(\Omega, \mathcal{A}, P)$  mit  $P(X \in U) = 1$  und Dichte  $f_X$ , so ist auch  $Y := \Psi(X)$  absolutstetig und eine Dichte  $f_Y$  von  $Y$  auf  $V$  ist gegeben durch

$$f_Y(y) = |\det(\Psi^{-1})'(y)| f_X(\Psi^{-1}(y)) \quad \forall y \in V$$

**Beweis** Seien  $A \subset V, A \in \mathfrak{B}^d$ . Mit Satz 3.5 folgt:

$$\begin{aligned} P(Y \in A) &= P(X \in \Psi^{-1}(A)) \\ &= \int_U \mathbf{1}_{\Psi^{-1}(A)}(x) f_X(x) dx \\ &= \int_V \mathbf{1}_{\Psi^{-1}(A)}(\Psi^{-1}(y)) f_X(\Psi^{-1}(y)) \cdot |\det(\Psi^{-1})'(y)| dy \\ &= \int_A |\det(\Psi^{-1})'(y)| f_X(\Psi^{-1}(y)) dy \end{aligned}$$

■

**Beispiel 3.3** (Box-Muller-Algorithmus zur Erzeugung von  $N(0, 1)$ -verteilten Zufallsvariablen)

Seien  $U_1, U_2 \sim U(0, 1)$  und unabhängig. Definiere:

$$X_1 := \sqrt{-2 \log(U_1)} \cos(2\pi U_2) = \Psi_1(U_1, U_2)$$

$$X_2 := \sqrt{-2 \log(U_1)} \sin(2\pi U_2) = \Psi_2(U_1, U_2)$$

Dann sind  $X_1, X_2 \sim N(0, 1)$  und unabhängig. Beweis mit Satz 3.6. Sei  $U = (0, 1)^2$

$$\begin{aligned}
 V &= \{(X_1, X_2) \in \mathbb{R}^2 \mid X_1 < 0 \text{ oder } X_2 \neq 0\} \\
 \Psi'(u) &= \begin{pmatrix} -(-2 \log(u_1))^{-\frac{1}{2}} \frac{\cos(2\pi u_2)}{u_1} & -(-2 \log u_1)^{\frac{1}{2}} 2\pi \sin(2\pi u_2) \\ -(-2 \log(u_1))^{-\frac{1}{2}} \frac{\sin(2\pi u_2)}{u_1} & (-2 \log u_1)^{\frac{1}{2}} 2\pi \cos(2\pi u_2) \end{pmatrix} \\
 \Rightarrow \det \Psi' &= -\frac{2\pi}{u_1} \text{ und} \\
 u_1 &= e^{-\frac{1}{2}(x_1^2 + x_2^2)} \\
 \Rightarrow f_X(x) &= \frac{1}{|\det \Psi'(\Psi^{-1}(x))|} \cdot 1 \\
 &= \frac{1}{2\pi} e^{-\frac{1}{2}(x_1^2 + x_2^2)} \\
 &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_1^2} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_2^2} \\
 \Rightarrow &\text{ Behauptung}
 \end{aligned}$$

### Satz 3.7

Sind  $X$  und  $Y$  unabhängige Zufallsvariablen mit Dichten  $f_X$  und  $f_Y$ , so ist auch die Zufallsvariable  $Z := X + Y$  absolutstetig und eine zugehörige Dichte ist gegeben durch:

$$f_Z(z) = \int f_X(x) \cdot f_Y(z - x) dx \quad \text{„Faltung“}$$

**Beweis** Verwende Satz 3.6 mit  $\Psi : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \Psi(x, y) = (x, x + y)$  ( $\Psi^{-1}(x, z) = (x, z - x)$ )

$$\Rightarrow f_{X,Z}(x, z) = f_{X,Y}(x, z - x) = f_X(x) \cdot f_Y(z - x)$$

Die „Randdichte“  $f_Z$  bekommt man durch Integration über  $x$ . ■

**Beispiel 3.4** a) Sind die Zufallsvariablen  $X_1, \dots, X_d$  unabhängig und  $X_i \sim \exp(\lambda), i = 1, \dots, d, \lambda > 0$ , so hat  $X_1 + \dots + X_d$  die Dichte

$$f_{X_1 + \dots + X_d}(z) = \frac{\lambda^d}{(d-1)!} z^{d-1} e^{-\lambda z} \mathbf{1}_{[0, \infty)}(z)$$

( $\rightarrow$  Gamma-Verteilung bzw. Erlang-Verteilung)

b) Sind  $X_1, \dots, X_d$  unabhängig und  $X_i \sim N(\mu_i, \sigma_i^2), a_i \in \mathbb{R}, i = 1, \dots, d$  so gilt falls  $\sum a_i^2 \neq 0$

$$\sum_{i=1}^d a_i X_i \sim N\left(\sum_{i=1}^d a_i \mu_i, \sum_{i=1}^d a_i^2 \sigma_i^2\right)$$

**Beispiel 3.5** (Gemeinsame Verteilung der Ordnungsstatistiken)

Es seien  $X_1, \dots, X_d$  unabhängige und identisch verteilte Zufallsvariablen mit Dichte  $f$ . Weiter sei  $(X_{1:d}, \dots, X_{d:d})$  eine Permutation von  $X_1, \dots, X_d$ , so dass

$$X_{1:d} < \dots < X_{d:d}$$

$X_{r:d}$  heißt  **$r$ -te Ordnungsstatistik** von  $X$ .

Sei  $S_d$  die Menge der Permutationen der Zahlen  $1, \dots, d$ . Dann gilt für  $\pi \in S_d$ :

$$(X_{1:d}, \dots, X_{d:d}) = (X_{\pi(1)}, \dots, X_{\pi(d)}), \text{ falls } X_{\pi(1)} < \dots < X_{\pi(d)}$$

Für jede messbare Funktion  $g: \mathbb{R}^d \rightarrow \mathbb{R}$  gilt:

$$g(X_{1:d}, \dots, X_{d:d}) = \sum_{\pi \in S_d} g(X_{\pi(1)}, \dots, X_{\pi(d)}) \cdot \mathbf{1}_{[X_{\pi(1)} < \dots < X_{\pi(d)}]}$$

Es gilt:

$$f_{X_{\pi(1)}, \dots, X_{\pi(d)}}(x_1, \dots, x_d) = \prod_{i=1}^d f(x_i) = f_X(x)$$

Also folgt:

$$\begin{aligned} Eg(X_{1:d}, \dots, X_{d:d}) &= \sum_{\pi \in S_d} \int_{x_1 < \dots < x_d} g(x) \prod_{i=1}^d f(x_i) dx_1 \dots dx_d \\ &= d! \int_{\mathbb{R}^d} g(x) \prod_{i=1}^d f(x_i) \mathbf{1}_{[x_1 < \dots < x_d]}(x) dx_1 \dots dx_d \end{aligned}$$

Sei  $g(x) = \mathbf{1}_B(x)$  mit  $B \in \mathfrak{B}^d$ , dann folgt:

$$f_{X_{1:d}, \dots, X_{d:d}}(x_1, \dots, x_d) = d! \prod_{i=1}^d f(x_i) \mathbf{1}_{[x_1 < \dots < x_d]}(x)$$

Konkrete Anwendung:

Gegeben 12 Trinkgläser. Lebensdauer unabhängig  $\exp(\lambda)$ -verteilt. Nach der vorigen Überlegung gilt

$$\begin{aligned} f_{(X_{1:d}, \dots, X_{d:d})}(x) &= \begin{cases} d! \lambda^d e^{-\lambda(x_1 + \dots + x_d)} & , \text{ falls } x_1 < \dots < x_d \\ 0 & , \text{ sonst} \end{cases} \\ \Rightarrow f_{(X_{1:d}, X_{2:d})}(x) &= \begin{cases} d(d-1) \lambda^2 e^{-(d-2)\lambda x_2} e^{-\lambda(x_1 + x_2)} & , \text{ falls } x_1 < x_2 \\ 0 & , \text{ sonst} \end{cases} \\ \xRightarrow{\text{Satz 3.6}} f_{(X_{2:d} - X_{1:d}, X_{1:d})}(y_1, y_2) &= \begin{cases} d(d-1) \lambda^2 e^{-d\lambda y_2} e^{-(d-1)\lambda y_1} & , \text{ falls } y_1, y_2 > 0 \\ 0 & , \text{ sonst} \end{cases} \\ \Rightarrow f_{X_{2:d} - X_{1:d}}(y_1) &= \begin{cases} (d-1) \lambda e^{-(d-1)\lambda y_1} & , \text{ falls } y_1 > 0 \\ 0 & , \text{ sonst} \end{cases} \\ f_{X_{1:d}}(y_2) &= \begin{cases} d \lambda e^{-d\lambda y_2} & , \text{ falls } y_2 > 0 \\ 0 & , \text{ sonst} \end{cases} \end{aligned}$$

also  $X_{1:d} \sim \exp(\lambda d)$ ,  $X_{2:d} - X_{1:d} \sim \exp(\lambda(d-1))$  und unabhängig.

$$\Rightarrow X_{k:d} - X_{(k-1):d} \sim \exp((d-k+1)\lambda)$$

Es folgt:

$$\begin{aligned}
 E[X_{k:d} - X_{(k-1):d}] &= \frac{1}{(d-k+1)\lambda} \\
 \Rightarrow \frac{E[X_{d:d} - X_{(d-1):d}]}{EX_{d:d}} &= \frac{\frac{1}{\lambda}}{\sum_{k=1}^d \frac{1}{(d-k+1)\lambda}} \\
 &= \left( \sum_{k=1}^d \frac{1}{k} \right)^{-1} \\
 &= (\log d)^{-1} + O(1)
 \end{aligned}$$

Für  $d = 12$  : 0.32

## 4 Das starke Gesetz der großen Zahlen

**Satz 4.1 (Borel-Cantelli Lemma)** Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum und  $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$  eine Folge von Ereignissen.

$$\limsup_{n \rightarrow \infty} A_n := \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

ist das Ereignis, dass unendlich viele der  $A_n$ 's eintreten.

a) Dann gilt:

$$\sum_{n=1}^{\infty} P(A_n) < \infty \implies P(\limsup_{n \rightarrow \infty} A_n) = 0.$$

b) Sind die Ereignisse  $A_n$ ,  $n \in \mathbb{N}$  stochastisch unabhängig, so gilt:

$$\sum_{n=1}^{\infty} P(A_n) = \infty \implies P(\limsup_{n \rightarrow \infty} A_n) = 1.$$

### Beweis

a) Sei  $B_n := \bigcup_{k=n}^{\infty} A_k$ ,  $n \in \mathbb{N} \Rightarrow P(B_n) \leq \sum_{k=n}^{\infty} P(A_k) \xrightarrow{n \rightarrow \infty} 0$ .  
Da  $B_n \downarrow \bigcap_{n=1}^{\infty} B_n$  folgt:

$$P(\limsup_{n \rightarrow \infty} A_n) = P\left(\bigcap_{n=1}^{\infty} B_n\right) = \lim_{n \rightarrow \infty} P(B_n) = 0.$$

b) Sei  $P_n := P(A_n)$ ,  $n \in \mathbb{N}$ .  $(A_n)$  stoch. unabh.  $\Rightarrow (A_n^c)$  stoch. unabh. Es gilt:

$$\begin{aligned} 0 \leq P\left(\bigcap_{n=1}^{\infty} A_n^c\right) &\stackrel{\text{stetig von oben}}{=} \lim_{N \rightarrow \infty} P\left(\bigcap_{n=1}^N A_n^c\right) \\ &\stackrel{\text{unabh.}}{=} \lim_{N \rightarrow \infty} \prod_{k=1}^N (1 - P_k) \\ &\leq \lim_{N \rightarrow \infty} \exp\left(-\sum_{k=1}^N P_k\right) \stackrel{\text{nach Vor.}}{=} 0 \end{aligned}$$

Somit:

$$0 \leq P((\limsup_{n \rightarrow \infty} A_n)^c) = P\left(\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^c\right) \leq \sum_{n=1}^{\infty} P\left(\bigcap_{k=n}^{\infty} A_k^c\right) = 0$$

■

**Definition** Es seien  $X, X_1, X_2, \dots$  ZV auf einem W-Raum  $(\Omega, \mathcal{A}, P)$ .

$X_n$  konvergiert  $P$ -fast sicher gegen  $X$ ,  $(X_n \xrightarrow{f.s.} X)$  wenn gilt:

$$P\left(\left\{\omega \in \Omega \mid \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\right\}\right) = 1.$$

**Bemerkung**  $\{\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\} \in \mathcal{A}$ , denn:

(i)  $\sup_{n \geq 1} X_n$  ist  $\mathcal{A}$ -messbar, da  $\{\sup_{n \geq 1} X_n \leq a\} = \bigcap_{n=1}^{\infty} \underbrace{\{X_n \leq a\}}_{\in \mathcal{A}} \in \mathcal{A}$ .

$\inf_{n \geq 1} X_n = -\sup_{n \geq 1} (-X_n)$  ist  $\mathcal{A}$ -mb.  $\Rightarrow \limsup_{n \rightarrow \infty} X_n = \inf_{n \geq 1} \sup_{k \geq n} X_k, \liminf_{n \rightarrow \infty} X_n$   $\mathcal{A}$ -messbar.

(ii)  $\{\lim_{n \rightarrow \infty} X_n = X\} = (\liminf_{n \rightarrow \infty} (X_n - X))^{-1}(\{0\}) \cap (\limsup_{n \rightarrow \infty} (X_n - X))^{-1}(\{0\}) \in \mathcal{A}$

Im Folgenden sei  $(X_n)_{n \in \mathbb{N}}$  eine Folge von ZV auf einem W-Raum  $(\Omega, \mathcal{A}, P)$ . Starke Gesetz der großen Zahlen sind Resultate der Form

$$\frac{1}{a_n} \left( \sum_{i=1}^n X_i - b_n \right) \xrightarrow{f.s.} 0$$

wobei  $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}} \subset \mathbb{R}$ . Der wichtigste Satz ist hier:

**Satz 4.2 (Starkes Gesetz der großen Zahlen)** Ist  $(X_n)_{n \in \mathbb{N}}$  eine Folge von u.i.v. ZV mit  $E|X_1| < \infty$ , so gilt:

$$\frac{1}{n} \underbrace{\sum_{i=1}^n X_i}_{=s_n} \xrightarrow{f.s.} EX_1.$$

**Beweis** Sei zunächst  $X_k \geq 0 \forall k \in \mathbb{N}$  und  $Y_k := X_k \cdot \mathbf{1}_{[X_k \leq k]}$  ( $Y_k$  entsteht aus  $X_k$  durch Abschneiden bei  $k$ ). Sei  $S_n^* := \sum_{k=1}^n Y_k$   $EY_k = E[X_k \cdot \mathbf{1}_{[X_k \leq k]}] = E[X_1 \cdot \mathbf{1}_{[X_1 \leq k]}] \xrightarrow{k \rightarrow \infty} EX_1$  mit S.2.1 (Monotone Konvergenz).  
Aus der Analysis: Sei  $(a_n)_{n \in \mathbb{N}} \subset \mathbb{R}$

$$\lim_{n \rightarrow \infty} a_n = a \Rightarrow \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n a_k = a.$$

Damit folgt:

$$\lim_{n \rightarrow \infty} \frac{1}{n} ES_n^* = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n EY_k = EX_1.$$

Die  $Y_n$ 's sind wieder unabhängig und es gilt:

$$\text{Var}(S_n^*) = \sum_{k=1}^n \text{Var}(Y_k) \leq \sum_{k=1}^n EY_k^2 \leq \sum_{k=1}^n E[X_k^2 \cdot \mathbf{1}_{[X_k \leq n]}] = n \cdot E[X_1^2 \cdot \mathbf{1}_{[0, n]}(X_1)] \quad (*)$$

Sei  $\alpha > 1$  und  $m_n := \lfloor \alpha^n \rfloor \forall n \in \mathbb{N}$ . Für  $x > 0$  sei  $\Psi(x) := \sum_{n=N(x)}^{\infty} \frac{1}{m_n}$  mit  $N(x) := \min\{n \mid m_n \geq x\}$

Für beliebige  $z \geq 1$  gilt:  $\lfloor z \rfloor \geq \frac{z}{2}$  und somit  $\frac{1}{m_n} = \frac{1}{\lfloor \alpha^n \rfloor} \leq \frac{2}{\alpha^n}$  und  $\alpha^{N(x)} \geq \lfloor \alpha^{N(x)} \rfloor = m_{N(x)} \geq x$ . Mit  $k := \frac{2\alpha}{\alpha-1}$  gilt:

$$\Psi(x) = \sum_{n=N(x)}^{\infty} \frac{1}{m_n} \leq 2 \cdot \sum_{n=N(x)}^{\infty} \frac{1}{\alpha^n} = 2 \cdot \alpha^{-N(x)} \cdot \frac{1}{1 - \frac{1}{\alpha}} \leq \frac{k}{x} \quad (**)$$

Die Ungleichung von Tschebyscheff liefert  $\forall \varepsilon > 0$ :

$$\begin{aligned} \sum_{n=1}^{\infty} P\left(\frac{1}{m_n} |S_{m_n}^* - ES_{m_n}^*| > \varepsilon\right) &\stackrel{(*)}{\leq} \sum_{n=1}^{\infty} \frac{1}{\varepsilon^2 m_n} E[X_1^2 \cdot \mathbf{1}_{[0, m_n]}(X_1)] \\ &\stackrel{\text{S.2.1}}{=} \frac{1}{\varepsilon^2} E[X_1^2 \sum_{n=1}^{\infty} \frac{1}{m_n} \cdot \mathbf{1}_{[0, m_n]}(X_1)] \\ &= \frac{1}{\varepsilon^2} E[X_1^2 \Psi(X_1)] \stackrel{(**)}{\leq} \frac{k}{\varepsilon^2} EX_1 \end{aligned}$$

$$\stackrel{\text{Üb}}{\implies} \frac{1}{m_n} (S_{m_n}^* - ES_{m_n}^*) \xrightarrow{f.s.} 0 \stackrel{\text{Üb}}{\implies} \frac{1}{m_n} S_{m_n}^* \xrightarrow{f.s.} EX_1.$$

Nächstes Ziel: \* weg bekommen.

Es gilt:

$$\begin{aligned} \sum_{n=1}^{\infty} P(X_n \neq Y_n) &= \sum_{n=1}^{\infty} P(X_1 > n) \\ &\leq \int_{[0, \infty]} P(X_1 > x) \mathbf{1}(x) \stackrel{\text{Bsp 3.1}}{=} EX_1 < \infty. \end{aligned}$$

$$\stackrel{\text{S.4.1a)}}{\implies} P(\underbrace{\{\omega \in \Omega \mid X_n(\omega) \neq Y_n(\omega) \text{ für unendlich viele } n\}}_{=: N_0}) = 0$$

$\forall \omega \notin N_0 \exists k(\omega) \in \mathbb{N}$  mit  $X_n(\omega) = Y_n(\omega) \forall n \geq k(\omega)$ .

Auf  $N_0^C$  gilt also:

$$\begin{aligned} \frac{1}{n} (S_n(\omega) - S_n^*(\omega)) &= \frac{1}{n} \left( \sum_{i=1}^{k(\omega)} X_i(\omega) - Y_i(\omega) \right) \xrightarrow{n \rightarrow \infty} 0 \\ \implies \frac{1}{n} (S_n - S_n^*) &\xrightarrow{f.s.} 0 \implies \frac{1}{m_n} S_{m_n} \xrightarrow{f.s.} EX_1 \quad (\Delta) \end{aligned}$$

Jetzt muss die Einschränkung auf die Teilfolge  $(m_n)_{n \in \mathbb{N}}$  weg.

Da  $S_n \geq 0$ , gilt für  $m_n \leq k \leq m_{n+1}$ :

$$\frac{m_n}{m_{n+1}} \cdot \frac{S_{m_n}}{m_n} \leq \frac{S_k}{k} \leq \frac{m_{n+1}}{m_n} \cdot \frac{S_{m_{n+1}}}{m_{n+1}}$$

Da  $\frac{m_{n+1}}{m_n} \xrightarrow{n \rightarrow \infty} \alpha$  folgt mit  $(\Delta)$ :

$$\frac{1}{\alpha} EX_1 \leq \liminf_{k \rightarrow \infty} \left( \frac{S_k}{k} \right) \leq \limsup_{k \rightarrow \infty} \left( \frac{S_k}{k} \right) \leq \alpha EX_1 \quad P\text{-f.s.}$$

Sei  $N_\alpha$  die Ausnahmемenge zu  $\alpha$  in der Konvergenz  $(\Delta)$ . Da  $\alpha > 1$  beliebig, gilt auf

$$\underbrace{\left(\bigcup_{j=1}^{\infty} N_{1+\frac{1}{j}}\right)^C}_{P\text{-Nullmenge}}$$

$$EX_1 \leq \liminf_{k \rightarrow \infty} \left(\frac{S_k}{k}\right) \leq \limsup_{k \rightarrow \infty} \left(\frac{S_k}{k}\right) \leq EX_1$$

$$\implies \overline{X_n} := \frac{1}{n} S_n \xrightarrow{f.s.} EX_1$$

Jetzt muss noch die Bedingung  $X_k \geq 0$  weg. Es folgt:

$$\overline{X_n} = \frac{1}{n} \sum_{k=1}^n X_k^+ - \frac{1}{n} \sum_{k=1}^n X_k^- \xrightarrow{f.s.} EX_1^+ - EX_1^- = EX_1.$$

■

#### Beispiel 4.1 (Wiederholte Spiele)

Gegeben 2 Spieler. Spieler A erzielt in Runde  $n$   $X_n$  Punkte und Spieler B  $Y_n$  Punkte. Die Zufallsvariablen seien alle unabhängig und identisch verteilt. Es sei  $D_n := X_n - Y_n$ . Spieler A gewinnt Runde  $n$ , falls  $D_n > 0$ .

Sei  $p_n = P(\sum_{k=1}^n D_k > 0)$  die Wahrscheinlichkeit, dass Spieler A nach  $n$  Runden mehr Punkte hat. Es gilt nach S.4.2:

$$\frac{1}{n} \sum_{k=1}^n \mathbf{1}_{[D_k > 0]} \xrightarrow{f.s.} E[\mathbf{1}_{[D_1 > 0]}] = p_1.$$

Ist  $p_1 > \frac{1}{2}$ , so gewinnt Spieler A langfristig mehr Runden als B. Dies gilt jedoch nicht, wenn die Punkte addiert werden! Beispiel dazu:

$$X_k := \begin{cases} n+1, & \text{mit Wahrscheinlichkeit } p_1 \\ 0, & \text{mit Wahrscheinlichkeit } 1-p_1 \end{cases}, \quad Y_k \equiv n \text{ mit Wahrscheinlichkeit } 1$$

$$\text{Sei } p_1 = 0,999, \quad n = 1000. \implies p_{1000} = (0,999)^{1000} \approx 0,37$$



# 5 Zentraler Grenzwertsatz von Lindeberg-Lévy

## 5.1 Charakteristische Funktionen

### Definition

Es sei  $X$  Zufallsvariable auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$ . Dann heißt

$$\phi_X(t) := Ee^{itX} = E \cos(tX) + iE \sin(tX)$$

die **charakteristische Funktion** zu  $X$ .

