Wahrscheinlichkeitstheorie
Gliederung zur Vorlesung
im Wintersemester 2005/06

Markus Reiß
Universität Heidelberg
reiss@statlab.uni-heidelberg.de

VORLÄUFIGE FASSUNG: 25. Juli 2006

Inhaltsverzeichnis

1 Maßtheorie

1.1 Wiederholung und Ergänzung

1.1 Definition. *Ein Wahrscheinlichkeitsraum ist ein Tupel* $(\Omega, \mathscr{F}, \mathbb{P})$*, bestehend aus einer Grundmenge* Ω*, einer* σ*-Algebra* $\mathscr{F}$ *über* Ω *und einem Wahrscheinlichkeitsmaß* $\mathbb{P}$ *auf* $\mathscr{F}$*.* $(\Omega, \mathscr{F})$ *heißt Messraum.*

Dabei heißt ein System $\mathscr{F}$ *von Teilmengen von* Ω σ*-Algebra, falls*

(a) $\Omega \in \mathscr{F}$;

(b) $A \in \mathscr{F} \Rightarrow A^C \in \mathscr{F}$;

(c) $A_n \in \mathscr{F}$, $n \in \mathbb{N} \Rightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathscr{F}$.

Eine Funktion $\mathbb{P} : \mathscr{F} \to [0, 1]$ *heißt Wahrscheinlichkeitsmaß, falls*

(a) $\mathbb{P}(\Omega) = 1$;

(b) $A_n \in \mathscr{F}$, $n \in \mathbb{N}$, *paarweise disjunkt* $\Rightarrow \mathbb{P}(\bigcup_n A_n) = \sum_n \mathbb{P}(A_n)$.

1.2 Lemma. *Sind* $(\mathscr{F}_i)_{i\in I}$ σ*-Algebren über* Ω *für eine beliebige nichtleere Indexmenge* I*, so ist auch* $\bigcap_{i\in I} \mathscr{F}_i$ *eine* σ*-Algebra.*

1.3 Korollar. *Zu jedem Mengensystem* $\mathscr{A} \subseteq \mathscr{P}(\Omega)$ *gibt es eine kleinste* σ*-Algebra, die* $\mathscr{A}$ *enthält („die von* $\mathscr{A}$ *erzeugt wird“).*

1.4 Definition. (a) *Die von* $\mathscr{A}$ *erzeugte* σ*-Algebra wird mit* $\sigma(\mathscr{A})$ *bezeichnet, und man nennt* $\mathscr{A}$ *den Erzeuger von* $\sigma(\mathscr{A})$*.*

(b) *Es sei* (S, d) *ein metrischer Raum (z.B.* $\mathbb{R}^d$ *mit der Euklidischen Metrik). Dann heißt die von den offenen Mengen in* S *erzeugte* σ*-Algebra Borel-*σ*-Algebra (Notation* $\mathfrak{B}_S$*) und ihre Elemente heißen Borelmengen.*

1.5 Definition. *Eine Abbildung* $g : F \to S$ *von dem Messraum* $(F, \mathscr{F})$ *in den Messraum* $(S, \mathscr{S})$ *heißt* $(\mathscr{F}, \mathscr{S})$*-messbar, falls*

$$\forall\, A \in \mathscr{S} :\ g^{-1}(A) \in \mathscr{F}$$

gilt. Ist $(\Omega, \mathscr{F}, \mathbb{P})$ *ein Wahrscheinlichkeitsraum, so heißt jede* $(\mathscr{F}, \mathscr{S})$*-messbare Abbildung* $X : \Omega \to S$ $(S, \mathscr{S})$*-wertige Zufallsvariable. Das induzierte Wahrscheinlichkeitsmaß (!) auf* $\mathscr{S}$

$$\mathbb{P}^X(A) := \mathbb{P}(X \in A) := \mathbb{P}(X^{-1}(A)) = \mathbb{P}(\{\omega \in \Omega \,|\, X(\omega) \in A\}), \quad A \in \mathscr{S},$$

heißt Verteilung von X *(oder Bildmaß von* $\mathbb{P}$ *unter* X*).*

1.6 Lemma. *Es sei* $\mathscr{A}$ *ein Erzeuger von* $\mathscr{S}$*. Dann ist* $g : F \to S$ *bereits* $(\mathscr{F}, \mathscr{S})$*-messbar, wenn* $g^{-1}(A) \in \mathscr{F}$ *für alle* $A \in \mathscr{A}$ *gilt.*

1.7 Korollar. *Sind* $(F, \mathfrak{B}_F)$ *und* $(S, \mathfrak{B}_S)$ *metrische Räume mit ihren Borel-*σ*-Algebren, so ist jede stetige Funktion* $g : F \to S$ *Borel-messbar.*

1.8 Lemma. *Die Komposition messbarer Funktion ist wiederum messbar.*

1.9 Lemma. *Eine Abbildung $g = (g_1, \dots, g_d) : F \to \mathbb{R}^d$ ist genau dann $(\mathscr{F}, \mathfrak{B}_{\mathbb{R}^d})$-messbar, wenn jede Koordinatenabbildung $g_i : F \to \mathbb{R}$ $(\mathscr{F}, \mathfrak{B}_{\mathbb{R}})$-messbar ist.*

1.10 Korollar. *Sind X, Y reellwertige Zufallsvariablen, so ist auch $f(X,Y)$ eine Zufallsvariable für jede Borel-messbare Funktion $f : \mathbb{R}^2 \to \mathbb{R}$, z.B. $X \pm Y$, XY, $|X|$, $\max(X,Y)$.*

1.11 Lemma. *Sind $g_n : F \to S$ $(\mathscr{F}, \mathfrak{B}_S)$-messbar für einen metrischen Raum (S,d), so ist auch der punktweise Grenzwert $\lim_{n\to\infty} g_n$ $(\mathscr{F}, \mathfrak{B}_S)$-messbar, sofern er existiert.*

1.2 Konstruktion allgemeiner Maße

1.12 Definition. *Ein Mengensystem $\mathscr{A} \subseteq \mathscr{P}(X)$ heißt Algebra über der Menge X, falls*

(a) $\Omega \in \mathscr{A}$;

(b) $A \in \mathscr{A} \Rightarrow A^C \in \mathscr{A}$;

(c) $A, B \in \mathscr{A} \Rightarrow A \cup B \in \mathscr{A}$.

1.13 Definition. *Eine Abbildung $\mu : \mathscr{A} \to [0,\infty]$ heißt Prämaß über der Algebra $\mathscr{A}$, falls*

$$A_n \in \mathscr{A},\ n \in \mathbb{N},\ \text{paarweise disjunkt mit} \bigcup_n A_n \in \mathscr{A} \Rightarrow \mu\Big(\bigcup_n A_n\Big) = \sum_n \mu(A_n)$$

gilt (Konvention: $a + \infty := \infty$, $a \in [0,\infty]$). Ist $\mathscr{A}$ eine σ-Algebra, so heißt μ Maß.

1.14 Definition. *Ein äußeres Maß auf X ist eine Abbildung $\eta : \mathscr{P}(X) \to [0,\infty]$ mit*

(a) $\eta(\varnothing) = 0$;

(b) $A \subseteq B$: $\eta(A) \leqslant \eta(B)$;

(c) $\eta(\bigcup_n A_n) \leqslant \sum_n \eta(A_n)$.

1.15 Definition. *Für ein äußeres Maß η auf X heißt $A \subseteq X$ η-messbar, wenn für alle $Q \subseteq X$ gilt*

$$\eta(Q) = \eta(Q \cap A) + \eta(Q \cap A^C).$$

1.16 Satz. *$\mathscr{A}_\eta := \{A \subseteq X \mid A\ \eta\text{-messbar}\}$ ist eine σ-Algebra und $\eta|_{\mathscr{A}_\eta}$ ein Maß auf dieser σ-Algebra.*

1.17 Satz. *(Maßfortsetzungssatz von Caratheodory) Es sei μ ein Prämaß auf einer Algebra $\mathscr{A}$ über X sowie*

$$\eta(A) := \inf\Big\{ \sum_{n\in\mathbb{N}} \mu(A_n) \,\Big|\, A_n \in \mathscr{A}, A \subseteq \bigcup_{n\in\mathbb{N}} A_n \Big\}, \quad A \subseteq X.$$

Dann ist η ein äußeres Maß, alle Mengen aus $\mathscr{A}$ sind η-messbar und $\eta|_{\mathscr{A}_\eta}$ ist eine Fortsetzung von μ zu einem Maß auf $\mathscr{A}_\eta \supseteq \sigma(\mathscr{A})$.

