Mitschrift zur Vorlesung

Wahrscheinlichkeitstheorie

von Wolfgang Wefelmeyer
im Wintersemester 1997/98

Martin R. Elsner

14. Oktober 1998

Einführung

Der Nobelpreis, der 1997 an Robert Merton und Myron Scholes vergeben wurde, zeugt von der großen Bedeutung der Wahrscheinlichkeitstheorie. Nicht nur in den Wirtschaftswissenschaften, wo z.B. der Bereich der Wertpapierentwicklung oder relativ neue stochastische Konjunkturtheorien mathematische Hilfe benötigen, sondern auch in den Naturwissenschaften ist die Theorie von der Wahrscheinlichkeit unerlässlich.

In der Physik seien nur die Stichworte Magnetisierung, Teilchensysteme, Supraleitung, Turbulenzen, Maxwellsche Geschwindigkeitsverteilung genannt.

Chemische Anwendungen finden sich z.B. im Bereich der Katalysatoren, der Proteinfaltung oder Polymere.

Stochastik in der Biologie begegnet uns in den wichtigen Bereichen der Evolutionstheorie und Epidemiologie, aber auch in der Neurologie, bei der stochastischen Resonanz und den Predator-Prey-Systemen.

Schließlich sei noch die Mineralogie genannt, in der das Kristallwachstum nicht ohne Wahrscheinlichkeitsbetrachtung denkbar wäre.

Dies stellt zwar nur einen Teil der interdisziplinären Anwendungen der Wahrscheinlichkeitstheorie dar, zeigt aber die Notwendigkeit der mathematischen Auseinandersetzung mit diesem Bereich.

Inhaltsverzeichnis

Kapitel I

Maßtheorie

1 Vereinigung und Durchschnitt von Mengen

Gegeben seien die Menge Ω und die Teilmengen $A_1, A_2, \dots$
Weiter werden die **De Morgan'schen Regeln** benötigt: $(\bigcup A_n)^c = \bigcap A_n^c, \quad (\bigcap A_n)^c = \bigcup A_n^c$.
Gilt $A_1 \subset A_2 \subset \dots$ und $A = \bigcup A_n$, dann schreibt man $A_n \uparrow A$ (aufsteigend);
analog für $A_1 \supset A_2 \supset \dots$ und $A = \bigcap A_n$ $A_n \downarrow A$ (absteigend).

Bemerkung 1.1 Mit De Morgan folgt:
$A_n \uparrow A \Rightarrow A_n^c \downarrow A^c$, $A_n \downarrow A \Rightarrow A_n^c \uparrow A^c$.

Für disjunkte A_n sei $\sum A_n := \bigcup A_n$.

Bemerkung 1.2 Jede Vereinigung lässt sich disjunkt schreiben:
$\bigcup A_n = \sum B_n, \quad B_n = A_n \cap \bigcap\limits_{i=1}^{n-1} A_i^c,$
also $\bigcup A_n = A_1 \cup (A_2 \cap A_1^c) \cup (A_3 \cap A_1^c \cap A_2^c) \cup \dots$

Desweiteren gelte $\limsup A_n = \bigcap\limits_n \bigcup\limits_{k=n}^{\infty} A_k, \quad \liminf A_n = \bigcup\limits_n \bigcap\limits_{k=n}^{\infty} A_k$.
Gilt $\limsup A_n = \liminf A_n$, so schreiben wir $A = \lim A_n, \quad A_n \to A$.

Bemerkung 1.3 vgl. $\limsup x_n = \inf_n \sup_{k \geq n} x_k, \quad \liminf x_n = \sup_n \inf_{k \geq n} x_k$.

Bemerkung 1.4 Wenn $A_n \uparrow$, dann gilt $\bigcup A_n = \sum B_n, \quad B_n = A_n - A_{n-1}$, wobei $A_0 = \emptyset$.

2 Algebra, σ-Algebra, Maß

Gegeben sei die Menge Ω, $\mathcal{F}$ sei eine Menge von Teilmengen von Ω.

Definition 2.1 *$\mathcal{F}$ heißt **Algebra**, wenn*

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

3. $A_1, A_2 \in \mathcal{F} \Rightarrow A_1 \cup A_2 \in \mathcal{F}$

Bemerkung 2.2 Eine Algebra ist abgeschlossen unter endlichen Vereinigungen und Durchschnitten.

Beweis: 1. induktiv: $(A_1 \cup A_2) \cup A_3$
2. $\bigcap A_k = (\bigcup A_k^c)^c$ $\square$

Definition 2.3 *$\mathcal{F}$ heißt σ-**Algebra**, wenn $\mathcal{F}$ Algebra und abgeschlossen unter abzählbaren Vereinigungen ist:*

$$A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \bigcup A_n \in \mathcal{F}$$

$(\Omega, \mathcal{F})$ heißt **messbarer** Raum, die Mengen in $\mathcal{F}$ heißen messbare Mengen.

Bemerkung 2.4 Sei $\mathcal{S}$ eine Menge von Teilmengen von Ω.
Dann ist $\sigma(\mathcal{S}) = \bigcap\limits_{\substack{\mathcal{F} \supset \mathcal{S} \\ \mathcal{F} \sigma - Algebra}} \mathcal{F}$ eine σ-Algebra. Sie heißt von $\mathcal{S}$ **erzeugt**.

Beweis: Sei $A \in \sigma(\mathcal{S})$, also $A \in \mathcal{F}$ für alle σ-Algebren $\mathcal{F}$ mit $\mathcal{F} \supset \mathcal{S}$.
$\Rightarrow A^c \in \mathcal{F} \Rightarrow A^c \in \sigma(\mathcal{S})$.
Entsprechend $\Omega \in \sigma(\mathcal{S})$, abzählbare Vereinigungen. $\square$

Bemerkung 2.5 Ist $\mathcal{F}$ eine Algebra und abgeschlossen unter aufsteigenden Limites, so ist $\mathcal{F}$ σ-Algebra.

Beweis: $\bigcup\limits_{k=1}^{n} A_k \uparrow \bigcup\limits_{k=1}^{\infty} A_k$ $\square$

Beispiel 2.6 $\Omega = \mathbb{R}, \quad \mathcal{S} = \{(a,b] : a, b \in \mathbb{R}\}$
Sei $\mathcal{A}$ die Familie der endlichen disjunkten Vereinigungen von Intervallen (eventuell unendlich), dann ist $\mathcal{A}$ Algebra. $\mathcal{B} = \sigma(\mathcal{A})$ ist σ-Algebra, die **Borel-Algebra**.

Bemerkung 2.7 $\mathcal{B}$ enthält sämtliche Arten von Intervallen:
$(a,b) = \bigcup (a, b - \frac{1}{n}], \quad (-\infty, b] = \bigcup (-n, b], \quad [a,b] = \bigcap (a - \frac{1}{n}, b]$

Definition 2.8 *Sei $\mathcal{A}$ Algebra. Eine Mengenfunktion $\mu : \mathcal{A} \mapsto (-\infty, \infty]$ heißt (endlich) **additiv**, wenn für disjunkte $A_1, A_2 \in \mathcal{A}$ gilt:*

$$\mu(A_1 + A_2) = \mu A_1 + \mu A_2$$

Bemerkung 2.9 Induktiv folgt $\mu \sum\limits_{i=1}^{n} A_i = \sum\limits_{i=1}^{n} \mu A_i$.

Definition 2.10 *Sei $\mathcal{A}$ Algebra. Eine Mengenfunktion $\mu : \mathcal{A} \mapsto (-\infty, \infty]$ heißt **σ-additiv** , wenn für disjunkte $A_1, A_2, \ldots \in \mathcal{A}$ mit $\sum_{n=1}^{\infty} A_n \in \mathcal{A}$ gilt:*

$$\mu \sum_{n=1}^{\infty} A_n = \sum_{n=1}^{\infty} \mu A_n$$

Definition 2.11 *Eine nichtnegative σ-additive Mengenfunktion auf einer σ-Algebra heißt **Maß**; gilt $\mu\Omega = 1$, heißt sie **Wahrscheinlichkeitsmaß** .*

Definition 2.12 *Eine Mengenfunktion heißt **endlich**, wenn sie nur endliche Werte annimmt. Eine nichtnegative, additive Mengenfunktion auf einer Algebra heißt **σ-endlich**, wenn*

$$\Omega = \bigcup_{n=1}^{\infty} A_n \quad \textit{mit } A_n \in \mathcal{A}, \quad \mu A_n < \infty$$

Beispiel 2.13 $\Omega = \mathbb{R}, \mathcal{A}$ Familie der endlichen disjunkten Vereinigungen von Intervallen $(a, b]$, $\mu(a, b] = b - a$ Länge des Intervalls.
$\mu \sum_{k=1}^{n} (a_k, b_k] = \sum_{k=1}^{n} \mu(a_k, b_k]$ σ-endlich auf $\mathcal{A}$.
Es stellen sich zwei Fragen:
1. Ist μ σ-additiv auf $\mathcal{A}$?
2. Ist μ fortsetzbar zu einem Maß auf $\mathcal{B} = \sigma(\mathcal{A})$?

Rechenregeln für additive Mengenfunktionen

Satz 2.14 *Sei μ additiv auf einer Algebra $\mathcal{A}$, und es existiere $A \in \mathcal{A}$ mit $\mu A < \infty$.*

1. $\mu\emptyset = 0$
2. $\mu(A \cup B) + \mu(A \cap B) = \mu A + \mu B, \quad A, B \in \mathcal{A}$
3. $\mu A = \mu B + \mu(A - B), \quad A, B \in \mathcal{A}, \quad B \subset A$
4. *Ist μ nichtnegativ, dann gilt*
$$\mu \bigcup_{k=1}^{n} A_k \leq \sum_{k=1}^{n} \mu A_k, \quad A_1, \ldots, A_n \in \mathcal{A}$$
5. *Ist μ Maß, dann gilt*
$$\mu \bigcup_{n=1}^{\infty} A_n \leq \sum_{n=1}^{\infty} \mu A_n, \quad A_1, A_2, \ldots \in \mathcal{A}, \quad \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$$

Beweis:
1. Sei $A \in \mathcal{A}$ mit $\mu A < \infty$. Es gilt $\mu A = \mu(A + \emptyset) = \mu A + \mu\emptyset$

2. $\mu A = \mu(A \cap B) + \mu(A - B), \quad \mu B = \mu(A \cap B) + \mu(B - A)$
   Also $\mu A + \mu B = \mu(A \cap B) + (\mu(A \cap B) + \mu(A - B) + \mu(B - A)) = \mu(A \cap B) + \mu(A \cup B)$
3. $A = B + (A - B)$
4. $\bigcup\limits_{k=1}^{n} A_k = \sum\limits_{k=1}^{n} \underbrace{(A_k \cap \bigcap\limits_{j=1}^{k-1} A_j^c)}_{\subset A_k}$ nach Bem. 1.2. Jetzt 3. anwenden.
5. wie 4. □

Satz 2.15 *Seien $\mathcal{F}$ σ-Algebra, μ σ-additiv und $A_1, A_2, \ldots \in \mathcal{F}$.*

1. $A_n \uparrow A \;\Rightarrow\; \mu A_n \to \mu A$
2. $A_n \downarrow A,\ \mu A_1 < \infty \;\Rightarrow\; \mu A_n \to \mu A$

Beweis:
1. a) Gilt $\mu A_n = \infty$ für ein n, dann auch $\forall k \geq n$ wegen $\mu A_k = \mu A_n + \mu(A_k - A_n)$
   b) Sei also μ endlich. Schreibe $A = A_1 + (A_2 - A_1) \ldots$ (Bem. 1.4). Es gilt: $\mu A = \mu A_1 + \mu A_2 - \mu A_1 + \ldots + \mu A_n - \mu A_{n-1} + \sum_{k>n} \mu(A_k - A_{k-1}) = \mu A_n + \sum_{k>n} \mu(A_k - A_{k-1})$
2. $A_1 - A_n \uparrow A_1 - A$. Jetzt 1. anwenden. □

Satz 2.16 *Sei μ additiv auf einer Algebra $\mathcal{A}$.*

1. *Ist μ aufsteigend stetig, also $\mu A_n \to \mu A$ für $A_1, A_2, \ldots \in \mathcal{A}$, $A_n \uparrow A \in \mathcal{A}$, dann ist μ σ-additiv.*
2. *Ist μ absteigend stetig gegen $\emptyset$, also $\mu A_n \to \mu\emptyset$ für $A_1, A_2, \ldots \in \mathcal{A}$, $A_n \downarrow \emptyset$, dann ist μ σ-additiv.*

Beweis: Seien $A_1, A_2, \ldots \in \mathcal{A}$ disjunkt, $A = \bigcup A_n \in \mathcal{A}$. Setze $B_n = \sum_{k=1}^{n} A_k$.
1. $B_n \uparrow A$, also $\sum_{k=1}^{n} \mu A_k = \mu B_n \to \mu A$
2. $\mu A = \mu B_n + \mu(A - B_n)$, $A - B_n \downarrow \emptyset$, also $\mu(A - B_n) \to 0$, also $\sum_{k=1}^{n} \mu A_k = \mu B_n \to \mu A$ □

In den folgenden Abschnitten 3 und 4 werden die Fragen 1 und 2 aus Beispiel 2.13 allgemeiner beantwortet, und zwar in umgekehrter Reihenfolge:

1. Maße lassen sich von Algebren auf σ-Algebren fortsetzen.
2. „Längenmaße" sind σ-additiv.

3 Maßerweiterung

Gibt es Maße auf σ-Algebren?

Satz 3.1 (Erweiterungssatz von Carathéodory) *Ein σ-endliches Maß auf einer Algebra $\mathcal{A}$ lässt sich eindeutig zu einem Maß auf $\sigma(\mathcal{A})$ fortsetzen.*

Zunächst zur Eindeutigkeit:

Definition 3.2 *$\mathcal{P}$ heißt π-System, wenn es abgeschlossen gegen Durchschnitte ist.*

Definition 3.3 *$\mathcal{L}$ heißt λ-System, wenn*

1. $\Omega \in \mathcal{L}$
2. $A, B \in \mathcal{L},\ A \subset B \Rightarrow B - A \in \mathcal{L}$
3. $A_n \in \mathcal{L},\ A_n \uparrow A \Rightarrow A \in \mathcal{L}$

Lemma 3.4 *Ist $\mathcal{L}$ π- und λ-System, dann ist $\mathcal{L}$ σ-Algebra.*

Beweis: 1. $\mathcal{L}$ ist Komplement-abgeschlossen, da $A^c = \Omega - A$
2. $\mathcal{L}$ ist Vereinigungs-abgeschlossen, da $A \cup B = (A^c \cap B^c)^c$
3. $\mathcal{L}$ ist abgeschlossen unter abzählbaren Vereinigungen: $\bigcup_{k=1}^{n} A_k \uparrow \bigcup_{k=1}^{\infty} A_k$ $\square$

Satz 3.5 (π-λ-Satz von Dynkin) *Ist $\mathcal{P}$ ein π-system und $\mathcal{L} \supset \mathcal{P}$ ein λ-System, dann gilt $\mathcal{L} \supset \sigma(\mathcal{P})$.*

Beweis: Sei $\lambda(\mathcal{P})$ das kleinste λ-System, in dem $\mathcal{P}$ enthalten ist.
a) Für $A \in \lambda(\mathcal{P})$ gilt: $\mathcal{L}_A := \{B : A \cap B \in \lambda(\mathcal{P})\}$ ist λ-System:
   1. $A \cap \Omega = A \in \lambda(\mathcal{P})$
   2. $A \cap (B - C) = A \cap B - A \cap C \in \lambda(\mathcal{P})$
   3. $B_n \uparrow B \Rightarrow B_n \cap A \uparrow B \cap A$
b) Für $A \in \mathcal{P}$ gilt $\mathcal{L}_A \supset \mathcal{P}$, also $\mathcal{L}_A \supset \lambda(\mathcal{P})$. Das heißt: für $A \in \mathcal{P}$ und $B \in \lambda(\mathcal{P})$ gilt $A \cap B \in \lambda(\mathcal{P})$ bzw. für $B \in \lambda(\mathcal{P})$ gilt: $\mathcal{L}_B \supset \mathcal{P}$. Also $\mathcal{L}_B \supset \lambda(\mathcal{P})$, d.h. $\lambda(\mathcal{P})$ ist Durchschnitts-abgeschlossen.
   $\Rightarrow \lambda(\mathcal{P})$ ist π-System, also nach Lemma 3.4 auch σ-Algebra. $\square$

Die Eindeutigkeit im Satz von Carathéodory ergibt sich aus folgendem

Satz 3.6 *Sei $\mathcal{P}$ π-System, μ_1, μ_2 Maße auf $\sigma(\mathcal{P})$, die auf $\mathcal{P}$ übereinstimmen. Es gebe $A_n \in \mathcal{P}, A_n \uparrow \Omega, \mu_1 A_n < \infty, \mu_2 A_n < \infty$.*
Dann stimmen μ_1 und μ_2 auf $\sigma(\mathcal{P})$ überein.