### Bemerkung

Ist  $X$  diskret mit Werten  $x_1, x_2, \dots$ , so gilt:

$$\phi_X(t) = \sum_{k=1}^{\infty} e^{itx_k} \cdot P(X = x_k)$$

Ist  $X$  absolutstetig mit Dichte  $f$ , so gilt:

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{itx} f(x) dx \quad (\text{Fourier-Transformation})$$

### Beispiel 5.1

a)  $X \sim B(n, p)$

$$\begin{aligned} \phi_X(t) &= \sum_{k=0}^n e^{itk} \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=0}^n \binom{n}{k} (pe^{it})^k (1-p)^{n-k} = (1-p + pe^{it})^n \end{aligned}$$

b)  $X \sim U(0, 1)$

$\phi_X(0) = 1$  und für  $t \neq 0$ :

$$\begin{aligned} \phi_X(t) &= \int_0^1 e^{itx} \cdot 1 dx = \int_0^1 \cos(tx) dx + i \int_0^1 \sin(tx) dx \\ &= \frac{1}{t} \sin(t) - \frac{i}{t} \cos(t) + \frac{i}{t} = \frac{1}{it} (e^{it} - 1) \end{aligned}$$

c)  $X \sim N(0, 1)$

$$\phi_X(t) = e^{-\frac{t^2}{2}} \quad \text{vgl. Stochastik 1}$$

**Satz 5.1** Sind  $X, Y$  unabhängige Zufallsvariablen mit charakteristischen Funktionen  $\phi_X$  und  $\phi_Y$ , so gilt für die charakteristische Funktion  $\phi_{X+Y}$  der Faltung:

$$\phi_{X+Y}(t) = \phi_X(t) \cdot \phi_Y(t) \quad \forall t \in \mathbb{R}$$

**Beweis** vgl. Stochastik 1, Satz 12.2. ■

**Lemma 5.1** Für alle  $m \in \mathbb{N}, t \in \mathbb{R}$  gilt:

$$\left| e^{it} - \sum_{k=0}^{m-1} \frac{(it)^k}{k!} \right| \leq \min \left\{ \frac{|t|^m}{m!}, \frac{2|t|^{m-1}}{(m-1)!} \right\}$$

**Beweis** vgl. Stochastik 1, Satz 13.2. ■

## 5.2 Umkehrsätze

Wir werden sehen, dass eine Verteilung eindeutig durch ihre charakteristische Funktion festgelegt ist. Hat man z.B. gezeigt, dass  $X$  die charakteristische Funktion  $(1 - p + pe^{it})^n$  hat, so ist  $X \sim B(n, p)$ .

Aus der Analysis ist die Integralsinusfunktion bekannt:

$$Si : \mathbb{R}_+ \rightarrow \mathbb{R}_+, \quad Si(x) := \int_0^x \frac{\sin(y)}{y} dy \quad \forall x > 0$$

Es gilt:  $\lim_{x \rightarrow \infty} (Si(x)) = \frac{\pi}{2}$

### Satz 5.2

Es sei  $X$  Zufallsvariable mit charakteristischer Funktion  $\phi_X$ . Dann gilt für alle  $-\infty < a < b < \infty$ :

$$\frac{1}{2}P(X = a) + P(a < X < b) + \frac{1}{2}P(X = b) = \lim_{T \rightarrow \infty} \left( \frac{1}{2\pi} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} \phi_X(t) dt \right)$$

### Beweis

Sei  $I(T) := \frac{1}{2\pi} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} \phi_X(t) dt$ . Definiere  $\psi : \mathbb{R} \times [-T, T] \rightarrow \mathbb{C}$  durch

$$\psi(t, x) := \begin{cases} \frac{e^{-it(a-x)} - e^{-it(b-x)}}{it}, & t \neq 0 \\ b - a, & t = 0 \end{cases}$$

Mit Lemma 5.1 folgt, dass  $\psi$  stetig ist und wegen

$$\left| \frac{e^{-ita} - e^{-itb}}{it} \right| = \left| \int_a^b e^{ity} dy \right| \leq b - a$$

ist  $|\psi| \leq b - a$ , also ist  $\psi P^X \otimes \lambda_{[-T,T]}$ -integrierbar. Mit Satz 3.1 (Fubini I) folgt:

$$\begin{aligned} I(T) &= \frac{1}{2\pi} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} \left( \int e^{itx} P^X(dx) \right) dt \\ &= \frac{1}{2\pi} \int \underbrace{\int_{-T}^T \frac{1}{it} (e^{-it(a-x)} - e^{-it(b-x)}) dt}_{=: \psi_{a,b,T}(x)} P^X(dx) \end{aligned}$$

Inneres Integral:

Da  $t \mapsto \frac{\cos(t(x-a))}{it}$  punktsymmetrisch ist, gilt:

$$\psi_{a,b,T}(x) = 2 \cdot \int_0^T \frac{1}{t} \sin((x-a)t) dt - 2 \cdot \int_0^T \frac{1}{t} \sin((x-b)t) dt$$

Es gilt weiterhin:

$$c \cdot \int_0^T \frac{1}{c \cdot t} \sin(ct) dt = \operatorname{sgn}(c) \cdot Si(T|c|) \quad \text{mit } \operatorname{sgn}(c) = \begin{cases} 1, & c > 0 \\ 0, & c = 0 \\ -1, & c < 0 \end{cases}$$

$$\Rightarrow \psi_{a,b,T}(x) = 2 \cdot \operatorname{sgn}(x-a) Si(T|x-a|) - 2 \cdot \operatorname{sgn}(x-b) Si(T|x-b|)$$

$$\Rightarrow \psi_{a,b}(x) := \lim_{T \rightarrow \infty} (\psi_{a,b,T}(x)) = \begin{cases} 0, & x < a \text{ oder } x > b \\ \pi, & x = a \text{ oder } x = b \\ 2\pi, & a < x < b \end{cases}$$

$\Rightarrow (\psi_{a,b,T})_{T \geq 0}$  besitzt eine (konstante) integrierbare Majorante. Mit dem Satz über die majorisierte Konvergenz gilt:

$$\begin{aligned} \lim_{T \rightarrow \infty} I(T) &= \frac{1}{2\pi} \int \psi_{a,b}(x) P^X(dx) \\ &= \frac{1}{2} P(X=a) + \frac{1}{2} P(X=b) + P(a < X < b) \end{aligned}$$

■

### Korollar 5.1

Sind  $X$  und  $Y$  Zufallsvariablen mit derselben charakteristischen Funktion, so haben  $X$  und  $Y$  dieselbe Verteilung.

**Beweis** Sei  $D = A(X) \cup A(Y)$  mit  $A(X) = \{x \in \mathbb{R} | P(X=x) > 0\}$ , analog  $A(Y)$ .  $A(X)$  ist abzählbar, da  $A(X) = \bigcup_{n=1}^{\infty} \{x \in \mathbb{R} | P(X=x) \geq \frac{1}{n}\}$  und  $|\{x \in \mathbb{R} | P(X=x) \geq \frac{1}{n}\}| \leq n \Rightarrow D$  abzählbar

$$\mathcal{D} := \{(a,b) | -\infty < a \leq b < \infty, a, b \notin D\}$$

ist ein durchschnittstabiles Erzeugendensystem von  $\mathfrak{B}(\mathbb{R})$ .  $\xrightarrow{\text{Sa5.2}} P^X$  und  $P^Y$  stimmen auf  $\mathcal{D}$  überein  $\xrightarrow{\text{Eindeutigkeitssatz}}$  Behauptung. ■

**Satz 5.3**

Sei  $X$  eine Zufallsvariable mit charakteristischer Funktion  $\phi$ . Gilt  $\int |\phi(t)| dt < \infty$ , so hat  $X$  eine stetige Dichte  $f$ , die gegeben ist durch

$$f(x) = \frac{1}{2\pi} \int e^{-itx} \phi(t) dt \quad \forall x \in \mathbb{R}$$

**Beweis** Wie in Beweis von Satz 5.2 gilt:

$$\left| \frac{e^{-ita} - e^{-itb}}{it} \right| \leq |b - a| \quad (*)$$

Da  $\phi$   $\lambda$ -integrierbar ist, ist  $|b - a||\phi|$  eine integrierbare Majorante für diesen Ausdruck in Satz 5.2. Es folgt:

$$\begin{aligned} \frac{1}{2}P(X = a) + P(a < X < b) + \frac{1}{2}P(X = b) &= \frac{1}{2\pi} \int \frac{e^{-ita} - e^{-itb}}{it} \phi(t) dt \\ \implies P(a < X < b) &\leq \frac{1}{2\pi} |b - a| \underbrace{\int |\phi(t)| dt}_{< \infty} \\ \implies P(X = x) &= \lim_{n \rightarrow \infty} P(x - \frac{1}{n} < X < x + \frac{1}{n}) \\ &= 0 \end{aligned}$$

Ist  $F$  die Verteilungsfunktion von  $X$ , so gilt:

$$F(b) - F(a) = \frac{1}{2\pi} \int \frac{e^{-ita} - e^{-itb}}{it} \phi(t) dt \quad \forall a < b$$

Wegen (\*) kann man den Satz von der majorisierten Konvergenz anwenden und bekommt:

$$\begin{aligned} \lim_{h \downarrow 0} \frac{F(x+h) - F(x)}{h} &= \frac{1}{2\pi} \int e^{-itx} \lim_{h \downarrow 0} \frac{1 - e^{-ith}}{ith} \phi(t) dt \\ &= \frac{1}{2\pi} \int e^{-itx} \phi(t) dt \\ &=: f(x) \end{aligned}$$

Außerdem folgt  $x \mapsto f(x)$  ist stetig. ■

## 5.3 Verteilungskonvergenz

**Definition**

- a) Gegeben sei der messbare Raum  $(\mathbb{R}, \mathfrak{B})$  mit Wahrscheinlichkeitsmaßen  $P, P_1, P_2, \dots$  und zugehörigen Verteilungsfunktionen  $F, F_1, F_2, \dots$ .  
 $P_n$  **konvergiert schwach** gegen  $P$  ( $P_n \xrightarrow{w} P$ ), wenn  $\lim_{n \rightarrow \infty} F_n(x) = F(x) \forall x \in \mathbb{R}$  an denen  $F$  stetig ist.
- b) Seien  $X, X_1, X_2, \dots$  Zufallsvariablen auf (unter Umständen verschiedenen) Wahrscheinlichkeitsräumen  $(\Omega, \mathcal{A}, P), (\Omega_1, \mathcal{A}_1, P_1), \dots$ .  
 $X_n$  **konvergiert in Verteilung** gegen  $X$  ( $X_n \xrightarrow{d} X$ ), wenn  $P^{X_n} \xrightarrow{w} P^X$ .

**Beispiel 5.2** Konvergenz in Verteilung bzw. schwache Konvergenz ist schwächer als f.s.-Konvergenz.

Sei z.B.  $X \sim N(0, 1)$  und  $X_{2n} = X, X_{2n+1} = -X \forall n \in \mathbb{N}$ .  $\implies P^{X_n} \equiv P^X = N(0, 1)$  und  $(X_n)$  konvergiert in Verteilung (gegen  $X$ ) jedoch  $X_n \not\xrightarrow{f.s.} X$

Jedoch gilt folgender nützlicher Satz:

**Satz 5.4 (Darstellungssatz von Skorohod)**

Es seien  $X, X_1, X_2, \dots$  Zufallsvariablen mit  $X_n \xrightarrow{d} X$ . Dann existiert ein Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$  und hierauf Zufallsvariablen  $X', X'_1, X'_2, \dots$  mit  $X' \stackrel{d}{=} X, X'_n \stackrel{d}{=} X_n \forall n \in \mathbb{N}$  derart, dass  $X'_n \xrightarrow{f.s.} X'$ .

**Beweis** Es seien  $F, F_1, F_2, \dots$  die Verteilungsfunktionen zu  $X, X_1, X_2, \dots$  und  $(\Omega, \mathcal{A}, P) = ((0, 1), \mathfrak{B}_{(0,1)}, \lambda_{(0,1)})$ . Weiter sei  $F^{-1} : (0, 1) \rightarrow \mathbb{R}, F^{-1}(y) := \inf\{x \in \mathbb{R} | F(x) \geq y\}$  die Quantilsfunktion zu  $F$ , analog (Quantilsfunktion)  $F_n^{-1}, n \in \mathbb{N}$ . Setze  $X' := F^{-1}, X'_n := F_n^{-1}$ .

Satz 5.7 (Stoch 1)  $\implies X' \stackrel{d}{=} X, X'_n \stackrel{d}{=} X_n, n \in \mathbb{N}$  ( $P(X' \leq x) = P(F^{-1}(\omega) \leq x) = \underbrace{P(\omega \leq F(x))}_{=\lambda(0,1)} = F(x)$ )

Es bleibt also zu zeigen, dass für  $P$ -fast alle  $\omega \in \Omega : \lim_{n \rightarrow \infty} X'_n(\omega) = X'(\omega)$ .

Sei  $\omega \in (0, 1)$ . Da  $X$  nur abzählbar viele Atome hat (vgl. Beweis von Korollar 5.1) existiert zu  $\varepsilon > 0$  ein  $x \in \mathbb{R}$  mit  $X'(\omega) - \varepsilon < x < X'(\omega)$  und  $P(X = x) = 0$ .

Es gilt (Lemma 5.6, Stoch 1):  $\forall y \in (0, 1), x \in \mathbb{R} :$

$$y \leq F(x) \iff F^{-1}(y) \leq x$$

Hier:  $\omega \leq F(x) \iff F^{-1}(\omega) = X'(\omega) \leq x$ . Wegen  $X'(\omega) > x$  folgt  $F(x) < \omega$ . Da  $F_n(x) \rightarrow F(x)$  für  $n \rightarrow \infty$  nach Voraussetzung,  $\exists n_0 \in \mathbb{N}$ , so dass  $\forall n \geq n_0 : F_n(x) < \omega$ . Also  $X'_n > x$ .

Mit  $\varepsilon \downarrow 0$  folgt:

$$\liminf_{n \rightarrow \infty} X'_n(\omega) \geq X'(\omega) \quad \forall \omega \in \Omega.$$

Ist  $\omega' > \omega$  und  $\varepsilon > 0$ , so  $\exists$  ein  $x$  mit  $X'(\omega') < x < X'(\omega') + \varepsilon$  und  $P(X = x) = 0$ . Da  $F$  rechtsseitig stetig, folgt  $F(F^{-1}(y)) \geq y \forall y \in (0, 1)$ , also mit der Monotonie von  $F : \omega < \omega' \leq F(X'(\omega')) \leq F(x)$ .

Wegen  $F_n(x) \rightarrow F(x)$  ( $n \rightarrow \infty$ ),  $\exists n_0 \in \mathbb{N}$  sodass  $\omega \leq F_n(x)$  (d.h.  $X'_n(\omega) \leq x$ )  $\forall n \geq n_0$  gilt mit  $\varepsilon \downarrow 0$  ergibt das

$$\limsup_{n \rightarrow \infty} X'_n(\omega) \leq X'(\omega') \quad \forall \omega' > \omega.$$

**Satz 5.5** Es sei  $C_b(\mathbb{R})$  die Menge aller stetigen und beschränkten Funktionen,  $h : \mathbb{R} \rightarrow \mathbb{R}$ . Dann gilt:

$$X_n \xrightarrow{d} X \iff Eh(X_n) \rightarrow Eh(X) \quad \forall h \in C_b(\mathbb{R})$$

**Beweis**

“ $\Rightarrow$ ”: Nach Satz 5.4 existieren ein Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$  und Zufallsvariablen  $X' \stackrel{d}{=} X$ ,  $X'_n \stackrel{d}{=} X_n \quad \forall n \in \mathbb{N}$  mit  $X'_n \xrightarrow{f.s.} X'$ . Es folgt:

$$\lim_{n \rightarrow \infty} (Eh(X_n)) = \lim_{n \rightarrow \infty} (Eh(X'_n)) \stackrel{h \text{ stetig}, X'_n \xrightarrow{f.s.} X'}{=} Eh(X') = Eh(X).$$

“ $\Leftarrow$ ”: Für  $a, b \in \mathbb{R}, a < b$  sei  $h_{a,b} : \mathbb{R} \rightarrow \mathbb{R}$  definiert durch

$$h_{a,b}(x) := \begin{cases} 1 & , x \leq a \\ \frac{b-x}{b-a} & , a < x < b \\ 0 & , x \geq b \end{cases}$$

$h_{a,b}$  ist stetig und beschränkt. Seien  $F, F_n$  die Verteilungsfunktionen zu  $X, X_n \quad \forall n \in \mathbb{N}$ . Dann gilt  $\forall y > x$ :

$$\begin{aligned} F_n(x) &= E[\mathbf{1}_{(-\infty, x)}(X_n)] \leq E[h_{x,y}(X_n)] \xrightarrow{n \rightarrow \infty} E[h_{x,y}(X)], \\ E[h_{x,y}(X)] &\leq E[\mathbf{1}_{(-\infty, y)}(X)] = F(y). \end{aligned}$$

Also folgt da  $F$  rechtsseitig stetig ist mit  $y \downarrow x$ :

$$\limsup_{n \rightarrow \infty} (F_n(x)) \leq F(x) \quad \forall x \in \mathbb{R}$$

Analog erhält man für  $y < x$ :

$$F_n(x) \geq E[h_{y,x}(X_n)] \xrightarrow{n \rightarrow \infty} E[h_{y,x}(X)] \geq F(y)$$

Mit  $y \uparrow x$ :  $\liminf_{n \rightarrow \infty} (F_n(x)) \geq F(x-)$   $\forall x \in \mathbb{R}$ .

Ist  $F$  in  $x$  stetig, so gilt  $F(x-) = F(x)$  und somit  $F_n(X) \xrightarrow{n \rightarrow \infty} F(x)$ . ■

**Satz 5.6 (“Continuous Mapping Theorem”)**

Es seien  $X, X_1, X_2, \dots$  Zufallsvariablen mit  $X_n \xrightarrow{d} X$ . Weiter sei  $f : \mathbb{R} \rightarrow \mathbb{R}$  eine Borel-messbare Funktion mit  $P(X \in \{x \in \mathbb{R} \mid f \text{ nicht stetig in } x\}) = 0$ .

Dann gilt auch  $f(X_n) \xrightarrow{d} f(X)$  für  $n \rightarrow \infty$ .

**Beweis** Übung. ■

**Satz 5.7 (Satz von Helly<sup>1</sup>)**

Zu jeder Folge  $(F_n)_{n \in \mathbb{N}}$  von Verteilungsfunktionen existieren eine Teilfolge  $(F_{n_k})_{k \in \mathbb{N}}$  und eine schwach monoton wachsende, rechtsseitig stetige Funktion  $G : \mathbb{R} \rightarrow [0, 1]$ , sodass  $\lim_{k \rightarrow \infty} (F_{n_k}(x)) = G(x) \quad \forall x \in \mathbb{R}$ , an denen  $G$  stetig ist.

**Beweis (Skizze)**

Für  $x \in \mathbb{R}$  ist  $(F_n(x))_{n \in \mathbb{N}} \subset [0, 1] \xrightarrow{\text{Bolzano-Weierstraß}} \exists$  Häufungspunkt. Sei  $(r_k)_{k \in \mathbb{N}}$  eine Abzählung von  $\mathbb{Q}$ . Wähle Teilfolgen  $(F_{n_{k,j}})_{j \in \mathbb{N}}$  mit  $F_{n_{k,j}} \xrightarrow{j \rightarrow \infty} G_0(r_k)$ , wobei  $(n_{k+1,j})_{j \in \mathbb{N}}$  eine Teilfolge von  $(n_{k,j})_{j \in \mathbb{N}}$  ist. (Definition der Funktion  $f_0$  auf  $\mathbb{Q}$ )  
Für die Diagonalfolge  $(n_{j,j})_{j \in \mathbb{N}}$  gilt dann:  $F_{n_{j,j}} \rightarrow G_0$  auf  $\mathbb{Q}$ . Sei  $G_0$  auf ganz  $\mathbb{R}$  durch  $G(x) := \inf\{G_0(r) \mid r \in \mathbb{Q}, r > x\}$  fortgesetzt.  
Rest:  $\epsilon - \delta$ -Argumente. ■

**Bemerkung**  $G$  aus Satz 5.7 muß keine Verteilungsfunktion sein.

Beispiel:  $F_n := \mathbf{1}_{[n, \infty]} \implies G \equiv 0$

**Definition** Eine Familie  $\mathcal{P}$  von Wahrscheinlichkeitsmaßen auf  $(\mathbb{R}, \mathfrak{B})$  heißt **straff**, wenn  $\forall \epsilon > 0 \exists$  kompaktes Intervall  $[a, b] \subset \mathbb{R}$  mit:

$$P([a, b]) \geq 1 - \epsilon \quad \forall P \in \mathcal{P}$$

**Bemerkung**

- (i) Ist  $\mathcal{P}$  straff, so auch jedes  $\mathcal{P}' \subset \mathcal{P}$ .
- (ii) Sind alle  $\mathcal{P}_i$  mit  $i \in \{1, \dots, n\}$  straff, so auch  $\bigcup_{i=1}^n \mathcal{P}_i$ .
- (iii) Ist  $|\mathcal{P}| = 1$ , so ist  $\mathcal{P}$  straff.

**Satz 5.8** Ist  $\{P_n \mid n \in \mathbb{N}\}$  eine straffe Familie von Wahrscheinlichkeitsmaßen auf  $(\mathbb{R}, \mathfrak{B})$ , so existieren eine Teilfolge  $(P_{n_k})_{k \in \mathbb{N}}$  von  $(P_n)_{n \in \mathbb{N}}$  und ein Wahrscheinlichkeitsmaß  $P$  derart, dass  $P_{n_k} \xrightarrow{w} P$  für  $k \rightarrow \infty$ .

**Beweis** Sei  $F_n$  die Verteilungsfunktion zu  $P_n \quad \forall n \in \mathbb{N}$ .

$\xrightarrow{\text{Satz 5.7}} \exists$  Folge  $(n_k)_{k \in \mathbb{N}}$  mit  $F_{n_k}(x) \rightarrow G(x)$  für  $k \rightarrow \infty \quad \forall x \in \mathbb{R}$  mit  $G$  stetig in  $x$ ;  $G$  ist wachsend und rechtsseitig stetig.

Bleibt zu zeigen:  $G$  ist Verteilungsfunktion, also  $\lim_{x \rightarrow -\infty} (G(x)) = 0$  und  $\lim_{x \rightarrow \infty} (G(x)) = 1$ . Ist dann  $P$  das Wahrscheinlichkeitsmaß zu  $G$ , so folgt  $P_{n_k} \xrightarrow{w} P$ .

Sei also  $\epsilon > 0$ . Da  $\{P_n \mid n \in \mathbb{N}\}$  straff ist  $\implies \exists a, b \in \mathbb{R}$  mit  $P_n([a, b]) \geq 1 - \epsilon \quad \forall n \in \mathbb{N}$ .  $\implies F_n(a) \leq \epsilon \quad \forall n \in \mathbb{N}$ .