1.3 Eindeutigkeit der Maßfortsetzung

1.18 Definition. *Ein Maß μ auf einer σ-Algebra $\mathscr{F}$ über X heißt*

(a) *σ-endlich, falls $A_n \in \mathscr{F}$ existieren mit $X = \bigcup_n A_n$ und $\mu(A_n) < \infty$, $n \in \mathbb{N}$.*

(b) *endlich, falls $\mu(X) < \infty$;*

(c) *Wahrscheinlichkeitsmaß, falls $\mu(X) = 1$.*

1.19 Satz. *(Eindeutigkeitssatz) Es seien μ, ν endliche Maße auf einer σ-Algebra $\mathscr{F}$ über X mit $\mu(X) = \nu(X)$. Ist $\mathscr{E}$ ein $\cap$-stabiler Erzeuger von $\mathscr{F}$ (d.h. $\sigma(\mathscr{E}) = \mathscr{F}$, $A, B \in \mathscr{E} \Rightarrow A \cap B \in \mathscr{E}$) mit $\mu|_{\mathscr{E}} = \nu|_{\mathscr{E}}$, so gilt bereits $\mu = \nu$.*

1.20 Definition. *Ein Mengensystem $\mathscr{D}$ heißt Dynkinsystem über X, falls*

(a) *$X \in \mathscr{D}$;*

(b) *$A, B \in \mathscr{D}$, $A \subseteq B \Rightarrow B \setminus A \in \mathscr{D}$;*

(c) *$A_n \in \mathscr{D}$, $n \in \mathbb{N}$, paarweise disjunkt $\Rightarrow \bigcup_n A_n \in \mathscr{D}$.*

1.21 Lemma. *Ist $\mathscr{E}$ ein $\cap$-stabiles Mengensystem und $\mathscr{D}$ das kleinste Dynkinsystem, das $\mathscr{E}$ enthält, so gilt $\sigma(\mathscr{E}) = \mathscr{D}$.*

1.22 Korollar. *Der Eindeutigkeitssatz gilt auch für σ-endliche Maße μ, ν, sofern es $A_n \in \mathscr{F}$ gibt mit $X = \bigcup_n A_n$ und $\mu(A_n) = \nu(A_n) < \infty$.*

1.23 Korollar. *Jede monoton wachsende, rechtsstetige Funktion $F : \mathbb{R} \to [0,1]$ mit $\lim_{x\to-\infty} F(x) = 0$, $\lim_{x\to\infty} F(x) = 1$ definiert ein eindeutiges Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $\mathfrak{B}_{\mathbb{R}}$ gemäß*

$$\mathbb{P}((a,b]) := F(b) - F(a), \quad a < b,$$

sowie

$$\mathbb{P}(B) := \inf\Big\{ \sum_{n\in\mathbb{N}} \mathbb{P}((a_n, b_n]) \,\Big|\, B \subseteq \bigcup_n (a_n, b_n] \Big\}, \quad B \in \mathfrak{B}_{\mathbb{R}}.$$

1.24 Lemma. *Ist $\mu : \mathscr{A} \to [0,\infty)$ eine additive Mengenfunktion, so ist μ σ-additiv genau dann, wenn gilt $A_n \uparrow X \Rightarrow \mu(A_n) \to \mu(X)$.*

2 Integrationstheorie

2.1 Integration von Treppenfunktionen

Im folgenden sei stets $(X, \mathscr{F}, \mu)$ ein Maßraum, $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ sowie $\bar{\mathfrak{B}} := \sigma(\mathfrak{B}_{\mathbb{R}}, \{[-\infty, a], [a, \infty] \,|\, a \in \mathbb{R}\})$ die zugehörige Borel-σ-Algebra.

2.1 Definition. *Eine Funktion $f : X \to \bar{\mathbb{R}}$ liegt*

(a) *in $\mathscr{T}$, falls $f = \sum_{k=1}^m a_k \mathbf{1}_{B_k}$ gilt mit $m \in \mathbb{N}$, $a_k \in \bar{\mathbb{R}}$, $B_k \in \bar{\mathfrak{B}}$ (f heißt Treppenfunktion);*

(b) *in* $\mathscr{T}^+$*, falls* $f \in \mathscr{T}$ *und* $f \geqslant 0$ *gilt;*

(c) *in* $\mathscr{M}$*, falls* f $(\mathscr{F}, \bar{\mathfrak{B}})$*-messbar ist;*

(d) *in* $\mathscr{M}^+$*, falls* $f \in \mathscr{M}$ *und* $f \geqslant 0$ *gilt.*

2.2 Definition. *Für* $f \in \mathscr{T}^+$ *mit* $f = \sum_{k=1}^m a_k \mathbf{1}_{B_k}$ *ist das Lebesgue-Integral bezüglich dem Maß* μ *definiert als*

$$\int f\,d\mu := \int f(x)\,\mu(dx) := \int f(x)\,d\mu(x) := \sum_{k=1}^m a_k \mu(B_k) \in [0,\infty],$$

wobei wir stets $\infty \bullet 0 := 0$*,* $a + \infty := \infty$ *setzen.*

2.2 Integration allgemeiner messbarer Funktionen

2.3 Satz. *Zu jeder Funktion* $f \in \mathscr{M}^+$ *existiert eine Folge* (f_n) *in* $\mathscr{T}^+$ *mit* $f_n(x) \uparrow f(x)$ *für jedes* $x \in X$*.*

2.4 Definition. *Für* $f \in \mathscr{M}^+$ *definiere das Lebesgue-Integral bezüglich* μ *als*

$$\int f\,d\mu := \sup\Big\{ \int g\,d\mu \,\Big|\, g \in \mathscr{T}^+,\ g \leqslant f \Big\} \in [0,\infty].$$

2.5 Lemma. *Es sei* (f_n) *eine beliebige Folge in* $\mathscr{T}^+$ *mit* $f_n \uparrow f$ *für ein* $f \in \mathscr{M}^+$*. Dann gilt* $\int f_n\,d\mu \uparrow \int f\,d\mu$*.*

2.6 Definition. *Eine Funktion* $f \in \mathscr{M}$ *heißt integrierbar, falls* $f^+ := \max(f,0)$ *und* $f^- := \max(-f,0)$ *jeweils endliche Integrale* $\int f^+ d\mu$*,* $\int f^- d\mu < \infty$ *besitzen. Man schreibt* $f \in \mathscr{L}^1 = \mathscr{L}^1(X,\mathscr{F},\mu)$ *und setzt*

$$\int f\,d\mu := \int f^+\,d\mu - \int f^-\,d\mu \in \mathbb{R}\,.$$

2.7 Lemma.

(a) *Für* $f,g \in \mathscr{M}^+$ *mit* $f \leqslant g$ *und* $g \in \mathscr{L}^1$ *gilt* $f \in \mathscr{L}^1$*.*

(b) $\mathscr{L}^1$ *ist ein Vektorraum und es gilt für* $f,g \in \mathscr{L}^1$*,* $\alpha \in \mathbb{R}$*:*

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

(c) *Es gilt* $f \in \mathscr{L}^1 \iff |f| \in \mathscr{L}^1$ *und* $|\int f\,d\mu| \leqslant \int |f|\,d\mu$*.*

(d) *Gilt* $f = g$ μ*-fast überall (d.h.* $\mu(f \neq g) = 0$*) und* $g \in \mathscr{L}^1$*, dann ist* $f \in \mathscr{L}^1$ *und* $\int f\,d\mu = \int g\,d\mu$*.*

(e) $\|f\|_{\mathscr{L}^1} := \int |f|\,d\mu$ *ist eine Seminorm auf* $\mathscr{L}^1$ *mit* $\|f\|_{\mathscr{L}^1} = 0 \iff f = 0$ μ*-fast überall.*

2.3 Der allgemeine Erwartungswert

2.8 Definition. *Ist $(\Omega, \mathscr{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und X eine Zufallsvariable in $\mathscr{L}^1(\Omega, \mathscr{F}, \mathbb{P})$, so setzt man*

$$\mathbb{E}[X] := \int X \, d\,\mathbb{P}$$

und bezeichnet dies als Erwartungswert von X (bezüglich $\mathbb{P}$). Entsprechend ist für $X \in \mathscr{M}^+(\Omega, \mathscr{F}, \mathbb{P})$ der Erwartungswert in $[0, \infty]$ erklärt.