Beweis: Sei $A \in \mathcal{P}$ mit $\mu_1 A (= \mu_2 A) < \infty$.

a) Wir zeigen: $\mathcal{L}_A = \{B \in \sigma(\mathcal{P}) : \mu_1(A \cap B) = \mu_2(A \cap B)\}$ ist λ-System.
   1. $\mu_i(A \cap \Omega) = \mu_i A$
   2. $\mu_i(A \cap (B - C)) = \mu_i(A \cap B) - \mu_i(A \cap C)$
   3. $B_n \uparrow B \Rightarrow \mu_i(B_n \cap A) \uparrow \mu_i(A \cap B)$ nach Satz 2.15.

b) Nach Satz 3.5 gilt $\mathcal{L}_A \supset \sigma(\mathcal{P})$
   Also für $B \in \sigma(\mathcal{P}) : \quad \mu_1(A \cap B) = \mu_2(A \cap B)$
   Mit $A = A_n$ folgt $\mu_1 B = \mu_2 B$. □

Jetzt zur Existenz:

Definition 3.7 *Zu $E \in \Omega$ sei*

$$\mu^* E = \inf \left\{ \sum \mu A_n : E \subset \bigcup A_n, A_n \in \mathcal{A} \right\}$$

Bemerkung 3.8 μ^* ist monoton und subadditiv:
1. $E \subset F \Rightarrow \mu^* E \leq \mu^* F$
2. $E \subset \bigcup E_n \Rightarrow \mu^* E \leq \sum \mu^* E_n$

Eine solche Mengenfunktion heißt **äußeres Maß**.
(Der Beweis ist offensichtlich.)

Definition 3.9 *Eine Menge $E \subset \Omega$ heißt **messbar**, wenn*

$$\mu^* F = \mu^*(F \cap E) + \mu^*(F \cap E^c), \quad \forall F \subset \Omega$$

$\mathcal{A}^*$ sei die Menge der messbaren Mengen.

Es soll nun gezeigt werden. dass $\mu^* | \sigma(\mathcal{A})$ die gesuchte Fortsetzung ist. Dazu ist zu zeigen:

$$\mu^* | \mathcal{A} = \mu | \mathcal{A}, \; \mathcal{A} \subset \mathcal{A}^*, \; \mathcal{A}^* \; \sigma\text{-Algebra}, \; \mu^* \; \sigma\text{-additiv auf } \mathcal{A}^*$$

Lemma 3.10 *Für $A \in \mathcal{A}$ gilt $\mu^* A = \mu A$ und $A \in \mathcal{A}^*$*

Beweis:
1. $\leq$: Wähle $A_1 = A, \; A_2 = \ldots = \emptyset$
   $\geq$: $\mu A \leq \mu \bigcup A_n \leq \sum \mu A_n$ nach Satz 2.14
2. $\leq$: $\mu^* F \leq \mu^*(F \cap A) + \mu^*(F \cap A^c)$ wegen Subadditivität
   $\geq$: o.E. $\mu^* F < \infty$; wähle $B_n \in \mathcal{A}$ mit $F \subset \bigcup B_n, \; \sum \mu B_n \leq \mu^* F + \varepsilon$
   Es gilt: $\mu B_n = \mu(B_n \cap A) + \mu(B_n \cap A^c) \stackrel{1.}{=} \mu^*(B_n \cap A) + \mu^*(B_n \cap A^c)$.
   Also $\mu^* F + \varepsilon \geq \sum \mu B_n = \sum \mu^*(B_n \cap A) + \sum \mu^*(B_n \cap A^c) \geq \mu^*(F \cap A) + \mu^*(F \cap A^c)$ □

Lemma 3.11 *$\mathcal{A}^*$ ist σ-Algebra, μ^* ist Maß auf $\mathcal{A}^*$.*

Beweis:
1. $\mathcal{A}^*$ ist Komplement-abgeschlossen, da die Def. von 'messbar' symmetrisch ist.

2. $\mathcal{A}^*$ ist Vereinigungs-abgeschlossen: Seien $E_1, E_2 \in \mathcal{A}^*$
   Wegen $E_1 \cup E_2 = E_1 \cup (E_2 \cap E_1^c)$ gilt:
   $\mu^*(G \cap (E_1 \cup E_2) + \mu^*(G \cap (E_1 \cup E_2)^c) \leq \mu^*(G \cap E_1) + \mu^*(G \cap E_1^c \cap E_2) + \mu^*(G \cap E_1^c \cap E_2^c) = \mu^*(G \cap E_1) + \mu^*(G \cap E_1^c) = \mu^* G$ wegen der Messbarkeit von E_1 und E_2.
   '$\geq$' gilt wegen der Subadditivität.
3. Für G bel. und disjunkte $E_1, E_2 \in \mathcal{A}^*$ gilt:
   $\mu^*(G \cap (E_1 + E_2)) = \mu^*(G \cap (E_1 + E_2) \cap E_1) + \mu^*(G \cap (E_1 + E_2) \cap E_1^c) = \mu^*(G \cap E_1) + \mu^*(G \cap E_2)$
4. $\mathcal{A}^*$ ist abgeschlossen unter abzählbaren Vereinigungen:
   Seien $E_1, E_2, \ldots \in \mathcal{A}$, $E = \bigcup E_n$. Wegen 1. und 2. seien sie o.E. disjunkt. Setze $F_n = \bigcup_{k=1}^n E_k \in \mathcal{A}^*$ nach 2.
   Also mit Monotonie und 3.:
   $\mu^* G = \mu^*(G \cap F_n) + \mu^*(G \cap F_n^c) \geq \mu^*(G \cap F_n) + \mu^*(G \cap E^c) \overset{3.}{=} \sum_{k=1}^n \mu^*(G \cap E_k) + \mu^*(G \cap E^c)$
   Mit $n \to \infty$ und Subadditivität folgt:
   $\mu^* G \geq \sum_{k=1}^\infty \mu^*(G \cap E_k) + \mu^*(G \cap E^c) \geq \mu^*(G \cap E) + \mu^*(G \cap E^c)$
   '$\leq$' gilt wegen der Subadditivität.
5. Seien $E_1, E_2, \ldots \in \mathcal{A}^*$ disjunkt, $E = \sum E_n$
   Setze $F_n = \sum_{k=1}^n E_k$ wie in 4.
   Mit Monotonie und 4. folgt:
   $\mu^* E \geq \mu^* F_n = \sum_{k=1}^n \mu^* E_k$; jetzt $n \to \infty$
   '$\leq$'gilt wegen der Subadditivität. □

4 Lebesgue-Stieltjes-Maße und Verteilungsfunktion

Definition 4.1 *Ein* ***Lebesgue-Stieltjes-Maß*** *auf $\mathbb{R}$ ist ein Maß μ auf $\mathcal{B}$ mit $\mu I < \infty$ für alle endlichen Intervalle I.*

Definition 4.2 *Eine* ***Verteilungsfunktion*** *auf $\mathbb{R}$ ist eine nichtfallende, rechtsstetige Funktion $F : \mathbb{R} \mapsto \mathbb{R}$.*

Satz 4.3 *Sei μ Lebesgue-Stieltjes-Maß. Definiere F (bis auf eine additive Konstante) durch $F(b) - F(a) = \mu(a, b]$, $a < b$. Dann ist F Verteilungsfunktion.*

Beweis: 1. F ist nichtfallend wegen $\mu(a, b] \geq 0$
2. Es gelte $x_n \downarrow x$. Dann gilt $F(x_n) - F(x) = \mu(x, x_n] \to 0$ nach Satz 2.15. □

Wir gehen nun von einer Verteilungsfunktion F auf $\mathbb{R}$ aus und setzen dann $\mu(a, b] = F(b) - F(a)$.
Sei $\mathcal{A}$ wieder die Algebra der disjunkten Vereinigungen von Intervallen, die linksoffen sind. Setze μ additiv auf $\mathcal{A}$ fort:

$$\mu \sum_{k=1}^n (a_k, b_k] = \sum_{k=1}^n \mu(a_k, b_k]$$

Beispiel 4.4 $F(x) = x,\ \mu(a,b] = b - a$

Lemma 4.5 *μ ist σ-additiv auf $\mathcal{A}$.*

Beweis: Setze F auf $\bar{\mathbb{R}}$ fort: $F(\infty) = \lim_{x\to\infty} F(x),\ F(-\infty) = \lim_{x\to-\infty} F(x)$
Setze $\mu(a,b]$ entsprechend auf $\bar{\mathcal{A}}$ fort, wobei $\bar{\mathcal{A}}$ von den linksoffenen Intervallen erzeugte Algebra in $\bar{\mathbb{R}}$ ist, insbes. $\mu[-\infty,b] = \mu(-\infty,b]$.

1. Sei $F(\infty) - F(-\infty) < \infty$, seien $A_1, A_2, \ldots \in \bar{\mathcal{A}},\ A_n \downarrow \emptyset$
   Wähle B_n mit Abschluss $\bar{B}_n \subset A_n$ und $\mu A_n - \mu B_n < \varepsilon 2^{-n}$. Dann gilt für hinreichend großes n:
   $\bigcap_{k=1}^n B_k = \emptyset$ und somit $\mu A_n = \mu(A_n - \bigcap_{k=1}^n B_k) + \mu \cap_{k=1}^n B_k = \mu(A_n - \bigcap_{k=1}^n B_k) \leq \mu \bigcup_{k=1}^n (A_k - B_k) \leq \sum_{k=1}^n \mu(A_k - B_k) < \varepsilon$
   Also $\mu A_n \to 0$. Jetzt Satz 2.16.
2. Sei $F(\infty) - F(-\infty) = \infty$.
   Setze $F_n(x) = \begin{cases} F(x) & ||x|| \leq n \\ F(n) & x > n \\ F(-n) & x < -n \end{cases}$
   Es gilt $\mu_n \leq \mu,\ \mu_n \to \mu$ auf $\bar{\mathcal{A}}$ $\quad (\mu_n A = \mu(A \cap [-n,n]) \to \mu A)$.
   Seien $A_1, A_2, \ldots \in \bar{\mathcal{A}}$ disjunkt, $A := \sum A_k \in \bar{\mathcal{A}}$
   Es gilt $\mu A \geq \sum \mu A_k$ (Übung 5).
   a) $\sum \mu A_k = \infty \Rightarrow$ fertig.
   b) $\sum \mu A_k < \infty$: $\sum_{k=1}^\infty \mu_n A_k \stackrel{1.}{=} \mu_n A \to \mu A$. Also mit obiger Ungleichung: $0 \leq \mu A - \sum \mu A_k = \lim_{n\to\infty} \sum_{k=1}^\infty (\mu_n A_k - \mu A_k) \leq 0$ □

Satz 4.6 *Sei F Verteilungsfunktion; definiere $\mu(a,b] = F(b) - F(a),\ a < b$.*
Dann hat μ eine eindeutige Fortsetzung zu einem Lebesgue-Stieltjes-Maß.

Beweis: Nach Lemma 4.5 hat μ eine σ-additive Fortsetzung auf $\mathcal{A}$.
μ ist σ-endlich wegen $\mu(-n,n] = F(n) - F(-n) < \infty$.
Nach Satz 3.1 hat μ eine eindeutige Erweiterung auf $\mathcal{B}$. Die Erweiterung ist n.Def. Lebesgue-Stieltjes-Maß. □
Damit haben wir eine große Klasse von Maßen auf $\mathbb{R}$ gefunden.

Beispiel 4.7 $F(x) = x$
Das zugehörige Lebesgue-Stieltjes-Maß heißt **Lebesgue-Maß** und wird mit λ bezeichnet.

5 Messbare Funktionen und Lebesgue-Integral

Gegeben seien die messbaren Räume $(\Omega_1, \mathcal{F}_1), (\Omega_2, \mathcal{F}_2)$ und $f : \Omega_1 \mapsto \Omega_2$.

Bemerkung 5.1 $f^{-1}A^c = (f^{-1}A)^c,\ f^{-1}\bigcup A_n = \bigcup f^{-1}A_n,\ f^{-1}\bigcap A_n = \bigcap f^{-1}A_n$

Definition 5.2 *f heißt **messbar**, wenn $f^{-1}\mathcal{F}_2 \subset \mathcal{F}_1$.*

Bemerkung 5.3 Gilt $\mathcal{F}_2 = \sigma(\mathcal{S})$ und $f^{-1}\mathcal{S} \subset \mathcal{F}_1$, dann ist f messbar.

Beweis: $\{A \in \mathcal{F}_2 : f^{-1}A \in \mathcal{F}_1\} \supset \mathcal{S}$ ist σ-Algebra. $\square$

Bemerkung 5.4 Sind $(\Omega_i, \mathcal{O}_i)$ topologische Räume, $\mathcal{F}_i = \sigma(\mathcal{O}_i)$, so gilt:
Ist f stetig, dann auch messbar.

Beweis: Bemerkung 5.3 mit $\mathcal{S} = \mathcal{O}_2$ $\square$

Bemerkung 5.5 Sind $f : (\Omega_1, \mathcal{F}_1) \mapsto (\Omega_2, \mathcal{F}_2)$ und $g : (\Omega_2, \mathcal{F}_2) \mapsto (\Omega_3, \mathcal{F}_3)$ messbar, so auch $g \circ f : (\Omega_1, \mathcal{F}_1) \mapsto (\Omega_3, \mathcal{F}_3)$.

Beweis: Für $A \in \mathcal{F}_3$ gilt: $(g \circ f)^{-1}A = f^{-1}(g^{-1}A) \in \mathcal{F}_1$ $\square$

Definition 5.6 *Eine Funktion $f : (\Omega, \mathcal{F}) \mapsto (\mathbb{R}, \mathcal{B})$ heißt* ***Borel-messbar****.*

Bemerkung 5.7 Gilt $\{f \leq c\} \in \mathcal{F},\ c \in \mathbb{R}$, so ist f Borel-messbar.

Beweis: $\mathcal{B}$ wird erzeugt von $(-\infty, c],\ c \in \mathbb{R}$; mit Bemerkung 5.3 folgt die Behauptung. $\square$

Satz 5.8 *Sind $f_1, f_2, \ldots$ Borel-messbar mit $f_n \to f$, dann ist f Borel-messbar.*

Beweis: Nach Bemerkung 5.7 genügt: $\{f > c\} = \{\lim f_n > c\} =$
$\{f_n > c + \frac{1}{r}$ schließlich für ein $r \in \mathbb{N}\} = \bigcup_{r \in \mathbb{N}} \bigcup_n \bigcap_{k \geq n} \{f_k > c + \frac{1}{r}\} \in \mathcal{F}$ $\square$

Definition 5.9 *f heißt* ***einfach****, wenn*

$$f = \sum_{k=1}^{n} a_k 1_{A_k}, \quad A_k \in \mathcal{F} \text{ disjunkt}$$

Satz 5.10 *Eine nichtnegative Borel-messbare Funktion ist Limes einer aufsteigenden Folge einfacher Funktionen.*

Beweis: $f_n = \begin{cases} \frac{k-1}{2^n} & \frac{k-1}{2^n} \leq f < \frac{k}{2^n},\ k = 1, \ldots, n2^n \\ n & f \geq n \end{cases}$ $\square$

Bemerkung 5.11 Setze $f^+ = \max\{f, 0\},\ f^- = \max\{-f, 0\}$.
Dann ist $f = f^+ - f^-$, f^+, f^- und $|f|$ sind messbar, und f ist Limes einfacher Funktionen.

Beweis: 1. $\{f^+ \in B\} = \{f \geq 0, f \in B\} \cup \{f < 0, 0 \in B\}$
2. $|f| = f^+ + f^-$
3. Wähle $f_n^+ \to f^+,\ f_n^- \to f^-$ wie in Satz 5.10. Dann gilt $f_n^+ - f_n^- \to f^+ - f^- = f$. $\square$

Satz 5.12 *Sind f, g Borel-messbar, dann auch $f+g, f-g, fg, f/g$ (wenn wohldefiniert).*

Beweis: Wähle $f_n \to f,\ g_n \to g$ wie in Satz 5.10.
Dann $f_n + g_n \to f + g,\ f_n g_n \to fg,\ \frac{f_n}{g_n + \frac{1}{n} 1_{\{g_n = 0\}}} \to \frac{f}{g}$, und die Folgen sind einfach. Also sind die Limites messbar nach Satz 5.8. □

Definition 5.13 *Sei μ ein Maß auf der σ-Algebra $\mathcal{F}$.*

1. *Für $f = 1_A$ sei $\mu f = \int 1_A d\mu = \mu A$*
2. *Für $f = \sum a_i 1_{A_i}$ einfach sei $\mu f = \sum a_i \mu A_i$*
3. *Für $f \geq 0$ messbar sei $\mu f = \sup\{\mu e : 0 \leq e \leq f, e \text{ einfach}\}$*
4. *Für f messbar sei $\mu f = \mu f^+ - \mu f^-$, wenn nicht $\mu f^+ = \mu f^- = \infty$.*

Ist μf endlich, so heißt f ***integrierbar****.*

Bemerkung 5.14 μf in 2. ist wohldefiniert:
Falls $f = \sum a_i 1_{A_i} = \sum b_j 1_{B_j}$ einfach, dann $f = \sum c_{ij} 1_{A_i \cap B_j}$, $c_{ij} = a_i = b_j$, also $\sum a_i \mu A_i = \sum_i a_i \sum_j \mu(A_i \cap B_j) = \sum_{i,j} c_{ij} \mu(A_i \cap B_j)$, entsprechend für $\sum b_j \mu B_j$.