$G$  hat höchstens abzählbar viele Unstetigkeitsstellen.  $\implies \exists c < a$ , in dem  $G$  stetig.  $\implies G(c) = \lim_{k \rightarrow \infty} (F_{n_k}(c)) \leq \epsilon \implies G(x) \leq \epsilon \quad \forall x \leq c$ .

Also:  $\forall \epsilon > 0 \quad \exists c \in \mathbb{R} : \quad \forall x \leq c \text{ gilt } 0 \leq G(x) \leq \epsilon \implies \lim_{x \rightarrow -\infty} (G(x)) = 0$  und  $\lim_{x \rightarrow \infty} (G(x)) = 1$ . ■

**Satz 5.9 (Stetigkeitssatz für charakteristische Funktionen)**

Es seien  $X, X_1, X_2, \dots$  Zufallsvariablen,  $\phi, \phi_1, \phi_2, \dots$  die zugehörigen charakteristischen Funktionen. Dann gilt:

$$X_n \xrightarrow{d} X \iff \phi_n(t) \rightarrow \phi(t) \quad \forall t \in \mathbb{R}$$

**Beweis**

“ $\Rightarrow$ ”: Sei  $t \in \mathbb{R}$ .  $x \mapsto \cos(tx)$ ,  $x \mapsto \sin(tx)$  sind stetig und beschränkt.

Satz 5.5  $\Rightarrow \phi_n(t) = E \cos(tX_n) + iE \sin(tX_n) \rightarrow E \cos(tX) + iE \sin(tX) = \phi(t)$ .

“ $\Leftarrow$ ”: Wir zeigen zunächst:  $\{P^{X_n}, n \in \mathbb{N}\}$  ist straff.  $\mathbb{C}$ -wertige Version von Fubini II liefert  $\forall \delta > 0$ .

$$\begin{aligned} \frac{1}{\delta} \int_{-\delta}^{\delta} (1 - \varphi_n(t)) dt &= \int \left( \frac{1}{\delta} \int_{-\delta}^{\delta} (1 - e^{itx}) dt \right) P^{X_n}(dx) \\ &= 2 \int \underbrace{\left( 1 - \frac{\sin(\delta x)}{\delta x} \right)}_{\geq 0} P^{X_n}(dx) \\ &\geq 2 \int_{|x| \geq \frac{2}{\delta}} \underbrace{\left( 1 - \frac{1}{|\delta x|} \right)}_{\geq \frac{1}{2}} P^{X_n}(dx) \\ &\geq P^{X_n}\left(\left[-\frac{2}{\delta}, \frac{2}{\delta}\right]^C\right) \end{aligned}$$

Sei  $\varepsilon > 0$ . Da  $\varphi$  in 0 stetig und  $\varphi(0) = 1$ ,  $\exists \delta > 0$ :

$$|1 - \varphi(t)| \leq \frac{\varepsilon}{4} \quad \forall |t| \leq \delta$$

$\Rightarrow \left| \frac{1}{\delta} \int_{-\delta}^{\delta} (1 - \varphi(t)) dt \right| \leq \frac{1}{\delta} 2\delta \frac{\varepsilon}{4} = \frac{\varepsilon}{2}$ . Da  $|\varphi_n| \leq 1$  folgt mit majorisierter Konvergenz:

$$\int_{-\delta}^{\delta} (1 - \varphi_n(t)) dt \xrightarrow{n \rightarrow \infty} \int_{-\delta}^{\delta} (1 - \varphi(t)) dt$$

$\Rightarrow \exists n_0 \in \mathbb{N}$ , so dass  $\frac{1}{\delta} \int_{-\delta}^{\delta} (1 - \varphi_n(t)) dt \leq \varepsilon \quad \forall n \geq n_0$ .  $\Rightarrow P^{X_n}\left(\left[-\frac{2}{\delta}, \frac{2}{\delta}\right]\right) \geq 1 - \varepsilon \quad \forall n \geq n_0$ .

Außerdem:  $\forall n \in \{1, \dots, n_0 - 1\} \exists a_n > 0$  mit  $P^{X_n}([-a_n, a_n]) \geq 1 - \varepsilon$  da  $P^{X_n}([-m, m]) \rightarrow 1$  für  $m \rightarrow \infty$ .

Insgesamt: Sei  $a := \max\{a_1, \dots, a_{n_0-1}, \frac{2}{\delta}\} \Rightarrow P^{X_n}([-a, a]) \geq 1 - \varepsilon \quad \forall n \in \mathbb{N} \Rightarrow \{P^{X_n}, n \in \mathbb{N}\}$  ist straff.

**Annahme:**  $X_n \xrightarrow{d} X$  gilt nicht.

$\Rightarrow \exists x \in \mathbb{R}$  mit  $P(X = x) = 0$  und  $P(X_n \leq x) \not\rightarrow P(X \leq x), n \rightarrow \infty$ .

d.h.  $\exists \varepsilon > 0$  und eine Teilfolge  $(X_{n_k})_{k \in \mathbb{N}}$  mit  $|P(X_{n_k} \leq x) - P(X \leq x)| \geq \varepsilon \quad \forall k \in \mathbb{N} (*)$ .

$\{P^{X_{n_k}}, k \in \mathbb{N}\}$  ist ebenfalls straff  $\xrightarrow{S.5.8} \exists$  Teilfolge  $(X_{n_{k_j}})_{j \in \mathbb{N}}$  und ein W'maß  $P_0$  mit  $P^{X_{n_{k_j}}} \xrightarrow{w} P_0$ .



Sei  $\varphi_0$  charakteristische Funktion zu  $P_0$ . Also folgt mit der Hinrichtung:  $\varphi_{n_{k_j}}(t) \rightarrow \varphi_0(t) = \varphi(t) \xrightarrow{Kor. 5.1} P_0 = P^X$ , also  $X_{n_{k_j}} \xrightarrow{d} X$  und damit  $P(X_{n_{k_j}} \leq x) \rightarrow P(X \leq x)$ .  
Wid zu (\*).  $\blacksquare$

Wir benötigen noch folgendes technisches Hilfslemma:

**Lemma 5.2**

Für alle  $z_1, \dots, z_n, w_1, \dots, w_n \in \{z \in \mathbb{C} \mid |z| \leq 1\}$  gilt:

$$\left| \prod_{k=1}^n z_k - \prod_{k=1}^n w_k \right| \leq \sum_{k=1}^n |z_k - w_k|$$

**Beweis**

$$\begin{aligned} \left| \prod_{k=1}^n z_k - \prod_{k=1}^n w_k \right| &\leq \left| \prod_{k=1}^n z_k - w_1 \prod_{k=2}^n z_k \right| + \left| w_1 \prod_{k=2}^n z_k - w_1 w_2 \prod_{k=3}^n z_k \right| + \dots + \left| w_1 \dots w_{n-1} z_n - \prod_{k=1}^n w_k \right| \\ &= |z_1 - w_1| \underbrace{\left| \prod_{k=2}^n z_k \right|}_{\leq 1} + |z_2 - w_2| \underbrace{|w_1| \left| \prod_{k=3}^n z_k \right|}_{\leq 1} + \dots + |z_n - w_n| \underbrace{\left| \prod_{k=1}^{n-1} w_k \right|}_{\leq 1} \end{aligned} \quad \blacksquare$$

Hauptsatz des Abschnitts:

**Satz 5.10 (Zentraler Grenzwertsatz von Lindeberg-Lévy)**

Für jedes  $n \in \mathbb{N}$  seien  $X_{nk}, k = 1, \dots, r_n$  unabhängige Zufallsvariablen (nicht notwendig identisch verteilt) auf einem Wahrscheinlichkeitsraum  $(\Omega_n, \mathcal{A}_n, P_n)$  mit  $\text{Var}(X_{nk}) = \sigma_{nk}^2 < \infty$  und  $EX_{nk} = \mu_{nk} < \infty$ . Es sei  $s_n^2 := \sum_{k=1}^{r_n} \sigma_{nk}^2 > 0$ . Ist dann für alle  $\varepsilon > 0$  die **Lindeberg-Bedingung**

$$(L) \quad \lim_{n \rightarrow \infty} \frac{1}{s_n^2} \sum_{k=1}^{r_n} \int_{|X_{nk} - \mu_{nk}| > \varepsilon s_n} (X_{nk} - \mu_{nk})^2 dP_n = 0$$

erfüllt, so gilt mit  $n \rightarrow \infty$ :

$$\frac{1}{s_n} \sum_{k=1}^{r_n} (X_{nk} - \mu_{nk}) \xrightarrow{d} Z, \quad Z \sim N(0, 1)$$

**Bemerkung 5.1** 1. Die Lindeberg-Bedingung schließt einen dominierenden Einfluss eines einzelnen Summanden  $X_{nk}$  auf die  $X_{n1} + \dots + X_{nr_n}$  aus. Insbesondere gilt:

$$\max\{\sigma_{nk}^2 \mid 1 \leq k \leq r_n\} = o(s_n^2) \text{ für } n \rightarrow \infty$$

2. Der ZGWS hat eine lange „Verbesserungsgeschichte“ hinter sich. Gelegentlich ist die Lyapunov-Bedingung einfacher zu verwenden:

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^{2+\delta}} \sum_{k=1}^{r_n} E(|X_{nk} - \mu_{nk}|^{2+\delta}) = 0 \text{ für ein } \delta > 0$$

3. Der Satz liefert eine Begründung für die „Allgegenwart“ der Normalverteilung.

Ein wichtiger Spezialfall ist

**Satz 5.11 (ZGWS St. I)**

Es seien  $Y_1, Y_2, \dots$  u.i.v. ZV mit  $EY_1 = \mu < \infty$  und  $0 < \text{Var}(Y_1) = \sigma^2 < \infty$ . Dann gilt:

$$\frac{Y_1 + \dots + Y_n - n\mu}{\sqrt{n}\sigma} \xrightarrow{d} Z \sim N(0, 1)$$

**Beweis** Sei  $X_{nk} := Y_k, r_n = n, (\Omega_n, \mathcal{A}_n, P_n) = (\Omega, \mathcal{A}, P)$ . Es gilt:  $s_n^2 = n\sigma^2$  und

$$\frac{1}{s_n^2} \sum_{k=1}^{r_n} \int_{|X_{nk} - \mu_{nk}| > \varepsilon s_n} (X_{nk} - \mu_{nk})^2 dP_n = \frac{1}{\sigma^2} \int_{|Y_1 - \mu_1| > \varepsilon \sqrt{n}\sigma} (Y_1 - \mu_1)^2 dP =: I_n$$

Da  $z_n := \mathbf{1}_{(\varepsilon\sqrt{n}\sigma, \infty)}(|Y_1 - \mu_1|)(Y_1 - \mu_1)^2 \leq (Y_1 - \mu_1)^2$  und  $\lim_{n \rightarrow \infty} z_n = 0$  folgt mit majorisierter Konvergenz, dass  $\lim_{n \rightarrow \infty} I_n = 0$ . Also ist die Lindeberg-Bedingung erfüllt und die Behauptung folgt mit Satz 5.10. ■

**Beweis** Beweis von Satz 5.10

O.B.d.A:  $\mu_{nk} = 0$  und  $s_n = 1$ . Anderfalls ersetze  $X_{nk}$  durch  $\frac{X_{nk} - \mu_{nk}}{s_n}$ .

**Idee:** Verwende S.5.9: Sei  $\varphi_{nk}$  die charakteristische Funktion von  $X_{nk}$  und  $\varphi_{s_n}$  die von  $\sum_{k=1}^{r_n} X_{nk}$ :  $\varphi_{s_n}(t) = \prod_{k=1}^{r_n} \varphi_{nk}(t) \rightarrow \varphi_z(t) = e^{-\frac{t^2}{2}}$   
Zu zeigen:

$$\prod_{k=1}^{r_n} \varphi_{nk}(t) \rightarrow \varphi_z(t) = e^{-\frac{t^2}{2}} \quad \forall t \in \mathbb{R}$$

Mit Lemma 5.1 ( $m = 3$ ):

$$\begin{aligned} \left| e^{itx} - \left(1 + itx - \frac{1}{2}t^2x^2\right) \right| &\leq \min\left\{\frac{|tx|^3}{3!}, |tx|^2\right\} \quad \forall x \in \mathbb{R} \\ &\leq \min\{|tx|^3, |tx|^2\} \end{aligned}$$

Integral über  $x$  liefert (beachte:  $EX = 0$ )

$$\left| \varphi_{n_k}(t) - \left(1 - \frac{1}{2}t^2\sigma_{n_k}^2\right) \right| \leq E \min\{|tX_{n_k}|^2, |tX_{n_k}|^3\} =: M_{n_k}$$

Sei  $\varepsilon > 0$  beliebig. Es gilt

$$\begin{aligned} M_{n_k} &\leq \int_{|X_{n_k}| \leq \varepsilon} |tX_{n_k}|^3 dP_n + \int_{|X_{n_k}| > \varepsilon} |tX_{n_k}|^2 dP_n \\ &\leq |t|^3 \varepsilon \sigma_{n_k}^2 + t^2 \int_{|X_{n_k}| > \varepsilon} X_{n_k}^2 dP_n \\ \Rightarrow \sum_{k=1}^{r_n} M_{n_k} &\leq |t|^3 \varepsilon + t^2 \sum_{k=1}^{r_n} \int_{|X_{n_k}| > \varepsilon} X_{n_k}^2 dP_n \xrightarrow{n \rightarrow \infty} \varepsilon |t|^3 + 0 \quad \text{folgt mit (L)} \end{aligned}$$

Mit  $\varepsilon \downarrow 0$  folgt:

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{r_n} \left| \phi_{n_k}(t) - \left(1 - \frac{1}{2}t^2\sigma_{n_k}^2\right) \right| = 0 \quad \forall t \in \mathbb{R} \quad (1)$$

Behauptung:  $\lim_{n \rightarrow \infty} \left| \prod_{k=1}^{r_n} \phi_{n_k}(t) - \prod_{k=1}^{r_n} \left(1 - \frac{1}{2}t^2\sigma_{n_k}^2\right) \right| = 0 \quad \forall t \in \mathbb{R} \quad (2)$

Beweis:  $\forall \varepsilon > 0$  gilt:

$$\begin{aligned} \sigma_{n_k}^2 &\leq \int_{|X_{n_k}| > \varepsilon} X_{n_k}^2 dP_n + \varepsilon^2 \\ \implies \limsup_{n \rightarrow \infty} \max\{\sigma_{n_k}^2 | 1 \leq k \leq r_n\} \\ &\leq \lim_{n \rightarrow \infty} \left( \varepsilon^2 + \sum_{k=1}^{r_n} \int_{|X_{n_k}| > \varepsilon} X_{n_k}^2 dP_n \right) \\ &\stackrel{(L)}{=} \varepsilon^2 + 0 \end{aligned}$$

Mit  $\varepsilon \downarrow 0$ :

$$\lim_{n \rightarrow \infty} \max\{\sigma_{n_k}^2 | 1 \leq k \leq r_n\} = 0 \quad (3)$$

$$\implies \forall t \in \mathbb{R}, \exists n_0 \in \mathbb{N}, \text{ so dass } \forall n \geq n_0 : |1 - \frac{1}{2}t^2\sigma_{n_k}^2| \leq 1 \quad \forall k \in \{1, \dots, r_n\}$$

$$\implies \text{Für } n \geq n_0 \text{ läßt sich das } \prod \text{ in (2) nach Lemma 5.2 durch die Summe in (1) abschätzen, d.h. (1) } \implies (2)$$

Es bleibt zu zeigen:

$$\lim_{n \rightarrow \infty} \left| \underbrace{\prod_{k=1}^{r_n} \exp\left(-\frac{1}{2}t^2\sigma_{n_k}^2\right)}_{=e^{-\frac{1}{2}t^2}} - \prod_{k=1}^{r_n} \left(1 - \frac{1}{2}t^2\sigma_{n_k}^2\right) \right| = 0 \quad \forall t \in \mathbb{R}$$

Behauptung folgt mit Lemma 5.2 falls

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{r_n} \left| \exp\left(-\frac{1}{2}t^2\sigma_{n_k}^2\right) - 1 + \frac{1}{2}t^2\sigma_{n_k}^2 \right| = 0 \quad (4)$$

Für  $x \in \mathbb{R}$  mit  $|x| \leq \frac{1}{2}$  gilt  $|e^x - 1 - x| \leq \frac{1}{2} \sum_{j=2}^{\infty} |x|^j \leq x^2$

$$\implies \sum_{k=1}^{r_n} \left| \underbrace{\exp\left(-\frac{1}{2}t^2\sigma_{n_k}^2\right) - 1 + \frac{1}{2}t^2\sigma_{n_k}^2}_{=x} \right| \leq \frac{1}{4}t^4 \sum_{k=1}^{r_n} \sigma_{n_k}^4$$

Wegen  $\sum_{k=1}^{r_n} \sigma_{n_k}^4 \leq \max\{\sigma_{n_k}^2 | 1 \leq k \leq r_n\} \cdot \underbrace{\sum_{k=1}^{r_n} \sigma_{n_k}^2}_{=1} \xrightarrow{n \rightarrow \infty, (3)} 0$

Also (3)  $\implies$  (4) ■

**Beispiel 5.3 (Rekorde)**

Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum und  $X_1, X_2, \dots$  eine Folge von unabhängigen identisch verteilten Zufallsvariablen darauf mit absolutstetiger Verteilungsfunktion  $F$ . Setze:

$$R_n := \begin{cases} 1, & \text{falls } X_n > X_i, i = 1, \dots, n-1 \\ 0, & \text{sonst} \end{cases}$$

$R_n = 1 \iff n\text{-ter Versuch ist ein Rekord. } F \text{ stetig} \implies P(X_i = X_j) = 0 \forall i \neq j$

$$\begin{aligned} \implies A &:= \{\omega \in \Omega \mid \exists i \neq j, X_i(\omega) = X_j(\omega)\} \\ &= \bigcup_{i,j \in \mathbb{N}, i \neq j} \{X_i = X_j\} \\ \implies P(A) &= 0 \end{aligned}$$

Sei  $S_n$  die Menge der Permutationen der Zahlen  $1, \dots, n$ . Sei  $\Psi_n : \Omega \rightarrow S_n$  gegeben durch

$$\Psi_n = \pi \iff X_{\pi(1)} < X_{\pi(2)} < \dots < X_{\pi(n)}$$

$\Psi_n$  ist messbar, da  $\Psi^{-1}(\{\pi\}) = \bigcap_{i=1}^n \underbrace{\{X_{\pi(i)} < X_{\pi(i+1)}\}}_{\in \mathcal{A}}$ . Beispiel 3.5  $\implies (X_{\pi(1)}, \dots, X_{\pi(n)}) \stackrel{d}{=}$

$(X_1, \dots, X_n) \forall \pi \in S_n$ .

Ist  $B := \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1 < \dots < x_n\}$  so gilt:

$$\begin{aligned} P(\Psi_n = \pi) &= P((X_{\pi(1)}, \dots, X_{\pi(n)}) \in B) \\ &= P((X_1, \dots, X_n) \in B) \\ &= P(\Psi_n = \text{id}) \quad \text{unabhängig von } \pi \end{aligned}$$

$$\begin{aligned} \implies P(\Psi_n = \pi) &= \frac{1}{n!} \quad \forall \pi \in S_n \text{ und} \\ P(R_n = 1) &= P(\Psi_n \in \{\pi \in S_n \mid \pi(n) = n\}) = \frac{1}{n} \\ \implies R_n &\sim B(1, \frac{1}{n}) \text{ sind also nicht identisch verteilt} \end{aligned}$$

Wegen  $\{R_{n+1} = 1\} \cap \{\Psi_n = \pi\} = \{\Psi_{n+1} = \tilde{\pi}\}$  mit

$$\tilde{\pi}(i) = \begin{cases} \pi(i) & , i \leq n \\ n+1 & , i = n+1 \end{cases}$$

folgt:

$$P(\Psi_n = \pi, R_{n+1} = 1) = \frac{1}{(n+1)!} = \underbrace{P(\Psi_n = \pi)}_{=\frac{1}{n!}} \underbrace{P(R_{n+1} = 1)}_{=\frac{1}{n+1}} \quad \forall \pi \in S_n$$

$$\begin{aligned} \implies \Psi_n \text{ und } R_{n+1} &\text{ sind unabhängig} \\ \implies \text{Da } (R_1, \dots, R_n) &= G(\Psi_n) \text{ sind } R_{n+1} \text{ und } (R_1, \dots, R_n) \text{ unabhängig} \\ \implies P(R_{i_1} = j_1, \dots, R_{i_n} = j_n) &= \\ P(R_{i_1} = j_1, \dots, R_{i_{n-1}} = j_{n-1}) \cdot P(R_{i_n} = j_n) &= \dots \\ P(R_{i_1} = j_1) \cdot \dots \cdot P(R_{i_n} = j_n) &\text{ für } i_1 < i_2 < \dots < i_n, j_1, \dots, j_n \in \{0, 1\}. \\ \implies \text{die Zufallsvariablen } (R_n)_{n \in \mathbb{N}} &\text{ sind unabhängig} \end{aligned}$$

Wie viele Rekorde gibt es unter den ersten  $n$  Versuchen?

$$S_n := \sum_{i=1}^n R_i$$

Es gilt:

$$ES_n = \sum_{k=1}^n ER_k = \sum_{k=1}^n \frac{1}{k}$$

$$\text{Var}(S_n) = \sum_{k=1}^n \text{Var}(R_k) = \sum_{k=1}^n \frac{1}{k} \left(1 - \frac{1}{k}\right) = \sum_{k=1}^n \frac{1}{k} - \sum_{k=1}^n \frac{1}{k^2}$$

Insbesondere:

$$\frac{ES_n}{\log n} \xrightarrow{n \rightarrow \infty} 1, \quad \frac{\text{Var}(S_n)}{\log n} \xrightarrow{n \rightarrow \infty} 1$$

Mit dem zentralen Grenzwertsatz (ZGWS) bekommen wir genauere Aussagen: Sei

$$X_{n_k} = R_k, r_n = n \implies s_n = (\text{Var}(S_n))^{\frac{1}{2}}$$

Überprüfen der Lyapunov-Bedingung ( $\delta = 1$ ):

$$E|R_k - \underbrace{ER_k}_{=\frac{1}{k}}|^3 = \underbrace{P(R_k=1)}_{=\frac{1}{k}} \left(1 - \frac{1}{k}\right)^3 + \underbrace{P(R_k=0)}_{=\frac{k-1}{k}} \left(\frac{1}{k}\right)^3 \leq \frac{2}{k}$$

$$\implies 0 \leq \frac{1}{s_n^3} \sum_{k=1}^n E|R_k - ER_k|^3 \leq \frac{1}{s_n^3} \sum_{k=1}^n \frac{2}{k} \rightarrow 0 \text{ für } n \rightarrow \infty$$

Der Zentrale Grenzwertsatz (Satz 5.10) liefert

$$\frac{S_n - ES_n}{\sqrt{\text{Var}(S_n)}} \xrightarrow{d} Z \sim N(0, 1)$$

bzw.

$$\frac{S_n - \log(n)}{\sqrt{\log(n)}} \xrightarrow{d} Z \sim N(0, 1)$$

Also für große  $n$ :  $P(\log(n) - 1,96\sqrt{\log(n)} \leq S_n \leq \log(n) + 1,96\sqrt{\log(n)}) \approx P(-1,96 \leq Z \leq 1,96) = 0,9$ .