2.9 Satz. *(Transformationssatz) Ist X eine $(S, \mathscr{S})$-wertige Zufallsvariable auf $(\Omega, \mathscr{F}, \mathbb{P})$ mit Verteilung $\mathbb{P}^X$, so gilt für alle $\varphi \in \mathscr{M}^+(S, \mathscr{S})$ und alle $\varphi \in \mathscr{L}^1(S, \mathscr{S}, \mathbb{P}^X)$:*

$$\mathbb{E}[\varphi(X)] := \int \varphi(X(\omega)) \, \mathbb{P}(d\omega) = \int \varphi(x) \, \mathbb{P}^X(dx).$$

2.10 Definition. *Man setzt $\int_A f \, d\mu := \int \mathbf{1}_A f \, d\mu$.*

2.11 Satz. *(Dichtesatz) Es sei $f \in \mathscr{L}^1(\Omega, \mathscr{F}, \mu)$ mit $f \geqslant 0$ und $\int f \, d\mu = 1$ gegeben, wobei μ ein beliebiges Maß ist. Dann wird durch*

$$\mathbb{P}(A) := \int_A f \, d\mu, \quad A \in \mathscr{F},$$

ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathscr{F})$ definiert. Es gilt für alle $X \in \mathscr{M}^+(\Omega, \mathscr{F})$ und alle $X \in \mathscr{L}^1(\Omega, \mathscr{F}, \mathbb{P})$

$$\mathbb{E}[X] := \int X \, d\,\mathbb{P} = \int X f \, d\mu.$$

2.12 Definition. *Die Funktion f im vorigen Satz heißt Dichte von $\mathbb{P}$ bezüglich μ, Notation $f = \frac{d\,\mathbb{P}}{d\mu}$.*

2.4 Integralkonvergenzsätze

2.13 Satz. *(monotone Konvergenz, Beppo Levi) Ist (f_n) eine Folge in $\mathscr{M}^+$ mit $f_n \uparrow f$, so gilt*

$$\int f_n \, d\mu \uparrow \int f \, d\mu.$$

2.14 Korollar. *Für jede Folge (f_n) in $\mathscr{M}^+$ gilt*

$$\int \sum_n f_n \, d\mu = \sum_n \int f_n \, d\mu.$$

2.15 Korollar. *(Lemma von Fatou) Für jede Folge (f_n) in $\mathscr{M}^+$ gilt*

$$\int \big(\liminf_{n \to \infty} f_n\big) \, d\mu \leqslant \liminf_{n \to \infty} \int f_n \, d\mu.$$

2.16 Satz. *(dominierte Konvergenz, Lebesgue) Es sei (f_n) eine Folge in $\mathscr{L}^1$ mit $|f_n| \leqslant g$ für ein $g \in \mathscr{L}^1$. Gilt $\lim_{n \to \infty} f_n = f$ μ-fast überall, so folgt*

$$\lim_{n \to \infty} \int f_n \, d\mu = \int f \, d\mu \text{ sowie } \lim_{n \to \infty} \|f - f_n\|_{\mathscr{L}^1} = 0.$$

3 Unabhängige Zufallsvariablen

3.1 Produktmaße

3.1 Definition. *Für beliebige Funktionen $g_i : X \to S_i$ mit Werten in Messräumen $(S_i, \mathscr{S}_i)$, $i \in I$, wird mit $\sigma(g_i,\, i \in I)$ die kleinste σ-Algebra über X bezeichnet, bezüglich der alle Funktionen g_i, $i \in I$, messbar sind.*

3.2 Definition. *Es sei $(X_i, \mathscr{F}_i)_{i \in I}$, I eine beliebige Indexmenge, eine Familie von Messräumen. Dann heißt die σ-Algebra $\sigma(\pi_j,\, j \in I)$ definiert durch die Koordinaten-Projektionen*

$$\pi_j : \textstyle\prod_{i \in I} X_i \to X_j, \quad \pi_j((x_i)_{i \in I}) = x_j$$

<u>Produkt-σ-Algebra</u> über $\prod_{i \in I} X_i$. Man setzt $\bigotimes_{i \in I} \mathscr{F}_i := \sigma(\pi_j,\, j \in I)$.

3.3 Definition. *Zu vorgegebenen Maßräumen $(X, \mathscr{F}, \mu)$, $(Y, \mathscr{G}, \nu)$ heißt ein Maß κ auf $\mathscr{F} \otimes \mathscr{G}$ <u>Produktmaß</u> von μ und ν, falls*

$$\forall\, A \in \mathscr{F},\, B \in \mathscr{G} :\ \kappa(A \times B) = \mu(A)\nu(B)$$

gilt. Man schreibt $\kappa = \mu \otimes \nu$.

3.4 Satz. *Sind μ und ν σ-endliche Maße, so existiert ein eindeutig bestimmtes Produktmaß $\mu \otimes \nu$.*

3.5 Satz. *Es seien $(\Omega_i, \mathscr{F}_i, \mathbb{P}_i)$, $i \in I$, Wahrscheinlichkeitsräume. Dann existiert genau ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\prod_{i \in I} \Omega_i, \bigotimes_{i \in I} \mathscr{F}_i)$ mit*

$$\mathbb{P}\Big(\prod_{i \in I} A_i\Big) = \prod_{i \in I} \mathbb{P}_i(A_i) \text{ für beliebige } A_i \in \mathscr{F}_i \text{ mit } A_i \neq \Omega_i \text{ nur endlich oft.}$$

3.6 Definition. *Das Wahrscheinlichkeitsmaß $\mathbb{P}$ im vorigen Satz heißt <u>Produktmaß</u> der $(\mathbb{P}_i)$, und man schreibt $\mathbb{P} = \bigotimes_{i \in I} \mathbb{P}_i$.*

3.7 Korollar. *Eine Familie $(S_i, \mathscr{S}_i)$-wertiger Zufallsvariablen $(X_i)_{i \in I}$ ist genau dann unabhängig, wenn für ihre Verteilungen gilt $\mathbb{P}^{(X_i)_{i \in I}} = \bigotimes_{i \in I} \mathbb{P}^{X_i}$.*

3.2 Der Satz von Fubini

3.8 Satz. *Es seien $(X, \mathscr{F}, \mu)$ und $(Y, \mathscr{G}, \nu)$ σ-endliche Maßräume.*

(a) (Tonelli) Für jedes $f \in \mathscr{M}^+(X \times Y, \mathscr{F} \otimes \mathscr{G})$ liegen die Funktionen

$$x \mapsto \int_Y f(x,y)\, \nu(dy), \quad y \mapsto \int_X f(x,y)\, \mu(dx)$$

in $\mathscr{M}^+(X, \mathscr{F})$ bzw. $\mathscr{M}^+(Y, \mathscr{G})$, und es gilt

$$\begin{aligned}\int_{X \times Y} f\, d(\mu \otimes \nu) &= \int_X \Big(\int_Y f(x,y)\, \nu(dy)\Big)\, \mu(dx) \\ &= \int_Y \Big(\int_X f(x,y)\, \mu(dx)\Big)\, \nu(dy).\end{aligned}$$

(b) *(Fubini) Ist $f \in \mathscr{L}^1(X \times Y, \mathscr{F} \otimes \mathscr{G}, \mu \otimes \nu)$, so ist $f(x, \bullet) \in \mathscr{L}^1(Y, \mathscr{G}, \nu)$ für μ-fast alle x und $f(\bullet, y) \in \mathscr{L}^1(X, \mathscr{F}, \mu)$ für ν-fast alle y. Es gilt*

$$\int_{X \times Y} f \, d(\mu \otimes \nu) = \int_{\{x \colon f(x,\bullet) \in \mathscr{L}^1\}} \Big(\int_Y f(x,y) \, \nu(dy) \Big) \, \mu(dx)$$
$$= \int_{\{y \colon f(\bullet,y) \in \mathscr{L}^1\}} \Big(\int_X f(x,y) \, \mu(dx) \Big) \, \nu(dy).$$

3.9 Lemma. *Für die Schnitte von $M \in \mathscr{F} \otimes \mathscr{G}$ gilt:*

$$M_a := \{ y \in Y \mid (a,y) \in M \} \in \mathscr{G}, \quad a \in X;$$
$$M^b := \{ x \in X \mid (x,b) \in M \} \in \mathscr{F}, \quad b \in Y.$$

3.10 Lemma. *(Prinzip von Cavalieri) Für $M \in \mathscr{F} \otimes \mathscr{G}$ ist $x \mapsto \nu(M_x)$ messbar, und es gilt*

$$\mu \otimes \nu(M) = \int_X \nu(M_x) \, \mu(dx).$$

3.3 Anwendungen

3.11 Definition. *Man setzt für Ereignisse $(A_n)_{n \in \mathbb{N}}$*

$$\limsup_{n \to \infty} A_n := \bigcap_{n \geqslant 1} \bigcup_{m \geqslant n} A_m = \text{„unendliche viele } A_n \text{ treten ein“};$$
$$\liminf_{n \to \infty} A_n := \bigcup_{n \geqslant 1} \bigcap_{m \geqslant n} A_m = \text{„alle bis auf endlich viele } A_n \text{ treten ein“}.$$

3.12 Satz. *(Borel-Cantelli)*

(a) *Aus $\sum_n \mathbb{P}(A_n) < \infty$ folgt $\mathbb{P}(\limsup_n A_n) = 0$.*

(b) *Ist $\sum_n \mathbb{P}(A_n) = \infty$ und sind die Ereignisse $(A_n)_n$ unabhängig, so folgt $\mathbb{P}(\limsup_n A_n) = 1$.*

3.13 Definition. *Mengen $(\mathscr{C}_i)_{i \in I}$ von Ereignissen heißen unabhängig, falls*

$$\mathbb{P}\Big(\bigcap_{j=1}^n C_{i_j} \Big) = \prod_{j=1}^n \mathbb{P}(C_{i_j}) \quad \forall n \in \mathbb{N}, \, i_1, \dots, i_n \in I, \, C_i \in \mathscr{C}_i.$$