6 Eigenschaften des Integrals

Satz 6.1 *Sind f, g integrierbar, so gilt:*

1. *$\mu f \leq \mu g$ für $f \leq g$*
2. *$\mu(cf) = c \mu f$*
3. *$|\mu f| \leq \mu |f|$*

Beweis:
1. Für $f, g \geq 0$ gilt $\mu f \leq \mu g$ nach Definition.
2. Sei $f \geq 0$. Dann ist $\mu(cf) = \sup\{\mu e : 0 \leq e \leq cf\} = \sup\{\mu(ce) : 0 \leq e \leq f\} = \sup\{c\mu e : 0 \leq e \leq f\} = c \sup\{\mu e : 0 \leq e \leq f\}$
3. $-|f| \leq f \leq |f|$, also $-\mu|f| \leq \mu f \leq \mu|f|$ □

Satz 6.2 *Ist f integrierbar, dann ist $\nu B = \mu 1_B f,\ B \in \mathcal{F}$, σ-additiv.*

Beweis:
1. Für $f = \sum a_i 1_{A_i}$ einfach: $\nu B = \sum a_i \mu(A_i \cap B)$ σ-additiv.
2. Sei $f \geq 0,\ B_i \in \mathcal{F}$ disjunkt, $B = \sum B_i$.
   a) Beh.: $\nu B \leq \sum \nu B_i$
   Sei $0 \leq s \leq f$, s einfach.
   $\mu 1_B s \overset{1.}{=} \sum \mu 1_{B_n} s \leq \sum \mu 1_{B_n} f = \sum \nu B_n$. Also $\nu B = \mu 1_B f \leq \sum \nu B_n$.

b) Beh.: $\nu B \geq \sum \nu B_i$
Zunächst: $\nu B_n = \mu 1_{B_n} f \leq \mu 1_B f = \nu B$; also o.E. $\nu B_n < \infty$
Zu n und ε wähle $0 \leq s \leq f$ einfach mit $\mu 1_{B_i} s \geq \mu 1_{B_i} f - \frac{\varepsilon}{n}$, $i = 1, \ldots, n$. Dann ist mit Satz 6.1 $\nu B \geq \nu(\sum_{i=1}^n B_i) = \mu 1_{\sum_{i=1}^n B_i} f \geq \mu 1_{\sum_{i=1}^n B_i} s = \sum_{i=1}^n \mu 1_{B_i} s \geq \sum_{i=1}^n \mu 1_{B_i} f - \varepsilon = \sum_{i=1}^n \nu B_i - \varepsilon$. $n \to \infty$, $\varepsilon \to 0$ liefert die Beh.

3. Schreibe $f = f^+ - f^-$. Dann $\nu B = \mu 1_B f^+ - \mu 1_B f^-$. Jetzt 2. anwenden. $\square$

Satz 6.3 (Satz von der monotonen Konvergenz) *Seien* $0 \leq f_1 \leq f_2 \leq \ldots$ *messbar,* $f = \lim f_n$. *Dann gilt* $\mu f_n \to \mu f$.

Beweis: 1. $\lim \mu f_n$ existiert, da $\mu f_n \uparrow$ nach Satz 6.1.
2. Ebenso gilt $\lim \mu f_n \leq \mu f$, da $\mu f_n \leq \mu f$ nach Satz 6.1.
3. $\lim \mu f_n \geq \mu f$:
Sei $0 \leq s \leq f$, s einfach. Sei $b \in (0,1)$.
Es gilt $B_n = \{f_n \geq bs\} \uparrow \Omega$. Es folgt $\lim \mu f_n \overset{Mon.}{\geq} \mu f_m \overset{6.1a}{\geq} \mu f_m 1_{B_m} \overset{6.1b}{\geq} b\mu s 1_{B_m} \overset{2.15}{\to} b\mu s$. Bildet man das sup über s für beliebige b, so folgt die Beh. $\square$

Satz 6.4 (Additivitätssatz) *Sind* f, g *integrierbar, so gilt* $\mu(f+g) = \mu f + \mu g$

Beweis: 1. Für f, g einfach folgt die Behauptung aus der Definition.
2. f, g nichtnegativ: Wähle $0 \leq s_n \uparrow f$, $0 \leq t_n \uparrow g$ einfach. Mit 1. und Satz 6.3 folgt die Beh.
3. Wenn $f = h - k$, $h, k \geq 0$ integrierbar, dann $\mu f = \mu h - \mu k$:
Es gilt: $f^+ + k = f^- + h$, also mit 2. $\mu f^+ + \mu k = \mu(f^+ + k) = \mu(f^- + h) = \mu f^- + \mu h$
4. Für beliebige f, g ist $f + g = f^+ + g^+ - (f^- + g^-)$. Jetzt 3. und 2. anwenden. $\square$

Beispiel 6.5 I.a. gilt nicht $(f+g)^+ = f^+ + g^+$; z.B. für $f = -g$.

Lemma 6.6 (Lemma von Fatou) *Seien* $f_n \geq 0$ *messbar. Dann gilt* $\liminf \mu f_n \geq \mu \liminf f_n$.

Beweis: $g_n = \inf_{k \geq n} f_k \uparrow \liminf f_n$, also $\mu g_n \to \mu \liminf f_n$ nach Satz 6.3. Außerdem $\mu g_n \leq \mu f_n$ (Satz 6.1 a). $\square$

Satz 6.7 (Satz von der dominierten Konvergenz) *Seien* $f_1, f_2, \ldots$ *messbar,* $f_n \to f$, $|f_n| \leq g$, g *integrierbar. Dann gilt* $\mu f_n \to \mu f$.

Beweis: $f_n + g \geq 0$. Nach Lemma 6.6 gilt $\liminf \mu(f_n + g) \geq \mu(f + g)$. Wegen der Additivität folgt $\liminf \mu f_n \geq \mu f$. Das gleiche Argument mit $-f$ liefert $\limsup \mu f_n \leq \mu f$. $\square$

7 Satz von Radón-Nikodým

Gegeben sei der messbare Raum $(\Omega, \mathcal{F})$.
Eine σ-additive Mengenfunktion auf einer σ-Algebra heißt **signiertes** Maß.

Satz 7.1 *Sei μ signiertes Maß auf $\mathcal{F}$. Dann existieren $C, D \in \mathcal{F}$ mit*

$$\mu C = \sup_{A \in \mathcal{F}} \mu A, \quad \mu D = \inf_{A \in \mathcal{F}} \mu A$$

Beweis: reicht für sup; inf mit $-\mu$.
O.E. sei $\mu < \infty$, denn sonst $\exists C$ mit $\mu C = \infty$.
Wähle $A_n \in \mathcal{F}$, $\mu A_n \to \sup \mu$. Schreibe $A = \bigcup A_n$ als Vereinigung von 2^n disjunkten Mengen: $A_n^d = A_1^{\delta_1} \cap \ldots \cap A_n^{\delta_n}$, $d = (\delta_1, \ldots, \delta_n)$ und $A_1^\delta = \begin{cases} A_1, & \delta=1 \\ A-A_1, & \delta=0 \end{cases}$.
Die Zerlegung wird mit $n \to \infty$ kleiner. Setze $B_n = \sum_{\mu A_n^d \geq 0} A_n^d$.
Dann gilt $\mu A_n \leq \mu B_n \leq \mu \bigcup_{k=n}^r B_k \to \mu \bigcup_{k=n}^\infty B_k$
Setze $C = \limsup B_n = \bigcap_n \bigcup_{k=n}^\infty B_k$.
Mit Satz 2.15 und $\mu < \infty$ folgt $\sup \mu = \lim \mu A_n \leq \lim \mu \bigcup_{k=n}^\infty B_k = \mu C \leq \sup \mu$ □

Satz 7.2 (Halm, Jordan) *Sei μ signiertes Maß auf $\mathcal{F}$,*

$$\mu^+ A = \sup\{\mu B : B \in \mathcal{F},\ B \subset A\}, \ \mu^- A = -\inf\{\mu B : B \in \mathcal{F},\ B \subset A\}$$

Dann sind μ^+, μ^- Maße mit $\mu = \mu^+ - \mu^-$.

Beweis: Wähle D nach Satz 7.1. Es gilt $-\infty < \mu D \leq 0$.
1. $\mu(A \cap D) \leq 0,\ \mu(A \cap D^c) \geq 0,\ A \in \mathcal{F}$
   a) Ann.: $\mu(A \cap D) > 0$. Dann $\mu(A^c \cap D) = \mu D - \mu(A \cap D) < \mu D$
   b) Ann.: $\mu(A \cap D^c) < 0$. Dann $\mu(D + (A \cap D^c)) = \mu D + \mu(A \cap D^c) < \mu D$
2. $\mu^+ A = \mu(A \cap D^c),\ \mu^- A = -\mu(A \cap D)$
   a) Für $B \in \mathcal{F},\ B \subset A$ gilt $\mu B = \mu(B \cap D) + \mu(B \cap D^c) \overset{1.a}{\leq} \mu(B \cap D^c) \leq \mu(B \cap D^c) + \mu((A - B) \cap D^c) = \mu(A \cap D^c)$; also $\mu^+ A \leq \mu(A \cap D^c) \leq \mu^+ A$
   b) Für $B \in \mathcal{F},\ B \subset A$ gilt $\mu B = \mu(B \cap D) + \mu(B \cap D^c) \overset{1.b}{\geq} \mu(B \cap D) \geq \mu(B \cap D) + \mu((A - B) \cap D) = \mu(A \cap D)$; also $-\mu^- A \geq \mu(A \cap D) \geq -\mu^- A$ □

Bemerkung 7.3 μ^+/μ^- heißen **Positiv-/Negativteil** von μ, $|\mu| = \mu^+ + \mu^-$ heißt **Totalvariation** von μ.

Definition 7.4 *Ist μ Maß und ν ein signiertes Maß auf $\mathcal{F}$, und gilt $\mu A = 0 \Rightarrow \nu A = 0$, dann heißt ν* ***absolut-stetig*** *bezüglich μ ($\nu \ll \mu$).*

Satz 7.5 (Radón-Nikodým) *Sei μ ein σ-endliches Maß und ν ein σ-endliches signiertes Maß auf $\mathcal{F}$ mit $\nu \ll \mu$.*
Dann existiert ein μ-fast sicher eindeutiges messbares f mit

$$\nu A = \mu 1_A f, \ A \in \mathcal{F}$$

*f wird auch als μ-**Dichte** von ν bezeichnet.*

Beweis:

1. Eindeutigkeit:
   Seien f, g zwei μ-Dichten von ν, sei $E \subset \{n \geq f \geq g\}$, $\mu E < \infty$.
   $\mu 1_E f = \mu 1_E g$, also $\mu 1_E (f - g) = 0$, also $f = g$ μ-f.s.
   ($f \geq 0$ *f.s.*, $\mu f = 0$; $\mu\{f \geq \frac{1}{n}\} \leq n\mu 1_{\{f \geq \frac{1}{n}\}} f \leq n\mu f = 0$, $\{f > 0\} = \bigcup \{f \geq \frac{1}{n}\} \Rightarrow f = 0$ *f.s.*)
2. Existenz:
   a) Seien μ, ν endliche Maße; sei $\mathcal{S} = \{f \geq 0, \mu - integrierbar, \mu 1_A f \leq \nu A, A \in \mathcal{F}\}$
   Partielle Ordnung: $f \leq g$, wenn $f \leq g$ μ-f.s.
   (i) Jede totalgeordnete Teilmenge von $\mathcal{S}$ hat eine obere Schranke:
   Sei $\mathcal{T} \subset \mathcal{S}$ totalgeordnet, $s := \sup_{f \in \mathcal{T}} \mu f$. Wähle $f_n \in \mathcal{T}$ mit $\mu f_n \uparrow s$, insbesondere $f_n \leq f_{n+1}$. Setze $f = \sup f_n$.
   Nach Satz 6.3 gilt dann $\mu f = s$ und $f \in \mathcal{S}$
   f ist eine obere Schranke von $\mathcal{T}$:
   Gilt $g \leq f_n$ für ein n, dann $g \leq f$; gilt $g \geq f_n$ $\forall n$, dann $g \geq f$, also $\mu g \geq \mu f = s \Rightarrow \mu g = \mu f$, also $g = f$.
   (ii) Nach dem Lemma von Zorn existiert eine obere Schranke f von $\mathcal{S}$.
   f erfüllt die Behauptung:
   $\rho A := \nu A - \mu 1_A f$. ρ ist ein Maß mit $\rho \ll \mu$, $\rho\Omega < \infty$. Z.z.: $\rho = 0$
   Ann.: $\rho\Omega > 0$
   Dann $\exists k: \ \mu\Omega - k\rho\Omega < 0$. Wende 1. aus Beweis Satz 7.2 an auf $\mu - k\rho$:
   $\exists D \in \mathcal{F}: \ \mu(A \cap D) \leq k\rho(A \cap D), \ \mu(A \cap D^c) \geq k\rho(A \cap D^c), \ A \in \mathcal{F}$
   Ann.: $\mu D = 0$
   Dann $\rho D = 0$, also $0 \leq \mu D^c - k\rho D^c = \mu\Omega - k\rho\Omega \Rightarrow$ Widerspruch, also $\mu D > 0$
   Sei $h = \frac{1}{k} 1_D$. Für $A \in \mathcal{F}$ gilt: $\mu 1_A h = \frac{1}{k}\mu(A \cap D) \leq \rho(A \cap D) \leq \rho A = \nu A - \mu 1_A f$.
   Also ist $\mu 1_A (f + h) \leq \nu A$, also ist f nicht maximal $\Rightarrow$ Widerspruch.
   b) Sei μ endliches Maß, ν σ-endliches Maß. Schreibe $\Omega = \sum A_n$, $\nu A_n < \infty$. Setze $\nu_n A = \nu(A \cap A_n)$. Wähle f_n mit $\nu_n A = \mu 1_A f_n$ nach a). Mit $f = \sum f_n$ folgt die Behauptung.
   c) Seien μ, ν σ-endliche Maße. Schreibe $\Omega = \sum A_n$, $\mu A_n < \infty$. Wähle $f_n | A_n$ mit $\nu(A \cap A_n) = \mu 1_{A \cap A_n} f_n$. Setze $f_n | A_n^c = 0$. Mit $f = \sum f_n$ folgt $\nu A = \sum \nu(A \cap A_n) = \sum \mu 1_{A \cap A_n} f_n = \mu 1_A f$
   d) Seien μ, ν beliebig. Zerlege $\nu = \nu^+ - \nu^-$, ν^- endlich. Wähle f^+, f^- mit $\nu^+ A = \mu 1_A f^+$, $\nu^- A = \mu 1_A f^-$ nach c). Wegen ν^- endlich ist f^- μ-integrierbar, also f^- μ-f.s. endlich, also ist $f = f^+ - f^-$ wohldefiniert und $\nu A = \mu 1_A f$ □

In den nächsten beiden Kapiteln stehen dynamische Prozesse im Vordergrund, die in der Vorlesung „Einführung in die Stochastik" auch als stochastische Prozesse oder gekoppelte Experimente bezeichnet wurden. Als Beispiel stelle man sich eine (endliche) Folge von täglichen Börsenschlusskursen vor. Diese bilden ein Ergebnis $\omega = (\omega_1, \ldots, \omega_n)$ aus dem $\mathbb{R}^n$.

8 Produktmaße, Satz von Fubini

Gegeben seien die messbaren Räume $(\Omega_k, \mathcal{F}_k)$, $k = 1, \ldots, n$.

Definition 8.1 *Sind $A_k \in \mathcal{F}_k$, so heißt $\times A_k$* ***messbares Rechteck****, die davon erzeugte σ-Algebra* ***Produkt-σ-Algebra*** $\bigotimes \mathcal{F}_k$.

Weiterhin sei für $F \in \Omega \times \Omega$ die *Schnittmenge* $F_{\omega_1} = \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in F\}$.

Satz 8.2 (Produktmaßsatz) *Sei $\mu_1|\mathcal{F}_1$ σ-endliches Maß. Für $\omega_1 \in \Omega_1$ sei $\mu(\omega_1, \cdot)|\mathcal{F}_2$ ein Maß, und $\mu(\cdot, A_2)$ sei messbar für $A_2 \in \mathcal{F}_2$. Es gelte $\Omega_2 = \bigcup B_n$ mit $\mu(\omega_1, B_n) \le k_n$ für $\omega_1 \in \Omega_1$.*
Dann existiert genau ein Maß μ auf $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$ mit $\mu(A \times B) = \int_A \mu(\omega_1, B)\mu_1(d\omega_1)$, $A \in \mathcal{F}_1$, $B \in \mathcal{F}_2$, nämlich $\mu F = \int \mu(\omega_1, F_{\omega_1})\mu_1(d\omega_1)$, $F \in \mathcal{F}$.