**Beispiel 5.4** (G. Polya, 1930: Eine Wahrscheinlichkeitsaufgabe zur Kundenwerbung, oder: "Coupon Collector's Problem")

Urne mit  $n$  verschiedenen Kugeln, Ziehen mit Zurücklegen

$S_n$  = Anzahl der Züge, bis  $r_n = [\phi \cdot n]$ ,  $0 < \phi < 1$  verschiedene Kugeln gezogen werden. Es sei  $X_{nk}$  = Anzahl der bis zum Erhalt einer neuen Kugel nötigen Züge, wenn bereits  $k-1$  verschiedene Kugeln gezogen (und zurückgelegt) wurden.  $X_{n1} := 1$ .

$X_{nk} \sim \text{Geo}(\frac{n-k+1}{n})$ , d.h.  $P(X_{nk} = j) = (\frac{k-1}{n})^{j-1} \cdot \frac{n-k+1}{n}$ ,  $j = 1, 2, \dots$

Falls  $Y \sim \text{Geo}(p)$ ,  $p \in (0, 1]$  mit  $Y \equiv 1$  bei  $p = 1$ , gilt:

$$EY = \frac{1}{p}, \quad \text{Var}(Y) = \frac{1-p}{p^2}, \quad EY^4 \leq \frac{24}{p^4}$$

Für  $S_n = X_{n1} + \dots + X_{nr_n}$  erhalten wir  $\mu_n := ES_n = \sum_{k=1}^{r_n} \frac{n}{n-k+1}$  und  $s_n^2 =$

$$\text{Var}(S_n) = \sum_{k=1}^{r_n} \frac{\frac{k-1}{n}}{(\frac{n-k+1}{n})^2} = n \sum_{k=1}^{r_n} \frac{k-1}{(n-k+1)^2}.$$

Wir prüfen die Lyapunov-Bedingung mit  $\delta = 2$ : Für  $Y \sim \text{Geo}(p)$  gilt:

$$E(Y - \frac{1}{p})^4 \leq E(\max\{Y, \frac{1}{p}\})^4 \leq EY^4 + \frac{1}{p^4} \leq \frac{25}{p^4}.$$

Insbesondere ist damit

$$\sum_{k=1}^{r_n} E|X_{nk} - \mu_{nk}|^4 \leq 25 \cdot \sum_{k=1}^{r_n} \frac{1}{(1 - \frac{k-1}{n})^4} \leq 25 \cdot [\phi \cdot n] \cdot \frac{1}{(1 - \phi)^4} = O(n).$$

Wegen  $s_n^2 \geq n \cdot \frac{1}{n^2} \sum_{k=1}^{r_n} (k-1) = \frac{1}{n} \cdot \frac{1}{2} (r_n - 1) \cdot r_n \geq \frac{1}{2n} (\phi n - 1) \phi n = \Theta(n)$  folgt damit

$$\lim_{n \rightarrow \infty} \left( \frac{1}{s_n^{2+2}} \sum_{k=1}^{r_n} E|X_{nk} - \mu_{nk}|^{2+2} \right) = 0.$$

Also folgt mit dem Zentralen Grenzwertsatz:

$$\frac{S_n - ES_n}{\sqrt{\text{Var}(S_n)}} \xrightarrow{d} Z \sim N(0, 1)$$

Weiter gilt:  $\frac{1}{n} ES_n = \sum_{k=1}^{r_n} \frac{1}{n} \cdot \frac{1}{1 - \frac{k-1}{n}} = \int_0^{\frac{r_n}{n}} \frac{1}{1-x} dx + O(\frac{1}{n}) = \int_0^{\frac{r_n}{n}} \frac{1}{1-x} dx + O(\frac{1}{n}) = \int_0^{\phi} \frac{1}{1-x} dx + O(\frac{1}{n}) = -\log(1 - \phi) + O(\frac{1}{n}).$

Analog:  $\lim_{n \rightarrow \infty} (\frac{1}{n} \text{Var}(S_n^2)) = \int_0^{\phi} \frac{x}{(1-x)^2} dx = \frac{\phi}{1-\phi} + \log(1 - \phi).$

Mit  $a(\phi) = -\log(1 - \phi)$ ,  $b(\phi) := \sqrt{\frac{\phi}{1-\phi} + \log(1 - \phi)}$  folgt:

$$\frac{S_n - a(\phi)n}{b(\phi)\sqrt{n}} \xrightarrow{d} Z \sim N(0, 1)$$

### Beispiel (Numerisches Beispiel)

Wie groß muss ihr Bekanntenkreis sein, damit mit einer Wahrscheinlichkeit von mindestens 0,95 an 180 Tagen im Jahr Geburtstag gefeiert werden kann?

$$\text{Also: } n = 365, \phi = \frac{180}{365}, S_n \leq k \iff \underbrace{\frac{S_n - a(\phi)n}{b(\phi)\sqrt{n}}}_{\approx Z} \leq \frac{k - a(\phi)n}{b(\phi)\sqrt{n}}$$

$$\underbrace{\Phi\left(\frac{k - a(\phi)n}{b(\phi)\sqrt{n}}\right)}_{=1,645} \geq 0,95 \iff k \geq a(\phi)n + 1,645 \cdot b(\phi)\sqrt{n} \implies k \geq 266.$$

Für  $\phi = 1$  kann man den Zentralen Grenzwertsatz nicht mehr anwenden:

$$r_n = n, \text{Var}(S_n) = n \sum_{k=1}^n \frac{k-1}{(n-k+1)^2} = n \sum_{k=1}^n \frac{n-k}{k^2} = n^2 \sum_{k=1}^n \frac{1}{k^2} - n \sum_{k=1}^n \frac{1}{k} = n^2 \cdot \frac{\pi^2}{6} + o(n^2).$$

$\text{Var}(X_{n,n}) = \frac{1 - \frac{1}{n}}{\frac{1}{n^2}} = n^2 + o(n^2) \implies$  bei großem  $n$  steckt etwa  $\frac{6}{\pi^2} \approx 0,61$  der Variabilität der Summe im letzten Summanden. Wir können jetzt eine andere Skalierung finden, allerdings ist die Grenzverteilung dann keine Normalverteilung mehr!

Sei  $A_{m,i}$  das Ereignis, dass die Kugel  $i$  in den ersten  $m$  Ziehungen nicht auftaucht

$$\Rightarrow \{S_n > m\} = \bigcup_{i=1}^n A_{m,i}.$$

( $S_n$  ist die Anzahl der Züge, bis alle  $n$  verschiedenen Kugeln mindestens einmal gezogen worden sind)

Mit der Siebformel:

$$P(S_n > m) = \sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} P\left(\bigcap_{l=1}^k A_{m,i_l}\right) = \sum_{k=1}^{n-1} (-1)^{k+1} \binom{n}{k} \left(1 - \frac{k}{n}\right)^m$$

Sei  $c \in \mathbb{R}$  fest,  $m_n = [n \log(n) + cn]$ . Für  $x > -1$  gilt  $\log(1+x) \leq x$ . Damit:  $\log\left(\binom{n}{k} \left(1 - \frac{k}{n}\right)^{m_n}\right) \leq k \log(n) - \log(k!) + \log\left(1 - \frac{k}{n}\right) (n \log(n) + cn - 1) \stackrel{\text{s.o.}}{\leq} k \log(n) - \log(k!) - \frac{k}{n} (n \log(n) + cn - 1) \leq -ck + \frac{k}{n} - \log(k!)$

$$\Rightarrow \left| (-1)^{k+1} \binom{n}{k} \left(1 - \frac{k}{n}\right)^{m_n} \right| \leq \frac{1}{k!} \exp\left(\frac{k}{n} - ck\right) \quad \forall c \in \mathbb{R}.$$

$$\text{Ähnlich: } \lim_{n \rightarrow \infty} (-1)^{k+1} \binom{n}{k} \left(1 - \frac{k}{n}\right)^{m_n} = (-1)^{k+1} \frac{1}{k!} e^{-ck} \quad \forall k \in \mathbb{N}.$$

Insgesamt:

$$\begin{aligned} \lim_{n \rightarrow \infty} \left( P\left(\frac{S_n - n \log(n)}{n} > c\right) \right) &= \lim_{n \rightarrow \infty} (P(S_n > m_n)) \\ &= \lim_{n \rightarrow \infty} \left( \sum_{k=1}^{n-1} (-1)^{k+1} \binom{n}{k} \left(1 - \frac{k}{n}\right)^{m_n} \right) \\ &= \lim_{n \rightarrow \infty} \left( \sum_{k=1}^{\infty} (-1)^{k+1} \binom{n}{k} \left(1 - \frac{k}{n}\right)^{m_n} \right) \\ &\stackrel{\text{maj. Konv.}}{=} \sum_{k=1}^{\infty} \lim_{n \rightarrow \infty} \left( (-1)^{k+1} \binom{n}{k} \left(1 - \frac{k}{n}\right)^{m_n} \right) \\ &= \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k!} (e^{-c})^k \\ &= 1 - e^{-e^{-c}} \end{aligned}$$

Das Wahrscheinlichkeitsmaß auf  $(\mathbb{R}, \mathfrak{B})$  mit der Verteilungsfunktion  $F(x) = e^{-e^{-x}} \forall x \in \mathbb{R}$  heißt **Gumbel-Verteilung**.

Also gilt:

$$\frac{S_n - n \log(n)}{n} \xrightarrow{d} Z \sim \text{Gumbel}.$$

### Beispiel (Variation des numerischen Beispiels von oben)

Der Bekanntenkreis soll jetzt so groß sein, dass mit einer Wahrscheinlichkeit von mindestens 0,95 täglich gefeiert werden kann.

$$\Rightarrow k \geq 365 \cdot \log(365) \cdot 365 \cdot 2,97 \approx 3237,51 \quad (?)$$





## 6 Zentraler Grenzwertsatz in $\mathbb{R}^n$

**Definition** Es sei  $X = (X_1, \dots, X_d)^T$  ein Zufallsvektor.

- a) Ist  $EX_i < \infty$ ,  $i = 1, \dots, d$ , so heißt  $EX := (EX_1, \dots, EX_d)$  **Erwartungswert** von  $X$ .
- b) Ist  $EX_i^2 < \infty$ ,  $i = 1, \dots, d$ , so heißt die  $d \times d$ -Matrix  $Cov(X) = (Cov(X_i, X_j))_{i,j=1,\dots,d}$  **Kovarianzmatrix** von  $X$ .  
 Beachte: Die Kovarianzmatrix ist symmetrisch, da  $Cov(X_i, X_j) = Cov(X_j, X_i)$ , und in der Diagonale steht die Varianz, denn  $Cov(X_i, X_i) = Var(X_i)$ , jeweils für  $i, j = 1, \dots, d$ .

**Bemerkung** a) Es gelten folgende Rechenregeln: Sei  $A \in \mathbb{R}^{s \times d}$ ,  $b \in \mathbb{R}^s$   
 $E(AX + b) = AEX + b$   
 $Cov(AX + b) = A \cdot Cov(X) A^T$

- b) Die 2. Rechenregel impliziert, dass Kovarianzmatrizen stets positiv semidefinit sind.

**Definition** Es sei  $X = (X_1, \dots, X_d)^T$  ein Zufallsvektor. Dann ist

$$\phi_X : \mathbb{R}^d \rightarrow \mathbb{C}, \quad \phi_X(t) = Ee^{it^T X}$$

die **charakteristische Funktion** zu  $X$ .

**Bemerkung**

- a) Es gilt für Zufallsvektoren  $X, Y$ :

$$X \stackrel{d}{=} Y \iff \phi_X(t) = \phi_Y(t) \quad \forall t \in \mathbb{R}^d \iff t^T X \stackrel{d}{=} t^T Y \quad \forall t \in \mathbb{R}^d$$

- b) Die **Verteilungskonvergenz** für Zufallsvektoren sei definiert durch  
 $X_n \xrightarrow{d} X : \iff Eh(X_n) \rightarrow Eh(X) \quad \forall h \in \mathbb{C}_b(\mathbb{R}^d)$ . (vgl. Satz 5.5)  
 Auch hier gelten  
 $X_n \xrightarrow{d} X \iff \phi_n(t) \rightarrow \phi(t) \quad \forall t \in \mathbb{R}^d$ . (vgl. Satz 5.9)  
 und das “Continuous Mapping Theorem”. (vgl. Satz 5.6)

**Satz 6.1 (Cramér-Wold-Technik)**

Es seien  $X, X_1, X_2, \dots$   $d$ -dimensionale Zufallsvektoren. Dann gilt:

$$X_n \xrightarrow{d} X \iff c^T X_n \xrightarrow{d} c^T X \quad \forall c \in \mathbb{R}^d$$

**Beweis**

“ $\Rightarrow$ ”: folgt aus dem “Continuous Mapping Theorem” mit  $h(x) := c^T x$ .

“ $\Leftarrow$ ”:  $c^T X_n \xrightarrow{d} c^T X \quad \forall c \in \mathbb{R}^d \xrightarrow{\text{S.5.9}} Ee^{itc^T X_n} \rightarrow Ee^{itc^T X} (n \rightarrow \infty) \quad \forall t \in \mathbb{R}, \forall c \in \mathbb{R}^d$ .

$\Rightarrow \phi_n(c) \rightarrow \phi(c) \quad \forall c \in \mathbb{R}^d \quad \Rightarrow X_n \xrightarrow{d} X.$  ■

**6.1 Mehrdimensionale Normalverteilung****Definition**

Der Zufallsvektor  $X = (X_1, \dots, X_n)^T$  besitzt eine  **$d$ -dimensionale Normalverteilung**, falls  $c^T X$  eine eindimensionale Normalverteilung besitzt  $\forall c \in \mathbb{R}^d$

**Bemerkung**

$X$  habe eine  $d$ -dimensionale Normalverteilung.

Setze  $c := e_i$  (Einheitsvektor) für ein  $i \in \{1, \dots, d\} \Rightarrow X_i$  ist normalverteilt.

$\Rightarrow \exists EX_i = \mu_i; \text{Var}(X_i) < \infty; EX_i^2 < \infty \Rightarrow \text{Cov}(X_i, X_j) \stackrel{\text{C.S.U.}}{<} \infty$ .

Sei  $\Sigma := \text{Cov}(X)$ . Weiter gilt:  $E(c^T X) = c^T \mu; \text{Var}(c^T X) = c^T \Sigma c$ .

$\Rightarrow c^T X \sim N(c^T \mu, c^T \Sigma c) \xrightarrow{\text{St.1, Bsp.12.3}} \phi_{c^T X}(t) = Ee^{itc^T X} = e^{ic^T \mu t - \frac{1}{2} c^T \Sigma c t^2} \quad \forall t \in \mathbb{R}.$

$\Rightarrow \phi_X(t) = Ee^{it^T X} = \phi_{t^T X}(1) = e^{it^T \mu - \frac{1}{2} t^T \Sigma t}, t \in \mathbb{R}.$

Wegen obiger Bemerkung, Teil a) folgt:

Die Normalverteilung ist durch  $\mu$  und  $\Sigma$  festgelegt. Schreibweise:  $X \sim N_d(\mu, \Sigma)$

**Lemma 6.1** Sei  $X \sim N_d(\mu, \Sigma)$ ,  $A \in \mathbb{R}^{s \times d}$ ,  $b \in \mathbb{R}^s$ . Dann gilt:

$Y := AX + b \sim N_s(A\mu + b, A\Sigma A^T)$

**Beweis**

$$\begin{aligned} \phi_Y(t) &= Ee^{it^T(AX+b)} \\ &= e^{it^T b} Ee^{it^T AX} \\ &= e^{it^T b} \phi_X(A^T t) \\ &= e^{it^T(b+A\mu) - \frac{1}{2} t^T (A\Sigma A^T) t} \end{aligned}$$

■

**Satz 6.2 (Existenzsatz)**

Sei  $\mu \in \mathbb{R}^d$  und  $\Sigma \in \mathbb{R}^{d \times d}$  eine beliebige symmetrische, positiv semidefinite Matrix. Dann existiert ein  $d$ -dimensionaler Zufallsvektor  $X$  mit  $X \sim N_d(\mu, \Sigma)$ .

**Beweis**

Sei  $Y = (Y_1, \dots, Y_d)$ , wobei  $Y_1, \dots, Y_d$  unabhängig und  $Y_k \sim N(0, 1)$ ,  $k = 1, \dots, d$ . Die Existenz dieser Konstruktion ist mit Satz 3.3 gegeben. Da  $c^T Y \sim N(0, c^T c)$ , ist  $Y \sim N_d(0, I_d)^1$

$\Sigma$  positiv semidefinit  $\Rightarrow \Sigma = AA^T$  mit einem  $A \in \mathbb{R}^{d \times d} \xrightarrow{\text{L.6.1}} X := AY + \mu \sim N_d(\mu, \Sigma).$  ■

<sup>1</sup>das ist die  $d$ -dimensionale Standardnormalverteilung

**Satz 6.3** Sei  $X \sim N_d(\mu, \Sigma)$  und  $\Sigma$  nicht singulär. Dann besitzt  $X$  eine Dichte der Form

$$f(x) = \frac{1}{(2\pi)^{\frac{d}{2}} |\det \Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right), \quad x \in \mathbb{R}^d$$

**Beweis** Sei  $\Sigma = AA^T$  und  $X = A \cdot Y + \mu$  mit  $Y \sim N_d(0, I_d)$ .  
Dichte von  $Y$ :

$$f_Y(y_1, \dots, y_d) = \prod_{j=1}^d \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y_j^2} = \frac{1}{(2\pi)^{\frac{d}{2}}} \exp\left(-\frac{1}{2}y^T y\right)$$

Sei  $\Psi(y) = Ay + \mu$ .  $\Psi$  ist bijektiv,  $\Sigma$  regulär.

$$\xrightarrow{\text{Satz 3.6}} \Rightarrow f_X(x) = \frac{1}{|\det A|} \cdot f_Y(A^{-1}(x - \mu))$$

Beachte:  $\det \Sigma = (\det A)^2$ ,  $\Sigma^{-1} = (A^{-1})^T (A^{-1})$ . ■

**Bemerkung** Ist  $\det \Sigma = 0 \Rightarrow \exists a \in \mathbb{R}^d$ ,  $a \neq 0$  mit  $a^T \Sigma a = 0 \Rightarrow \text{Var}(a^T X) = 0$ .  
 $N(\mu, \Sigma)$  ist dann auf  $H = \{x \in \mathbb{R}^d \mid a^T x = a^T \mu\}$  konzentriert, d.h.  $P^X(H) = 1$ .  
Wegen  $\lambda^d(H) = 0$  folgt mit dem Satz von Radon-Nikodym:  $\nexists$  Dichte.

## 6.2 Zentraler Grenzwertsatz in $\mathbb{R}^d$

**Satz 6.4** Es sei  $(X_n)_{n \in \mathbb{N}}$  eine Folge von unabh. u. identisch verteilten  $d$ -dim Zufallsvektoren mit Erwartungsvektor  $\mu$  und Kovarianzmatrix  $\Sigma$ . Dann gilt für  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ :

$$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} Z, \quad Z \sim N_d(0, \Sigma).$$

**Beweis** Sei  $Z_n := \sqrt{n}(\bar{X}_n - \mu)$ .

Nach Satz 6.1 ist z.z.  $c^T Z_n \xrightarrow{d} c^T Z \quad \forall c \in \mathbb{R}^d$ .

Wegen  $\text{Var}(c^T Z_n) = \frac{1}{n} \sum_{i=1}^n \text{Var}(c^T X_i) = c^T \Sigma c$ ,  $E c^T Z_n = 0$  können wir o.B.d.A.  $c^T \Sigma c > 0$  annehmen (andernfalls ist  $c^T Z_n \equiv 0$ ).

$$\begin{aligned} 1 - \dim \text{ ZGWS : } \quad & \frac{c^T Z_n}{\sqrt{c^T \Sigma c}} = \frac{\sum_{j=1}^n c^T X_j - n c^T \mu}{\sqrt{n c^T \Sigma c}} \xrightarrow{d} Z_0, \quad Z_0 \sim N(0, 1) \\ \Rightarrow \quad & c^T Z_n \xrightarrow{d} \sqrt{c^T \Sigma c} \cdot Z_0 \sim N(0, c^T \Sigma c) \end{aligned} \quad \blacksquare$$

**Beispiel 6.1 ( $\chi^2$ -Anpassungstest)** Es seien  $X_1, X_2$  unabh. u. identisch verteilte,  $d$ -dim. Zufallsvektoren mit

$$P(X_1 = e_k) = p_k, \quad k = 1, \dots, d, \quad \sum_{k=1}^d p_k = 1.$$

Dann hat  $S_n = \sum_{k=1}^n X_k$  eine Multinomialverteilung (vgl. Sto. I) mit Zähldichte:

$$P(S_n = (k_1, \dots, k_d)) = \frac{n!}{k_1! \dots k_d!} p_1^{k_1} \dots p_d^{k_d}$$

für  $k_1, \dots, k_d \in \mathbb{N}_0$ ,  $k_1 + \dots + k_d = n$ .

Weiter gilt:  $EX_1 = p := (p_1, \dots, p_d)^T$ ,  $\text{Cov}(X_1) = \Sigma$  mit

$$(\Sigma)_{ij} = \begin{cases} p_i(1-p_i), & i = j \\ -p_i p_j, & i \neq j \end{cases} \Rightarrow \Sigma = \text{diag}(p) - pp^T.$$

ZGWS (Satz 6.4):

$$\frac{1}{\sqrt{n}}(S_n - np) \xrightarrow{d} Z, \quad Z \sim N_d(0, \Sigma)$$

Anmerkung: Wir kennen  $p_1, \dots, p_d$  nicht, nur die Realisierungen von  $X_1, \dots, X_n$ . Betrachte die Testgröße  $T_n := \sum_{i=1}^d \frac{1}{np_i} (S_{n,i} - np_i)^2$ .

Aufgabe: Zu  $(p_1, \dots, p_d)$ ,  $X, n$  gegeben, bestimme  $c_\alpha$  mit  $P(T_n > c_\alpha) = \alpha$ . Also: Bestimme Verteilung von  $T_n$ .

Lösung: Approximativ. Sei  $h: \mathbb{R}^d \rightarrow \mathbb{R}$ ,  $h(x_1, \dots, x_d) := \sum_{j=1}^d \frac{x_j^2}{p_j}$ .

$$h \text{ stetig} \xrightarrow{\text{Cont. mapping}} \Rightarrow T_n = h\left(\frac{1}{\sqrt{n}}(S_n - np)\right) \xrightarrow{d} h(Z), \quad Z \sim N_d(0, \Sigma)$$

Welche Verteilung hat  $h(Z)$ ?