3.14 Lemma. *Sind $\mathscr{E}_i$ $\cap$-stabile, unabhängige Erzeuger der σ-Algebren $\mathscr{F}_i$, $i \in I$, so sind auch $(\mathscr{F}_i)_{i \in I}$ unabhängig.*

3.15 Lemma. *Es seien $\mathscr{F}_i$, $i \in I$, unabhängige σ-Algebren und $I = \bigcup_{j \in J} I_j$ eine Partition von I. Dann sind auch $\mathscr{F}_{I_j} := \sigma(\mathscr{F}_i, \, i \in I_j)$, $j \in J$, unabhängig.*

3.16 Definition. *Für Zufallsvariablen $(X_n)_{n \in \mathbb{N}}$ setze $\mathscr{F}_n^\infty := \sigma(X_k, \, k \geqslant n)$, $\mathscr{F}^\infty := \bigcap_{n \in \mathbb{N}} \mathscr{F}_n^\infty$. Die Elemente von $\mathscr{F}^\infty$ heißen terminale Ereignisse.*

3.17 Satz. *(0-1-Gesetz von Kolmogorov) Ist (X_n) eine Folge unabhängiger Zufallsvariablen und $A \in \mathscr{F}^\infty$, so gilt $\mathbb{P}(A) = 0$ oder $\mathbb{P}(A) = 1$.*

4 Konvergenz von Zufallsvariablen und ihren Verteilungen

4.1 Konvergenz von Zufallsvariablen

Im folgenden sei (S,d) stets ein separabler metrischer Raum mit Borel-σ-Algebra $\mathscr{S}$.

4.1 Definition. *Für $(S,\mathscr{S})$-wertige Zufallsvariablen X, $(X_n)_{n\in\mathbb{N}}$ auf dem Wahrscheinlichkeitsraum $(\Omega,\mathscr{F},\mathbb{P})$ bedeutet*

(a) $\lim_{n\to\infty} X_n = X$ *$\underline{\mathbb{P}\text{-fast sicher}}$, dass $\mathbb{P}(\lim_{n\to\infty} X_n = X) = 1$ (Notation: $X_n \to X$ $\mathbb{P}$-f.s.);*

(b) $\lim_{n\to\infty} X_n = X$ *$\underline{\mathbb{P}\text{-stochastisch (oder in } \mathbb{P}\text{-Wahrscheinlichkeit)}}$, dass*

$$\forall\,\varepsilon > 0: \lim_{n\to\infty} \mathbb{P}(d(X_n,X) > \varepsilon) = 0.$$

(Notation: $X_n \xrightarrow{\mathbb{P}} X$).

4.2 Satz. *Es gilt $X_n \xrightarrow{\mathbb{P}} X$ genau dann, wenn jede Teilfolge $(X_{n_k})_{k\geqslant 1}$ eine Teilteilfolge $(X_{n_{k_l}})$ besitzt mit $X_{n_{k_l}} \to X$ $\mathbb{P}$-f.s. für $l\to\infty$.*

4.3 Korollar. *$\mathbb{P}$-fast sichere Konvergenz impliziert $\mathbb{P}$-stochastische Konvergenz.*

4.4 Satz. *Setze für $(S,\mathscr{S})$-wertige Zufallsvariablen X und Y*

$$\alpha(X,Y) := \mathbb{E}[\min(1,d(X,Y))].$$

Dann ist α eine Metrik auf dem Raum L^0 der $(S,\mathscr{S})$-wertigen Zufallsvariablen auf einem festen Wahrscheinlichkeitsraum (mit Identifizierung fast sicher gleicher Zufallsvariablen). Es gilt

$$\alpha(X_n,X) \to 0 \iff X_n \xrightarrow{\mathbb{P}} X,$$

d.h. α metrisiert $\mathbb{P}$-stochastische Konvergenz.

4.5 Satz. *Ist (T,d') ein weiterer separabler metrischer Raum und ist $f: S\to T$ stetig, so gilt:*

(a) $X_n \to X$ *$\mathbb{P}$-f.s.* $\Rightarrow f(X_n) \to f(X)$ *$\mathbb{P}$-f.s.;*

(b) $X_n \xrightarrow{\mathbb{P}} X \Rightarrow f(X_n) \xrightarrow{\mathbb{P}} f(X)$.

4.6 Definition. *Es sei $(\Omega,\mathscr{F},\mathbb{P})$ ein Wahrscheinlichkeitsraum. Man setzt für $p>0$*

$$\mathscr{L}^p := \mathscr{L}^p(\Omega,\mathscr{F},\mathbb{P}) = \{X:\Omega\to\bar{\mathbb{R}} \text{ messbar} \mid \mathbb{E}[|X|^p] < \infty\}$$

sowie $\|X\|_p := \mathbb{E}[|X|^p]^{1/p}$. Der Quotientenraum $L^p = \mathscr{L}^p/\|\bullet\|_p$ entsteht durch Identifizierung $\mathbb{P}$-fast sicher gleicher Zufallsvariablen.

4.7 Satz. *$\mathscr{L}^p$ ist ein Vektorraum. Für $p\geqslant 1$ ist $\|\bullet\|_p$ eine Seminorm auf $\mathscr{L}^p$ und eine Norm auf L^p, bezüglich der L^p vollständig, d.h. ein Banachraum ist. Für $p>q>0$ gilt $\|X\|_p \geqslant \|X\|_q$. Für alle $p>0$ gilt*

$$\|X_n - X\|_p \to 0 \Rightarrow X_n \xrightarrow{\mathbb{P}} X.$$

4.2 Schwache Konvergenz von Wahrscheinlichkeitsmaßen

4.8 Definition. *Wahrscheinlichkeitsmaße $\mathbb{P}_n$ auf $(S, \mathscr{S})$ konvergieren schwach gegen ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(S, \mathscr{S})$, falls*

$$\forall\, f \in C_b(S): \lim_{n\to\infty} \int f \, d\,\mathbb{P}_n = \int f \, d\,\mathbb{P}\,.$$

($C_b(S) := \{f : S \to \mathbb{R} \mid f$ stetig, beschränkt$\}$) In Kurzform: $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$.

$(S, \mathscr{S})$-wertige Zufallsvariablen X_n konvergieren in Verteilung gegen eine $(S, \mathscr{S})$-wertige Zufallsvariable X, falls $\mathbb{P}^{X_n} \xrightarrow{w} \mathbb{P}^X$, d.h.

$$\forall\, f \in C_b(S): \lim_{n\to\infty} \mathbb{E}[f(X_n)] = \mathbb{E}[f(X)].$$

Man schreibt $X_n \xrightarrow{d} X$ bzw. $X_n \xrightarrow{d} \mathbb{P}^X$.

4.9 Satz. *Für reellwertige Zufallsvariablen (X_n), X sind äquivalent:*

(a) $X_n \xrightarrow{d} X$;

(b) die Verteilungsfunktionen erfüllen $F^{X_n}(x) \to F^X(x)$ für alle $x \in \mathbb{R}$, an denen F^X stetig ist (Stetigkeitspunkte).

4.10 Satz. *Aus $X_n \xrightarrow{\mathbb{P}} X$ folgt $X_n \xrightarrow{d} X$.*

4.11 Satz. *(Portmanteau) Für Wahrscheinlichkeitsmaße $(\mathbb{P}_n)_{n\in\mathbb{N}}$, $\mathbb{P}$ auf $(S, \mathscr{S})$ sind äquivalent:*

(a) $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$;

(b) $\forall\, U \subseteq S$ offen : $\liminf_{n\to\infty} \mathbb{P}_n(U) \geqslant \mathbb{P}(U)$;

(c) $\forall\, F \subseteq S$ abgeschlossen : $\limsup_{n\to\infty} \mathbb{P}_n(F) \leqslant \mathbb{P}(F)$;

(d) $\forall\, A \in \mathscr{S}$ mit $\mathbb{P}(\partial A) = 0 : \lim_{n\to\infty} \mathbb{P}_n(A) = \mathbb{P}(A)$.

4.12 Satz. *Ist $g : S \to T$ stetig, T ein separabler metrischer Raum, so gilt: $X_n \xrightarrow{d} X \Rightarrow g(X_n) \xrightarrow{d} g(X)$.*

4.13 Satz. *$\mathbb{P}_n \xrightarrow{w} \mathbb{P}$ gilt bereits, wenn $\int f \, d\,\mathbb{P}_n \to \int f \, d\,\mathbb{P}$ für alle beschränkten, Lipschitz-stetigen Funktionen f gilt.*

4.14 Satz. *(Slutsky) Für Zufallsvektoren (X_n), (Y_n) im $\mathbb{R}^d$ gilt:*

$$X_n \xrightarrow{d} X,\ X_n - Y_n \xrightarrow{\mathbb{P}} 0 \Rightarrow Y_n \xrightarrow{d} X.$$

4.15 Korollar. *Für Zufallsvektoren (X_n) im $\mathbb{R}^d$, (Y_n) im $\mathbb{R}^{d'}$ und $c \in \mathbb{R}^{d'}$ gilt:*

$$X_n \xrightarrow{d} X,\ Y_n \xrightarrow{\mathbb{P}} c \Rightarrow (X_n, Y_n) \xrightarrow{d} (X, c).$$

Insbesondere folgt also $X_n + Y_n \xrightarrow{d} X + c$, $X_n Y_n \xrightarrow{d} cX$.