Beweis:

1. Existenz
   a) Sei $\mu(\omega_1, \cdot)$ endlich.
      (i) Ist $F \in \mathcal{F}$, so ist $F_{\omega_1} \in \mathcal{F}_2$: Sei $\mathcal{S} = \{F \in \mathcal{F} : F_{\omega_1} \in \mathcal{F}_2\}$
      – $\mathcal{S}$ ist σ-Algebra: $(\bigcup F_n)_{\omega_1} = \bigcup F_{n,\omega_1}$, $(F^c)_{\omega_1} = (F_{\omega_1})^c$
      – $\mathcal{S}$ enthält alle messbaren Rechtecke: $(A \times B)_{\omega_1} = \begin{cases} B & \omega_1 \in A \\ \emptyset & sonst \end{cases}$
      Also $\mathcal{S} = \mathcal{F}$.
      (ii) Ist $F \in \mathcal{F}$, so ist $\omega_1 \mapsto \mu(\omega_1, F_{\omega_1})$ $\mathcal{F}_1$-messbar: Sei $\mathcal{S} = \{F \in \mathcal{F} : \mu(\omega_1, F_{\omega_1})\ \mathcal{F}_1 - messbar\}$
      – $\mathcal{S}$ enthält alle messbaren Rechtecke: $\mu(\omega_1, (A \times B)_{\omega_1}) = \begin{cases} \mu(\omega_1, B) & \omega_1 \in A \\ \mu(\omega_1, \emptyset) = 0 & sonst \end{cases}$
      $= \mu(\omega_1, B)1_A(\omega_1)$
      – $\mathcal{S}$ enthält alle endlichen disjunkten Vereinigungen messbarer Rechtecke: Seien $C_1, \ldots, C_n$ disjunkte messbare Rechtecke, dann $\mu(\omega_1, (\sum C_i)_{\omega_1}) = \mu(\omega_1, \sum C_{i,\omega_1}) = \sum \mu(\omega_1, C_{i,\omega_1})$
      – Wenn $F_n \in \mathcal{S}$, $F_n \uparrow F$, dann $F_{n,\omega_1} \uparrow F_{\omega_1}$, also $\mu(\omega_1, F_{n,\omega_1}) \uparrow \mu(\omega_1, F_{\omega_1})$
      Also $\mathcal{S} = \mathcal{F}$.
      (iii) Nach (ii) existiert $\mu F = \int \mu(\omega_1, F_{\omega_1})\mu_1(d\omega_1)$, $F \in \mathcal{F}$.
      – μ ist Maß: Seien $F_1, F_2, \ldots \in \mathcal{F}$ disjunkt. Dann $\mu \sum F_n = \int \mu(\omega_1, (\sum F_n)_{\omega_1})\mu_1(d\omega_1)$
      $\overset{Übg.17}{=} \sum \int \mu(\omega_1, F_{n,\omega_1})\mu_1(d\omega_1) = \sum \mu F_n$
      – μ ist wie verlangt: $\mu(A \times B) = \int \mu(\omega_1, (A \times B)_{\omega_1})\mu_1(d\omega_1) = \int \mu(\omega_1, B)1_A(\omega_1)\mu_1(d\omega_1)$
   b) Sei $\mu(\omega_1, \cdot)$ wie vorausgesetzt, B_n disjunkt. Setze $\mu_n(\omega_1, B) = \mu(\omega_1, B \cap B_n)$. Dann folgt mit (a) $\mu = \sum \mu_n$ wie verlangt.
2. Eindeutigkeit
   Sei $\mathcal{A}$ die Algebra der endlichen disjunkten Vereinigungen messbarer Rechtecke. μ ist σ-endlich auf $\mathcal{A}$:
   μ_1 ist σ-endlich, also $\Omega_1 = \bigcup A_m$, $A_m \in \mathcal{F}_1$, $\mu_1 A_m < \infty$, also $\Omega_1 \times \Omega_2 = \sum A_m \times B_n$ mit $\mu(A_m \times B_n) = \int_{A_m} \mu(\omega_1, B_n)\mu_1(d\omega_1) \le k_n \mu_1 A_m < \infty$.

Sei ν ebenso wie verlangt, m.a.W. $\nu|\mathcal{A} = \mu|\mathcal{A}$. Dann gilt $\mu = \nu$ auf $\mathcal{F} = \sigma(\mathcal{A})$ nach Satz 3.1. $\square$

Korollar 8.3 (Klassischer Produktmaßsatz) *Seien $\mu_1|\mathcal{F}_1$, $\mu_2|\mathcal{F}_2$ σ-endliche Maße. Dann existiert genau ein Maß μ auf $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$ mit $\mu(A_1 \times A_2) = \mu_1 A_1 \cdot \mu_2 A_2$, $A_i \in \mathcal{F}_i$, nämlich $\mu F = \int \mu_2(F_{\omega_1})\mu_1(d\omega_1)$, $F \in \mathcal{F}$. Es wird auch als **Produktmaß** $\mu = \mu_1 \otimes \mu_2$ bezeichnet.*

Beweis: Setze $\mu(\omega_1, \cdot) = \mu_2$ in Satz 8.2. $\square$

Satz 8.4 (Satz von Fubini) *Es gelten die Voraussetzungen von Satz 8.2, f sei auf $\Omega_1 \times \Omega_2$ integrierbar.*
Dann ist $\int f(\cdot, \omega_2)\mu(\cdot, d\omega_2)$ μ_1-f.s. integrierbar, und $\mu f = \iint f(\omega_1, \omega_2)\mu(\omega_1, d\omega_2)\mu_1(d\omega_1)$.

Beweis: $f(\omega_1, \cdot)$ ist messbar: $f(\omega_1, \cdot)^{-1}B = (f^{-1}B)_{\omega_1} \in \mathcal{F}_2$

a) Sei $f = 1_F$, $F \in \mathcal{F}$:
$\int 1_F(\omega_1, \omega_2)\mu(\omega_1, d\omega_2) = \int 1_{F_{\omega_1}}(\omega_2)\mu(\omega_1, d\omega_2) = \mu(\omega_1, F_{\omega_1})$; $\mu 1_F = \mu F$
$\overset{\text{Satz 8.2}}{=} \int \mu(\omega_1, F_{\omega_1})\mu_1(d\omega_1) = \iint 1_F(\omega_1, \omega_2)\mu(\omega_1, d\omega_2)\mu_1(d\omega_1)$

b) Sei $f = \sum_{i=1}^n a_i 1_{F_i}$, F_i disjunkt, $a_i \geq 0$.
$\int f(\omega_1, \omega_2)\mu(\omega_1, d\omega_2) = \sum a_i \mu(\omega_1, F_{i,\omega_1})$,
$\mu f = \sum a_i \mu F_i = \sum a_i \iint 1_{F_i}(\omega_1, \omega_2)\mu(\omega_1, d\omega_2)\mu_1(d\omega_1) = \iint f(\omega_1, \omega_2)\mu(\omega_1, d\omega_2)\mu_1(d\omega_1)$

c) Sei $f \geq 0$. Wähle f_n einfach mit $f_n \uparrow f$.
$\int f(\omega_1, \omega_2)\mu(\omega_1, d\omega_2) = \lim \int f_n(\omega_1, \omega_2)\mu(\omega_1, d\omega_2)$. Mit Satz 6.3 folgt $\mu f = \lim \mu f_n$
$= \lim \iint f_n(\omega_1, \omega_2)\mu(\omega_1, d\omega_2)\mu_1(d\omega_1) = \iint f(\omega_1, \omega_2)\mu(\omega_1, d\omega_2)\mu_1(d\omega_1)$.

d) Sei f integrierbar. Dann $f = f^+ - f^-$
Nach c): $\iint f^\pm(\omega_1, \omega_2)\mu(\omega_1, d\omega_2)\mu_1(d\omega_1) - \mu f^\mp < \infty$, also $\int f^\pm(\omega_1, \omega_2)\mu(\omega_1, d\omega_2) < \infty$ μ_1-f.a. ω_1.
Also gilt für μ_1-f.a. ω_1: $\int f(\omega_1, \omega_2)\mu(\omega_1, d\omega_2) = \int f^+(\omega_1, \omega_2)\mu(\omega_1, d\omega_2) - \int f^-(\omega_1, \omega_2)\mu(\omega_1, d\omega_2)$. Jetzt integrieren und c) anwenden. $\square$

Korollar 8.5 (Klassischer Satz von Fubini) *Seien $\mu_1|\mathcal{F}_1$, $\mu_2|\mathcal{F}_2$ σ-endlich, f auf $\Omega_1 \times \Omega_1$ $\mu_1 \otimes \mu_2$-integrierbar. Dann gilt*

$$\mu f = \iint f(\omega_1, \omega_2)\mu_2(d\omega_2)\mu_1(d\omega_1)$$

Beweis: Setze $\mu(\omega_1, \cdot) = \mu_2$ in Satz 8.4. $\square$

9 Maße auf abzählbaren Produkträumen

Gegeben seien die messbaren Räume $(\Omega_1, \mathcal{F}_1), (\Omega_2, \mathcal{F}_2), \ldots, \Omega = \times\Omega_n$.

Definition 9.1 *Für $B^n \subset \Omega_1 \times \ldots \times \Omega_n$ sei $B_n = \{\omega : (\omega_1, \ldots, \omega_n) \in B^n\}$ der Zylinder mit Basis B^n. Für $B^n = A_1 \times \ldots \times A_n$ heißt B_n* ***Rechteck****, und* ***messbar****, wenn $A_k \in \mathcal{F}_k,\ k = 1, \ldots, n$.*

Bemerkung 9.2 Die messbaren Zylinder bilden eine Algebra. Die endlichen disjunkten Vereinigungen messbarer Rechtecke bilden eine Algebra. Beide erzeugen die gleiche σ-Algebra $\mathcal{F} = \otimes \mathcal{F}_n$.

Satz 9.3 *Sei P_1 Wahrscheinlichkeitsmaß auf $\mathcal{F}_1$, $P(\omega_1, \ldots, \omega_{n-1}, \cdot)$ Wahrscheinlichkeitsmaß auf $\mathcal{F}_n$. Für $B^n \subset \mathcal{F}_1 \otimes \ldots \otimes \mathcal{F}_n$ sei*
$P_n B^n = \int P_1(d\omega_1) \int P(\omega_1, d\omega_2) \ldots \int 1_{B^n}(\omega_1, \ldots, \omega_n) P(\omega_1, \ldots, \omega_{n-1}, d\omega_n)$.
Dann existiert ein eindeutiges Wahrscheinlichkeitsmaß P auf $\mathcal{F}$, das alle P_n fortsetzt: $PB_n = P_n B^n,\ B^n \in \mathcal{F}_1 \otimes \ldots \otimes \mathcal{F}_n$

Beweis:

a) P ist wohldefiniert auf den messbaren Zylindern:
Sei $B_n = C_m$, o.E. $m < n$. Dann $B^n = C^m \times \Omega_{m+1} \times \ldots \times \Omega_n$. Also $P_n B^n = P_m C^m$.

b) P ist offensichtlich additiv auf der Algebra der messbaren Zylinder. Nach Satz 3.1 genügt es, die σ-Additivität zu zeigen. Nach Satz 2.16 genügt dafür aber schon die Stetigkeit in $\emptyset$.
Seien $B_{n_1}, B_{n_2}, \ldots$ messbare Zylinder, $B_{n_i} \downarrow \emptyset$, o.E. $n_i = i$.
Ann.: $\lim PB_n > 0$
Schreibe $PB_n = \int P_1(d\omega_1) g_n^1(\omega_1)$. Es gilt $1_{B^{n+1}}(\omega_1, \ldots, \omega_n) \leq 1_{B^n}(\omega_1, \ldots, \omega_n)$, also $g_n^1 \downarrow$.
Sei $h_1 = \lim g_n^1$. Dann $PB_n \to P_1 h_1$ nach Satz 6.7. Also $h_1(\omega_1') > 0$ für ein $\omega_1' \in B^1$ (sonst $1_{B^n}(\omega_1', \omega_2, \ldots, \omega_n) = 0\ \forall n$, also $g_n^1(\omega_1') = 0\ \forall n$, also $h_1(\omega_1') = 0$).
Schreibe $g_n^1(\omega_1') = \int P(\omega_1', d\omega_2) g_n^2(\omega_2)$. Wieder $g_n^2 \downarrow h_2$, also $g_n^1(\omega_1) \to \int P(\omega_1', d\omega_2) h_2(\omega_2)$.
Also $h_2(\omega_2') > 0$ mit $(\omega_1', \omega_2') \in B^2$.
Insgesamt: $(\omega_1, \omega_2, \ldots) \in \bigcap B_n \Rightarrow$ Widerspruch. □

Korollar 9.4 (Klassischer Produktmaßsatz für abzählbare Produkträume)
Seien $(\Omega_n, \mathcal{F}_n, P_n)$ Wahrscheinlichkeitsräume, $n = 1, 2, \ldots$, $\Omega = \times \Omega_n$, $\mathcal{F} = \otimes \mathcal{F}_n$.
Dann gibt es genau ein Wahrscheinlichkeitsmaß P auf $\mathcal{F}$ mit $P\{\omega : \omega_1 \in A_1, \ldots, \omega_n \in A_n\} = \prod_{i=1}^{n} P_i A_i$. Es wird $\otimes P_n$ geschrieben und heißt ***Produktmaß****.*

Beweis:

1. Existenz:
Setze $P(\omega_1, \ldots, \omega_{n-1}, \cdot) = P_n$ in Satz 9.3 Dann $P_n(A_1 \times \ldots \times A_n) = \prod P_i A_i$, also hat das Wahrscheinlichkeitsmaß P aus Satz 9.3 die gewünschten Eigenschaften.
2. Eindeutigkeit:
P ist eindeutig auf der Algebra der endlichen disjunkten Vereinigungen messbarer Rechtecke, also eindeutig nach Satz 3.1. □

Kapitel II

Wahrscheinlichkeitstheorie

10 Zufallsvariable und Erwartungswert

Definition 10.1 *Sei $P|\mathcal{B}$ ein Wahrscheinlichkeitsmaß. Dann heißt $EP = \int xP(dx)$* ***Mittelwert*** *von P, $VarP = \int (x - EP)^2 P(dx)$* ***Varianz*** *von P.*

Im Folgenden sei immer der Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ gegeben.

Definition 10.2 *Eine Borel-messbare Funktion $X : \Omega \mapsto \mathbb{R}^n$ heißt (n-dimensionaler)* ***Zufallsvektor****, für $n = 1$* ***Zufallsvariable****. Die durch X* ***induzierte Verteilung*** *$P^X|\mathcal{B}^n$ ist definiert durch $P^X B = P(X \in B) = PX^{-1}B,\ B \in \mathcal{B}^n$.*

Satz 10.3 *Sei X n-dimensionaler Zufallsvektor, $g : \mathbb{R}^n \mapsto \mathbb{R}$ messbar.*
Dann gilt $P^X g = Pg \circ X$, wenn eine der beiden Seiten existiert.

Beweis: 1. $g = 1_B,\ B \in \mathcal{B}^n :\quad P^X 1_B = P(X \in B) = P1_B \circ X$. 2. $g = \sum_{i=1}^n a_i 1_{B_i},\ B_i \in \mathcal{B}^n$ disjunkt: $P^X \sum a_i 1_{B_i} = \sum a_i P1_{B_i} \circ X = P\sum a_i 1_{B_i} \circ X$ 3. $g \geq 0 :$ Wähle $g_1, g_2, \ldots$ einfach, $g_n \uparrow g$. Dann monotone Konvergenz (Satz 6.3). 4. g beliebig: $g = g^+ - g^-$ $\square$

Definition 10.4 *Sei X Zufallsvariable. Der Mittelwert von P^X wird EX geschrieben und* ***Erwartungswert*** *genannt. Die Varianz von P^X wird $VarX$ geschrieben: $VarX = \int (x - EX)^2 P^X(dx) = E(X - EX)^2$*

11 L_p

Gegeben sei ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$.