Sei  $\tilde{Z} = \text{diag}\left(\frac{1}{\sqrt{p_1}}, \dots, \frac{1}{\sqrt{p_d}}\right) \cdot Z$ .  $\xrightarrow{\text{Lemma 6.1}} \Rightarrow \tilde{Z} \sim N_d(0, \tilde{\Sigma})$  wobei

$$\begin{aligned} \tilde{\Sigma} &= \text{diag}\left(\frac{1}{\sqrt{p_1}}, \dots, \frac{1}{\sqrt{p_d}}\right) \cdot (\text{diag}(p) - pp^T) \cdot \text{diag}\left(\frac{1}{\sqrt{p_1}}, \dots, \frac{1}{\sqrt{p_d}}\right) \\ &= I_d - \underbrace{(\sqrt{p_1}, \dots, \sqrt{p_d})^T \cdot (\sqrt{p_1}, \dots, \sqrt{p_d})}_{=: r} \\ &= I_d - rr^T \end{aligned}$$

Es gilt:  $\|r\| = 1 \Rightarrow \exists$  orthogonale Matrix  $A = (r, *) \in \mathbb{R}^{d \times d}$ . Sei  $Y := A^T \tilde{Z} \Rightarrow Y \sim N_d(0, \Sigma_Y)$ , wobei  $\Sigma_Y = A^T \tilde{\Sigma} A = I_d - \text{diag}(1, 0, \dots, 0) = \text{diag}(0, 1, \dots, 1)$ .

$\Rightarrow Y^T Y \stackrel{d}{=} \sum_{i=1}^{d-1} W_i^2$ ,  $W_i \sim N(0, 1)$  unabh.  $\Rightarrow h(Z) = \tilde{Z}^T \tilde{Z} = Y^T Y \sim \chi_{d-1}^2$ , Chi<sup>2</sup>-Verteilung mit  $d-1$  Freiheitsgraden.

### Zahlenbeispiel:

Würfel wird 189 mal geworfen.

Ergebnis	1	2	3	4	5	6
	30	37	26	29	29	38

Ist der Würfel fair?

D.h.  $p_1 = \dots = p_6 = \frac{1}{6}$ .

$T_n = 3, 37$ ,  $d-1 = 5$ ,  $\alpha = 0,05$ ,  $p = (\frac{1}{6}, \dots, \frac{1}{6})$

$P(T_n > c_\alpha) \stackrel{!}{=} 0,05 \Leftrightarrow 1 - F_{\chi_5^2}(c_\alpha) \stackrel{!}{=} 0,05 \Rightarrow c_\alpha = 11,1$ . d.h. Nullhypothese „Würfel fair“ kann nicht abgelehnt werden.

## 7 Bedingte Erwartungswerte und Bedingte Verteilungen

Sei  $(\Omega, \mathcal{A}, P)$  ein W'Raum,  $(\Omega', \mathcal{A}')$  ein Messraum,  $Y : \Omega \rightarrow \Omega'$  sei  $(\mathcal{A}, \mathcal{A}')$ -messbar und nehme die Werte  $y_1, \dots, y_n \in \Omega'$  an.  $Y^{-1}(y_k) = \{\omega \in \Omega \mid Y(\omega) = y_k\} =: A_k \Rightarrow \Omega = A_1 + \dots + A_n$  und  $\sigma(Y) = \{\sum_{k \in I} A_k \mid I \subset \{1, \dots, n\}\}$ .

**Definition** Sei  $X : \Omega \rightarrow \mathbb{R}$  eine ZV mit  $E|X| < \infty$ . Dann ist der bedingte Erwartungswert von  $X$  unter der Bedingung  $Y = y_k$  definiert durch:

$$E[X|Y = y_k] := \frac{1}{P(A_k)} \int_{A_k} X dP, \quad k = 1, \dots, n$$

Falls  $X$  diskret mit  $x_1, \dots, x_m$ :

$$\begin{aligned} E[X|Y = y_k] &= \frac{1}{P(Y = y_k)} \sum_{j=1}^m x_j \cdot P(X = x_j, Y = y_k) \\ &= \sum_{j=1}^m x_j \cdot P(X = x_j | Y = y_k) \end{aligned}$$

**Definition** Der *bedingte Erwartungswert von  $X$  gegeben  $Y$*  ist  $E[X|Y] : \Omega \rightarrow \mathbb{R}$  mit

$$E[X|Y](\omega) := \sum_{k=1}^n E[X|Y = y_k] \cdot \mathbf{1}_{[Y=y_k]}(\omega)$$

**Bemerkung** a) Offenbar ist  $E[X|Y]$   $(\sigma(Y), \mathfrak{B})$ -messbar.

b) Sei  $Z := E[X|Y]$ . Dann gilt

$$\begin{aligned} \int_{A_k} Z dP &= \int_{\Omega} \mathbf{1}_{A_k} Z dP \\ &= E[X, Y = y_k] \cdot P(A_k) \\ &= \int_{A_k} X dP \end{aligned}$$

Wegen der Struktur von  $\sigma(Y)$  folgt auch

$$\int_A Z dP = \int_A X dP \quad \forall A \in \sigma(Y)$$

c)  $E[X|Y] = g(Y)$  mit

$$g(y) = \sum_{k=1}^n E[X|Y = y_k] \cdot \mathbf{1}_{\{y_k\}}(y)$$

- d) Offenbar hängt die Definition von  $E[X|Y]$  nur davon ab, auf welchen Mengen  $A_k$   $Y$  die verschiedenen Werte annimmt, nicht aber welche Werte das genau sind.

Deshalb schreibt man auch:

$$E[X|Y] = E[X|\sigma(Y)]$$

**Beispiel 7.1** Sei  $([0, 1), \mathfrak{B}_{[0,1)}, \underbrace{\lambda_{[0,1)}}_{=:P}), X(\omega) = \omega$

- Hier fehlt ein Bild -

$$A_k = \left[ \frac{k-1}{n}, \frac{k}{n} \right), k = 1, \dots, n, \quad \mathfrak{F} := \left\{ \sum_{k \in I} A_k \mid I \subset \{1, \dots, n\} \right\}$$

$$\begin{aligned} E[X, A_k] &= \frac{1}{P(A_k)} \int_{A_k} \omega P(d\omega) \\ &= n \int_{\frac{k-1}{n}}^{\frac{k}{n}} \omega d\omega \\ &= \frac{1}{2} \frac{2k-1}{n} \end{aligned}$$

$E[X, \mathfrak{F}]$  ist also eine „Approximation“ oder „Vergröberung“ von  $X$ . Bezüglich einer beliebigen Sub- $\sigma$ -Algebra  $\mathfrak{F} \subset \mathcal{A}$  wird der bedingte Erwartungswert wie folgt definiert:

**Definition** Sei  $X$  eine Zufallsvariable mit  $E|X| < \infty$  und  $\mathfrak{F} \subset \mathcal{A}$  eine Sub- $\sigma$ -Algebra von  $\mathcal{A}$ . Dann heißt  $Z : \Omega \rightarrow \mathbb{R}$  **eine Version des bedingten Erwartungswertes  $E[X|\mathfrak{F}]$  von  $X$  unter  $\mathfrak{F}$** , wenn gilt

(i)  $Z$  ist  $\mathfrak{F}$ -messbar

(ii)  $\int_A Z dP = \int_A X dP \quad \forall A \in \mathfrak{F}$

**Satz 7.1**

Der bedingte Erwartungswert existiert und ist bis auf Nullmengen eindeutig.

**Beweis** Sei  $X \geq 0$ . Durch

$$Q(A) := \int_A X(\omega) P(d\omega) \quad \forall A \in \mathfrak{F}$$

wird ein Maß auf  $(\Omega, \mathfrak{F})$  definiert (Satz 2.7).

Sei  $P_{\mathfrak{F}}$  die Einschränkung von  $P$  auf  $\mathfrak{F}$ . Offenbar  $Q \ll P_{\mathfrak{F}}$ . Satz von Radon-Nikodym  $\implies Q$  besitzt eine Dichte  $Z$  bzgl.  $P_{\mathfrak{F}}$  und  $Z$  ist nach Definition  $\mathfrak{F}$ -messbar.

Falls  $X$  beliebig:  $X = X^+ - X^-$

P-f.s. Eindeutigkeit: Seien  $Z, \tilde{Z}$  Versionen von  $E[X, \mathfrak{F}]$ .

$$\implies \int_A (Z - \tilde{Z}) dP = 0 \quad \forall A \in \mathfrak{F}$$

Wegen  $\{Z > \tilde{Z}\} \in \mathfrak{F}, \{Z < \tilde{Z}\} \in \mathfrak{F}$  folgt:

$$E|Z - \tilde{Z}| = \int_{\{Z > \tilde{Z}\}} (Z - \tilde{Z}) dP - \int_{\{Z < \tilde{Z}\}} (Z - \tilde{Z}) dP = 0$$

$\implies Z = \tilde{Z}$  P-f.s. ■

**Bemerkung** Der bedingte Erwartungswert ist also eigentlich die Äquivalenzklasse

$$E[X|\mathfrak{F}] = \left\{ Z \in L^1(\Omega, \mathfrak{F}, P) \mid \int_A Z dP = \int_A X dP \ \forall A \in \mathfrak{F} \right\}$$

Ein Element davon nennt man „Version“. Oft wird  $E[X|\mathfrak{F}]$  mit einer Version identifiziert.

**Definition** Sei  $A \in \mathfrak{F}$ . Eine Version von  $E[\mathbf{1}_A|\mathfrak{F}]$  bezeichnet man als **Version der bedingten Wahrscheinlichkeit**  $P(A|\mathfrak{F})$ .

**Bemerkung** Es gilt für  $B \in \mathfrak{F}$ :

$$\int_B P(A|\mathfrak{F}) dP \stackrel{(ii)}{=} \int_B \mathbf{1}_A dP = P(A \cap B)$$

### Satz 7.2

Sei  $X \in L^2(\Omega, \mathcal{A}, P)$  mit  $\|X\|^2 = EX^2$ . Dann gilt:

$$\|X - E[X|\mathfrak{F}]\|^2 = \inf \{ \|X - Y\|^2 \mid Y \in L^2(\Omega, \mathfrak{F}, P) \}$$

**Beweis** siehe Henze Stochastik II, S.214 ■

### Satz 7.3 (Rechenregeln für bedingte Erwartungswerte)

Es seien  $X, Y \in L^1(\Omega, \mathcal{A}, P)$ ,  $\mathfrak{F}, \mathfrak{F}_1, \mathfrak{F}_2$  Sub- $\sigma$ -Algebren von  $\mathcal{A}$ . Dann gilt:

a)  $E[aX + bY|\mathfrak{F}] = aE[X|\mathfrak{F}] + bE[Y|\mathfrak{F}]$  P-f.s.  $a, b \in \mathbb{R}$

b)  $E[E[X|Y]] = EX$

c)  $X \leq Y \implies E[X|\mathfrak{F}] \leq E[Y|\mathfrak{F}]$  P-f.s.

d) Für  $\mathfrak{F}_1 \subset \mathfrak{F}_2$  gilt  $E[E[X|\mathfrak{F}_2]|\mathfrak{F}_1] = E[X|\mathfrak{F}_1]$

Für  $\mathfrak{F}_1 \supset \mathfrak{F}_2$  gilt  $E[E[X|\mathfrak{F}_2]|\mathfrak{F}_1] = E[X|\mathfrak{F}_2]$

e) Falls  $Y$   $\mathfrak{F}$ -messbar und  $EXY < \infty$  gilt:

$$E[XY|\mathfrak{F}] = YE[X|\mathfrak{F}]$$

f) Falls  $X$  von  $\mathfrak{F}$  unabhängig ist (d.h. falls die  $X$  und  $\mathbf{1}_A \ \forall A \in \mathfrak{F}$  unabhängig sind), dann gilt:

$$E[X|\mathfrak{F}] = EX$$

**Bemerkung** Aus Satz 7.3 bekommt man:

1.  $X \equiv c \in \mathbb{R} \xrightarrow{\text{f)}} E[c|\mathfrak{F}] = c$
2.  $\mathfrak{F} = \{\emptyset, \Omega\} \xrightarrow{\text{f)}} E[X|\mathfrak{F}] = EX$
3.  $X \text{ } \mathfrak{F}\text{-messbar} \xrightarrow{\text{e)}} E[X|\mathfrak{F}] = X$
4.  $X \geq 0 \xrightarrow{\text{c)}} E[X|\mathfrak{F}] \geq 0 \text{ P-f.s.}$

**Beweis** von Satz 7.3:

a)

$$\begin{aligned}
 \int_A E[aX + bY|\mathfrak{F}]dP &= \int_A aX + bYdP \\
 &\stackrel{\text{Linearität}}{=} a \int_A XdP + b \int_A YdP \\
 &= a \int_A E[X|\mathfrak{F}]dP + b \int_A E[Y|\mathfrak{F}]dP \\
 &= \int_A (aE[X|\mathfrak{F}] + bE[Y|\mathfrak{F}])dP \quad \forall A \in \mathfrak{F}
 \end{aligned}$$

$\implies$  Behauptung, da  $aE[X|\mathfrak{F}] + bE[Y|\mathfrak{F}]$   $\mathfrak{F}$ -messbar und Radon-Nikodym-Dichte  $P$ -f.s. eindeutig.

b)

$$E[E[X|\mathfrak{F}]] = \int_{\Omega} E[X|\mathfrak{F}]dP = \int_{\Omega} XdP = EX$$

c)

$$\begin{aligned}
 A &:= \{\omega \in \Omega \mid E[X|\mathfrak{F}](\omega) > E[Y|\mathfrak{F}](\omega)\} \in \mathfrak{F} \\
 &= \bigcup_{n \in \mathbb{N}} \underbrace{\left\{ \omega \in \Omega \mid E[X|\mathfrak{F}](\omega) > E[Y|\mathfrak{F}](\omega) + \frac{1}{n} \right\}}_{A_n}
 \end{aligned}$$

Annahme:  $P(A) > 0 \implies \exists n \in \mathbb{N}$  mit  $P(A_n) > 0$

$$\begin{aligned}
 \implies 0 &\leq \int_{A_n} (Y - X)dP \\
 &= \int_{A_n} E[Y|\mathfrak{F}]dP - \int_{A_n} E[X|\mathfrak{F}]dP \\
 &= \int_{A_n} (E[Y|\mathfrak{F}] - E[X|\mathfrak{F}])dP \\
 &\leq -\frac{1}{n} \cdot P(A_n) \\
 &< 0 \quad \text{Widerspruch!}
 \end{aligned}$$



d) Z.z. Für  $\mathfrak{F}_1 \subset \mathfrak{F}_2$  gilt:  $E[E[X|\mathfrak{F}_2]|\mathfrak{F}_1] = E[X|\mathfrak{F}_1]$ . Sei  $A \in \mathfrak{F}_1 \implies A \in \mathfrak{F}_2$  und

$$\int_A E[X|\mathfrak{F}_1]dP = \int_A XdP = \int_A E[X|\mathfrak{F}_2]dP = \int_A E[E[X|\mathfrak{F}_2]|\mathfrak{F}_1]dP$$

$\implies$  Behauptung, da Radon-Nikodym-Dichte eindeutig.

Für  $\mathfrak{F}_1 \supset \mathfrak{F}_2$  ähnlich.

e) Mit algebraischer Induktion:

– Sei  $Y = \mathbf{1}_B$ ,  $B \in \mathfrak{F}$  und  $A \in \mathfrak{F}$  beliebig.

$$\int_A Y \cdot E[X|\mathfrak{F}]dP = \int_{A \cap B} E[X|\mathfrak{F}]dP = \int_{A \cap B} XdP = \int_A YXdP$$

Außerdem ist  $Y \cdot E[X|\mathfrak{F}]$   $\mathfrak{F}$ -messbar  $\implies$  Behauptung, da Radon-Nikodym-Dichte  $P$ -f.s. eindeutig.

– Linearität des Integrals + Teil a)  $\implies$  Aussage für  $Y \in \mathcal{E}, Y \geq 0$  :  
Bedingte Version des Satzes von der monotonen Konvergenz ( $\rightarrow$  Übung).

– Dann  $Y = Y^+ - Y^-$

f)

$$\begin{aligned} \int_A E[X|\mathfrak{F}]dP &= \int_A XdP \\ &= \int_{\Omega} \mathbf{1}_A XdP \\ &\stackrel{\text{unabh.}}{=} \int \mathbf{1}_A dP \cdot \underbrace{\int XdP}_{=EX} \\ &= \int_A EXdP \end{aligned}$$

$\implies$  Behauptung, da  $EX$   $\mathfrak{F}$ -messbar. ■

#### Satz 7.4 (Faktorisierungssatz)

Es seien  $(\Omega, \mathcal{A}), (\Omega', \mathcal{A}')$  Messräume und  $Y : \Omega \rightarrow \Omega'$  ein Zufallsgröße. Ist  $X : \Omega \rightarrow \mathbb{R}$  eine  $(\sigma(Y), \mathfrak{B})$ -messbare Zufallsvariable. Dann gibt es eine  $\mathfrak{B}$ -messbare Funktion  $g : \Omega' \rightarrow \mathbb{R}$  mit

$$X = g \circ Y.$$

**Beweis** Algebraische Induktion:

- (i) Sei  $X = \sum_{j=1}^n a_j \mathbf{1}_{A_j} \in \mathcal{E}$  mit  $a_j \geq 0, A_j \in \sigma(Y)$ .  
 $\implies A_j = Y^{-1}(A'_j), A'_j \in \mathcal{A}'$ . Wähle  $g = \sum_{j=1}^n a_j \mathbf{1}_{A'_j}$   
 $\implies X = g \circ Y$   
 $\implies$  Behauptung

- (ii) Sei  $X \geq 0$  und  $(\sigma(Y), \mathfrak{B})$ -messbar.  $\implies \exists (X_n) \subset \mathcal{E}, 0 \leq X_n \uparrow X$  und wegen  
 (i)  $\exists (\mathcal{A}', \mathfrak{B})$ -messbare Funktion  $g_n$  mit  $X_n = g_n \circ Y, n \in \mathbb{N}$ .

$$\implies X = \sup_{n \in \mathbb{N}} X_n = \sup_{n \in \mathbb{N}} (g_n \circ Y) = (\sup_{n \in \mathbb{N}} g_n) \circ Y$$

Wähle also  $g = \sup_{n \in \mathbb{N}} g_n$

- (iii)  $X = X^+ - X^- \xrightarrow{(ii)} X = g_1 \circ Y - g_2 \circ Y$ . Wähle  $g = g_1 - g_2$ . ■

**Bemerkung** Statt  $E[X|\sigma(Y)]$  schreiben wir auch  $E[X|Y]$  und wegen Satz 7.4  $\exists g : \Omega' \rightarrow \mathbb{R}$   $(\mathcal{A}', \mathfrak{B})$ -messbar mit  $E[X|Y] = g \circ Y$   $P$ -f.s.. Die Funktion  $g$  ist  $P^Y$ -f.s. eindeutig.

**Definition** Ist  $E[X|Y] = g \circ Y$  wie oben, so heißt  $E[X|Y = y] = g(y)$  (ein) **bedingter Erwartungswert von  $X$  unter der Bedingung  $Y = y$** .

**Satz 7.5**

Für alle  $A' \in \mathcal{A}'$  gilt:

$$\int_{A'} E[X|Y = y] P^Y(dy) = \int_{Y^{-1}(A')} X dP$$

**Beweis**

$$\int_{A'} E[X|Y = y] P^Y(dy) = \int_{A'} g dP^Y \stackrel{\text{Sa. 2.4}}{=} \int_{Y^{-1}(A')} g \circ Y dP = \int_{Y^{-1}(A')} X dP.$$

**Bemerkung** Für  $A \in \mathcal{A}$  heißt  $P(A|Y = y) := E[\mathbf{1}_A|Y = y]$  (eine) **bedingte Wahrscheinlichkeit von  $A$  unter der Bedingung  $Y = y$** . Bedingte Wahrscheinlichkeiten treten oft bei gekoppelten Zufallsexperimenten auf. Die folgende Sichtweise ist konstruktiver:

**Definition** Es seien  $(\Omega_1, \mathcal{A}_1), (\Omega_2, \mathcal{A}_2)$  messbare Räume. Eine Abbildung  $Q : \Omega_1 \times \mathcal{A}_2 \rightarrow [0, 1]$  mit

(i)  $\omega_1 \mapsto Q(\omega_1, A_2)$  ist  $\mathcal{A}_1$ -messbar  $\forall A_2 \in \mathcal{A}_2$ .

(ii)  $A_2 \mapsto Q(\omega_1, A_2)$  ist ein Wahrscheinlichkeitsmaß auf  $(\Omega_2, \mathcal{A}_2) \forall \omega_1 \in \Omega_1$

nennt man **Übergangskern** oder **Kern** von  $(\Omega_1, \mathcal{A}_1)$  nach  $(\Omega_2, \mathcal{A}_2)$ .