4.3 Straffheit

4.16 Definition. *Eine Familie* $(\mathbb{P}_i)_{i\in I}$ *von Wahrscheinlichkeitsmaßen auf* $(S, \mathscr{S})$ *heißt* <u>*(schwach) relativ kompakt*</u>*, falls jede Folge* $(\mathbb{P}_{i_k})_{k\geqslant 1}$ *eine schwach konvergente Teilfolge besitzt. Die Familie* $(\mathbb{P}_i)_{i\in I}$ *heißt* <u>*(gleichmäßig) straff*</u>*, falls es zu jedem* $\varepsilon > 0$ *eine kompakte Menge* $K_\varepsilon \subseteq S$ *gibt mit* $\mathbb{P}_i(K_\varepsilon) \geqslant 1 - \varepsilon$ *für alle* $i \in I$*.*

4.17 Satz. *Ist* (S, d) *ein vollständiger separabler metrischer Raum, so ist jede relativ kompakte Familie von Wahrscheinlichkeitsmaßen auf* $(S, \mathscr{S})$ *straff.*

4.18 Korollar. *Ist* $\mathbb{P}$ *ein Wahrscheinlichkeitsmaß auf einem vollständigen separablen metrischen Raum, so gibt es zu jedem* $\varepsilon > 0$ *eine kompakte Menge* K_ε *mit* $\mathbb{P}(K_\varepsilon) \geqslant 1 - \varepsilon$*.*

4.19 Satz. *Ist* (S, d) *ein separabler metrischer Raum, so ist jede straffe Familie von Wahrscheinlichkeitsmaßen auf* $(S, \mathscr{S})$ *relativ kompakt.*

4.20 Korollar. *(Prohorov) Auf einem vollständigen separablen metrischen Raum ist eine Familie von Wahrscheinlichkeitsmaßen genau dann relativ kompakt, wenn sie straff ist.*

4.21 Definition. *Setze für jede beschränkte Lipschitz-Funktion* $f : S \to \mathbb{R}$

$$\|f\|_{Lip} := \|f\|_\infty + \sup_{x\neq y} \frac{|f(x) - f(y)|}{d(x, y)}.$$

Die <u>*duale Lipschitz-Metrik*</u> *für Wahrscheinlichkeitsmaße* $\mathbb{P}$ *und* $\mathbb{Q}$ *ist definiert als*

$$\beta(\mathbb{P}, \mathbb{Q}) := \sup_{\|f\|_{Lip}\leqslant 1} \Big| \int f\, d\,\mathbb{P} - \int f\, d\,\mathbb{Q} \Big|.$$

4.22 Satz. β *ist eine Metrik auf den Wahrscheinlichkeitsmaßen in* $(S, \mathscr{S})$*. Es gilt* $\mathbb{P}_n \xrightarrow{w} \mathbb{P} \iff \beta(\mathbb{P}_n, \mathbb{P}) \to 0$*, d.h.* β *metrisiert schwache Konvergenz.*

5 Charakteristische Funktionen

5.1 Definition und erste Eigenschaften

5.1 Definition. *Die* <u>*charakteristische Funktion*</u> *eines Wahrscheinlichkeitsmaßes* $\mathbb{P}$ *auf* $(\mathbb{R}^d, \mathfrak{B}_{\mathbb{R}^d})$ *ist gegeben durch*

$$\varphi(u) := \varphi_{\mathbb{P}}(u) := \int e^{i\langle x,u\rangle}\, \mathbb{P}(dx), \quad u \in \mathbb{R}^d.$$

Die charakteristische Funktion eines d-dimensionalen Zufallsvektors X ist die charakteristische Funktion seiner Verteilung:

$$\varphi(u) := \varphi_X(u) := \int e^{i\langle x,u\rangle}\, \mathbb{P}^X(dx) = \mathbb{E}\left[e^{i\langle X,u\rangle}\right], \quad u \in \mathbb{R}^d.$$

5.2 Lemma. *Die charakteristische Funktion ist stetig und erfüllt* $|\varphi(u)| \leqslant 1$*,* $\varphi(0) = 1$*.*

5.3 Satz. *Ist X eine reellwertige Zufallsvariable in L^k, $k \in \mathbb{N}$, so ist φ_X k-mal stetig differenzierbar, und es gilt $\varphi^{(r)}(0) = i^r\,\mathbb{E}[X^r]$, $r = 1, \dots, k$.*

5.4 Lemma. *Sind $X_1, \dots, X_n$ unabhängige d-dimensionale Zufallsvektoren, so gilt $\varphi_{X_1+\cdots+X_n} = \prod_{j=1}^n \varphi_{X_j}$.*

5.5 Satz. *Für die eindimensionale Standardnormalverteilung gilt $\varphi_{N(0,1)}(u) = e^{-u^2/2}$, $u \in \mathbb{R}$.*

5.6 Korollar. *Für die Normalverteilung gilt allgemeiner $\varphi_{N(\mu,\sigma^2)}(u) = \exp(i\mu u - \sigma^2 u^2/2)$, $u \in \mathbb{R}$, mit $\mu \in \mathbb{R}$ und $\sigma > 0$ sowie im Mehrdimensionalen $\varphi_{N(0,\mathbb{E}_d)}(u) = \exp(-|u|^2/2)$, $u \in \mathbb{R}^d$.*

5.2 Eindeutigkeit

5.7 Lemma. *Sind X und Y unabhängige d-dimensionale Zufallsvektoren, so hat $X+Y$ die Verteilung $\mathbb{P}^X * \mathbb{P}^Y$, wobei die Faltung allgemein definiert ist über*

$$\mathbb{P}_1 * \mathbb{P}_2(A) := \int_{\mathbb{R}^d} \mathbb{P}_1(A-y)\;\mathbb{P}_2(dy), \quad A \in \mathfrak{B}_{\mathbb{R}^d}.$$

5.8 Lemma. *Für ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\mathbb{R}^d, \mathfrak{B}_{\mathbb{R}^d})$ und $\sigma > 0$ setze $\mathbb{P}^{(\sigma)} := \mathbb{P} * N(0, \sigma^2\,\mathbb{E}_d)$. Dann besitzt $\mathbb{P}^{(\sigma)}$ die Lebesguedichte*

$$f^{(\sigma)}(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \varphi_{\mathbb{P}}(u) \exp(-i\langle x, u\rangle - \sigma^2 |u|^2/2)\, du, \quad x \in \mathbb{R}^d.$$

5.9 Lemma. *Es gilt $\mathbb{P}^{(\sigma)} \xrightarrow{w} \mathbb{P}$ für $\sigma \to 0$.*

5.10 Satz. *(Eindeutigkeitssatz) Sind $\mathbb{P}$ und $\mathbb{Q}$ Wahrscheinlichkeitsmaße mit identischer charakteristischer Funktion $\varphi_{\mathbb{P}}(u) = \varphi_{\mathbb{Q}}(u)$, $u \in \mathbb{R}^d$, so folgt $\mathbb{P} = \mathbb{Q}$.*

5.11 Satz. *(inverse Fouriertransformation) Hat $\mathbb{P}$ die Lebesguedichte f und gilt $\int |\varphi_{\mathbb{P}}(u)|\, du < \infty$, so gilt für Lebesgue-fast alle $x \in \mathbb{R}^d$*

$$f(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \varphi_{\mathbb{P}}(u) e^{-i\langle u, x\rangle}\, du.$$

5.12 Satz. *(Konvergenzsatz) Es sei $(\mathbb{P}_n)_{n\in\mathbb{N}}$ eine straffe Folge von Wahrscheinlichkeitsmaßen auf $(\mathbb{R}^d, \mathfrak{B}_{\mathbb{R}^d})$, so dass $\lim_{n\to\infty} \varphi_{\mathbb{P}_n}(u) =: \varphi(u)$ für alle $u \in \mathbb{R}^d$ existiert. Dann ist φ die charakteristische Funktion eines Wahrscheinlichkeitsmaßes $\mathbb{P}$, und es gilt $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$.*

5.3 Der zentrale Grenzwertsatz

5.13 Satz. *Sind $(X_k)_{k\in\mathbb{N}}$ unabhängige, identisch verteilte Zufallsvektoren im $\mathbb{R}^d$, für die $\mathbb{E}[|X_k|^2]$ endlich ist, so gilt*

$$S_n^* := \frac{1}{\sqrt{n}} \sum_{k=1}^n (X_k - \mu) \xrightarrow{d} \mathbb{P}_\Sigma,$$

wobei $\mu := (\mathbb{E}[X_{k,1}], \dots, \mathbb{E}[X_{k,d}])$ der Mittelwertvektor der (X_k), $\Sigma := (\operatorname{Cov}(X_{k,i}, X_{k,j}))_{1\leqslant i,j\leqslant d}$ die Kovarianzmatrix und $\mathbb{P}_\Sigma$ das Wahrscheinlichkeitsmaß mit charakteristischer Funktion $\varphi_\Sigma(u) = \exp(-\langle \Sigma u, u\rangle)$ ist.