Definition 11.1 *Für $p > 0$ ist $\mathbf{L_p}$ die Menge der Zufallsvariablen X mit $E|X|^p < \infty$. Es sei $||X||_p = (E|X|^p)^{\frac{1}{p}}$.*

Die nächsten zwei Lemmata werden für die Höldersche Ungleichung benötigt:

Lemma 11.2 *Für $c, d, r \geq 0$ gilt*

$$c^r d^{1-r} \leq rc + (1-r)d$$

Beweis: äquivalent: $r \log c + (1-r) \log d \leq \log(rc + (1-r)d)$. Dies ist aber erfüllt, weil log konkav ist. $\square$

Lemma 11.3 *Sei $1 < p < \infty$. Definiere q durch $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt für $a, b > 0$*

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q$$

Beweis: Wende Lemma 11.2 an auf $r = \frac{1}{p},\ c = a^p,\ d = b^q$. $\square$

Satz 11.4 (Hölder-Ungleichung) *Sei $1 < p < \infty$. Definiere q durch $\frac{1}{p} + \frac{1}{q} = 1$. Für $X \in L_p,\ Y \in L_q$ gilt*

$$XY \in L_1 \quad \textit{und} \quad ||XY||_1 \leq ||X||_p ||Y||_q$$

Beweis: Wende Lemma 11.3 an auf $a = \frac{|X|}{||X||_p},\ b = \frac{|Y|}{||Y||_q}$. Dann gilt $\frac{|XY|}{||X||_p ||Y||_q} \leq \frac{|X|^p}{pE|X|^p} + \frac{|Y|^q}{qE|Y|^q}$. Bildung der Erwartungswerte liefert $\frac{||XY||_1}{||X||_p ||X||_q} \leq \frac{1}{p} + \frac{1}{q} = 1$ $\square$

Korollar 11.5 (Cauchy-Schwarz) *Für $X, Y \in L_2$ gilt $XY \in L_1$ und $||XY||_1 \leq ||X||_2 ||Y||_2$*

Folgendes Lemma dient dem Beweis des nächsten Satzes:

Lemma 11.6 *Für $p \geq 1$ und $a, b \geq 0$ gilt $(a+b)^p \leq 2^p(a^p + b^p)$*

Beweis: $(a+b)^p \leq (2\max\{a,b\})^p = 2^p \max\{a^p, b^p\} \leq 2^p(a^p + b^p)$ $\square$

Satz 11.7 (Minkowski-Ungleichung) *Ist $p \geq 1$ und $X, Y \in L_p$, dann gilt $X + Y \in L_p$ und $||X+Y||_p \leq ||X||_p + ||Y||_p$*

Beweis: Trivial für $p = 1$. Nach Lemma 11.6 gilt $|X+Y|^p \leq (|X| + |Y|)^p \leq 2^p(|X|^p + |Y|^p)$. Sei $p > 1$ und $q = \frac{p}{p-1}$. Nach Satz 11.4 gilt dann $||X+Y||_p^p = E|X+Y|^p \leq E|X|\ |X+Y|^{p-1} + E|Y|\ |X+Y|^{p-1} \leq (||X||_p + ||Y||_p)(E|X+Y|^{(p-1)q})^{\frac{1}{q}} = (||X||_p + ||Y||_p)||X+Y||_p^{\frac{p}{q}}$. Wegen $p - \frac{p}{q} = 1$ folgt die Behauptung. $\square$

Satz 11.8 (Tschebyscheff-Ungleichung) *Sei $X \geq 0$ Zufallsvariable. Dann gilt für $p, \varepsilon > 0$:*

$$P(X \geq \varepsilon) \leq \frac{1}{\varepsilon^p} EX^p$$

Beweis: $1_{(X \geq \varepsilon)} \leq \frac{1}{\varepsilon^p} X^p$ $\square$

Definition 11.9 *Seien $X, Y \in L_2$. Die **Kovarianz** von X und Y ist*

$$Cov(X,Y) = E(X - EX)(Y - EY)$$

*Der **Korrelations-Koeffizient** ist $\rho = \frac{Cov(X,Y)}{(VarX\ VarY)^{\frac{1}{2}}}$. X und Y heißen **unkorreliert**, wenn $\rho = 0$.*

Satz 11.10 *Seien $X_1, \ldots, X_n \in L_2$. Dann gilt*

$$Var \sum X_i = \sum VarX_i + 2 \sum_{i<j} Cov(X_i, X_j)$$

Sind $X_1, \ldots, X_n$ paarweise unkorreliert, so gilt $Var \sum X_i = \sum VarX_i$.

Beweis: $Var \sum X_i = E(\sum(X_i - EX_i))^2 = \sum E(X_i - EX_i)^2 + \sum_{i \neq j} E(X_i - EX_i)(X_j - EX_j)$ □

12 Unabhängige Zufallsvariablen

Definition 12.1 *Zufallsvariablen $X_1, \ldots, X_n$ heißen **unabhängig**, wenn*

$$P \bigcap (X_i \in B_i) = \prod P(X_i \in B_i), \;\; B_i \in \mathcal{B}$$

Satz 12.2 *$X_1, \ldots, X_n$ sind unabhängig genau dann, wenn $P^{(X_1,\ldots,X_n)}$ ein Produktmaß ist, nämlich $\otimes P^{X_i}$.*

Beweis: $P \bigcap (X_i \in B_i) = P^{(X_1,\ldots,X_n)}(\times B_i)$ □

Beispiel 12.3 Sei $P = \otimes P_i$ Produktmaß, $X_i(\omega) = \omega_i$ Projektion. Dann $P^{(X_1,\ldots,X_n)} = P = \otimes P_i = \otimes P^{X_i}$

Satz 12.4 *$X_1, \ldots, X_n$ sind unabhängig genau dann, wenn*

$$P \bigcap (X_i \leq t_i) = \prod P(X_i \leq t_i)$$

Beweis: $\times(-\infty, t]$ erzeugen $\mathcal{B}^n$ (vgl. Übung 25) und sind durchschnittsabgeschlossen. Also ist $P^{(X_1,\ldots,X_n)}$ Produktmaß nach Satz 3.6. Jetzt Satz 12.2 anwenden.

Definition 12.5 *$A_1, \ldots, A_n \in \mathcal{F}$ heißen unabhängig, wenn $1_{A_1}, \ldots, 1_{A_n}$ unabhängig sind.*

Satz 12.6 *$1_{A_1}, \ldots, 1_{A_n}$ sind unabhängig genau dann, wenn*

$$P \bigcap_{j=1}^{k} A_{i_j} = \prod_{j=1}^{k} PA_{i_j} \;\; \forall \{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}$$

Beweis:
$\Rightarrow$: $\bigcap_{j=1}^{k} A_{i_j} = \{1_{A_{i_j}} = 1,\ j = 1, \ldots, k,\ 1_{A_i} \in \Omega\ sonst\}$

$\Leftarrow$: $\{1_A \geq t\} = \begin{cases} \emptyset & t > 1 \\ A & 0 < t \leq 1 \\ \Omega & t \leq 0 \end{cases}$

Also $\bigcap(1_{A_i} \geq t_i) = \bigcap_{0<t_i\leq 1} A_i$. Jetzt Satz 12.4 anwenden. □

Satz 12.7 *Sind $X_1, \ldots, X_n$ unabhängige Zufallsvektoren und $f_1, \ldots, f_n : I\!R^{n_j} \mapsto I\!R$ messbar, dann sind $f_1 \circ X_1, \ldots, f_n \circ X_n$ unabhängig.*

Beweis: $P\bigcap(f_i \circ X_i \in B_i) = P\bigcap(X_i \in f_i^{-1}B_i) = \prod P(X_i \in f_i^{-1}B_i) = \prod P(f_i \circ X_i \in B_i)$ □

Satz 12.8 *Sind $X_1, \ldots, X_n$ unabhängige Zufallsvariablen, $1 \leq n_1 < n_2 < \ldots < \ldots < n_k = n$ $(n_0 = 0)$, und $f_j : I\!R^{n_j - n_{j-1}} \mapsto I\!R$ messbar, dann sind $f_j \circ (X_{n_{j-1}+1}, \ldots, X_{n_j})$, $j = 1, \ldots, k$, unabhängig.*

Beweis: Nach Satz 12.2 gilt: $P^{(X_1,\ldots,X_n)} = \otimes P^{X_i}$. Also $P^{((X_1,\ldots,X_n),\ldots,(X_{n_{k-1}+1},\ldots,X_n))}$
$= \otimes_{j=1}^{k} P^{(X_{n_{j-1}+1},\ldots,X_{n_j})}$. Jetzt Satz 12.7 anwenden. □

Satz 12.9 *Sind X und Y unabhängig mit endlichem Erwartungswert, dann gilt*

$$EXY = EX \cdot EY$$

Beweis: $EXY \overset{10.3}{=} \int xyP^{(X,Y)}(dx, dy) = \int xyP^X(dx)P^Y(dy) \overset{8.5}{=} \int x \int yP^Y(dy)P^X(dx)$
$= \int xP^X(dx) \int yP^Y(dy) = EX \cdot EY$ □

Korollar 12.10 *Sind $X_1, \ldots, X_n$ unabhängig mit endlichem Erwartungswert, dann*

$$E\prod X_i = \prod EX_i$$

Beweis: Mit Induktion aus Satz 12.9: $\prod_{i=1}^{k} X_i$, $\prod_{i=k+1}^{n} X_i$ unabhängig nach Satz 12.8. □

Satz 12.11 (Schwaches Gesetz der großen Zahl) *Seien $X_1, \ldots, X_n \in L_2$ unabhängig und identisch verteilt. Dann gilt für $\varepsilon > 0$:*

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - EX_1\right| \geq \varepsilon\right) \leq \frac{VarX_1}{n\varepsilon^2}$$

Beweis: Satz 11.8 für $X = \frac{1}{n}\sum X_i$, $EX = EX_1$, $VarX = \frac{1}{n}VarX_1$ nach Satz 11.10. (Unabhängigkeit impliziert Unkorreliertheit nach Übung 39.b) □

Gibt es Folgen unabhängiger Zufallsvariablen?

Satz 12.12 *Seien* $(\Omega_n, \mathcal{F}_n, P_n)$, $n = 1, 2, \ldots$, *Wahrscheinlichkeitsräume. Dann existiert ein Wahrscheinlichkeitsraum* $(\Omega, \mathcal{F}, P)$ *und unabhängige Zufallsvariablen* $X_n : (\Omega, \mathcal{F}) \mapsto (\Omega_n, \mathcal{F}_n)$ *mit* $P^{X_n} = P_n$.

Beweis: Setze $\Omega = \times\Omega_n$, $\mathcal{F} = \otimes\mathcal{F}_n$, $P = \otimes P_n$ wie in Korollar 9.4. Setze $X_n(\omega) = \omega_n$ (Projektion). Dann X_n messbar: $\{\omega : X_n(\omega) \in B\} = \{\omega_n \in B\}$. Außerdem $P^{X_n}B = P_nB$ und $P(X_1 \in B_1, \ldots, X_n \in B_n) = \prod P_iB_i = \prod P^{X_i}B_i$, also sind $X_1, \ldots, X_n$ unabhängig. □

13 Konvergenz von Zufallsvariablen

Gegeben seien ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ und Zufallsvariablen $X, X_1, X_2, \ldots$.

Definition 13.1 *Seien* $X, X_1, X_2, \ldots \in L_p$, $p > 0$.

1. $X_n \to X$ ***in*** $\mathbf{L_p}$ *(oder* ***im p-ten Mittel****), wenn* $||X_n - X||_p \to 0$.
2. $X_n \to X$ (P) *(oder* ***in Wahrscheinlichkeit****), wenn* $\forall\varepsilon > 0 \quad P(|X_n - X| > \varepsilon) \to 0$.
3. $X_n \to X$ ***f.s.****, wenn ein* $N \subset \Omega$ *existiert mit* $PN = 0$*, so dass für* $\omega \notin N : \ X_n(\omega) \to X(\omega)$.

Satz 13.2 *Konvergenz in* L_p *impliziert Konvergenz in Wahrscheinlichkeit.*

Beweis: Satz 11.8 (Tschebyscheff-Ungleichung) mit $X := |X_n - X|$ □

Satz 13.3 $X_n \to X$ *f.s. genau dann, wenn* $\forall\varepsilon > 0$

$$P(|X_k - X| > \varepsilon \textit{ für ein } k \geq n) \to 0$$

Beweis: Setze $B_{k\varepsilon} = \{|X_k - X| > \varepsilon\}$, $B_\varepsilon = \limsup B_{k\varepsilon}$. Dann $P\bigcup_{k\geq n} B_{k\varepsilon} \to PB_\varepsilon$ (Satz 2.15). Es gilt $\{X_k \not\to X\} = \bigcup B_\varepsilon = \bigcup B_{\frac{1}{m}}$, also $X_n \to X$ *f.s.* $\Leftrightarrow PB_\varepsilon = 0,\ \varepsilon > 0 \Leftrightarrow P\bigcup_{k\geq n} B_{k\varepsilon} \to 0,\ \varepsilon > 0$. □

Satz 13.4 *Konvergenz f.s. impliziert Konvergenz in Wahrscheinlichkeit.*

Beweis: Satz 13.3. □

Beispiel 13.5 Konvergenz in Wahrscheinlichkeit impliziert nicht Konvergenz f.s.:
$P = \lambda|\mathcal{B} \cap (0,1)$, $X_{2^n+k-1} = 1_{((k-1)2^{-n}, k2^{-n})}$, $k = 1, \ldots, 2^n$.

Beispiel 13.6 Konvergenz in Wahrscheinlichkeit impliziert nicht Konvergenz in L_p:
$P = \lambda|\mathcal{B} \cap (0,1)$, $X_n = 2^n 1_{(0,\frac{1}{n})}$; $\quad E|X_n|^p = \frac{2^{np}}{n} \to \infty,\ p > 0$

14 Starkes Gesetz der großen Zahl

Nach Satz 12.11 gilt $\frac{1}{n}\sum X_i \to EX_1$ (P), wenn $X_i \in \mathrm{L}_2$ unabhängig und identisch verteilt sind. Gilt dies auch *fast sicher*?

Lemma 14.1 *Sei $(a_{ij})_{i,j=1,\dots}$ unendliche Matrix mit $a_{nj} \to 0$ für $n \to \infty$ und alle j, und $\sum_{j=1}^{\infty}|a_{nj}| \le c$ für alle n. Sei x_n Folge. Dann gilt*

1. $x_n \to 0 \;\Rightarrow\; \sum_j a_{nj}x_j \to 0$
2. $\sum_j a_{nj} \to 1$ *und* $x_n \to x \;\Rightarrow\; \sum_j a_{nj}x_j \to x$

Beweis:
a) Sei $\varepsilon > 0$. Wähle N mit $|x_j| \le \frac{\varepsilon}{c}$ für $j > N$. Dann $\sum_{j=N+1}^{\infty}|a_{nj}||x_j| \le \frac{\varepsilon}{c}c = \varepsilon$. Außerdem $\sum_{j=1}^{N}|a_{nj}||x_j| \to 0$ wegen $a_{nj} \to 0$.
b) $\sum_j a_{nj}(x_j - x) \to 0$ nach a); $\sum_j a_{nj}x \to x$ n.Vor. $\Rightarrow \sum_j a_{nj}x_j \to x$ □

Lemma 14.2 (Toeplitz) *Sei (a_n) nichtnegative Folge und $0 < b_n = \sum_{j=1}^{n} a_j \to \infty$. Gilt $x_n \to x$, dann*

$$\frac{1}{b_n}\sum_{j=1}^{n} a_j x_j \to x$$

Beweis:

Die Matrix $(a_{nj}) = \begin{pmatrix} a_1/b_1 & 0 & \cdots & & \\ \vdots & \ddots & \ddots & & \\ a_1/b_n & \cdots & a_n/b_n & 0 & \cdots \\ \vdots & & \vdots & \ddots & \ddots \end{pmatrix}$ erfüllt $\sum_{j=1}^{\infty} a_{nj} = 1$.

Jetzt Lemma 14.1 anwenden.

Satz 14.3 (Kronecker-Lemma) *Seien (b_n) und (x_n) Folgen, $0 < b_n \uparrow \infty$ und $\sum_{n=1}^{\infty} x_n = x$. Dann*

$$\frac{1}{b_n}\sum_{j=1}^{n} b_j x_j \to 0$$

Beweis: Setze $b_0 = 0$, $s_0 = 0$, $s_n = \sum_{j=1}^{n} x_j$. Dann $\sum_{j=1}^{n} b_j x_j = \sum_{j=1}^{n} b_j(s_j - s_{j-1}) = b_n s_n - b_0 s_0 - \sum_{j=1}^{n}(b_j - b_{j-1})s_{j-1}$, also $\frac{1}{b_n}\sum_{j=1}^{n} b_j x_j = s_n - \frac{1}{b_n}\sum_{j=1}^{n}(b_j - b_{j-1})s_{j-1}$. Nach Voraussetzung gilt $s_n \to x$. Außerdem $\frac{1}{b_n}\sum(b_j - b_{j-1})s_{j-1} \to x$ nach Lemma 14.2 mit $a_j = b_j - b_{j-1}$.
□

Im Weiteren sei $S_n = X_1 + \ldots + X_n$.

