Ruprecht-Karls-Universität Heidelberg
Institut für Angewandte Mathematik
Professor Dr. Jan JOHANNES

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

Skript zur Vorlesung

WAHRSCHEINLICHKEITSTHEORIE 1

Sommersemester 2021

Fassung Stand 9. Juli 2021

Falls Sie **Fehler im Skript** finden, teilen Sie mir diese bitte
per eMail an johannes@math.uni-heidelberg.de mit.

MATHEMATIKON, Im Neuenheimer Feld 205, 69120 Heidelberg
Telefon: +49 6221 54.14.190 – Fax: +49 6221 54.14.101
eMail: johannes@math.uni-heidelberg.de
Webseite: sip.math.uni-heidelberg.de

Inhaltsverzeichnis

Kapitel 1

Maß- und Integrationstheorie

§01 Maßtheorie

§01|01 Mengensysteme

§01.01 **Definition.** Ein Mengensystem $\mathscr{E} \subseteq 2^\Omega$ heißt

$\cap$-stabil (sprich: *schnittstabil*), falls für je zwei Mengen $A, B \in \mathscr{E}$ gilt, dass auch $A \cap B \in \mathscr{E}$;

σ-$\cap$-stabil (sprich: *sigma-schnittstabil*), falls für jede Folge $(A_n)_{n\in\mathbb{N}}$ von Teilmengen aus $\mathscr{E}$ gilt, dass auch $\cap_{n\in\mathbb{N}} A_n \in \mathscr{E}$;

$\cup$-stabil (sprich: *vereinigungsstabil*), falls für je zwei Mengen $A, B \in \mathscr{E}$ gilt, dass auch $A \cup B \in \mathscr{E}$;

σ-$\cup$-stabil (sprich: *sigma-vereinigungsstabil*), falls für jede Folge $(A_n)_{n\in\mathbb{N}}$ von Teilmengen aus $\mathscr{E}$ gilt, dass auch $\cup_{n\in\mathbb{N}} A_n \in \mathscr{E}$;

$\setminus$-stabil (sprich: *differenzmengenstabil*), falls je zwei Mengen $A, B \in \mathscr{E}$ gilt, dass auch $A \setminus B \in \mathscr{E}$;

komplementstabil, falls mit jeder Menge $A \in \mathscr{E}$ auch $A^c \in \mathscr{E}$ gilt. □

§01.02 **Bemerkung.**

(i) Für ein komplementstabiles Mengensystem $\mathscr{E} \subseteq 2^\Omega$ folgen aus den de Morgan'schen Regeln die Äquivalenzen von $\cup$-stabil und $\cap$-stabil als auch von σ-$\cup$-stabil und σ-$\cap$-stabil.

(ii) Sei $\mathscr{E} \subseteq 2^\Omega$ $\setminus$-stabil. Dann ist $\mathscr{E}$ auch $\cap$-stabil. Falls $\mathscr{E}$ σ-$\cup$-stabil ist, dann ist $\mathscr{E}$ auch σ-$\cap$-stabil. Jede abzählbare Vereinigung von Mengen aus $\mathscr{E}$ lässt sich als abzählbare, disjunkte Vereinigung von Mengen in $\mathscr{E}$ schreiben. □

§01.03 **Definition.** Ein Mengensystem $\mathscr{E} \subseteq 2^\Omega$ heißt

Semiring, falls (i) $\emptyset \in \mathscr{E}$, (ii) für je zwei Mengen $A, B \in \mathscr{E}$ ist $A \setminus B$ endliche Vereinigung von paarweise disjunkten Mengen aus $\mathscr{E}$, und (iii) $\mathscr{E}$ ist $\cap$-stabil;

Ring, falls (R1) $\emptyset \in \mathscr{E}$, (R2) $\mathscr{E}$ ist $\setminus$-stabil, und (R3) $\mathscr{E}$ ist $\cup$-stabil;
σ-Ring, falls $\mathscr{E}$ ein σ-$\cup$-stabiler Ring ist;

Algebra, falls (A1) $\Omega \in \mathscr{E}$, (A2) $\mathscr{E}$ ist $\setminus$-stabil, und (A3) $\mathscr{E}$ ist $\cup$-stabil;
σ-Algebra, falls $\mathscr{E}$ eine σ-$\cup$-stabile Algebra ist;

Dynkin-System, falls (D1) $\Omega \in \mathscr{E}$; (D2) $\mathscr{E}$ ist komplementstabil; (D3) für paarweise disjunkte Mengen $(A_n)_{n\in\mathbb{N}}$ aus $\mathscr{E}$ gilt $\biguplus_{n\in\mathbb{N}} A_n \in \mathscr{E}$. □

§01.04 **Bemerkung.**

(i) Für $\Omega \neq \emptyset$ sind $\{\emptyset, \Omega\}$ und 2^Ω triviale Beispiele für Algebren und Dynkin-Systeme. Triviale Beispiele für Semiringe und Ringe sind $\{\emptyset\}$ und 2^Ω.