5.14 Korollar. *Sind $(X_k)_{k\in\mathbb{N}}$ reellwertige Zufallsvariablen in L^2, die unabhängig und identisch verteilt sind, so gilt*

$$\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\frac{X_k-\mu}{\sigma}\xrightarrow{d} N(0,1)$$

mit $\mu=\mathbb{E}[X_k]$, falls $\sigma^2:=\operatorname{Var}(X_k)>0$ ist.

5.15 Definition. *Die* <u>*d-dimensionale (allgemeine) Normalverteilung $N(\mu,\Sigma)$*</u> *mit Mittelwertvektor $\mu\in\mathbb{R}^d$ und mit symmetrischer, positiv semi-definiter Kovarianzmatrix $\Sigma\in\mathbb{R}^{d\times d}$ ist definiert als die Verteilung des Zufallsvektors $Y=\Sigma^{1/2}X+\mu$ für standardnormalverteiltes $X\sim N(0,\mathbb{E}_d)$.*

5.16 Lemma. *Es gilt für $Y\sim N(\mu,\Sigma)$:*

(a) $\mathbb{E}[Y_i]=\mu_i$, $\operatorname{Cov}(Y_i,Y_j)=\Sigma_{ij}$, $i,j=1,\dots,d$;

(b) $\varphi_Y(u)=\varphi_{N(\mu,\Sigma)}(u)=\exp(i\langle\mu,u\rangle-\langle\Sigma u,u\rangle)$, $u\in\mathbb{R}^d$;

(c) *Ist Σ strikt positiv definit, so hat Y (also $N(\mu,\Sigma)$) die Lebesguedichte*

$$f_{N(\mu,\Sigma)}(x)=\frac{1}{(2\pi)^{d/2}\sqrt{\det(\Sigma)}}\exp\big(-\langle\Sigma^{-1}(x-\mu),x-\mu\rangle/2\big),\quad x\in\mathbb{R}^d.$$

5.17 Korollar. *(zentraler Grenzwertsatz) Sind $(X_k)_{k\in\mathbb{N}}$ unabhängige, identisch verteilte Zufallsvektoren im $\mathbb{R}^d$, für die $\mathbb{E}[|X_k|^2]$ endlich ist, so gilt*

$$S_n^*:=\frac{1}{\sqrt{n}}\sum_{k=1}^{n}(X_k-\mu)\xrightarrow{d} N(0,\Sigma)$$

mit Mittelwertvektor $\mu:=(\mathbb{E}[X_{k,1}],\dots,\mathbb{E}[X_{k,d}])$ und Kovarianzmatrix $\Sigma:=(\operatorname{Cov}(X_{k,i},X_{k,j}))_{1\leqslant i,j\leqslant d}$. Falls Σ strikt positiv definit ist, gilt auch

$$\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\Sigma^{-1/2}(X_k-\mu)\xrightarrow{d} N(0,\mathbb{E}_d).$$

5.18 Satz. *(Cramér-Wold) Ein Zufallsvektor X ist genau dann normalverteilt, wenn die reellwertigen Zufallsvariablen $\langle X,u\rangle$ für jedes $u\in\mathbb{R}^d$ normalverteilt sind. Für die Parameter gilt in diesem Fall: $X\sim N(\mu,\Sigma)\iff\langle X,u\rangle\sim N(\langle\mu,u\rangle,\langle\Sigma u,u\rangle)$, $u\in\mathbb{R}^d$.*

5.19 Satz. *Es seien X und Y d_1- bzw. d_2-dimensionale Zufallsvektoren, die gemeinsam normalverteilt sind ($\mathbb{P}^{(X,Y)}\sim\mathbb{N}(\mu,\Sigma)$). Dann sind X und Y genau dann unabhängig, wenn sie unkorreliert sind, d.h. $\operatorname{Cov}(X_i,Y_j)=0$ für alle $i=1,\dots,d_1$, $j=1,\dots,d_2$ gilt.*

5.20 Satz. *(Lindeberg) Für jedes $n\in\mathbb{N}$ seien $X_{n,1},\dots,X_{n,k(n)}$ unabhängige, reellwertige Zufallsvariablen mit $\mathbb{E}[X_{n,j}]=0$, $j=1,\dots,k(n)$, und $\sum_{j=1}^{k(n)}\mathbb{E}[X_{n,j}^2]=1$. Ist die* <u>*Lindeberg-Bedingung*</u>

$$\forall\,\varepsilon>0:\ \lim_{n\to\infty}\sum_{j=1}^{k(n)}\mathbb{E}\left[X_{n,j}^2\mathbf{1}_{\{|X_{n,j}|>\varepsilon\}}\right]=0$$

erfüllt, so folgt $S_n:=\sum_{j=1}^{k(n)}X_j\xrightarrow{d} N(0,1)$.

6 Bedingte Erwartungen

6.1 Orthogonalprojektionen

6.1 Satz. *Es sei L ein abgeschlossener Unterraum des Hilbertraums H. Dann gibt es zu jedem $x \in H$ ein eindeutig bestimmtes $y_x \in L$ mit $\|x-y_x\| = \operatorname{dist}_L(x)$.*

6.2 Definition. *Für einen abgeschlossenen Unterraum L des Hilbertraums H ist die Orthogonalprojektion $P_L : H \to L$ gegeben durch $P_L(x) = y_x$ mit y_x aus vorigem Satz.*

6.3 Lemma. *Es gilt:*

(a) $P_L \circ P_L = P_L$ (Projektionseigenschaft);

(b) $\forall\, x \in H :\ (x - P_L x) \in L^\perp$ (Orthogonalität).

6.4 Korollar. *Für jedes $x \in H$ ist $x = P_L x + (x - P_L x)$ die eindeutige Zerlegung als Summe eines Elements aus L und aus $L^\perp$.*

6.5 Korollar. *P_L ist linear und symmetrisch: $\langle P_L x, y\rangle = \langle x, P_L y\rangle$.*

6.2 Konstruktion der bedingten Erwartung

6.6 Satz. *Es seien X eine $(S, \mathcal{S})$-wertige und Y eine reellwertige Zufallsvariable. Dann ist Y genau dann $\sigma(X)$-messbar, wenn es eine $(\mathcal{S}, \mathfrak{B}_\mathbb{R})$-messbare Funktion $\varphi : S \to \mathbb{R}$ gibt mit $Y = \varphi(X)$.*

6.7 Lemma. *Ist $\mathcal{G}$ eine Unter-σ-Algebra von $\mathcal{F}$, so ist $L^2(\Omega, \mathcal{G}, \mathbb{P})$ als abgeschlossener Unterraum in $L^2(\Omega, \mathcal{F}, \mathbb{P})$ eingebettet.*

6.8 Definition. *Es sei X eine $(S, \mathcal{S})$-wertige Zufallsvariable auf $(\Omega, \mathcal{F}, \mathbb{P})$. Dann ist für $Y \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ die bedingte Erwartung von Y gegeben X durch die Orthogonalprojektion von Y auf $L^2(\Omega, \sigma(X), \mathbb{P})$ definiert: $\mathbb{E}[Y \mid X] = P_{L^2(\Omega,\sigma(X),\mathbb{P})} Y$.*

Ist $\mathcal{G}$ allgemein eine Unter-σ-Algebra von $\mathcal{F}$, so ist die bedingte Erwartung von $Y \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ gegeben $\mathcal{G}$ definiert als $\mathbb{E}[Y \mid \mathcal{G}] = P_{L^2(\Omega,\mathcal{G},\mathbb{P})} Y$.

6.9 Lemma. *$\mathbb{E}[Y \mid \mathcal{G}]$ ist als Zufallsvariable in L^2 eindeutig durch folgende Eigenschaften bestimmt:*

(a) $\mathbb{E}[Y \mid \mathcal{G}]$ ist $\mathcal{G}$-messbar (modulo Nullmengen);

(b) $\forall\, G \in \mathcal{G} :\ \mathbb{E}[\mathbb{E}[Y \mid \mathcal{G}]\mathbf{1}_G] = \mathbb{E}[Y\mathbf{1}_G]$.