Satz 14.4 (Kolmogorov-Ungleichung) *Seien $X_1, \dots, X_n \in L_2$ unabhängig. Dann gilt für $\varepsilon > 0$:*

$$P(\max_{1\le j\le n} |S_j - ES_j| \ge \varepsilon) \le \frac{VarS}{\varepsilon^2}$$

Beweis: O.E. sei $ES_j = 0$. Setze $A_k = \{|S_j| < \varepsilon, j = 1, \dots, k-1; |S_k| \ge \varepsilon\}$. Setze $A = \{\max\limits_{1\le j\le n} |S_j| \ge \varepsilon\}$. Dann gilt $A = \sum_{k=1}^n A_k$ und $VarS_n \ge E1_A S_n^2 = \sum E1_{A_k} S_n^2$. Mit $Y_k = S_n - S_k = \sum_{j=k+1}^n X_j$ gilt $E1_{A_k}S_n^2 = E1_{A_k}S_k^2 + E1_{A_k}Y_k^2 + 2E1_{A_k}S_kY_k$ und $E1_{A_k}S_kY_k = 0$ nach Satz 12.9, also $E1_{A_k}S_n^2 \ge E1_{A_k}S_k^2 \ge \varepsilon^2 PA_k$, also $VarS_n \ge \varepsilon^2 \sum PA_k = \varepsilon^2 PA$. □

Satz 14.5 (Erstes Borel-Cantelli-Lemma) *Sind $A_1, A_2, \dots \in \mathcal{F}$ und $\sum PA_n < \infty$. dann gilt*

$$P \limsup A_n = 0$$

Beweis: $P\limsup A_n \le P\bigcup_{k\ge n} A_k \le \sum_{k=n}^\infty PA_k \to 0$ für $n \to \infty$. □

Satz 14.6 (Zweites Borel-Cantelli-Lemma) *Sind $A_1, A_2, \dots \in \mathcal{F}$ unabhängig und $\sum PA_n = \infty$, dann gilt*

$$P \limsup A_n = 1$$

Beweis: $P\limsup A_n = P\bigcap_n \bigcup_{k\ge n} A_k = \lim_n P\bigcup_{k\ge n} A = \lim_n \lim_m P\bigcup_{k=n}^m A_k$ und

$P(\bigcup_{k=n}^m A_k)^c = P\bigcap_{k=n}^m A_k^c = \prod_{k=n}^m PA_k^c \overset{1-x\le e^{-x}}{\le} \prod_{k=n}^m e^{-PA_k} = e^{-\sum_{k=n}^m PA_k} \to 0$ für $m \to \infty$
□

Satz 14.7 *Sind $X_1, X_2, \dots$ unabhängig mit $\sum VarX_n < \infty$, dann konvergiert $\sum\limits_{i=1}^n (X_i - EX_i)$ f.s. gegen eine Zufallsvariable X.*

Beweis: O.E. sei $EX_i = 0$. Nach Übung 45 gilt: $S_n = \sum_{i=1}^n X_i$ ist Cauchy f.s. genau dann, wenn $P\bigcup_{j,k\ge m}(|S_j - S_k| > \varepsilon) \to 0$ für $m \to \infty$.
Es gilt $P\bigcup_{k=1}^\infty(|S_{m+k} - S_m| > \varepsilon) = \lim_n P\bigcup_{k=1}^n(|S_{m+k} - S_m| > \varepsilon) = \lim_n P(\max_{1\le k\le n} |S_{m+k} - S_m| > \varepsilon) \overset{14.4}{\le} \lim_n \frac{1}{\varepsilon^2} Var(S_{m+n} - S_m) \overset{11.8}{=} \lim_n \frac{1}{\varepsilon^2}\sum_{j=m+1}^{m+n} VarX_j = \frac{1}{\varepsilon^2}\sum_{j=m+1}^\infty VarX_j \to 0$ für $m \to \infty$.
Das reicht wegen $|S_j - S_k| \le |S_j - S_m| + |S_k - S_m|$. □

Bemerkung 14.8 Betrachte einen Münzwurf: $X_1, X_2, \dots$ sind unabhängig, $P(X_n = 1) = P(X_n = -1) = \frac{1}{2}$. S_n sei das Vermögen zum Zeitpunkt n bei Einsatz 1.
S_n konvergiert *nicht*: $\limsup S_n = \infty$, $\liminf S_n = -\infty$
Reduziert man den Einsatz ($P(X_n = a_n) = P(X_n = -a_n) = \frac{1}{2}$), dann ist $EX_n = 0$, $VarX_n = a_n^2$. Ist nun $\sum a_n^2 < \infty$, dann konvergiert S_n f.s.

Satz 14.9 (Starkes Gesetz der großen Zahl von Kolmogorov) *Seien $X_1, X_2, \ldots \in L_2$ unabhängig und (b_n) eine Folge mit $0 < b_n \uparrow \infty$.*
Wenn $\sum \frac{VarX_n}{b_n^2} < \infty$, dann

$$\frac{S_n - ES_n}{b_n} \to 0 \ f.s.$$

Beweis: Satz 14.7 auf $\frac{X_n}{b_n}$ anwenden: Dann konvergiert $\sum_{i=1}^n \frac{X_i - EX_i}{b_i}$ f.s. Nun Satz 14.3 auf $x_j = \frac{X_j - EX_j}{b_j}$ anwenden:
$\frac{S_n - ES_n}{b_n} = \frac{1}{b_n} \sum_{j=1}^n b_j \frac{X_j - EX_j}{b_j} \to 0$ f.s. □

15 Schwache Konvergenz

Wenn $X_1, X_2, \ldots$ unabhängig und identisch verteilt sind mit $EX_1 = \mu$, $VarX_1 = \sigma^2$, dann gilt $\frac{1}{n} \sum_{i=1}^n (X_i - \mu) \to 0$ nach dem Gesetz der großen Zahl. „Bläst" man dies mit $n^{-1/2}$ auf, so erhält man $En^{-1/2} \sum_{i=1}^n (X_i - \mu) = 0$ und $Var(n^{-1/2} \sum_{i=1}^n (X_i - \mu)) = \sigma^2$.
In Kapitel 18 zeigen wir den Zentralen Grenzwertsatz

$$P(a < n^{-1/2} \sum_{i=1}^n (X_i - \mu) < b) \to N_{0,\sigma^2}(a, b) \text{ für } n \to \infty$$

(wobei N_{μ,σ^2} die Dichte $\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ hat).

Definition 15.1 *Seien $P, P_1, P_2, \ldots$ Wahrscheinlichkeitsmaße auf $\mathcal{B}$.*
*Dann konvergiert P_n **schwach** gegen P ($P_n \Rightarrow P$), wenn für alle stetigen und beschränkten f $P_n f \to Pf$ gilt.*
*Sind $X, X_1, X_2, \ldots$ Zufallsvariablen mit $P^{X_n} \Rightarrow P^X$, so sagen wir: X_n konvergiert gegen X **in Verteilung** und schreiben $X_n \Rightarrow X$.*

Manchmal gilt Konvergenz sogar mengenweise:

Satz 15.2 (Scheffé) *Seien $P, P_1, P_2, \ldots$ Wahrscheinlichkeitsmaße mit Lebesgue-Dichten $p, p_1, p_2, \ldots$ Gilt $p_n \to p$ f.s., dann gilt*

$$\sup_{A \in \mathcal{B}} |P_n A - PA| = \frac{1}{2} \lambda |p_n - p| \to 0$$

Beweis: Es gilt $\lambda(p_n - p) = 0$.
1. Für $A \in \mathcal{B}$: $2|P_n A - PA| = |\lambda 1_A (p_n - p)| + |\lambda 1_{A^c}(p_n - p)| \leq \lambda |p_n - p|$. Für $A = \{p_n - p \geq 0\}$ gilt die Gleichheit.
2. Es gilt $0 \leq (p_n - p)^+ \leq p$, $(p_n - p)^+ \to 0$ f.s. Also $\lambda(p_n - p)^+ \to 0$ und somit $\lambda |p_n - p| = 2\lambda(p_n - p)^+ \to 0$ nach Satz 6.7 (dominante Konvergenz). □

Definition 15.3 *Eine Funktion $f : \mathbb{R} \mapsto \mathbb{R}$ heißt* ***unterhalbstetig (oberhalbstetig)****, wenn $\{f > a\}$ ($\{f < a\}$) für alle $a \in \mathbb{R}$ offen ist.*

Satz 15.4 *Eine Funktion f ist unterhalbstetig genau dann, wenn*

$$\lim \inf_{x_n \to x} f(x_n) \geq f(x)$$

Beweis: Es gilt $\{f > a\}$ offen genau dann, wenn für x mit $f(x) > a$ gilt: $x_n \to x \;\Rightarrow\; f(x_n) > a$ schließlich. □

Satz 15.5 *Ist f unterhalbstetig, so gibt es stetige f_n mit $f_n \uparrow f$.*

Beweis:

1. Gelte $f(\mathbb{R}) \subset (0, \pi)$. Für $t > 0$ definiere $f_t(x) = \inf\{f(z) + t|x - z| : z \in \mathbb{R}\}$.
   a) f_t ist stetig:
   Für $x, y \in \mathbb{R}$ gilt $f(z) + t|x-z| \leq f(z) + t|y-z| + t|x-y|$. Also $f_t(x) \leq f_t(y) + t|x-y|$. Mit symmetrischem Argument folgt $|f_t(x) - f_t(y)| \leq t|x-y|$.
   b) $f_n \uparrow f$: Es gilt $f_n \leq f$, $f_n \uparrow$.
   Sei $\varepsilon > 0$. Für jedes n ex. ein z_n mit $f(x) + \varepsilon \geq f_n(x) + \varepsilon > f(z_n) + n|x - z_n| > n|x - z_n|$. Also $|x - z_n| \to 0$, somit $\liminf f(z_n) \geq f(x)$ nach Satz 15.4, d.h. $f(z_n) \geq f(x) - \varepsilon$ schließlich; also $f_n(x) > f(z_n) - \varepsilon \geq f(x) - 2\varepsilon$ schließlich.
2. f beliebig.
   Betrachte $g \circ f$ mit $g(x) = \frac{1}{2}\pi + \arctan x$ (g ist homöomorph und ordnungserhaltend). □

Definition 15.6 *Für $A \subset \mathbb{R}$ heißt $\partial A = \bar{A} - \mathring{A}$* ***Rand*** *von A.*

Satz 15.7 (Portmanteau-Satz) *Seien $P, P_1, P_2, \ldots$ Wahrscheinlichkeitsmaße auf $\mathcal{B}$. Dann sind äquivalent:*

a) $P_n \Rightarrow P$

b) $\limsup P_n f \leq Pf$ *für alle oberhalbstetigen beschränkten f*

c) $\limsup P_n A \leq PA$ *für alle abgeschlossenen A*

d) $P_n A \to PA$ *für alle $A \in \mathcal{B}$ mit $P\partial A = 0$*

Beweis:

a) $\Rightarrow$ b) Sei f oberhalbstetig. Sei $g \leq f$ und g stetig. Dann $\liminf P_n f \geq \liminf P_n g = Pg$. Nach Satz 15.5 ist jede unterhalbstetige Funktion limes stetiger Funktionen. Satz 6.7 (dom. Konv.) anwenden für $-f$ statt f.

b) $\Rightarrow$ c) Ist A abgeschlossen, so ist 1_A oberhalbstetig.

c) $\Rightarrow$ d) $\limsup P_n A \leq \limsup P_n \bar{A} \overset{c)}{\leq} P\bar{A} = PA = P\mathring{A} \overset{c)}{\leq} \liminf P_n \mathring{A} \leq \liminf P_n A$

d) $\Rightarrow$ a) Sei f stetig und beschränkt.
$B = \{c : P\{f = c\} > 0\}$ ist abzählbar. Wähle M mit $|f| < M$ und Zerlegung $-M = t_0 < t_1 < \ldots < t_j = M$ und $t_i \notin B$.
Für $A_i = \{t_i < f \leq t_{i+1}\}$ gilt $P\partial A_i = 0$, also nach d) $\sum_{i=0}^{j-1} t_i P_n A_i \to \sum_{i=0}^{j-1} t_i P A_i$.
Schreibe $|P_n f - Pf| \leq |P_n f - \sum t_i P_n A_i| + |\sum t_i P_n A_i - \sum t_i P A_i| + |\sum t_i P A_i - Pf|$.
1. Summand: $|\sum P_n 1_{A_i}(f - t_i)| \leq \max(t_{i+1} - t_i)$
2. Summand: konvergiert gegen 0.
3. Summand: wie der erste.
Jetzt die Zerlegung feiner machen. □

Im Folgenden gelte für jede Verteilungsfunktion F $F(-\infty) = 0$.

Satz 15.8 *Seien $P, P_1, P_2, \ldots$ Wahrscheinlichkeitsmaße auf $\mathcal{B}$ mit Verteilungsfunktionen $F, F_1, F_2, \ldots$. Dann sind äquivalent:*

1. $P_n \Rightarrow P$
2. $F_n(t) \to F(t)$, *wenn F stetig in t*

Beweis:
$\Rightarrow$ Sei F stetig in t.
Dann gilt $P\partial(-\infty, t] = P\{t\} = F(t) - F(t-) = 0$. Mit Satz 15.7 d) folgt $F_n(t) \to F(t)$.
$\Leftarrow$ Sei A offen, also abzählbare Vereinigung von offenen disjunkten Intervallen I_k. Nach Lemma 6.6 (Fatou) gilt: $\liminf_n P_n A = \liminf_n \sum_k P_n I_k \geq \sum_k \liminf_n P_n I_k$.
Zu $\varepsilon > 0$ wähle $J_k = (a_k, b_k] \subset I_k$, so dass a_k, b_k Stetigkeitspunkte von F sind und $PJ_k \geq PI_k - \varepsilon 2^{-k}$. Dann gilt nach Vor.: $\liminf P_n I_k \geq \liminf P_n J_k = PJ_k$, also $\liminf P_n A \geq \sum PJ_k \geq \sum PI_k - \varepsilon = PA - \varepsilon$. Jetzt Satz 15.7 c) anwenden. □

16 Schwache Kompaktheit

Sind F_n Verteilungsfunktionen von Wahrscheinlichkeitsmaßen, und gilt $F_n(t) \to F(t)$ für alle Stetigkeitspunkte von F, so braucht F nicht Verteilungsfunktion eines Wahrscheinlichkeitsmaßes zu sein; als Beispiel diene $F_n = 1_{[n,\infty)} \to 0$.

Satz 16.1 (Helly) *Seien $F, F_1, F_2, \ldots$ Verteilungsfunktionen mit $F_n(-\infty) = 0$, $F_n(\infty) \leq 1$. Dann existiert eine Verteilungsfunktion F und eine Teilfolge (F_{n_k}) mit $F_{n_k}(t) \to F(t)$, wenn F stetig ist.*

Wie können wir „Massenverlust“ vermeiden?

Definition 16.2 *Eine Familie $\mathcal{P}$ von Wahrscheinlichkeitsmaßen auf $\mathcal{B}$ heißt* ***straff****, wenn für jedes $\varepsilon > 0$ ein endliches Intervall I existiert mit $PI^c < \varepsilon$ für alle $P \in \mathcal{P}$.*

Definition 16.3 *Eine Familie $\mathcal{P}$ von Wahrscheinlichkeitsmaßen auf $\mathcal{B}$ heißt* ***relativ kompakt****, wenn jede Folge eine schwach konvergente Teilfolge besitzt.*

Satz 16.4 (Prohorov) *Eine Familie von Wahrscheinlichkeitsmaßen ist straff genau dann, wenn sie relativ kompakt ist.*

17 Charakteristische Funktionen

Definition 17.1 *Für ein Wahrscheinlichkeitsmaß $P|\mathcal{B}$ ist die* ***charakteristische Funktion***

$$\varphi(t) = \int e^{itx} P(dx)$$

Ist X Zufallsvariable, so bezeichnet $\varphi^X = Ee^{itX}$ die charakteristische Funktion der induzierten Verteilung.

Bemerkung 17.2 Es gilt $e^{ix} = \cos x + i \sin x$.

Lemma 17.3 *Sind $X_1, \ldots, X_n$ unabhängige Zufallsvariablen, so gilt*

$$\varphi^{\Sigma X_i} = \prod \varphi^{X_i}$$

Beweis: $Ee^{it\Sigma X_j} = E \prod e^{itX_j} = \prod Ee^{itX_j}$ □

Bemerkung 17.4 Es gilt $\varphi^{aX+b}(t) = e^{itb}\varphi^X(at)$.

Beispiel 17.5 Die charakteristische Funktion von N_{μ,σ^2} ist $\varphi(t) = e^{it\mu}e^{-\frac{1}{2}t^2\sigma^2}$.

Beweis: Nach Bem. 17.4 ist z.z.: $N_{0,1}$ hat die char. Funktion $e^{-\frac{1}{2}t^2}$.
$\varphi(t) = \frac{1}{\sqrt{2\pi}} \int e^{itx}e^{-\frac{1}{2}x^2}dx = \frac{1}{\sqrt{2\pi}} \int cos(tx)e^{-\frac{1}{2}x^2}dx$
Jetzt differenzieren und partiell integrieren:
$\varphi'(t) = -\frac{1}{\sqrt{2\pi}} \int x \sin(tx)e^{-\frac{1}{2}x^2}dx = -\frac{1}{\sqrt{2\pi}} \int t\cos(tx)e^{-\frac{1}{2}x^2}dx = -t\varphi(t)$ □

Satz 17.6 (Inversionsformel) *Sei $P|\mathcal{B}$ ein Wahrscheinlichkeitsmaß mit charakteristischer Funktion φ. Dann gilt für $a < b$:*

$$\lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t)dt = P(a,b) + \frac{1}{2}P\{a,b\}$$

Satz 17.7 (Lévy, Stetigkeitssatz) *Seien $P, P_1, P_2 \ldots$ Wahrscheinlichkeitsmaße mit charakteristischen Funktionen $\varphi, \varphi_1, \varphi_2, \ldots$*

a) Wenn $P_n \Rightarrow P$, dann $\varphi_n \to \varphi$

b) Wenn $\varphi_n \to \varphi$ und φ stetig in 0, dann $P_n \Rightarrow P$

Beispiel 17.8 1. Exponentialverteilung E_a (Dichte: $\frac{1}{a}e^{-\frac{x}{a}},\ x > 0$): $\varphi(t) = (1 - ait)^{-1}$

2. Cauchy-Verteilung C_a (Dichte: $\frac{a}{\pi(a^2+x^2)}$): $\varphi(t) = e^{-a|t|}$

3. Gamma-Verteilung $\Gamma_{a,b}$ (Dichte: $\frac{1}{a^b\Gamma(b)}x^{b-1}e^{-\frac{x}{a}},\ x \geq 0$): $\varphi(t) = (1 - ait)^{-b}$

18 Zentraler Grenzwertsatz

Lemma 18.1

$$\left| e^{ix} - \sum_{m=0}^{n} \frac{(ix)^m}{m!} \right| \leq \min\left\{ \frac{|x|^{n+1}}{(n+1)!}, \frac{2|x|^n}{n!} \right\}$$

Korollar 18.2

$$\left| Ee^{itX} - \sum_{m=0}^{n} \frac{1}{m!}E(itX)^m \right| \leq \frac{|t|^n}{(n+1)!} E\min\{|t||X|^{n+1}, 2(n+1)|X|^n\}$$

Satz 18.3 *Sei $EX = 0,\ EX^2 = \sigma^2 < \infty$. Dann gilt für $t \to 0$*

$$Ee^{itX} = 1 - \frac{1}{2}t^2\sigma^2 + o(t^2)$$

Beweis: Es gilt $\min\{|t||X|^3, 6|X|^2\} \begin{cases} \leq 6X^2 \\ \to 0,\ t \to 0 \end{cases}$.
Also mit Satz 6.7 und 18.2 $|Ee^{itx} - 1 + \frac{1}{2}t^2\sigma^2| = o(t^2)$. □

Satz 18.4 (Zentraler Grenzwertsatz) *Sind $X_1, X_2, \ldots$ unabhängig und identisch verteilte Zufallsvariablen mit Erwartungswert μ und endlicher Varianz σ^2, dann gilt*

$$n^{-\frac{1}{2}} \sum_{i=1}^{n} (X_i - \mu) \Rightarrow N_{0,\sigma^2}$$

Beweis: O.E. $\mu = 0$. Nach Satz 18.3 gilt: $Ee^{itX_1} = 1 - \frac{1}{2}t^2\sigma^2 + o(t^2)$, also $Ee^{itn^{-1/2}\Sigma X_j} = (1 - \frac{t^2\sigma^2}{2n} + o(\frac{1}{n}))^n \to e^{-\frac{1}{2}t^2\sigma^2}$. Mit Satz 17.7(b) folgt die Behauptung. □

19 Bedingte Wahrscheinlichkeiten und bedingte Erwartungswerte

Im Folgenden sei immer der Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ gegeben.
Eine $\mathcal{F}$-messbare Funktion ist eine Funktion $X : (\Omega, \mathcal{F}) \mapsto (\mathbb{R}, \mathcal{B})$.