**Satz 7.6**

Es seien  $(\Omega_1, \mathcal{A}_1, P_1)$  ein Wahrscheinlichkeitsraum,  $(\Omega_2, \mathcal{A}_2)$  ein Messraum und  $Q$  ein Übergangskern von  $(\Omega_1, \mathcal{A}_1)$  nach  $(\Omega_2, \mathcal{A}_2)$ . Dann wird durch

$$P(A) := \int_{\Omega_1} \left( \int_{\Omega_2} \mathbf{1}_A(\omega_1, \omega_2) Q(\omega_1, d\omega_2) \right) P_1(d\omega_1)$$

ein Wahrscheinlichkeitsmaß  $P =: P_1 \otimes Q$  auf  $\mathcal{A}_1 \otimes \mathcal{A}_2$  definiert.  $P$  heißt **Koppelung** und ist das einzige Wahrscheinlichkeitsmaß auf  $\mathcal{A}_1 \otimes \mathcal{A}_2$  mit der Eigenschaft

$$P(A_1 \times A_2) = \int_{A_1} Q(\omega_1, A_2) P_1(d\omega_1) \quad (*)$$

## Beweis

1. Ähnlich wie in §3 zeigt man: für  $f : \Omega_1 \times \Omega_2 \rightarrow \mathbb{R}_+$ ,  $f$   $(\mathcal{A}_1 \otimes \mathcal{A}_2)$ -messbar ist  $\omega_1 \mapsto \int_{\Omega_2} f(\omega_1, \omega_2) Q(\omega_1, d\omega_2)$   $\mathcal{A}_1$ -messbar.
2. Für  $A = A_1 \times A_2$  ist  $\mathbf{1}_A(\omega_1, \omega_2) = \mathbf{1}_{A_1}(\omega_1) \mathbf{1}_{A_2}(\omega_2) \implies (*)$ .
3.  $P(\Omega_1 \times \Omega_2) = 1$  wegen  $(*)$ .  $P \geq 0$  ist klar.

$$\begin{aligned}
 P\left(\sum_{n=1}^{\infty} A_n\right) &= \int_{\Omega_1} \left( \int_{\Omega_2} \underbrace{\mathbf{1}_{\sum_{n=1}^{\infty} A_n}(\omega_1, \omega_2)}_{=\sum_{n=1}^{\infty} \mathbf{1}_{A_n}(\omega_1, \omega_2)} Q(\omega_1, d\omega_2) \right) P_1(d\omega_1) \\
 &= \sum_{n=1}^{\infty} \left( \int_{\Omega_1} \left( \int_{\Omega_2} \mathbf{1}_{A_n}(\omega_1, \omega_2) Q(\omega_1, d\omega_2) \right) P_1(d\omega_1) \right) \\
 &= \sum_{n=1}^{\infty} P(A_n).
 \end{aligned}$$

4. Eindeutigkeitssatz für Maße. ■

**Satz 7.7** Es seien  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum,  $(\Omega_1, \mathcal{A}_1)$  ein messbarer Raum,  $Y : \Omega \rightarrow \Omega_1$   $(\mathcal{A}, \mathcal{A}_1)$ -messbar und  $X$  ein  $d$ -dimensionaler Zufallsvektor. Dann existiert ein Kern  $Q$  von  $(\Omega_1, \mathcal{A}_1)$  nach  $(\mathbb{R}^d, \mathfrak{B}^d)$  derart, dass

$$P^{X,Y} = P^Y \otimes Q.$$

$Q$  ist eine Version der bedingten Verteilung von  $X$  unter  $Y$ . Schreibweise:

$$Q(y, \cdot) = P^X(\cdot | Y = y).$$

**Beweis** - ohne Beweis - ■

**Bemerkung** Für  $A \in \mathcal{A}$ ,  $B \in \mathfrak{B}^d$  gilt:

$$P(X \in B, Y \in A) = \int_A Q(y, B) P^Y dy = \int_A P^X(B | Y = y) P^Y(dy)$$

## Satz 7.8

Es seien  $\mu$  und  $\nu$   $\sigma$ -endliche Maße auf  $\mathcal{A}_1$  bzw.  $\mathfrak{B}^d$ .  $P^{(Y,X)}$  besitze eine Dichte  $f$  bezüglich  $\mu \otimes \nu$ . Es sei  $f_Y(y) := \int_{\mathbb{R}^d} f(x, y) \nu(dx)$  die (Rand-)Dichte von  $P^Y$  bzgl.  $\mu$ . Weiterhin sei

$$f(x|y) := \frac{f(x, y)}{f_Y(y)} \quad \text{und} \quad \frac{0}{0} := 0.$$

So wird durch

$$P^X(B | Y = y) := \int_B f(x|y) \nu(dx) \quad \forall B \in \mathfrak{B}^d, y \in \Omega_1$$

eine bedingte Verteilung von  $X$  unter der Bedingung  $Y = y$  definiert.

$f(\cdot | y)$  heißt **bedingte  $\nu$ -Dichte von  $X$  unter der Bedingung  $Y = y$** .

**Beweis**

$y \mapsto \int_B f(x|y)\nu(dx)$  ist messbar  $\forall B \in \mathfrak{B}^d$  (Satz von Tonelli),

$B \mapsto \int_B f(x|y)\nu(dx)$  ist ein Wahrscheinlichkeitsmaß  $\forall y \in \Omega_1$ .

Für  $A \in \mathcal{A}_1, B \in \mathfrak{B}^d$  gilt:

$$\begin{aligned} P^{(Y,X)}(A \times B) &= \int_{A \times B} f d(\mu \otimes \nu) \\ &= \int_A \left( \int_B f(x, y) \nu(dx) \right) \mu(dy) \\ &= \int_A \left( \int_B f(x|y) \nu(dx) \right) f_Y(y) \mu(dy) \\ &\stackrel{!}{=} \int_A P^X(B|Y=y) \underbrace{P^Y(dy)}_{=f_Y(y)\mu(dy)} \end{aligned}$$

■

**Satz 7.9**

Es seien  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum,  $Y : \Omega \rightarrow \mathbb{R}^d$  ein Zufallsvektor und  $X$  eine Zufallsvariable mit  $E|X| < \infty$ . Dann ist

$$h(y) := \int_{\mathbb{R}} x P^X(dx|Y=y)$$

ein bedingter Erwartungswert von  $X$  unter der Bedingung  $Y = y$ .

**Beweis** Nach 7.5:

$$\int_B E[X|Y=y] P^Y(dy) = \int_{Y^{-1}(B)} X dP.$$

Für  $B \in \mathfrak{B}^d$  und  $T(Y, X) := X \cdot (\mathbf{1}_B \circ Y)$  gilt:

$$\begin{aligned} \int_{Y^{-1}(B)} X dP &= \int T(Y, X) dP \\ &\stackrel{2.4}{=} \int T(y, x) P^{(Y,X)}(dy, dx) \\ &= \int x \mathbf{1}_B(y) P^{(Y,X)}(dy, dx) \\ &= \int_B \left( \int_{\mathbb{R}} x P^X(dx|Y=y) \right) P^Y(dy) \end{aligned}$$

$\stackrel{7.5}{\implies}$  Beh.

■

**Beispiel 7.2**

$U$  und  $V$  seien unabhängig und  $U(0, 1)$ -verteilt und entsprechen den zufälligen Seitenlängen eines Rechtecks. Es sei  $X = \text{Flächeninhalt des Rechtecks}$  und  $Y = \text{Umfang des Rechtecks}$ . Klar:  $X$  und  $Y$  sind nicht unabhängig.

Weiter ist  $f_{U,V}(u,v) = \begin{cases} 1 & 0 < u < 1 \text{ und } 0 < v < 1 \\ 0 & \text{sonst} \end{cases}$  die gemeinsame Dichte von  $U$

und  $V$ .

$\implies$  (Transformationssatz für Dichten)  $f_{X,Y}(x,y) = \frac{2}{\sqrt{y^2-16x}}$  für  $0 < x < 1$  und

$4\sqrt{x} < y < 2 + 2x$ ;  $f_X(x) = -\log x$  für  $0 < x < 1$ .

$\implies f(y|x) = -\frac{2}{\log x \sqrt{y^2-16x}}$  für  $4\sqrt{x} < y < 2 + 2x$ .

$\implies E[Y|X=x] = \int y \cdot f(y|x) dy = -\frac{4(1-x)}{\log x}$ .

### Beispiel 7.3 (Buffonsches Nadelproblem)

Wir werfen eine Nadel der Länge 1 zufällig auf einen unendlich langen Streifen der Breite 1. Wie groß ist die Wahrscheinlichkeit, dass die Nadel mindestens eine Wand des Korridors schneidet?

$X$  = Abstand der Nadelmitte von der linken Wand

$Y$  = Winkel der Nadel zum Lot

Annahme:  $X \sim U(0,1)$ ,  $Y \sim U(-\frac{\pi}{2}, \frac{\pi}{2})$  und  $X, Y$  unabhängig.

$A$  = Nadel schneidet die Wand =  $\{\omega \mid (X,Y)(\omega) \in B\}$  mit

$B = \{(x,y) \mid |y| < \frac{\pi}{2}, x \in [0, \frac{1}{2} \cos y] \cup [1 - \frac{1}{2} \cos y, 1]\}$

- hier fehlt eine Skizze -

Es ergibt sich:

$$\begin{aligned} P(A) &= P^{X,Y}(B) \\ &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^1 \mathbf{1}_B(x,y) P^X(dx|Y=y) P^Y(dy) \\ &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} P^X\left([0, \frac{\cos y}{2}] \cup [1 - \frac{\cos y}{2}, 1] \mid Y=y\right) P^Y(dy) \\ &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos y \cdot \frac{1}{\pi} dy \\ &= \frac{2}{\pi} \end{aligned}$$

So lässt sich zum Beispiel auch  $\pi$  näherungsweise bestimmen.



## 8 Martingale und Stoppzeiten

**Definition** Sei  $I \neq \emptyset$  eine beliebige Indexmenge und  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum.

- a) Eine Familie von Zufallsvariablen  $(X_t)_{t \in I}$  auf  $(\Omega, \mathcal{A}, P)$  heißt **stochastischer Prozess** ( $I \subset \mathbb{R}$ )
- b) Eine Familie von  $\sigma$ -Algebren  $(\mathfrak{F}_t)_{t \in I}$ , mit  $\mathfrak{F}_t \subset \mathcal{A}$  und  $\mathfrak{F}_s \subset \mathfrak{F}_t$ , für  $s \leq t$  heißt **Filtration**. Ein stochastischer Prozess  $(X_t)_{t \in I}$  heißt  $(\mathfrak{F}_t)_{t \in I}$ -**adaptiert**, falls  $X_t$   $\mathfrak{F}_t$ -messbar  $\forall t \in I$ .

**Bemerkung** Oft wird  $\mathfrak{F}_t := \sigma(\{X_s, s \leq t\})$  gewählt. Dann ist  $(\mathfrak{F}_t)_{t \in I}$  eine Filtration und  $X_t$  ist  $\mathfrak{F}_t$ -messbar.

**Definition** Gegeben sei ein Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$ ,  $I \subset \mathbb{R}$ , eine Filtration  $(\mathfrak{F}_t)_{t \in I}$  und ein dazu adaptierter stochastischer Prozess  $(X_t)_{t \in I}$ . Ist  $E|X_t| < \infty \forall t \in I$ , so heißt  $(X_t)_{t \in I}$  ein  $(\mathfrak{F}_t)_{t \in I}$ -**Martingal**, falls  $E[X_t | \mathfrak{F}_s] = X_s \forall s, t \in I, s \leq t$ .

Ist  $X_s \leq E[X_t | \mathfrak{F}_s]$  bzw.  $X_s \geq E[X_t | \mathfrak{F}_s]$ , so nennt man  $(X_t)_{t \in I}$  ein  $(\mathfrak{F}_t)_{t \in I}$ -**Submartingal** bzw.  $(\mathfrak{F}_t)_{t \in I}$ -**Supermartingal**.

**Bemerkung** a) Beim Martingal gilt:  $EX_s = E[E[X_t | \mathfrak{F}_s]] = EX_t \forall t \in I$ , d.h. der Erwartungswert ist konstant (wachsend beim Submartingal, fallend beim Supermartingal).

- b) Ist  $I = \mathbb{N}$ , so genügt z.z.:

$$E[X_{t+1} | \mathfrak{F}_t] = X_t \forall t \in \mathbb{N}$$

- c) Ist  $(F_t)_{t \in I}$  die natürliche Filtration, so sagt man oft nur  $(X_t)_{t \in I}$  ist ein Martingal.

**Beispiel 8.1** Sei  $I = \mathbb{N}$ ,  $(X_n)_{n \in \mathbb{N}}$  eine Folge von unabhängigen und identisch verteilten Zufallsvariablen mit Erwartungswert  $\mu$ . Sei  $S_n := \sum_{k=1}^n X_k \forall n \in \mathbb{N}$  und  $\mathfrak{F}_n := \sigma(S_1, \dots, S_n)$ . Dann gilt  $\forall n \in \mathbb{N} : E[S_{n+1} | \mathfrak{F}_n] = E[S_n | \mathfrak{F}_n] + E[X_{n+1} | \mathfrak{F}_n] = S_n + \mu$ .

Also:  $\mu = 0 \implies (S_n)$  ist Martingal  
 $\mu \leq 0 \implies (S_n)$  ist Supermartingal  
 $\mu \geq 0 \implies (S_n)$  ist Submartingal

**Beispiel 8.2** Sei  $(\mathfrak{F}_t)_{t \in I}$  eine Filtration und  $X$  eine Zufallsvariable mit  $E|X| < \infty$ . Sei  $X_t := E[X | \mathfrak{F}_t]$ . dann ist  $(X_t)_{t \in I}$   $(\mathfrak{F}_t)_{t \in I}$ -adaptiert und  $\forall s, t \in I, s \leq t$ :

$$E[X_t | \mathfrak{F}_s] = E[E[X | \mathfrak{F}_t] | \mathfrak{F}_s] \stackrel{S.7.3a)}{=} E[X | \mathfrak{F}_s] = X_s$$

$\implies (X_t)_{t \in I}$  ist ein  $(\mathfrak{F}_t)_{t \in I}$ -Martingal.

**Satz 8.1**

Ist  $(X_t)_{t \in I}$  ein  $(\mathfrak{F}_t)_{t \in I}$ -Martingal und  $\Phi : \mathbb{R} \rightarrow \mathbb{R}$  eine konvexe Funktion mit  $E|\Phi(X_t)| < \infty \forall t \in I$ , so ist  $(\Phi(X_t))_{t \in I}$  ein  $(\mathfrak{F}_t)_{t \in I}$ -Submartingal.

**Beweis** Sei  $s, t \in I, s \leq t : E[\Phi(X_t)|\mathfrak{F}_s] \stackrel{\text{Jensen}}{\geq} \Phi(\underbrace{E[X_t|\mathfrak{F}_s]}_{=X_s})$  ■

Im Folgenden:  $I = \{1, 2, \dots, n\}$  und  $X^* := \max_{1 \leq i \leq n} X_i$

**Satz 8.2 (Submartingal-Ungleichung von Doob)**

Ist  $(X_i)_{i=1, \dots, n}$  ein  $(\mathfrak{F}_i)_{i=1, \dots, n}$ -Submartingal, so gilt  $\forall c > 0 :$

$$c \cdot P(X^* > c) \leq \int_{\{X^* > c\}} X_n dP \leq EX_n^+$$

**Beweis** Sei  $A := \{X^* > c\}, A_i := \{X_1 \leq c, \dots, X_{i-1} \leq c, X_i > c\}, i = 1, \dots, n$

$$\implies A = A_1 + \dots + A_n, A_i \in \mathfrak{F}_i \text{ und } X_i > c \text{ auf } A_i, i = 1, \dots, n.$$

$$\implies \int_{A_i} X_n dP \stackrel{\text{bed. EW}}{=} \int_{A_i} E[X_n|\mathfrak{F}_i] dP \stackrel{\text{Sub-M.}}{\geq} \int_{A_i} X_i dP \geq cP(A_i), i = 1, \dots, n$$

Summation über  $i = 1, \dots, n \implies 1.$  Ungleichung

2. Ungleichung:  $X_n \cdot \mathbf{1}_A \leq X_n^+$  ■

**Satz 8.3 ( $L^p$ -Ungleichung von Doob)**

Es sei  $p > 1$  und  $(X_i)_{i=1, \dots, n}$  ein nicht-negatives  $(\mathfrak{F}_i)_{i=1, \dots, n}$ -Submartingal mit der Eigenschaft  $\sup_{i=1, \dots, n} EX_i^p < \infty$ . Dann gilt:

$$E(X^*)^p \leq \left( \frac{p}{p-1} \right)^p EX_n^p$$

**Beweis**

$$\begin{aligned} E(X^*)^p &= E \int_0^{X^*} p \cdot y^{p-1} dy \\ &= E \int_0^\infty p \cdot y^{p-1} \mathbf{1}_{[X^* \geq y]} dy \\ &\stackrel{\text{Fubini}}{=} \int_0^\infty p y^{p-1} \cdot P(X^* \geq y) dy \\ &\stackrel{\text{S.8.2}}{\leq} \int_0^\infty p \cdot y^{p-2} E[X_n \cdot \mathbf{1}_{[X^* \geq y]}] dy \\ &\stackrel{\text{Fubini}}{=} E \left[ X_n \int_0^{X^*} p y^{p-2} dy \right] \\ &= \frac{p}{p-1} E[X_n (X^*)^{p-1}] \\ &\stackrel{\text{Hölder}}{\leq} \frac{p}{p-1} (EX_n^p)^{\frac{1}{p}} \left( E((X^*)^{p-1})^q \right)^{\frac{1}{q}} \\ &= \frac{p}{p-1} (EX_n^p)^{\frac{1}{p}} \cdot (E(X^*)^p)^{1-\frac{1}{p}} \end{aligned}$$



Teile Ungleichung durch  $(E(X^*)^p)^{1-\frac{1}{p}}$  (falls  $E(X^*)^p = 0$  ist Aussage richtig) und nehme  $p$ -te Potenz  $\implies$  Behauptung. ■

**Bemerkung** a) Ist  $\frac{1}{p} + \frac{1}{q} = 1$ , so lässt sich Satz 8.3 schreiben als  $\|X^*\|_p \leq q \cdot \|X_n\|_p$

b) Ein stochastischer Prozess  $(X_t)_{t \in I}$  mit  $\sup_{t \in I} \|X_t\|_p < \infty$  heißt  $L^p$ -**beschränkt**.

c) Ist  $(X_i)_{i=1, \dots, n}$  ein Martingal, so ist  $(|X_i|)_{i=1, \dots, n}$  ein nicht negatives Submartingal (Satz 8.1)

**Beispiel 8.3** Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum und  $(X_n)_{n \in \mathbb{N}_0}$  ein stochastischer Prozess. Interpretation von  $(X_n)$ :

$X_0 \equiv$  Anfangskapital des Spielers

$X_n - X_{n-1} \equiv$  Gewinn pro gesetzter Geldeinheit in der  $n$ -ten Runde

Wird immer eine Geldeinheit pro Runde gesetzt, so ist also  $X_n = X_0 + \sum_{k=1}^n (X_k - X_{k-1})$  das Kapital des Spielers nach  $n$  Runden. Es sei

$$\mathfrak{F}_n = \sigma(X_0, X_1 - X_0, \dots, X_n - X_{n-1}) = \sigma(X_0, X_1, \dots, X_n)$$

Das entspricht der Information nach  $n$  Runden.

$$\implies E[X_{n+1} - X_n | \mathfrak{F}_n] = E[X_{n+1} | \mathfrak{F}_n] - X_n$$

Das entspricht dem erwarteten Gewinn pro gesetzter Geldeinheit bei Kenntnis des bisherigen Spielverlaufs.

Offenbar gilt:	$X$ Martingal	$\iff$	Spiel fair
	$X$ Supermartingal	$\iff$	Spiel nachteilig
	$X$ Submartingal	$\iff$	Spiel vorteilhaft

#### Beispiel 8.4

$X_n - X_{n-1}$  sei der Gewinn pro gesetzter Geldeinheit (GE) in der  $n$ -ten Runde.

Jetzt: In Runde  $n$  werden  $c_n$  GE gesetzt mit  $c_n \mathfrak{F}_{n-1}$ -messbar.

$\mathfrak{F}_n = \sigma(X_0, X_1 - X_0, \dots, X_n - X_{n-1})$ , d.h.  $(c_n)_{n \in \mathbb{N}}$  ist vorhersagbar.

Kapital nach  $n$  Spielen:

$$X_0 + \sum_{k=1}^n c_k (X_k - X_{k-1})$$

#### Satz 8.4

Es seien  $(c_n)_{n \in \mathbb{N}}$  ein vorhersagbarer Prozess und  $X = (X_n)_{n \in \mathbb{N}}$  ein Prozess mit  $E|c_n(X_n - X_{n-1})| < \infty \quad \forall n \in \mathbb{N}$ . Wir setzen

$$Y_n := X_0 + \sum_{k=1}^n c_k (X_k - X_{k-1}), \quad Y = (Y_n)_{n \in \mathbb{N}}.$$

Dann gilt:

a) Ist  $X$  ein Martingal, so auch  $Y$ .

- b) Ist  $X$  ein Sub- bzw. Supermartingal und  $c_n \geq 0 \quad \forall n$ , so ist auch  $Y$  ein Sub- bzw. Supermartingal.

**Beweis**

$$E[Y_{n+1} - Y_n | \mathfrak{F}_n] = E[c_{n+1}(X_{n+1} - X_n) | \mathfrak{F}_n] \stackrel{c_{n+1}\mathfrak{F}_n\text{-m.b.}}{=} c_{n+1} \cdot E[X_{n+1} - X_n | \mathfrak{F}_n].$$

**Definition**

Eine Abbildung  $\tau : \Omega \rightarrow \mathbb{N}_0 \cup \{\infty\}$  heißt **Stoppzeit** bezüglich einer Filtration  $(\mathfrak{F}_n)_{n \in \mathbb{N}}$ , wenn

$$\{\tau \leq n\} \in \mathfrak{F}_n \quad \forall n \in \mathbb{N}_0.$$

**Bemerkung**

- a) Stoppzeiten kann man analog für  $\tau : \Omega \rightarrow \mathbb{R}_+ \cup \{\infty\}$  definieren.  
b)  $\tau : \Omega \rightarrow \mathbb{N}_0 \cup \{\infty\}$  ist Stoppzeit  $\iff \{\tau = n\} \in \mathfrak{F}_n \quad \forall n \in \mathbb{N}_0$ . (Übung)

**Beispiel 8.5**

- a)  $\tau \equiv n_0$  ist Stoppzeit, da  

$$\{\tau \leq n\} = \begin{cases} \Omega & n \geq n_0 \\ \emptyset & n < n_0 \end{cases} \in \mathfrak{F}_n$$
  
b) Sei  $(X_n)_{n \in \mathbb{N}_0}$  ein zu  $(\mathfrak{F}_n)_{n \in \mathbb{N}_0}$  adaptierter reellwertiger Prozess und  $A \in \mathfrak{B}$ . Sei  $\tau_A : \Omega \rightarrow \mathbb{N} \cup \{\infty\}$  definiert durch

$$\tau_A(\omega) := \inf \{n \in \mathbb{N}_0 \mid X_n(\omega) \in A\} \quad (\inf \{\emptyset\} := \infty)$$

$\tau_A$  heißt **Eintrittszeit** in  $A$ .

$\tau_A$  ist Stoppzeit, da

$$\{\tau_A \leq n\} = \bigcup_{i=1}^n \underbrace{\{X_i \in A\}}_{\in \mathfrak{F}_i} \in \mathfrak{F}_n.$$

**Lemma 8.1**

- a) Für eine Stoppzeit ist

$$\mathfrak{F}_\tau := \{A \in \mathcal{A} \mid A \cap \{\tau \leq n\} \in \mathfrak{F}_n \quad \forall n \in \mathbb{N}_0\}$$

eine  $\sigma$ -Algebra, die  $\sigma$ -**Algebra der  $\tau$ -Vergangenheit**.