6.10 Satz. *Es sei $Y \in \mathcal{M}^+(\Omega, \mathcal{F})$ oder $Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ und $\mathcal{G}$ eine Unter-σ-Algebra von $\mathcal{F}$. Dann gibt es bis auf $\mathbb{P}$-fast sichere Gleichheit genau ein Element $\mathbb{E}[Y \mid \mathcal{G}]$ in $\mathcal{M}^+(\Omega, \mathcal{G})$ bzw. $L^1(\Omega, \mathcal{G}, \mathbb{P})$ mit*

$$\forall\, G \in \mathcal{G} :\ \mathbb{E}[\mathbb{E}[Y \mid \mathcal{G}]\mathbf{1}_G] = \mathbb{E}[Y\mathbf{1}_G].$$

6.11 Definition. *Für $Y \in \mathcal{M}^+(\Omega, \mathcal{F})$ oder $Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ und eine Unter-σ-Algebra $\mathcal{G}$ von $\mathcal{F}$ wird die allgemeine bedingte Erwartung von Y gegeben $\mathcal{G}$ definiert wie $\mathbb{E}[Y \mid \mathcal{G}]$ im vorigen Satz. Man setzt $\mathbb{E}[Y \mid (X_i)_{i\in I}] := \mathbb{E}[Y \mid \sigma(X_i,\ i \in I)]$ für Zufallsvariablen X_i, $i \in I$.*

6.3 Eigenschaften der bedingten Erwartung

6.12 Satz. *Es sei $Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ und $\mathcal{G}$ eine Unter-σ-Algebra von $\mathcal{F}$. Dann gilt (jeweils $\mathbb{P}$-f.s.):*

(a) $\mathbb{E}[\mathbb{E}[Y \mid \mathcal{G}]] = \mathbb{E}[Y]$;

(b) Y *$\mathcal{G}$-messbar* $\Rightarrow \mathbb{E}[Y \mid \mathcal{G}] = Y$;

(c) $\alpha \in \mathbb{R}$, $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P})$: $\mathbb{E}[\alpha Y + Z \mid \mathcal{G}] = \alpha\, \mathbb{E}[Y \mid \mathcal{G}] + \mathbb{E}[Z \mid \mathcal{G}]$;

(d) $Y \geqslant 0 \Rightarrow \mathbb{E}[Y \mid \mathcal{G}] \geqslant 0$;

(e) $0 \leqslant Y_n \uparrow Y \Rightarrow \mathbb{E}[Y_n \mid \mathcal{G}] \uparrow \mathbb{E}[Y \mid \mathcal{G}]$ *(monotone Konvergenz);*

(f) $Y_n \geqslant 0 \Rightarrow \mathbb{E}[\liminf_n Y_n \mid \mathcal{G}] \leqslant \liminf_n \mathbb{E}[Y_n \mid \mathcal{G}]$ *(Fatou);*

(g) $Y_n \to Y$, $|Y_n| \leqslant Z$ *mit* $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P})$: $\mathbb{E}[Y_n \mid \mathcal{G}] \to \mathbb{E}[Y \mid \mathcal{G}]$ *(dom. Konv.);*

(h) $\mathcal{H} \subseteq \mathcal{G} \Rightarrow \mathbb{E}[\mathbb{E}[Y \mid \mathcal{G}] \mid \mathcal{H}] = \mathbb{E}[Y \mid \mathcal{H}]$ *(Projektionseigenschaft);*

(i) Z *$\mathcal{G}$-messbar, beschränkt:* $\mathbb{E}[ZY \mid \mathcal{G}] = Z\, \mathbb{E}[Y \mid \mathcal{G}]$;

(j) Y *unabhängig von* $\mathcal{G}$: $\mathbb{E}[Y \mid \mathcal{G}] = \mathbb{E}[Y]$.

6.13 Satz. *(Jensensche Ungleichung) Ist $\varphi : \mathbb{R} \to \mathbb{R}$ konvex und sind Y, $\varphi(Y)$ in L^1, so gilt $\varphi(\mathbb{E}[Y \mid \mathcal{G}]) \leqslant \mathbb{E}[\varphi(Y) \mid \mathcal{G}]$ für jede Unter-σ-Algebra $\mathcal{G}$ von $\mathcal{F}$. Insbesondere gilt für die L^p-Normen $\|X\|_p \leqslant \|X\|_{p'}$ wenn $p \leqslant p'$.*

7 Martingaltheorie

7.1 Martingale, Sub- und Supermartingale

7.1 Definition. *Eine Folge $\mathcal{F}_n$ von Unter-σ-Algebren von $\mathcal{F}$ heißt Filtration, falls $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$, $n \geqslant 0$, gilt. $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n))$ heißt filtrierter Wahrscheinlichkeitsraum.*

7.2 Definition. *Eine Folge von Zufallsvariablen (ein Prozess) $(M_n)_{n \geqslant 0}$ auf einem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n))$ heißt Martingal (Submartingal, Supermartingal), falls:*

(a) $M_n \in L^1$, $n \geqslant 0$;

(b) M_n *ist* $\mathcal{F}_n$*-messbar,* $n \geqslant 0$ *(adaptiert);*

(c) $\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] = M_n$ *(bzw.* $\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] \geqslant M_n$ *für Submartingal bzw.* $\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] \leqslant M_n$ *für Supermartingal).*

Gilt $\mathcal{F}_n = \sigma(M_0, \ldots, M_n)$ so wird $(\mathcal{F}_n)$ natürliche Filtration von M genannt, Notation $(\mathcal{F}_n^M)$.

7.3 Definition. *Ein Martingal (M_n) heißt abschließbar, falls es ein $X \in L^1$ gibt mit $M_n = \mathbb{E}[X \mid \mathcal{F}_n]$, $n \geqslant 0$.*

7.4 Definition. *Ein Prozess* $(X_n)_{n\geqslant 1}$ *heißt (bezüglich* $(\mathcal{F}_n)$*)* *vorhersehbar, falls jedes* X_n $\mathcal{F}_{n-1}$*-messbar ist,* $n \geqslant 1$. *Für einen vorhersehbaren Prozess* (X_n) *und ein Martingal* (M_n) *wird das diskrete stochastische Integral (oder die Martingaltransformation)* $((X \bullet M)_n)_{n\geqslant 0}$ *definiert als* $(X \bullet M)_0 := 0$, $(X \bullet M)_n := \sum_{k=1}^{n} X_k (M_k - M_{k-1})$.

7.5 Lemma. *Ist* (X_n) *zusätzlich beschränkt, so ist* $((X \bullet M)_n)_{n\geqslant 0}$ *wiederum ein Martingal.*

7.6 Lemma. *Es seien* (M_n) *ein Martingal und* $\varphi : \mathbb{R} \to \mathbb{R}$ *konvex mit* $\varphi(M_n) \in L^1$, $n \geqslant 0$. *Dann ist* $\varphi(M_n)$ *ein Submartingal. Insbesondere ist also* (M_n^2) *ein Submartingal für ein Martingal* (M_n) *in* L^2.

7.7 Satz. *(Doob-Zerlegung) Zu jedem Submartingal* (X_n) *existiert ein Martingal* (M_n) *und ein vorhersehbarer wachsender Prozess* (A_n)*, so dass*

$$X_n = X_0 + M_n + A_n, \quad n \geqslant 1; \qquad M_0 = A_0 = 0.$$

Diese Zerlegung ist $\mathbb{P}$*-fast sicher eindeutig und* $A_n = \sum_{k=1}^{n} \mathbb{E}[X_k - X_{k-1} \,|\, \mathcal{F}_{k-1}]$.

7.8 Definition. *Der vorhersehbare Prozess* (A_n) *in der Doob-Zerlegung von* (X_n) *heißt Kompensator von* (X_n). *Für ein Martingal* (M_n) *in* L^2 *heißt der Kompensator von* (M_n^2) *quadratische Variation von* (M_n)*, Notation:* $\langle M \rangle_n$.

7.9 Lemma. *Es gilt* $\langle M \rangle_n = \sum_{k=1}^{n} \mathbb{E}[(M_k - M_{k-1})^2 \,|\, \mathcal{F}_{k-1}]$, $n \geqslant 1$.

7.2 Stoppzeiten

7.10 Definition. *Eine Abbildung* $\tau : \Omega \to \{0, 1, \ldots, +\infty\}$ *heißt Stoppzeit, falls* $\{\tau = n\} \in \mathcal{F}_n$ *für alle* $n \geqslant 0$ *gilt.*

7.11 Lemma. *Jede deterministische Zeit* $\tau = n_0$ *ist Stoppzeit. Sind* σ *und* τ *Stoppzeiten, so auch* $\sigma \wedge \tau$, $\sigma \vee \tau$ *und* $\sigma + \tau$.

7.12 Satz. *(Optional Stopping) Ist* (M_n) *ein (Sub-)Martingal und* τ *eine Stoppzeit, so ist der gestoppte Prozess* $(M_n^\tau) = (M_{n\wedge\tau})$ *wiederum ein (Sub-)Martingal.*

7.13 Definition. *Für jede Stoppzeit* τ *ist die* σ*-Algebra der* τ*-Vergangenheit definiert als* $\mathcal{F}_\tau := \{A \in \mathcal{F} \,|\, , \forall\, n \geqslant 0 : \; A \cap \{\tau \leqslant n\} \in \mathcal{F}_n\}$.