Definition 19.1 *Seien $A, B \in \mathcal{F}$ mit $PB > 0$. Die **bedingte Wahrscheinlichkeit** von A gegeben B ist*

$$P(A|B) = \frac{P(A \cap B)}{PB}$$

Bemerkung 19.2 $P(\cdot|B)$ ist ein Wahrscheinlichkeitsmaß auf $\mathcal{F}$.

Definition 19.3 *Sei $B \in \mathcal{F}$ mit $PB > 0$. Ist X $\mathcal{F}$-messbar und integrierbar, so ist der* ***bedingte Erwartungswert*** *von X gegeben B definiert als*

$$E(X|B) = \frac{1}{PB} E1_B X = \int X(\omega) P(d\omega|B)$$

Bemerkung 19.4 Sei $\Omega = \sum B_n,\ PB_n > 0$. Die σ-Algebra $\mathcal{G} = \sigma\{B_1, B_2, \ldots\}$ besteht aus Mengen der Form $C = \sum_{n \in I} B_n,\ I \subset \mathbb{N}$. Die Mengen B_n heißen *Atome* von $\mathcal{G}$.

a) Sei $A \in \mathcal{F}$. Die bedingte Wahrscheinlichkeit von A gegeben $\mathcal{G}$ ist $P(A|\mathcal{G})(\omega) = P(A|B_n)$, wenn $\omega \in B_n$. Dann ist $P(A|\mathcal{G})$ auf B_n konstant, also $\mathcal{G}$-messbar.
Für $C = \sum_{n \in I} B_n$ gilt:
$E(1_C 1_A) = P(C \cap A) = \sum_{n \in I} P(B_n \cap A) = \sum P(B_n) P(A|B_n) = E1_C P(A|\mathcal{G})$

b) Sei X $\mathcal{F}$-messbar und integrierbar. Der bedingte Erwartungswert von X gegeben $\mathcal{G}$ sei $E(X|\mathcal{G})(\omega) = E(X|B_n)$, wenn $\omega \in B_n$.
Für $C = \sum_{n \in I} B_n \in \mathcal{G}$ gilt
$E1_C X = \sum_{n \in I} E1_{B_n} X = \sum_{n \in I} PB_n E(X|B_n) = E(1_C E(X|\mathcal{G}))$
Eine $\mathcal{G}$-messbare Funktion ist durch diese Beziehung f.s. eindeutig bestimmt.
Diese Charakterisierung lässt sich auf den „atomlosen“ Fall verallgemeinern:

Satz 19.5 *Sei $\mathcal{G}$ eine Sub-σ-Algebra von $\mathcal{F}$. Sei X $\mathcal{F}$-messbare Zufallsvariable und integrierbar. Dann existiert eine f.s. eindeutige $\mathcal{G}$-messbare Funktion $E(X|\mathcal{G})$, der* ***bedingte Erwartungswert*** *von X gegeben $\mathcal{G}$ mit*

$$E1_C X = E1_C E(X|\mathcal{G}),\ C \in \mathcal{G}$$

Beweis: $\nu C = E1_C X$ ist ein endliches und signiertes Maß auf $\mathcal{G}$ nach Satz 6.2, und es gilt $\nu \ll P$ auf C. Nach Satz 7.5 existiert eine P-Dichte g von ν mit $\nu C = P1_C g$. Setze $E(X|\mathcal{G}) = g$. □

Beispiel 19.6 a) $E(X|\{\emptyset, \Omega\}) = EX$
b) $E(X|\sigma\{A\}) = E(X|A)1_A + E(X|A^c)1_{A^c}$ f.s.
c) $E(X|\mathcal{F}) = X$ f.s.

Beispiel 19.7 $\Omega = \Omega_1 \times \Omega_2,\ \mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2,\ P(d\omega) = P_1(d\omega_1) P_2(\omega_1, d\omega_2)$
Sei $\mathcal{G} = \{B_1 \times \Omega_2 : B_1 \in \mathcal{F}_1\}$; es gilt $P_1 B_1 = P(B_1 \times \Omega_2)$.
Sei X $\mathcal{F}$-messbar und integrierbar. Setze $E(X|\mathcal{G})(\omega) = \int X(\omega_1, \omega_2) P_2(\omega_1, d\omega_2)$.
Dann ist $E(X|\mathcal{G})$ $\mathcal{G}$-messbar, und nach Satz 8.4 gilt für $C = B_1 \times \Omega_2 \in \mathcal{G}$:
$E1_C X = \int 1_B(\omega_1) (\int X(\omega_1, \omega_2) P_2(\omega_1, d\omega_2)) P_1(d\omega_1)$
$= \int 1_B(\omega_1) E(X|\mathcal{G}) P_1(d\omega_1) \int P_2(\omega_1, d\omega_2) = E(1_C E(X|\mathcal{G}))$

Definition 19.8 *Ist $Y : (\Omega, \mathcal{F}) \mapsto (\Omega', \mathcal{F}')$ messbar, so heißt $\mathcal{F}(Y) = Y^{-1}\mathcal{F}'$ die von Y erzeugte σ-Algebra.*

Lemma 19.9 *Ist $Z : \Omega \mapsto \mathbb{R}$ $\mathcal{F}(Y)$-messbar, dann gilt $Z = f \circ Y$ für eine Borel-messbare Funktion $f : \Omega' \mapsto \mathbb{R}$.*

Beweis: 1. $Z = 1_A,\ A = Y^{-1}B,\ B \in \mathcal{F}'$: Mit $f = 1_B$ gilt $f \circ Y = 1_B \circ Y = 1_A$.
2. $Z = \sum a_i 1_{A_i}$: Setze $f = \sum a_i 1_{B_i},\ A_i = Y^{-1}B_i$.
3. Z beliebig: Wähle $Z_n$wie in 2. mit $Z_n \to Z$ und f als Limes der zugehörigen f_n. □

Satz 19.10 *Der bedingte Erwartungswert $E(X|\mathcal{F}(Y))$ lässt sich schreiben als $E(X|Y=y) \circ Y$. Er ist charakterisiert durch*

$$E1_{(Y\in B)}X = \int_B E(X|Y=y)P^Y(dy)$$

Beweis: a) Lemma 19.9 für $Z = E(X|\mathcal{F}(Y))$
b) $E1_{(Y\in B)}X = E1_{Y^{-1}B}X = E1_{Y^{-1}B}E(X|\mathcal{F}(Y)) = E1_{B\circ Y}E(X|Y) \overset{10.3}{=} \int_B E(X|Y=y)P^Y(dy)$
□

Bemerkung 19.11 Sei Y unabhängig von X. Dann gilt

$$E(X|Y) = EX \ f.s.$$

Beweis: $E1_{(Y\in B)}X = P(Y \in B)EX = P^Y BEX = \int_B EXP^Y(dy)$ □

20 Eigenschaften des bedingten Erwartungswerts

Im Folgenden seien $X, X_1, X_2, \ldots$ integrierbare und $\mathcal{F}$-messbare Zufallsvariablen, $\mathcal{G} \subset \mathcal{F}$ eine Sub-σ-Algebra. Die folgenden Aussagen gelten f.s.

Satz 20.1 *a)* $E(aX_1 + bX_2|\mathcal{G}) = aE(X_1|\mathcal{G}) + bE(X_2|\mathcal{G})$

b) $X_1 \leq X_2 \ \Rightarrow\ E(X_1|\mathcal{G}) \leq E(X_2|\mathcal{G})$

c) $|E(X|\mathcal{G})| \leq E(|X|\ |\mathcal{G})$

Beweis: a) Für $C \in \mathcal{G}$ gilt: $E1_C(aX_1 + bX_2) = aE1_CX_1 + bE1_CX_2 = aE1_CE(X_1|\mathcal{G}) + bE1_CE(X_2|\mathcal{G}) = E1_C(aE(X_1|\mathcal{G}) + bE(X_2|\mathcal{G}))$
b) Für $C \in \mathcal{G}$ gilt: $E1_CX_1 \leq E1_CX_2 \ \Rightarrow\ E1_CE(X_1|\mathcal{G}) \leq E1_CE(X_2|\mathcal{G})$
c) Wegen $-|X| \leq X \leq |X|$ aus b). □

Satz 20.2 (Monotone Konvergenz) *Wenn $0 \leq X_n \uparrow X$, dann $E(X_n|\mathcal{G}) \uparrow E(X|\mathcal{G})$.*

Beweis: $E(X_n|\mathcal{G}) \uparrow$ nach Satz 20.1 b). Der Limes Z ist $\mathcal{G}$-messbar.
Für $C \in \mathcal{G}$ gilt nach Satz 6.3 $E1_CE(X_n|\mathcal{G}) \uparrow E1_CZ$.
Andererseits $E1_CE(X_n|\mathcal{G}) = E1_CX_n \uparrow E1_CX$.
Also $E1_CZ = E1_CX$, also $Z = E(X|\mathcal{G})$. □

Satz 20.3 (Dominierte Konvergenz) *Wenn $X_n \to X$ und $|X| \leq Z$, Z integrierbar, dann gilt*

$$E(X_n|\mathcal{G}) \to E(X|\mathcal{G})$$

Beweis: Setze $Y_n = \sup_{k \geq n} |X_k - X|$. Es gilt nach Satz 20.1 c)
$|E(X_n|\mathcal{G}) - E(X|\mathcal{G})| \leq E(Y_n|\mathcal{G})$
z.z.: $E(Y_n|\mathcal{G}) \downarrow 0$
Es gilt $Y_n \downarrow 0$. Also $E(Y_n|\mathcal{G}) \downarrow$ nach Satz 20.1 b). Der Limes W ist nichtnegativ und $\mathcal{G}$-messbar.
Wegen $Y_n \leq 2Z$ folgt mit Satz 6.7:
$0 \leq EW \leq EE(Y_n|\mathcal{G}) = EY_n \to 0$, also $W = 0$ f.s. □

Satz 20.4 *Ist Y $\mathcal{G}$-messbar und XY integrierbar, dann gilt*

$$E(XY|\mathcal{G}) = Y E(X|\mathcal{G})$$

Beweis:
1. $Y = 1_B,\ B \in \mathcal{G}:\ E1_C XY = E1_{C\cap B}X = E1_{C\cap B}E(X|\mathcal{G}) = E1_C Y E(X|\mathcal{G}),\ C \in \mathcal{G}$
2. $Y = \sum b_i 1_{B_i},\ B_i \in \mathcal{G}$: Satz 20.1 a) und 1.
3. $Y \geq 0$: Wähle $Y_n \uparrow Y$, Y_n wie in 2. Dann $E(XY_n|\mathcal{G}) \to E(XY|\mathcal{G})$ nach Satz 20.3. Andererseits $E(XY_n|\mathcal{G}) = Y_n E(X|\mathcal{G}) \to Y E(X|\mathcal{G})$ □

Lemma 20.5 *Sei $I \subset \mathbb{R}$ offenes Intervall und $f : I \mapsto \mathbb{R}$ konvex. Dann existieren Folgen $(a_n), (b_n)$ mit*

$$f(x) = \sup_n (a_n x + b_n)$$

Satz 20.6 (Jensen-Ungleichung) *Sei $I \subset \mathbb{R}$ offenes Intervall und $f : I \mapsto \mathbb{R}$ konvex. Sei $X : \Omega \mapsto I$ $\mathcal{F}$-messbar und integrierbar. Dann gilt*

$$E(f \circ X|\mathcal{G}) \geq f \circ E(X|\mathcal{G})$$

Beweis:
1. $E(X|\mathcal{G}) \in I$: Ist z.B. $X > a$, dann ist $0 \geq E1_{(E(X|\mathcal{G})\leq a)}E(X-a|\mathcal{G}) = E1_{(E(X|\mathcal{G})\leq a)}(X-a) \geq 0$
2. $f(x) = \sup_n(a_n x + b_n) = \sup f_n(x)$ wie in Lemma 20.5, also $E(f \circ X|\mathcal{G}) \geq E(a_n x + b_n|\mathcal{G}) = a_n E(X|\mathcal{G}) + b_n = f_n \circ E(X|\mathcal{G})$. Jetzt sup über n. □

Satz 20.7 *Sind $\mathcal{G}_1 \subset \mathcal{G}_2$ Sub-σ-Algebren von $\mathcal{F}$, dann gilt*

$$E(X|\mathcal{G}_1) = E(E(X|\mathcal{G}_2)|\mathcal{G}_1)$$

Beweis: Für $C \in \mathcal{G}_1$ gilt $C \in \mathcal{G}_2$, also $E1_C E(X|\mathcal{G}_1) = E1_C X = E1_C E(X|\mathcal{G}_2)$ □

21 Martingale; Optional Skipping

Gegeben sei im Folgenden immer $(\Omega, \mathcal{F}, P)$.

Definition 21.1 *Seien $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots$ Sub-σ-Algebren von $\mathcal{F}$. Seien $X_1, X_2, \dots$ integrierbare Zufallsvariablen, X_n $\mathcal{F}_n$-messbar. Dann heißt die Folge $(X_n, \mathcal{F}_n)$ ein* ***Martingal****, wenn*

$$E(X_{n+1}|\mathcal{F}_n) = X_n \quad (f.s.)$$

Gilt '$\geq$' bzw. '$\leq$', so heißt die Folge ***Sub-*** *bzw.* ***Supermartingal****.*

Bemerkung 21.2 Ist $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$, dann ist nach Satz 19.11 die Definition des Martingals äquivalent zu $E(X_{n+1}|X_1 = x_1, \dots, X_n = x_n) = x_n$ (für $P^{X_1 \dots X_n}$-f.a. $x_1, \dots, x_n$).

Beispiel 21.3 Irrfahrt:
Seien $Y_1, Y_2, \dots$ unabhängig mit Erwartungswert 0. Setze $X_n = \sum_{i=1}^n Y_i$, $\mathcal{F}_n = \sigma(Y_1, \dots, Y_n)$.
Dann ist $(X_n, \mathcal{F}_n)$ ein Martingal, denn es gilt:
$E(X_{n+1}|\mathcal{F}_n) = E(X_n + Y_{n+1}|Y_1, \dots, Y_n) = X_n + E(Y_{n+1}|Y_1, \dots, Y_n) = X_n + EY_{n+1} = X_n$.

Beispiel 21.4 Seien $Y_1, Y_2, \dots$ unabhängig, $EY_i = a_i \neq 0$. Setze $X_n = \prod_{i=1}^n \frac{Y_i}{a_i}$, $\mathcal{F}_n = \sigma(Y_1, \dots, Y_n)$.
Dann ist $(X_n, \mathcal{F}_n)$ ein Martingal, denn es gilt:
$E(X_{n+1}|\mathcal{F}) = E(X_n \frac{Y_{n+1}}{a_{n+1}}|Y_1, \dots, Y_n) = X_n E(\frac{Y_{n+1}}{a_{n+1}}|Y_1, \dots, Y_n) = X_n \frac{EY_{n+1}}{a_{n+1}} = X_n$.

Beispiel 21.5 Sei Y $\mathcal{F}$-messbar und integrierbar. Seien $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots$ Sub-σ-Algebren von $\mathcal{F}$.
Setze $X_n = E(Y|\mathcal{F}_n)$. Dann ist $(X_n|\mathcal{F}_n)$ ein Martingal, denn es gilt:
$E(X_{n+1}|\mathcal{F}) = E(E(Y|\mathcal{F}_{n+1})|\mathcal{F}_n) \overset{20.7}{=} E(Y|\mathcal{F}_n) = X_n$.

Bemerkung 21.6 $(X_n, \mathcal{F}_n)$ ist ein Submartingal genau dann, wenn $E1_A X_{n+1} \geq E1_A X_n$, $A \in \mathcal{F}_n$.