- b) Sind  $\tau_1, \tau_2$  Stoppzeiten mit  $\tau_1 \leq \tau_2$ , so gilt  $\mathfrak{F}_{\tau_1} \subset \mathfrak{F}_{\tau_2}$ .

c) Ist  $\tau$  eine Stoppzeit, so ist  $X_\tau^* : \Omega \rightarrow \mathbb{R}$  mit

$$X_\tau^*(\omega) := \begin{cases} X_{\tau(\omega)}(\omega) & \text{wenn } \tau(\omega) < \infty \\ 0 & \text{sonst} \end{cases} \quad \mathfrak{F}_\tau\text{-messbar}$$

### Beweis

a) Übung.

b) Sei  $A \in \mathfrak{F}_{\tau_1}$  beliebig.  $\forall n \in \mathbb{N}$  gilt:

$$\{\tau_2 \leq n\} \subset \{\tau_1 \leq n\} \implies A \cap \{\tau_2 \leq n\} = \underbrace{A \cap \{\tau_1 \leq n\}}_{\in \mathfrak{F}_n} \cap \underbrace{\{\tau_2 \leq n\}}_{\in \mathfrak{F}_n} \in \mathfrak{F}_n.$$

$\implies$  Beh.

c) zu zeigen:  $\{X_\tau^* \in A\} \in \mathfrak{F}_\tau \quad \forall A \in \mathfrak{B}$

zeige also:  $\{X_\tau^* \in A\} \cap \{\tau \leq n\} \in \mathfrak{F}_n \quad \forall n \in \mathbb{N}_0$

Es gilt:

$$\{X_\tau^* \in A\} \cap \{\tau \leq n\} = \bigcup_{k=0}^n \underbrace{\{X_k \in A\}}_{\in \mathfrak{F}_k} \cap \underbrace{\{\tau = k\}}_{\in \mathfrak{F}_k} \in \mathfrak{F}_k,$$

$$\text{da } \{\tau = k\} = \underbrace{\{\tau \leq k\}}_{\in \mathfrak{F}_k} \cap \underbrace{\{\tau \leq k-1\}^C}_{\in \mathfrak{F}_k} \in \mathfrak{F}_k.$$

$\implies$  Beh. ■

### Bemerkung

a)  $\mathfrak{F}_\tau \equiv$  Information, die bis zur zufälligen Zeit  $\tau$  vorhanden ist.

b) Falls  $\tau$   $P$ -f.s. endlich, schreibt man  $X_\tau$  statt  $X_\tau^*$ .

c) Ist  $\tau$  eine Stoppzeit und  $(X_n)_{n \in \mathbb{N}_0}$  ein stochastischer Prozess, so ist  $X^\tau = (X_n^\tau)_{n \in \mathbb{N}_0}$  mit  $X_n^\tau := X_{\tau \wedge n} \quad \forall n \in \mathbb{N}_0$  der **gestoppte Prozess**.

Da  $\tau \wedge n$  eine Stoppzeit ist, ist wegen Lemma 8.1c)  $X_{\tau \wedge n} \mathfrak{F}_{\tau \wedge n}$ -messbar und  $(X_n^\tau)$  ist  $(\mathfrak{F}_{\tau \wedge n})$ -adaptiert.

### Satz 8.5

Ist  $X$  ein (Sub-, Super-) Martingal und ist  $\tau$  eine Stoppzeit, so ist auch  $X^\tau$  ein (Sub-, Super-) Martingal.

### Beweis

Sei  $c_n := \mathbf{1}_{\{\tau \geq n\}} \implies \{\tau \geq n\} = \{\tau \leq n-1\}^C \in \mathfrak{F}_{n-1}$ .

$\implies (c_n)_{n \in \mathbb{N}_0}$  ist vorhersagbar. Da  $X_0 + \sum_{k=1}^n c_k(X_k - X_{k-1}) = X_{\tau \wedge n}$ , folgt die Behauptung mit Satz 8.4. ■

**Bemerkung**

Ist  $X$  ein Martingal, so auch  $X^\tau$  und damit gilt  $EX_{\tau \wedge n} = EX_0$ .

Betrachte Bsp 8.4 mit  $\tau := \inf\{k \in \mathbb{N}_0 \mid X_k \geq X_0 + c\}$  und  $c_n := \mathbf{1}_{\{\tau \geq n\}}$ :

Solange  $c$  nicht erreicht ist, wird eine Geldeinheit gesetzt, danach aufgehört. Spielt man maximal  $n$ -mal, so ist  $X_{\tau \wedge n}$  das Kapital am Ende. Im Mittel kann man das Kapital bei einem fairen Spiel nicht erhöhen.

**Beispiel 8.6 (Kartenspiel)**

Sei

- $S_0$  die Anzahl der schwarzen Karten und
- $R_0$  die Anzahl der roten Karten und
- $N := S_0 + R_0$  die Gesamtzahl an Karten.
- $(R_n, S_n)$  die Anzahl der roten / schwarzen Karten im Stapel, nachdem  $n$  Karten aufgedeckt wurden.
- $Z_n$  die Farbe der  $n$ -ten aufgedeckten Karte.
- $\mathfrak{F}_n = \sigma(Z_1, \dots, Z_n)$  und
- $X_n := \frac{S_n - R_n}{S_n + R_n}$ .

Behauptung:  $(X_n)$  ist  $(\mathfrak{F}_n)$ -Martingal!

$$\begin{aligned}
 E[X_{n+1} \mid \mathfrak{F}_n] &= E\left[\frac{S_{n+1} - R_{n+1}}{S_{n+1} + R_{n+1}} \mid Z_1, \dots, Z_n\right] \\
 &= \frac{S_n}{S_n + R_n} \left[\frac{S_n - 1 - R_n}{S_n - 1 + R_n}\right] + \frac{R_n}{S_n + R_n} \left[\frac{S_n - R_n + 1}{S_n + R_n - 1}\right] \\
 &= \frac{(R_n + S_n - 1)(S_n - R_n)}{(S_n + R_n)(S_n + R_n - 1)} \\
 &= \frac{S_n - R_n}{S_n + R_n}
 \end{aligned}$$

Sei  $\tau$  eine Stoppzeit ( $\leq N$ ). Erwarteter Gewinn:

$$\begin{aligned}
 &E[\mathbf{1}_{[Z_{\tau+1} = \text{schwarz}]} - \mathbf{1}_{[Z_{\tau+1} = \text{rot}]}] \\
 &= E\left[\sum_{k=1}^N (\mathbf{1}_{[Z_{k+1} = \text{schwarz}]} - \mathbf{1}_{[Z_{k+1} = \text{rot}]} ) \mathbf{1}_{[\tau=k]}\right] \\
 &= \sum_{k=1}^N E\left[E[(\mathbf{1}_{[Z_{k+1} = \text{schwarz}]} - \mathbf{1}_{[Z_{k+1} = \text{rot}]} ) \mathbf{1}_{[\tau=k]} \mid \mathfrak{F}_k]\right] \\
 &= \sum_{k=1}^N E\left[\mathbf{1}_{[\tau=k]} \underbrace{E[\mathbf{1}_{[Z_{k+1} = \text{schwarz}]} - \mathbf{1}_{[Z_{k+1} = \text{rot}]} \mid \mathfrak{F}_k]}_{= \frac{S_k - R_k}{S_k + R_k} = X_k}\right]
 \end{aligned}$$

$$= E[X_\tau] = EX_0 = \frac{S_0 - R_0}{S_0 + R_0}$$

$EX_\tau = EX_0$  gilt nur unter einer Bedingung, wie dieses Beispiel zeigt.

**Beispiel 8.7** Sei  $(Y_n)_{n \in \mathbb{N}}$  eine Folge von u.i.v. ZVen mit

$$P(Y_n = -1) = P(Y_n = 1) = \frac{1}{2}, \quad X_0 \equiv 0$$

$Y_n$  = Ergebnis Münzwurf in Runde  $n$ .

Der Spieler setzt  $2^{n-1}$  GE in der  $n$ -ten Runde, bei Gewinn erhält er  $2^n$  GE, d.h.  $Y_n \cdot 2^{n-1}$  ist der Geldzu-/abgang in der  $n$ -ten Runde.

Kapital nach  $n$  Runden:

$$X_n := \sum_{i=1}^n 2^{i-1} Y_i$$

Sei  $\mathfrak{F}_n := \sigma(X_0, \dots, X_n)$  und  $\tau := \inf\{n \in \mathbb{N} \mid Y_n = 1\}$  d.h. gestoppt wird, wenn erstmals  $Y_n = 1$  ( $\rightarrow$  Martingalstrategie).  $(X_n)_{n \in \mathbb{N}}$  ist ein  $(\mathfrak{F}_n)_{n \in \mathbb{N}}$ -Martingal (s. Bsp. 8.1).

Es gilt:

$$P(\tau > k) = \left(\frac{1}{2}\right)^k \quad \forall k \in \mathbb{N} \Rightarrow P(\tau < \infty) = 1$$

und

$$X_\tau = \sum_{k=1}^{\infty} X_k \mathbf{1}_{\tau=k} = \sum_{k=1}^{\infty} \underbrace{\left(-\sum_{i=1}^{k-1} 2^{i-1} + 2^{k-1}\right)}_{=1} \mathbf{1}_{\tau=k} \equiv 1$$

Also ist hier  $EX_\tau = 1 \neq EX_0 = 0$ .

Vorsicht bei der Nachahmung!

Das benötigte Kapital beträgt  $-X_{\tau-1}$  GE und

$$\begin{aligned} E(-X_{\tau-1}) &= E\left(\sum_{k=1}^{\tau-1} 2^{k-1}\right) \\ &= E\left(\sum_{k=1}^{\infty} 2^{k-1} \mathbf{1}_{[\tau > k]}\right) \\ &= \sum_{k=1}^{\infty} 2^{k-1} \underbrace{P(\tau > k)}_{=2^{-k}} = \infty \end{aligned}$$

### Satz 8.6 (Optional Stopping Theorem OST)

Es sei  $X = (X_n)_{n \in \mathbb{N}}$  ein Supermartingal und  $\tau$  eine Stoppzeit. Jede der folgenden Bedingungen impliziert, dass  $E|X_\tau| < \infty$  und  $EX_\tau \leq EX_1$  gilt:

1.  $\tau$  ist f.s. beschränkt, also  $P(\tau < c) = 1$  für ein  $c \in \mathbb{R}$ .

2.  $\tau$  ist f.s. endlich und  $X$  ist f.s. beschränkt, d.h.  $P(\tau < \infty) = 1$  und es gibt ein  $c \in \mathbb{R}$  mit  $P(|X_n| \leq c) = 1 \ \forall n \in \mathbb{N}_0$ .
3.  $E\tau < \infty$  und  $X$  hat f.s. beschränkte Zuwächse, d.h.  $\exists c \in \mathbb{R}$  mit  $P(|X_n - X_{n-1}| \leq c) = 1 \ \forall n \in \mathbb{N}$ .
4.  $P(\tau < \infty) = 1, E|X_\tau| < \infty$  und  $\int_{\{\tau > n\}} |X_n| dP \rightarrow 0$  für  $n \rightarrow \infty$ .

Ist eine dieser Bedingungen erfüllt und  $X$  ein Martingal, so gilt:  $EX_\tau = EX_1$ .

**Beweis** 1. Ist klar, da hier  $X_\tau = X_{\tau \wedge n}$  für ein  $n \in \mathbb{N}$  groß ( $n > c$ ). Die Behauptung folgt aus Satz 8.5.

2. Satz 8.5 und majorisierte Konvergenz.
3. Verwende  $|X_1| + c(\tau - 1)$  als integrierbare Majorante.
4. Wir zeigen die Aussage für  $X$  ist Martingal:

$$\begin{aligned}
 |EX_\tau - EX_{\tau \wedge n}| &= \left| \int X_\tau dP - \int_{\{\tau \leq n\}} X_\tau dP - \int_{\{\tau > n\}} X_n dP \right| \\
 &\leq \left| \int_{\{\tau > n\}} X_\tau dP \right| + \left| \int_{\{\tau > n\}} X_n dP \right| \\
 &\leq \underbrace{\int_{\{\tau > n\}} |X_\tau| dP}_{\rightarrow 0(n \rightarrow \infty)} + \underbrace{\int_{\{\tau > n\}} |X_n| dP}_{\rightarrow 0(n \rightarrow \infty)} \rightarrow 0 \ (n \rightarrow \infty) \quad \blacksquare
 \end{aligned}$$

**Beispiel 8.8 (Ruinspiel, vgl. Stochastik I, Bsp 10.4)** Spieler I besitze  $n$  GE ( $n \in \mathbb{N}$ ), Spieler II  $N - n$  GE ( $N - n \in \mathbb{N}$ ). Pro Runde gewinnt Spieler I von Spieler II 1 GE mit W'keit  $p$  und verliert eine GE an Spieler II mit W'keit  $1 - p$ . Spielrunden sind unabhängig. Seien  $(Y_n)_{n \in \mathbb{N}}$  u.i.v. ZV mit

$$P(Y_n = 1) = p, \ P(Y_n = -1) = 1 - p.$$

$X_n := \sum_{k=1}^n Y_k$  ist dann der Gewinn (Verlust) von Spieler I nach  $n$  Runden.

Sei

$$\tau := \inf\{n \in \mathbb{N} \mid X_n = N - n \text{ oder } X_n = -n\}$$

$P(X_\tau = -n)$  = Ruinwahrscheinlichkeit von Spieler I.

Sei  $\mu = EY_1 = 2p - 1$ . Nach Beispiel 8.1  $\mu = 0 \Rightarrow (X_n)$  Martingal.  $\mu \leq 0 \Rightarrow (X_n)$

Supermartingal.  $\mu \geq 0 \Rightarrow (X_n)$  Submartingal.

**Behauptung:**  $\exists a > 0, 0 < \gamma < 1$ , sodass  $P(\tau > j) \leq a\gamma^j \ \forall j \in \mathbb{N}$ .

**Beweis:** Sei  $k \in \mathbb{N}$ .

$$\begin{aligned}
 P(\tau > Nk) &\leq P((Y_1, \dots, Y_n) \neq (1, \dots, 1), \\
 &\quad (Y_{N+1}, \dots, Y_{2N}) \neq (1, \dots, 1), \dots, (Y_{(k-1)N+1}, \dots, Y_{kN}) \neq (1, \dots, 1)) \\
 &\stackrel{(Y_n) \text{ unabh.}}{=} \prod_{v=0}^{k-1} P((Y_{vN+1}, \dots, Y_{(v+1)N}) \neq (1, \dots, 1)) \\
 &= (1 - p^N)^k
 \end{aligned}$$

Für  $j > N$  gilt:

$$P(\tau > j) \leq P(\tau > \lfloor \frac{j}{N} \rfloor N) \leq (1 - p^N)^{\lfloor \frac{j}{N} \rfloor} \leq \underbrace{\left( (1 - p^N)^{\frac{1}{N}} \right)^j}_{=: \gamma^j} \underbrace{(1 - p^N)^{-1}}_{=: a}$$

■

Also folgt:  $P(\tau < \infty) = 1$ ,  $E\tau = \sum_{j=1}^{\infty} P(\tau \geq j) < \infty$  und  $1 = P(\tau < \infty) = P(X_\tau = N - n) + P(X_\tau = -n)$ .

Sei nun  $M_n := \sum_{k=1}^n (Y_k - \underbrace{EY_k}_{=\mu})$ ,  $n \in \mathbb{N}_0$ ,  $M_0 = 0$  und  $\mathfrak{F}_n := \sigma(Y_1, \dots, Y_n)$ .

Dann ist  $(M_n)_{n \in \mathbb{N}_0}$  ein  $(\mathfrak{F}_n)$ -Martingal. Das OST ist anwendbar, da (iii) erfüllt ist.

$$\begin{aligned} \Rightarrow 0 &= EM_\tau = P(X_\tau = N - n)(N - n - E\tau\mu) + P(X_\tau = -n)(-n - E\tau\mu) \\ &= P(X_\tau = N - n)(N - n) - P(X_\tau = -n)n - E\tau\mu. \end{aligned}$$

Fall 1:  $\mu = 0$  (d.h.  $p = \frac{1}{2}$ , faires Spiel)

$$\Rightarrow 0 = (1 - P(X_\tau = -n))(N - n) - P(X_\tau = -n)n \Rightarrow P(X_\tau = -n) = \frac{N - n}{N}$$

Fall 2:  $p \neq \frac{1}{2}$

Sei  $\Theta := \log(\frac{1-p}{p}) \neq 0$  und  $L_0 := 1$ ,  $L_n := \prod_{k=1}^n e^{\Theta Y_k} = e^{\Theta X_n}$ .  
 $(L_n)_{n \in \mathbb{N}}$  ist ein  $(\mathfrak{F}_n)_{n \in \mathbb{N}}$ -Martingal, da

$$E[L_{n+1} \mid \mathfrak{F}_n] = \prod_{k=1}^n e^{\Theta Y_k} \cdot \underbrace{E[e^{\Theta Y_{n+1}}]}_{pe^{\Theta} + (1-p)e^{-\Theta} = 1} = L_n$$

Das Optional Stopping Theorem 8.6 ist anwendbar, da (iv) erfüllt  
 $E|L_\tau| = Ee^{\Theta X_\tau} \leq e^{|\Theta|N} < \infty$  und

$$\int_{\{\tau > n\}} |L_n| dP \leq e^{|\Theta|N} \underbrace{P(\tau > n)}_{\rightarrow 0 \text{ (} n \rightarrow \infty)}$$

$$\begin{aligned} \Rightarrow 1 &= EL_0 = EL_\tau = P(X_\tau = N - n)e^{\Theta(N-n)} + P(X_\tau = -n)e^{-\Theta n} \\ \Rightarrow 1 &= (1 - P(X_\tau = -n)) \cdot \left(\frac{1-p}{p}\right)^{N-n} + P(X_\tau = -n) \left(\frac{p}{1-p}\right)^n \\ \Rightarrow P(X_\tau = -n) &= \frac{\phi^N - \phi^n}{\phi^N - 1}, \quad \phi = \frac{1-p}{p} \end{aligned}$$

### Optimales Stoppen

Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum und  $X = (X_n)_{n=1, \dots, N}$  ein stochastischer Prozess adaptiert an eine Filtration  $(\mathfrak{F}_n)_{n=1, \dots, N}$ . Es sei  $E|X_k| < \infty \quad \forall k = 1, \dots, N$ . Betrachte das Optimierungsproblem

$$v := \sup_{\tau \text{ ist Stoppzeit } \leq N} \{EX_\tau\} = EX_{\tau_0}$$

$v$  = maximaler Wert,

$\tau_0$  = optimale Stoppzeit (falls existent). Wegen

$$E|X_\tau| = \sum_{n=1}^N E(|X_n| \cdot \mathbf{1}_{\{\tau=n\}}) \leq \sum_{n=1}^N E|X_n| < \infty$$

nach Voraussetzung ist  $v < \infty$ . Ist  $(X_n)_{n=1,\dots,N}$  ein  $(\mathfrak{F}_n)_{n=1,\dots,N}$  Supermartingal, so folgt mit Satz 8.6:  $EX_1 \geq EX_\tau \quad \forall$  Stoppzeiten  $\tau \leq N$ .

Also:  $\tau_0 \equiv 1$  ist optimal (sofort aufhören).

### Definition

Der Prozess  $Z = (Z_n)_{n=1,\dots,N}$  mit

$$Z_N := X_N, \quad Z_n := \max\{X_n, E[Z_{n+1} | \mathfrak{F}_n]\}, \quad n = N-1, \dots, 1$$

heißt **Snell-Einhüllende** von  $X$ .

**Satz 8.7** Mit den obigen Bezeichnungen gilt:

- a)  $Z$  ist ein  $(\mathfrak{F}_n)_{n=1,\dots,N}$ -Supermartingal mit  $Z_n \geq X_n$  für  $n = 1, \dots, N$ .
- b)  $Z$  ist das kleinste  $(\mathfrak{F}_n)$ -Supermartingal, welches  $X$  dominiert, d.h. ist  $(Y_n)_{n=1,\dots,N}$  ein weiteres  $(\mathfrak{F}_n)$ -Supermartingal mit  $Y_n \geq X_n$ ,  $n = 1, \dots, N$  so gilt:  $Y_n \geq Z_n$  für  $n = 1, \dots, N$ .

### Beweis

- a) Aus der Definition:  $Z_n \geq X_n \quad \forall n$ ,  $Z_n \geq E[Z_{n+1} | \mathfrak{F}_n]$ , also  $(Z_n)$  Supermartingal.
- b) Rückwärtsinduktion:  
 $(n = N): Y_N \geq X_N = Z_N$   
 $(n \rightarrow n-1): Y_{n-1} \stackrel{Y \text{ Supermartingal}}{\geq} E[Y_n | \mathfrak{F}_{n-1}] \stackrel{\text{I.H.}}{\geq} E[Z_n | \mathfrak{F}_{n-1}]$  und  $Y_{n-1} \geq X_{n-1}$   
 $\implies Y_{n-1} \geq \max\{X_{n-1}, E[Z_n | \mathfrak{F}_{n-1}]\} = Z_{n-1}$  ■

### Satz 8.8

Mit den obigen Bezeichnungen und  $\tau_0 = \min\{n \in \{1, \dots, N\} \mid X_n = Z_n\}$  gilt:

- a)  $\tau_0$  ist eine Stoppzeit.
- b)  $(Z_n^{\tau_0})_{n=1,\dots,N}$  ist ein  $(\mathfrak{F}_n)_{n=1,\dots,N}$ -Martingal.
- c)  $EX_{\tau_0} = \sup_{\tau \text{ Stoppzeit}} \{EX_\tau\}$

### Beweis



a) Wegen  $Z_N = X_N$  ist  $\tau_0 \leq N$ . Es gilt:

$$\{\tau_0 \leq n\} = \bigcup_{i=1}^n \underbrace{\{Z_i = X_i\}}_{\in \mathfrak{F}_i} \in \mathfrak{F}_n$$

b) Es gilt:

$$\underbrace{Z_{n+1}^{\tau_0}}_{=Z_{(n+1) \wedge \tau_0}} - \underbrace{Z_n^{\tau_0}}_{=Z_n \wedge \tau_0} = \mathbf{1}_{\{\tau_0 \geq n+1\}} (Z_{n+1} - E[Z_{n+1} | \mathfrak{F}_n]) \quad (*)$$

da

Fall 1:  $\tau_0 \geq n+1$

linke Seite =  $Z_{n+1} - Z_n$ ,

rechte Seite =  $Z_{n+1} - \underbrace{E[Z_{n+1} | \mathfrak{F}_n]}_{=Z_n}$ , da  $X_n < Z_n$  auf  $\{\tau_0 \geq n+1\}$ . (stimmt)

Fall 2:  $\tau_0 \leq n$

$0 = 0$  (stimmt)

Wende nun  $E[\cdot | \mathfrak{F}_n]$  auf (\*) an:

Da  $\{\tau_0 \geq n+1\} = \{\tau_0 \leq n\}^C \in \mathfrak{F}_n$  folgt

$$E[Z_{n+1}^{\tau_0} - Z_n^{\tau_0} | \mathfrak{F}_n] = \mathbf{1}_{\{\tau_0 \geq n+1\}} E[Z_{n+1} - E[Z_{n+1} | \mathfrak{F}_n] | \mathfrak{F}_n] = 0$$

$\implies (Z_n^{\tau_0})$  ist  $(\mathfrak{F}_n)$ -Martingal.

c) Wegen b) und Satz 8.6:

$$EZ_1 = EZ_1^{\tau_0} = EZ_N^{\tau_0} = EZ_{\tau_0} = EX_{\tau_0}$$

Für eine beliebige Stoppzeit  $\tau$  gilt:

$EZ_1 \geq EZ_\tau$ , da  $Z$  Supermartingal. Und weiterhin:

$$EX_{\tau_0} = EZ_1 \geq EZ_1 \geq EZ_\tau \geq EX_\tau \implies \text{Beh.} \quad \blacksquare$$

**Beispiel 8.9 (Das Sekretärinnenproblem)**  $N$  Bewerber(innen) um eine Stelle stellen sich nacheinander vor. Nach jedem Interview muss entschieden werden, ob die Person die Stelle bekommt.

Annahme: Die Bewerber lassen sie linear anordnen und erscheinen in beliebiger Reihenfolge. ( $N!$  mögliche Reihenfolgen)

Welche Strategie maximiert die Wahrscheinlichkeit, dass die beste Person die Stelle bekommt?