7.14 Lemma. $\mathcal{F}_\tau$ *ist in der Tat eine* σ*-Algebra und* τ *ist* $\mathcal{F}_\tau$*-messbar.*

7.15 Lemma. *Für Stoppzeiten* σ *und* τ *mit* $\sigma \leqslant \tau$ *gilt* $\mathcal{F}_\sigma \subseteq \mathcal{F}_\tau$.

7.16 Lemma. *Ist* (X_n) *ein adaptierter Prozess und* τ *eine endliche Stoppzeit, so ist* X_τ $\mathcal{F}_\tau$*-messbar.*

7.17 Satz. *(Optional Sampling) Es seien* (M_n) *ein Martingal (Submartingal) und* σ, τ *beschränkte Stoppzeiten mit* $\sigma \leqslant \tau$. *Dann gilt* $\mathbb{E}[M_\tau \,|\, \mathcal{F}_\sigma] = M_\sigma$ *(bzw.* $\mathbb{E}[M_\tau \,|\, \mathcal{F}_\sigma] \geqslant M_\sigma$*).*

7.18 Korollar. *Es sei (M_n) ein Martingal und τ eine endliche Stoppzeit. Dann gilt $\mathbb{E}[M_\tau] = \mathbb{E}[M_0]$ unter einer der folgenden Bedingungen:*

(a) τ ist beschränkt;

(b) $(M_n)_{n \geqslant 0}$ ist gleichmäßig beschränkt;

(c) $\mathbb{E}[\tau] < \infty$ und $\mathbb{E}[|M_{n+1} - M_n| \,|\, \mathcal{F}_n]$, $n \geqslant 0$, ist gleichmäßig beschränkt.

7.19 Korollar. *(Waldsche Identität) Es seien $(X_k)_{k \geqslant 1}$ adaptierte, identisch verteilte Zufallsvariablen mit $\mathbb{E}[X_k] = \mu \in \mathbb{R}$, so dass X_k jeweils unabhängig von $\mathcal{F}_{k-1}$ ist. Dann gilt für die Partialsummen $S_n := \sum_{k=1}^n X_k$, $S_0 = 0$ und jede Stoppzeit τ mit $\mathbb{E}[\tau] < \infty$: $\mathbb{E}[S_\tau] = \mu\, \mathbb{E}[\tau]$.*

7.3 Martingalungleichungen

7.20 Satz. *(Maximalungleichung) Für ein Martingal (M_n) und $\alpha > 0$ gilt*

$$\mathbb{P}\Big(\sup_{0 \leqslant k \leqslant n} |M_k| \geqslant \alpha\Big) \leqslant \tfrac{1}{\alpha}\, \mathbb{E}[|M_n|], \quad n \geqslant 0.$$

7.21 Definition. *Die Anzahl der aufsteigenden Überquerungen ("upcrossings") eines Intervalls $[a, b]$ durch einen Prozess (X_k) bis zur Zeit n ist definiert als $U_n^{[a,b]} := \sup\{k \geqslant 1 \,|\, \tau_k \leqslant n\}$, wobei induktiv $\tau_0 := 0$, $\sigma_{k+1} := \inf\{\ell \geqslant \tau_k \,|\, X_\ell \leqslant a\}$, $\tau_{k+1} := \inf\{\ell \geqslant \sigma_k \,|\, X_\ell \geqslant b\}$.*

7.22 Satz. *(Upcrossing Inequality) Für die aufsteigenden Überquerungen eines Submartingals (X_n) gilt $\mathbb{E}[U_n^{[a,b]}] \leqslant \frac{1}{b-a}\, \mathbb{E}[(M_n - a) \vee 0]$.*

7.4 Martingalkonvergenzsätze

7.23 Satz. *(1. Martingalkonvergenzsatz) Es sei (M_n) ein Submartingal mit $\sup_n \mathbb{E}[|M_n|] < \infty$. Dann existiert $M_\infty := \lim_{n\to\infty} M_n$ $\mathbb{P}$-fast sicher, und es gilt $M_\infty \in L^1$.*

7.24 Korollar. *Jedes nicht-negative Supermartingal konvergiert $\mathbb{P}$-fast sicher.*

7.25 Satz. *Es sei (M_n) ein Martingal in L^2. Dann existiert $\lim_{n\to\infty} M_n(\omega)$ für $\mathbb{P}$-fast alle ω, für die $\lim_{n\to\infty} \langle M \rangle_n(\omega) < \infty$ gilt.*

7.26 Korollar. *(Starkes Gesetz der großen Zahlen für Martingale) Ist (M_n) ein Martingal in L^2, so gilt für jedes $\alpha > 1/2$*

$$\lim_{n\to\infty} \frac{M_n(\omega)}{(\langle M \rangle_n(\omega))^\alpha} = 0$$

für $\mathbb{P}$-fast alle ω, für die $\lim_{n\to\infty} \langle M \rangle_n(\omega) = \infty$ gilt.

7.27 Definition. *Eine Familie $(X_i)_{i \in I}$ von Zufallsvariablen in L^1 heißt gleichgradig integrierbar, falls*

$$\lim_{R\to\infty} \sup_{i \in I} \mathbb{E}[|X_i| \mathbf{1}_{\{|X_i| > R\}}] = 0.$$

7.28 Satz. *(2. Martingalkonvergenzsatz) Ist (M_n) ein gleichgradig integrierbares Martingal, so konvergiert (M_n) $\mathbb{P}$-fast sicher und in L^1 gegen ein $M_\infty \in L^1$. (M_n) ist abschließbar mit $M_n = \mathbb{E}[M_\infty \,|\, \mathcal{F}_n]$.*
Andererseits ist jedes abschließbare Martingal gleichgradig integrierbar.

7.29 Korollar. *Jedes Martingal, das gleichmäßig in L^p für ein $p > 1$ beschränkt ist, konvergiert $\mathbb{P}$-fast sicher und in L^1 (ohne Beweis: sogar in L^p).*

7.30 Definition. *Ein Prozess $(M_{-n})_{n\geqslant 0}$ heißt Rückwärtsmartingal bezüglich $(\mathcal{F}_{-n})_{n\geqslant 0}$ mit $\mathcal{F}_{-n-1} \subseteq \mathcal{F}_{-n}$, falls $M_{-n} \in L^1$, M_{-n} $\mathcal{F}_{-n}$-messbar und $\mathbb{E}[M_{-n} \,|\, \mathcal{F}_{-n-1}] = M_{-n-1}$ für alle $n \geqslant 0$ gilt.*

7.31 Satz. *Jedes Rückwärtsmartingal $(M_{-n})_{n\geqslant 0}$ konvergiert für $n \to \infty$ $\mathbb{P}$-fast sicher und in L^1.*

7.32 Korollar. *(Starkes Gesetz der großen Zahlen, Kolmogorov) Für unabhängige und identisch verteilte Zufallsvariablen $(X_k)_{k\geqslant 1}$ in L^1 gilt*

$$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{\mathbb{P}\text{-f.s.}} \mathbb{E}[X_1].$$

8 Der Satz von Radon-Nikodym

8.1 Definition. *Es seien μ und ν Maße auf dem Messraum $(\Omega, \mathcal{F})$. Dann heißt μ absolutstetig bezüglich ν, in Zeichen $\mu \ll \nu$, falls $\forall A \in \mathcal{F} : \nu(A) = 0 \Rightarrow \mu(A) = 0$ gilt. μ und ν heißen äquivalent, in Zeichen $\mu \sim \nu$, falls $\mu \ll \nu$ und $\nu \ll \mu$ gilt. Gibt es ein $A \in \mathcal{F}$ mit $\nu(A) = 0$ und $\mu(A^C) = 0$, so heißen μ und ν singulär, in Zeichen $\mu \perp \nu$.*

8.2 Lemma. *Für ein Wahrscheinlichkeitsmaß $\mathbb{P}$ und ein allgemeines Maß μ mit $\mu \ll \mathbb{P}$ gilt*

$$\forall \varepsilon > 0 \,\exists \delta > 0 \,\forall A \in \mathcal{F} :\ \mathbb{P}(A) < \delta \Rightarrow \mu(A) < \varepsilon.$$

8.3 Satz. *(Radon-Nikodym) Ist ν ein σ-endliches Maß und $\mu \ll \nu$ endlich, so existiert ein (ν-fast überall eindeutig bestimmtes) $f \in L^1(\Omega, \mathcal{F}, \nu)$ mit*

$$\mu(A) = \int_A f \, d\nu \text{ für alle } A \in \mathcal{F}.$$

f heißt Radon-Nikodym-Ableitung oder Dichte oder in der Statistik auch Likelihood von μ bezüglich ν, in Zeichen $f = \frac{d\mu}{d\nu}$.