Beweis: $E(X_{n+1}|\mathcal{F}_n) \geq X_n$ genau dann, wenn $E1_A E(X_{n+1}|\mathcal{F}_n) \geq E1_A X_n$, $A \in \mathcal{F}_n$. Die linke Seite ist aber gerade $E1_A X_{n+1}$. □

Bemerkung 21.7 Ist $(X_n, \mathcal{F}_n)$ ein Submartingal, dann gilt $E(X_{n+k}|\mathcal{F}_n) \geq X_n$, $k \in \mathbb{N}$.

Beweis: Für $k = 2$ gilt: $E(X_{n+2}|\mathcal{F}_n) = E(E(X_{n+2}|\mathcal{F}_{n+1})|\mathcal{F}_n) \geq E(X_{n+1}|\mathcal{F}_n) \geq X_n$ □

Bemerkung 21.8 Ist $(X_n, \mathcal{F}_n)$ ein Submartingal, so ist $(X_n, \sigma(X_1, \dots, X_n))$ Submartingal.

Satz 21.9 *Ist $(X_n, \sigma(X_1, \dots, X_n))$ ein Martingal und gilt $EX_n^2 < \infty\ \forall n$, dann sind die Martingal-Differenzen $X_1, X_2 - X_1, \dots$ orthogonal.*

Beweis: Setze $X_0 = 0$. Für $j < k$ gilt $E(X_j - X_{j-1})(X_k - X_{k-1}) = EE((X_j - X_{j-1})(X_k - X_{k-1})|\mathcal{F}_j) \overset{20.4}{=} E((X_j - X_{j-1})E(X_k - X_{k-1}|\mathcal{F}_j))$ und $E(X_k - X_{k-1}|\mathcal{F}_j) = X_j - X_j = 0$ □

Satz 21.10 *Sei $(X_n, \mathcal{F}_n)$ ein (Sub-)Martingal und $f : I\!R \mapsto I\!R$ konvex (und nichtfallend). Ist $f \circ X_n$ integrierbar $\forall n$, dann ist $(f \circ X_n, \mathcal{F}_n)$ ein Submartingal.*

Beweis: Nach Satz 20.6 gilt $E(f \circ X_{n+1}|\mathcal{F}_n) \geq f \circ E(X_{n+1}|\mathcal{F}_n)$
a) Submartingal: $E(X_{n+1}|\mathcal{F}_n) \geq X_n$, f nichtfallend, also $f \circ E(X_{n+1}|\mathcal{F}_n) \geq f \circ X_n$
b) Martingal: Wegen $E(X_{n+1}|\mathcal{F}_n) = X_n$ gilt $f \circ E(X_{n+1}|\mathcal{F}_n) = f \circ X_n$. □

Beispiel 21.11 1. Ist (X_n) Submartingal, dann auch (X_n^+)
2. Ist (X_n) Martingal, $p \geq 1$ und $X_n \in L_p$, dann ist $|X_n|^p$ ein Submartingal.

Satz 21.12 (optional skipping) *Sei $(X_n, \mathcal{F}_n)$ Submartingal.*
Für $B_n \in \mathcal{B}^n$ setze $\varepsilon_n = 1_{((X_1,\ldots,X_n)\in B_n)}$, $Y_n = X_1 + \sum_{i=1}^{n-1} \varepsilon_i (X_{i+1} - X_i)$.
Dann ist $(Y_n, \mathcal{F}_n)$ Submartingal, und $EY_n \leq EX_n$.

Beweis:
a) $E(Y_{n+1}|\mathcal{F}_n) = E(Y_n + \varepsilon_n(X_{n+1} - X_n)|\mathcal{F}_n) = Y_n + \varepsilon_n E(X_{n+1} - X_n|\mathcal{F}_n) \geq Y_n + \varepsilon_n(X_n - X_n) = Y_n$
b) $EX_1 = EY_1$. Schreibe $X_{n+1} - Y_{n+1} = X_{n+1} - Y_n - \varepsilon_n(X_{n+1} - X_n) = (1 - \varepsilon_n)(X_{n+1} - X_n) + X_n - Y_n$.
Dann $E(X_{n+1} - Y_{n+1}|\mathcal{F}_n) = (1 - \varepsilon_n)E(X_{n+1} - X_n|\mathcal{F}_n) + E(X_n - Y_n|\mathcal{F}) \geq E(X_n - Y_n|\mathcal{F}) = X_n - Y_n$.
Also $E(X_{n+1} - Y_{n+1}) \geq E(X_n - Y_n) \geq \ldots \geq E(X_1 - Y_1) = 0$ □

22 Stoppzeiten; Optional Sampling

Gegeben seien $(\Omega, \mathcal{F}, P)$ und $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \ldots$ als Sub-σ-Algebren von $\mathcal{F}$.

Definition 22.1 *Eine **Stoppzeit** ist eine Zufallsvariable $T : \Omega \mapsto I\!N_\infty$ $(= I\!N \cup \{\infty\})$ mit $\{T \leq n\} \in \mathcal{F}_n$, $n \in I\!N$.*

Bemerkung 22.2 T ist Stoppzeit genau dann, wenn $\{T = n\} \in \mathcal{F}_n$, $n \in I\!N$.

Beweis: $\Rightarrow$: $\{T = n\} = \{T \leq n\} - \{T \leq n-1\} \in \mathcal{F}_n$
$\Leftarrow$: $\{T \leq n\} = \bigcup_{i=1}^n \{T = i\} \in \mathcal{F}_n$ □

Beispiel 22.3 Seien $X_1, X_2, \ldots$ Zufallsvariablen. Die **Eintreffzeit** in $B \in \mathcal{B}$ ist $T = \inf\{n : X_n \in B\}$; T ist Stoppzeit bzgl. $\mathcal{F}_n = \sigma\{X_1, \ldots, X_n\}$.

Beweis: $\{T \leq n\} = \bigcup_{k=1}^n \{X_k \in B\} \in \mathcal{F}_n$.

Definition 22.4 *Sei T Stoppzeit und $A \in \mathcal{F}$. Das Ereignis A **tritt bis** T **ein**, wenn $A \cap \{T \leq n\} \in \mathcal{F}_n$ für $n \in I\!N$.*

Bemerkung 22.5 äquivalent: $A \cap \{T = n\} \in \mathcal{F}_n$ für $n \in \mathbb{N}$.

Bemerkung 22.6 Die Menge der bis T eingetretenen Ereignisse ist eine σ-Algebra, sie heißt $\mathcal{F}_T$.

Satz 22.7 (optional sampling) *Sei $(X_n, \mathcal{F}_n)$ ein Submartingal und $T_1 \leq T_2 \leq \ldots$ Stoppzeiten mit*

1. X_{T_n} *integrierbar*
2. $\liminf\limits_k E1_{(T_n > k)}|X_k| = 0$

Dann ist $(X_{T_n}, \mathcal{F}_{T_n})$ ein Submartingal.

Beweis:
1. $\mathcal{F}_{T_1} \subset \mathcal{F}_{T_2} \subset \ldots$; X_{T_n} $\mathcal{F}_{T_n}$-messbar
2. $E1_A X_{T_{n+1}} \geq E1_A X_{T_n}$ für $A \in \mathcal{F}_{T_n}$:
   Wegen $A = \bigcup_j A \cap \{T_n = j\}$ genügt es, $D_j = A \cap \{T_n = j\}$ statt A zu betrachten.
   Für $k > j$ gilt: $E1_{D_j} X_{T_{n+1}} = \sum_{i=j}^k E1_{D_j \cap \{T_{n+1} = i\}} X_{T_{n+1}} + E1_{D_j \cap \{T_{n+1} > k\}} X_{T_{n+1}} =$
   $\sum_{i=j}^k E1_{D_j \cap \{T_{n+1} = i\}} X_i + E1_{D_j \cap \{T_{n+1} > k\}} X_k - E1_{D_j \cap \{T_{n+1} > k\}} (X_k - X_{T_{n+1}})$
   Nun gilt für $i = k$: $E1_{D_j \cap \{T_{n+1} = k\}} X_k + E1_{D_j \cap \{T_{n+1} > k\}} X_k = E1_{D_j \cap \{T_{n+1} \geq k\}} X_k \geq$
   $E1_{D_j \cap \{T_{n+1} \geq k\}} X_{k-1}$ wegen $\{T_{n+1} \geq k\} = \{T_{n+1} \leq k-1\}^c \in \mathcal{F}_{k-1}$ und $D_j \in \mathcal{F}_j \subset \mathcal{F}_{k-1}$.
   Ebenso für $i = k-1$ u.s.w.
   Insgesamt folgt $E1_{D_j} X_{T_{n+1}} \geq E1_{D_j \cap \{T_{n+1} \geq j\}} X_j - \underbrace{E1_{D_j \cap \{T_{n+1} > k\}} (X_k - X_{T_{n+1}})}_{\to 0,\ k \to \infty}$ n. Vor., und
   es gilt $D_j \cap \{T_{n+1} \geq j\} = D_j$ wegen $T_n = j$ auf D_j. □

Bemerkung 22.8 Die Voraussetzungen an X_n in Satz 22.7 gelten, wenn eine der folgenden Bedingungen erfüllt ist:

a) Die Stoppzeiten sind beschränkt

b) $\sup |X_n|$ ist integrierbar

Beweis:
a) 1. Sei $T_n \leq K_n$. Dann $E|X_{T_n}| = \sum_{i \leq K_n} E1_{(T_n = i)} |X_i| \leq \sum_{i \leq K_n} E|X_i| < \infty$
   2. $\{T > k\} = \emptyset$ für $k \geq K_n$
b) 1. $E|X_{T_n}| \leq E \sup |X_n| < \infty$
   2. $E1_{(T_n > k)} |X_k| \leq E1_{(T_n > k)} \sup |X_n| \to 0,\ k \to \infty$ □

23 Submartingal-Konvergenzsatz

Sei $(X_1, \mathcal{F}_1), \ldots, (X_n, \mathcal{F}_n)$ ein Submartingal; es gelte $a < b$.
Setze die Stoppzeiten $T_1 = \min\{j : X_j \leq a\}$ (erstes Downcrossing), $T_2 = \min\{j > T_1 : X_j \geq b\}$ (erstes Upcrossing), $T_3 = \min\{j > T_2 : X_j \leq a\}$ (zweites Downcr.) etc. bis T_N.
Die Anzahl der Upcrossings ist $U_{nab} = [\frac{N}{2}]$.

Satz 23.1 (Upcrossing-Ungleichung; Doob) *Ist $(X_1, \mathcal{F}_1) \ldots, (X_n, \mathcal{F}_n)$ ein Submartingal und $a < b$, so gilt*

$$EU_{nab} \leq \frac{E(X_n - a)^+}{b-a}$$

Beweis:

1. Gelte $X_1, \ldots, X_n \geq 0,\ a = 0$
   Dann $X_{T_j} = 0$ für j ungerade, $1 < T_j < n$. Setze $T_0 = 1,\ T_j = n$ für $j > N$. Schreibe $X_n - X_1 = X_{T_n} - X_{T_0} = \sum_{j=0}^{n-1}(X_{T_{j+1}} - X_{T_j})$.
   Für ungerade j gilt
   $$j \left\{ \begin{array}{c} < \\ = \\ > \end{array} \right\} N: \quad X_{T_{j+1}} \left\{ \begin{array}{l} \geq b > 0 = X_{T_j} \\ = X_n \geq 0 = X_{T_j} \\ = X_n = X_{T_j} \end{array} \right.$$
   $$\sum_{j \text{ ungerade}} (X_{T_{j+1}} - X_{T_j}) \geq \sum_{\substack{j \text{ ungerade} \\ j<N}} (X_{T_{j+1}} - X_{T_j}) \geq bU_{n0b}$$
   Nach Satz 22.7 ist $(X_{T_1}, \mathcal{F}_{T_1}), \ldots (X_{T_n}, \mathcal{F}_{T_n})$ ein Submartingal, also gilt $E(X_{T_{j+1}} - X_{T_j}) \geq 0$, also $E(X_n - X_1) \geq E \sum_{j \text{ ungerade}} (X_{T_{j+1}} - X_{T_j}) \geq bEU_{n0b}$
2. Allgemeiner Fall:
   Die Anzahl der Upcrossings von X_n auf $[a, b]$ ist gleich der Anzahl der Upcrossings von $(X_n - a)^+$ auf $[0, b-a]$ und $(X_n - a)^+$ ist nach Satz 21.10 ein Submartingal. Mit 1. folgt die Behauptung. □

Satz 23.2 (Submartingal-Konvergenzsatz) *Sei $(X_n, \mathcal{F}_n),\ n = 1, 2, \ldots$ ein Submartingal und $\sup EX_n^+ < \infty$. Dann existiert ein integrierbares X_∞ mit $X_n \to X_\infty$ f.s.*

Beweis:

1. Existenz:
   Sei $U_{ab} = \sup_n U_{nab}$. Nach Satz 23.1 gilt $EU_{nab} \leq \frac{E(X_n-a)^+}{b-a} \leq \frac{C+|a|}{b-a}$
   Also $EU_{ab} < \infty$ mit Satz 6.3 und somit $U_{ab} < \infty$ f.s.; also $P(X_n$ konvergiert nicht$)$ $= P \bigcup_{\substack{a<b \\ a,b \in \mathbb{Q}}} \{\liminf X_n < a < b < \limsup X_n\} = 0$.
2. Integrierbarkeit:
   $|X| = X^+ + X^- = 2X^+ - (X^+ - X^-) = 2X^+ - X$, somit $E|X_n| = 2EX_n^+ - EX_n \leq 2C - EX_1$, also mit Lemma 6.6 $E|X_\infty| \leq \sup E|X_n| < \infty$. □

Anhang A

Die wichtigsten Verteilungen

Verteilung	Bezeichnung	$P(X=k)$	Dichte
Gleich	U_A	$\frac{1}{n}$ für $\|A\|=n$	$\frac{1}{PA}1_A(x)$
Bernoulli	B_{1p}	$1:\ p,\ 0:\ 1-p$	
Binomial	B_{np}	$\binom{n}{k}p^k(1-p)^{n-k}$	
neg. Binomial	B^-_{np}	$\binom{k+n-1}{n-1}p^n(1-p)^k,\ k\geq 0$	
Geometrische	G_p	$p(1-p)^{k-1}$	
Hypergeom.	$H_{N,K,n}$	$\frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$	
Poisson	P_λ	$e^{-\lambda}\frac{\lambda^k}{k!},\ k\geq 0$	
Exponential	E_a		$\frac{1}{a}e^{-x/a},\ x>0$
Gamma	Γ_{ab}		$\frac{1}{a^b\Gamma(b)}x^{b-1}e^{-x/a},\ x>0$
Normal	N_{μ,σ^2}		$\frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{(x-\mu)^2}{2\sigma^2})$
Cauchy	C_a		$\frac{a}{\pi(a^2+x^2)}$

Verteilung	Erwartungswert	Varianz	Char. Funktion
U_A	PA	$PA-(PA)^2$	$\frac{\sin at}{at}$ für $A=(-a,a)$
B_{1p}	p	$p(1-p)$	$1+p(e^{it}-1)$
B_{np}	np	$np(1-p)$	$(1+p(e^{it}-1))^n$
B^-_{np}	$\frac{n(1-p)}{p}$	$\frac{n(1-p)}{p^2}$	$(\frac{p}{1-(1-p)e^{it}})^n$
G_p	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^{it}}{1-(1-p)e^{it}}$
$H_{N,K,n}$	$n\frac{K}{N}$	$n\frac{K}{N}\frac{K-N}{N}\frac{N-n}{N-1}$	
P_λ	λ	λ	$e^{\lambda(e^{it}-1)}$
E_a	a	a^2	$(1-ait)^{-1}$
Γ_{ab}	ab	a^2b	$(1-ait)^{-b}$
N_{μ,σ^2}	μ	σ^2	$e^{it\mu}e^{-\frac{1}{2}t^2\sigma^2}$
C_a	/	/	$e^{-a\|t\|}$

Anhang B

Literatur

Literatur zur Vorlesung:

[1] Ash, R.B.: *Real Analysis and Probability*, Academie Press 1972

[2] Chung, K.L.: *A Course in Probability Theory*, Academie Press 1974

[3] Fristedt,B. und Gray, L.: *A Modern Approach to Probability Theory*, Birkhäuser 1996

Anwendungen der Wahrscheinlichkeitstheorie:

[4] Adler, R.J., Müller, P. und Rozovskii, B.L.: *Stochastic Modelling in Physical Oceanography*, Progress in Probability 39, Birkhäuser 1996

[5] Beltrami, E.: *Von Krebsen und Kriminellen*, Vieweg 1993

[6] Gardiner, C.W.: *Handbook of Stochastik Methods for Physics, Chemistry and Natural Sciences*, Springer Series in Synergetics 13, Springer 1985

[7] Kree, P. und Wedig, W.: *Probabiltistic Methods in Applied Physics*, Lecture Notes in Physics 451, Springer 1985

[8] Merriam, D.F.: *Random Processes in Geology*, Springer 1976

[9] Molchanov, S.A. und Woyczynski, W.A.: *Stochastik Models in Geosystems*, IMA Volumes in Mathematics and its Applications 85, Springer-Verlag 1996

[10] Musiela, M. und Rutkowski, M.: *Martingale Methods in Financial Modelling*, Applications of Mathematics 36, Springer 1997

Index