- $A_n$  = absoluter Rang des  $n$ -ten Kandidaten unter allen  $N$ .
- $R_n$  = dessen relativer Rang unter den ersten  $N$ .  $R_n = \{1 \leq m \leq n \mid A_m \leq A_n\}$ .

Es gibt eine Bijektion zwischen den  $A$ -Werten und den  $R$ -Werten. Somit gilt  $\forall r_1, \dots, r_N$ ,  $1 \leq r_i \leq i$ ,  $1 \leq i \leq N$ :

$$P(R_1 = r_1, \dots, R_N = r_N) = \frac{1}{N!}$$

Bestimme Randverteilungen:

$$P(R_n = l) = \frac{1}{n} \quad \text{für } l = 1, \dots, n \quad \forall n \in \{1, \dots, N\}$$

und  $R_1, \dots, R_N$  unabhängig. Sei nun

$$\bar{X}_n := \begin{cases} 1, & A_n = 1 \\ 0, & \text{sonst} \end{cases}, \quad \mathfrak{F}_n = \sigma(R_1, \dots, R_n)$$

und  $X_n = E[\bar{X}_n | \mathfrak{F}_n]$ .  $(X_n)$  ist  $(\mathfrak{F}_n)$ -adaptiert.  $P(\bar{X}_\tau = 1) \rightarrow \max$ .

$$\begin{aligned} P(\bar{X}_\tau = 1) &= \sum_{n=1}^N P(\bar{X}_n = 1, \tau = n) = \sum_{n=1}^N E \mathbf{1}_{[\tau=n, \bar{X}_n=1]} \\ &= \sum_{n=1}^N \int_{\{\tau=n\}} \bar{X}_n dP = \sum_{n=1}^N \int_{\{\tau=n\}} \underbrace{E[\bar{X}_n | \mathfrak{F}_n]}_{=X_n} dP \\ &= EX_\tau \end{aligned}$$

Also maximiere  $EX_\tau$  mit Satz 8.8.

$$\begin{aligned} P(R_1 = r_1, \dots, R_{n-1} = r_{n-1}, A_n = 1) &= P(R_1 = r_1, \dots, R_{n-1} = r_{n-1}, R_n = 1, R_{n+1} > 1, \dots, R_N > 1) \\ &= \frac{1}{N!} \cdot 1 \cdot \dots \cdot 1 \cdot n \cdot (n+1) \cdot \dots \cdot (N-1) = \frac{n}{N} \cdot \frac{1}{n!} \end{aligned}$$

$$\begin{aligned} \Rightarrow P(A_n = 1 \mid R_1 = r_1, \dots, R_{n-1} = r_{n-1}, R_n = 1) &= \frac{P(R_1 = r_1, \dots, R_{n-1} = r_{n-1}, R_n = 1, A_n = 1)}{P(R_1 = r_1, \dots, R_{n-1} = r_{n-1}, R_n = 1)} \\ &= \frac{\frac{n}{N} \cdot \frac{1}{n!}}{\frac{1}{n!}} = \frac{n}{N} \end{aligned}$$

$$\Rightarrow X_n = E[\mathbf{1}_{\{1\}}(A_n) | \mathfrak{F}_n] = \begin{cases} \frac{n}{N}, & \text{falls } R_n = 1 \\ 0, & \text{sonst} \end{cases} \quad (*)$$

**Behauptung:**  $\exists (c_n)_{n=1, \dots, N} \subset \mathbb{R}$ ,  $c_n \downarrow$ ,  $c_N = \frac{1}{N}$  und  $E[Z_n | \mathfrak{F}_{n-1}] \equiv c_n$  für  $n = 1, \dots, N$ , wobei  $Z$  die Snell-Einhüllende von  $X$  ist.

**Beweis:** Rückwärtsinduktion:

$n = N$ :

$$\begin{aligned} E[Z_N | \mathfrak{F}_{N-1}] &= E[X_N | \mathfrak{F}_{N-1}] \stackrel{A_N = R_N}{=} E[\mathbf{1}_{\{1\}}(R_N) | \mathfrak{F}_{N-1}] \\ &\stackrel{R_N, \mathfrak{F}_N \text{ unabh.}}{=} P(R_N = 1) = \frac{1}{N} = c_N. \end{aligned}$$

$n+1 \rightsquigarrow n$ :

$$\begin{aligned}
E[Z_n | \mathfrak{F}_{n-1}] &= E[\max\{X_n, E[Z_{n+1} | \mathfrak{F}_n]\} | \mathfrak{F}_{n-1}] \\
&\stackrel{(*)}{=} E[\max\{\frac{n}{N} \cdot \mathbf{1}_{\{1\}}(R_n), c_{n+1}\} | \mathfrak{F}_{n-1}] \\
&= E[\mathbf{1}_{\{1\}}(R_n) \cdot \max\{\frac{n}{N}, c_{n+1}\} + (1 - \mathbf{1}_{\{1\}}(R_n))c_{n+1} | \mathfrak{F}_{n-1}] \\
&\stackrel{R_n, \mathfrak{F}_{n-1} \text{ unabh.}}{=} P(R_n = 1) \cdot \max\{\frac{n}{N}, c_{n+1}\} + (1 - P(R_n = 1)) \cdot c_{n+1} \\
&= \frac{1}{n} \max\{\frac{n}{N}, c_{n+1}\} + (1 - \frac{1}{n})c_{n+1} \\
&\Rightarrow c_n = c_{n+1} + \underbrace{\max\{\frac{1}{N}, \frac{c_{n+1}}{n}\} - \frac{c_{n+1}}{n}}_{\geq 0} \Rightarrow c_n \geq c_{n+1}
\end{aligned}$$

$$\tau^* := \inf\{n \mid Z_n = X_n\}$$

Stoppregel nach Satz 8.8:  $\tau^* = \min\{n \mid X_n = Z_n\}$ .

- gestoppt wird vor  $N$  nur, wenn  $R_n = 1$ .
- die Werte  $X_n \neq 0$  sind wachsend.
- die Werte  $E[Z_{n+1} | \mathfrak{F}_n] = c_{n+1}$  fallend.

$$\begin{aligned}
\tau^* &= \min\{1 \leq n \leq N-1 \mid R_n = 1, \frac{n}{N} \geq c_{n+1}\} \wedge N \\
&= \min\{n \geq k_n \mid R_n = 1\} \wedge N.
\end{aligned}$$

Wir bestimmen jetzt noch  $k_N$ .

Sei  $\tau_k := \inf\{n \geq k \mid R_n = 1\} \wedge N$ . Bestimme  $EX_{\tau_k}$ .

$k_N$  ist dann der  $k$ -Wert, bei dem  $EX_{\tau_k}$  maximal ist. Es gilt

$$\begin{aligned}
EX_{\tau_k} &= \sum_{l=k}^N E[X_l \cdot \mathbf{1}_{\{l\}}(\tau_k)] \\
&= \sum_{l=k}^N \frac{l}{N} \underbrace{P(R_m > 1 \text{ für } m = k, \dots, l-1, R_l = 1)}_{=P(\tau_k=l)} \\
&= \sum_{l=k}^N \frac{l}{N} \underbrace{\left( \prod_{m=k}^{l-1} \frac{m-1}{m} \right)}_{=P(R_m > 1)} \cdot \underbrace{\frac{1}{l}}_{=P(R_l=1)} \\
&\stackrel{\text{Teleskop. Prod.}}{=} \frac{k-1}{N} \sum_{l=k}^N \frac{1}{l-1}
\end{aligned}$$

$\Phi(k) := \frac{k-1}{N} \sum_{l=k}^N \frac{1}{l-1}$  wird maximal in  $k_N := \inf\{k \mid \frac{1}{k} + \frac{1}{k+1} + \dots + \frac{1}{N-1} \leq 1\}$ .

Beachte:  $\lim_{N \rightarrow \infty} \frac{k_N}{N} \stackrel{!}{=} \frac{1}{e}$ .

Bei einem großen Bewerberkreis wird man etwa 37 Prozent der Bewerber passieren lassen und dann den ersten nehmen, der besser als alle vorangegangenen ist.



## 9 Konvergenzsätze für Martingale

Im Folgenden sei  $(\Omega, \mathcal{A}, P)$  ein W'Raum,  $(\mathfrak{F}_n)_{n \in \mathbb{N}}$  eine Filtration und  $\mathfrak{F}_\infty := \sigma(\cup_{n \in \mathbb{N}} \mathfrak{F}_n)$ .

**Definition 9.1** Sei  $X_1, \dots, X_n$  eine Folge von ZV und  $-\infty < a < b < \infty$ .

$U_n[a, b]$  sei die **Anzahl der aufsteigenden Überschreitungen** des Intervalls  $[a, b]$  durch  $X_1, \dots, X_n$  also

$$U_n[a, b] = \max\{k \leq \lfloor \frac{n}{2} \rfloor \mid \exists \text{ Indizes } 1 \leq i_1 < \dots < i_{2k} \leq n \text{ mit } X_{i_{2j-1}} \leq a, b \leq X_{i_{2j}} \text{ für } j = 1, \dots, k\}$$

**Bemerkung 9.1** Wegen

$$\{U_n[a, b] \geq k\} = \bigcup_{1 \leq i_1 < \dots < i_{2k} \leq n} \bigcap_{j=1}^k \{X_{i_{2j-1}} \leq a\} \cap \{X_{i_{2j}} \geq b\} \in \mathfrak{F}_n$$

ist  $U_n[a, b]$  eine ZV.

**Lemma 9.1** Sei  $(X_n)_{n \in \mathbb{N}}$  ein Supermartingal. Dann gilt:

$$EU_n[a, b] \leq \frac{1}{b-a} E(X_n - a)^-$$

**Beweis**

Sei  $p := \lfloor \frac{n}{2} \rfloor + 1$ ,  $\tau_0 \equiv 1$  und für  $k = 1, \dots, p$ :

$$\tau_{2k-1} := \min\{j \geq \tau_{2k-2} \mid X_j \leq a\} \wedge n$$

$$\tau_{2k} := \min\{j \geq \tau_{2k-1} \mid X_j \geq b\} \wedge n$$

Die  $(\tau_k)$  sind Stoppzeiten mit  $1 \leq \tau_1 \leq \tau_2 \leq \dots \leq \tau_{2p} = n$  und

falls  $\tau_{2k-1} < n$ , ist  $\tau_{2k-1} < \tau_{2k}$ .

Sei  $k_0 := U_n[a, b]$ , d.h.  $X_{\tau_{2k}} - X_{\tau_{2k-1}} \geq b - a$  für  $k = 1, \dots, k_0$

$$X_{\tau_{2k_0+2}} - X_{\tau_{2k_0+1}} \neq 0 \implies X_{\tau_{2k_0+1}} \leq a, X_{\tau_{2k_0+2}} = X_n$$

$$\implies X_{\tau_{2k_0+2}} - X_{\tau_{2k_0+1}} \geq X_n - a \geq \min\{X_n - a, 0\} = -(X_n - a)^-$$

$$\implies \sum_{k=1}^p (X_{\tau_{2k}} - X_{\tau_{2k-1}}) \geq (b-a) \cdot U_n[a, b] - (X_n - a)^-$$

Wir zeigen jetzt:  $E(X_{\tau_{2k}} - X_{\tau_{2k-1}}) \leq 0$ .

Sei  $c_j := \mathbf{1}_{\{\tau_{2k-1} < j \leq \tau_{2k}\}}$ .  $(c_j)_{j \geq 2}$  ist vorhersehbar.

$$\{c_j = 1\} = \{\tau_{2k-1} \leq j-1\} \cap \{\tau_{2k} \leq j-1\}^C \in \mathfrak{F}_{j-1}$$

Sei  $Y_n = X_1 + \sum_{j=2}^n c_j(X_j - X_{j-1})$ ,  $n \in \mathbb{N}$ ;  $Y_1 := X_1$ .

Satz 8.4  $\implies (Y_n)$  ist ein Supermartingal.

$$\begin{aligned}
\Rightarrow EY_n &= E[X_1 + \sum_{j=1}^n c_j(X_j - X_{j-1})] \\
&= EX_1 + \underbrace{E[X_{\tau_{2k}} - X_{\tau_{2k-1}}]}_{\leq 0} \\
&\leq EY_1 = EX_1
\end{aligned}$$

$\Rightarrow$  Beh. ■

**Satz 9.1 (Vorwärtskonvergenzsatz von Doob)**

Sei  $(X_n)_{n \in \mathbb{N}}$  ein  $(\mathfrak{F}_n)_{n \in \mathbb{N}}$ -Supermartingal mit der Eigenschaft  $\sup_{n \in \mathbb{N}} \{E|X_n|\} < \infty$ .

Dann existiert eine  $\mathfrak{F}_\infty$ <sup>1</sup>-messbare Zufallsvariable  $X_\infty$  mit  $E|X_\infty| < \infty$  und  $\lim_{n \rightarrow \infty} X_n = X_\infty$   $P$ -f.s.

**Beweis**

Sei  $N := \{\omega \in \Omega \mid \liminf_{n \rightarrow \infty} \{X_n(\omega)\} < \limsup_{n \rightarrow \infty} \{X_n(\omega)\}\}$  und

$U_\infty[a, b] := \lim_{n \rightarrow \infty} \{U_n[a, b]\}$  (existiert, da  $U_n[a, b]$  wachsend)

$\Rightarrow N = \cup_{a, b \in \mathbb{Q}, a < b} \{\omega \in \Omega \mid U_\infty[a, b](\omega) = \infty\}$

Lemma 9.1  $\Rightarrow (b - a)EU_n[a, b] \leq E(X_n - a)^- \leq |a| + E|X_n| \quad \forall n \in \mathbb{N}$

Mit der Voraussetzung und monotoner Konvergenz:  $EU_\infty[a, b] < \infty$ .

$\Rightarrow P(U_\infty[a, b] = \infty) = 0 \Rightarrow P(N) = 0$ , da  $N$  abzählbare Vereinigung von  $P$ -Nullmengen. Außerdem:  $N \in \mathfrak{F}_\infty$ .

Sei  $\tilde{X}_\infty(\omega) := \begin{cases} \lim_{n \rightarrow \infty} \{X_n(\omega)\} & \omega \in N^C \text{ (evtl. } \tilde{X}_\infty(\omega) = \infty) \\ 0 & \omega \in N \end{cases}$

$$\begin{aligned}
\Rightarrow E|\tilde{X}_\infty| &= E\left[\liminf_{n \rightarrow \infty} \{|X_n|\}\right] \\
&\stackrel{\text{Lem. von Fatou}}{\leq} \limsup_{n \rightarrow \infty} \{E|X_n|\} \\
&\leq \sup_{n \in \mathbb{N}} \{E|X_n|\} \\
&< \infty
\end{aligned}$$

Sei  $\tilde{N} := \{\omega \in \Omega \mid \tilde{X}_\infty \in \{-\infty, \infty\}\} \Rightarrow P(\tilde{N}) = 0$

folgt  $X_\infty := \tilde{X}_\infty \cdot \mathbf{1}_{\tilde{N}^C}$  erfüllt die Bedingung. ■

**Bemerkung**

(i)  $(X_n)_{n \in \mathbb{N}}$  mit der Eigenschaft  $\sup_{n \in \mathbb{N}} \{E|X_n|\} < \infty$  heißt  **$L^1$ -beschränkt**.

(ii) Bei Supermartingalen folgt die  $L^1$ -Beschränktheit aus  $\sup_{n \in \mathbb{N}} \{EX_n^-\} < \infty$ , also z.B. falls  $X_n \geq 0$ .

---

<sup>1</sup> $\mathfrak{F}_\infty = \sigma(\cup_{n \in \mathbb{N}} \mathfrak{F}_n)$

**Beispiel 9.1 (Verzweigungsprozesse)**

Es sei  $\{Y_{nk} \mid n, k \in \mathbb{N}\}$  eine Familie von unabhängigen und identisch verteilten  $\mathbb{N}_0$ -wertigen Zufallsvariablen.

$$P_j := P(Y_{nk} = j) \quad \forall j \in \mathbb{N}_0$$

Sei  $(Z_n)$  definiert durch

$$Z_1 := 1, \quad Z_{n+1} := \sum_{k=1}^{Z_n} Y_{nk} \quad \forall n \in \mathbb{N}_0 \text{ und } \mu := \sum_{k=1}^{\infty} kp_k < \infty$$

$$\mathfrak{F}_n = \sigma(\{Y_{mk} \mid k \in \mathbb{N}, m \leq n-1\})$$

Es gilt:

$$\begin{aligned} E[Z_{n+1} \mid \mathfrak{F}_n] &= E\left[\sum_{k=1}^{Z_n} Y_{nk} \mid \mathfrak{F}_n\right] \\ &= E\left[\sum_{l=0}^{\infty} \left(\sum_{k=1}^l Y_{nk}\right) \cdot \mathbf{1}_{\{Z_n=l\}} \mid \mathfrak{F}_n\right] \\ &= \sum_{l=0}^{\infty} \underbrace{E\left[\sum_{k=1}^l Y_{nk} \mid \mathfrak{F}_n\right]}_{=E[\sum_{k=1}^l Y_{nk}]} \cdot \mathbf{1}_{\{Z_n=l\}} \\ &= \sum_{l=0}^{\infty} l \cdot \mu \cdot \mathbf{1}_{\{Z_n=l\}} = \mu \cdot Z_n \end{aligned}$$

Sei  $X_n := \frac{Z_n}{\mu^n} \implies (X_k)_{k \in \mathbb{N}}$  ist ein  $(\mathfrak{F}_n)_{n \in \mathbb{N}}$ -Martingal.

Insbesondere gilt:

$$EZ_n = \mu^n \cdot EX_n = \mu^n EX_1 = \mu^{n-1} EZ_1 = \mu^{n-1} \quad (*)$$

$(X_n)_{n \in \mathbb{N}}$  ist  $L^1$ -beschränkt, da  $E|X_n| = EX_n = \frac{1}{\mu} \quad \forall n \in \mathbb{N}$ .

Satz 9.1  $\implies \exists X_{\infty}$  mit  $X_n \rightarrow X_{\infty}$   $P$ -f.s. .

Falls  $\mu < 1$ :  $\xrightarrow{(*)} P(Z_n \geq \epsilon) \rightarrow 0 \quad (n \rightarrow \infty) \implies X_{\infty} \equiv 0$

Falls  $\mu = 1$ :  $X_n$  ganzzahlig  $\implies$  Folge irgendwann konstant. Wenn  $P_1 \neq 1 \implies X_{\infty} \equiv 0$

Falls  $\mu > 1$ :  $X_{\infty}$  ist nicht degeneriert.  $P(X_{\infty} = 0)$  ist Lösung von  $g(z) = z$ , wobei  $g$  erzeugende Funktion von  $Y$  ist.





# Stichwortverzeichnis

- $L^1$ -beschränkt, 86
- $L^p$ -Ungleichung, 72
- $\mu$ -Dichte, 19
- $\mu$ -Integral, 7, 10, 11
- $\mu$ -Nullmenge, 18
- $\mu$ -fast überall, 18
- $\mu$ -integrierbar
  - $p$ -fach, 22
- $\mu$ -stetig, 19
- $\sigma$ -Algebra
  - der  $\tau$ -Vergangenheit, 74
  - Produkt-, 25
- $d$ -dimensionale Normalverteilung, 58
- $p$ -fach  $\mu$ -integrierbar, 22
  
- abzählendes Maß, 6
- adaptiert, 71
- algebraische Induktion, 18
  
- bedingte Dichte, 67
- bedingter Erwartungswert, 61, 62, 66
- beschränkt
  - $L^1$ -, 86
  - $L^p$ -, 73
- Bildmaß, 16
- Borel-Cantelli Lemma, 37
  
- charakteristische Funktion, 41
- charakteristische Funktion (Zufallsvektor), 57
- Continuous Mapping Theorem, 46
  
- Darstellungssatz von Skorohod, 45
- Dirac-Maß, 5
  
- Eindeutigkeitssatz für Maße, 6
- Einpunktmaß, 5
- Eintrittszeit, 74
- Elementarfunktion, 7
  
- Erlang-Verteilung, 34
- Erwartungswert
  - bedingt, 61, 66
  - Version des bedingten, 62
- Erwartungswert (Zufallsvektor), 57
  
- Faktorisierungssatz, 65
- Faltung, 34
- fast überall, 18
- Filtration, 71
  
- Gamma-Verteilung, 34
- gemeinsame Verteilung, 31
- gestoppter Prozess, 75
- Gumbelverteilung, 55
  
- Höldersche Ungleichung, 22
  
- Jensensche Ungleichung, 21
  
- Kern, 66
- Konvergenz
  - in Verteilung, 45
  - schwache, 45
- konvex, 21
- Koppelung, 66
- Kovarianzmatrix, 57
  
- Lebesgue-Maß, 6
- Lebesgue-Stieltjes-Maß, 6
- Lemma
  - Borel-Cantelli, 37
  - von Fatou, 15
- Lindeberg-Bedingung, 49
  
- Maß
  - Produkt-, 27
- Martingal, 71
- Sub-, 71

- Super-, 71
- Maß, 5
  - $\sigma$ -endlich, 5
  - endlich, 5
  - Lebesgue-Stieltjes, 6
- Maß mit Dichte, 19
- maßdefinierende Funktion, 6
- Maßraum, 5
- Maßtransport, 16
- Minkowskische Ungleichung, 22
- Normalverteilung
  - $d$ -dimensionale, 58
- Nullmenge, 18
- Optional Stopping Theorem, 77
- Ordnungsstatistik, 35
- OST, 77
- Produkt- $\sigma$ -Algebra, 25
- Produktmaß, 27
- Projektion, 25
- Prozess
  - gestoppter, 75
- quasi-integrierbar, 11
- Randdichte, 34
- Satz
  - Faktorisierungs-, 65
  - Integration bezüglich des Bildmaßes, 17
  - Transformations-, 17, 33
  - Transformationssatz, 32
  - von der majorisierten Konvergenz, 15
  - von Doob, 72
  - von Helly, 47
  - von Lebesgue, 15
  - von Radon-Nikodym, 20
- schwache Konvergenz, 45
- Snell-Einhüllende, 80
- Stetigkeitssatz für charakteristische Funktionen, 48
- stochastisch unabhängig, 31
- stochastischer Prozess, 71
- Stoppzeit, 74
- straff, 47
- Submartingal, 71
- Submartingal-Ungleichung, 72
- Supermartingal, 71
- Theorem
  - Optional Stopping-, 77
- Transformationssatz, 17
- Übergangskern, 66
- unabhängig
  - stochastisch, 31
- Ungleichung
  - $L^p$ -, 72
  - Höldersche, 22
  - Jensensche, 21
  - Minkowskische, 22
  - Submartingal-, 72
- Verteilung
  - Erlang-, 34
  - Gamma-, 34
  - gemeinsame, 31
  - Gumbel, 55
- Verteilungskonvergenz, 57
- Wahrscheinlichkeitsmaß
  - straffes, 47
- Zentraler Grenzwertsatz von Lindeberg-Lévy, 49
- Zufallsgröße, 7
- Zufallsvariable, 7
- Zylindermengen, 25