(ii) Ein (Mengen-)Ring $\mathscr{R}$ bildet mit der symmetrischen Differenz Δ als Addition und dem Durchschnitt $\cap$ als Multiplikation einen Ring $(\mathscr{R}, \Delta, \cap)$ im Sinne der Algebra.

(iii) Ein Mengensystem $\mathscr{A} \subseteq 2^\Omega$ ist eine Algebra genau dann wenn $\mathscr{A}$ komplementstabil und $\cap$-stabil mit $\Omega \in \mathscr{A}$.

(iv) Ein Mengensystem $\mathscr{A} \subseteq 2^\Omega$ mit $\Omega \in \mathscr{A}$, das komplementstabil und σ-$\cup$-stabil ist, ist somit eine σ-Algebra.

(v) Sei $\mathscr{D} \subseteq 2^\Omega$ ein Dynkinsystem. Die Bedingung (D2), das $\mathscr{D}$ komplementstabil ist; kann äquivalent ersetzt werden durch die scheinbar stärkere Bedingung (D2') für alle $A, B \in \mathscr{D}$ mit $A \subseteq B$ gilt $B \setminus A \in \mathscr{D}$. Da jedes Dynkin-System auch (D2') erfüllt. Denn für $A, B \in \mathscr{D}$ mit $A \subseteq B$ sind A und B^c disjunkt und es gilt $B \setminus A = (A \uplus B^c)^c \in \mathscr{D}$.

(vi) Jede σ-Algebra ist ein Dynkin-System. Die Umkehrung gilt nicht, da (D3) nur für Folgen paarweise disjunkter Ereignisse gefordert ist. Zum Beispiel für $\Omega = \{1, 2, 3, 4\}$ ist $\mathscr{D} = \{\emptyset, \{1, 2\}, \{1, 4\}, \{2, 3\}, \{3, 4\}, \Omega\}$ ein Dynkin-System aber keine σ-Algebra. Der Unterschied ist allerdings nicht sehr groß. □

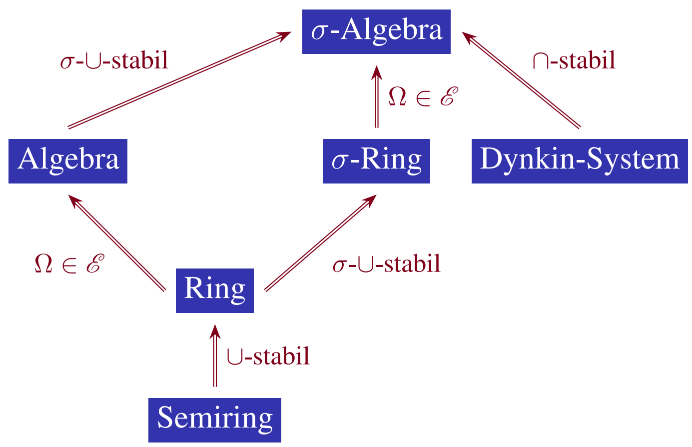
§01.05 **Skizze.**

(i) Jede σ-Algebra ist ein Dynkin-System, eine Algebra und ein σ-Ring.

(ii) Jeder σ-Ring ist auch ein Ring, jeder Ring ein Semiring.

(iii) Jede Algebra ist auch ein Ring. Eine Algebra auf einer endlichen Menge Ω ist auch eine σ-Algebra.

Zusammenhang zwischen den Mengensystemen $\mathscr{E} \subseteq 2^\Omega$.



Die Abbildung wurde auf der Grundlage von Klenke [2020, Abb.1.1, S.7] erstellt. □

§01.06 **Lemma**. *Ein Dynkin-System $\mathscr{D} \subseteq 2^\Omega$ ist genau dann $\cap$-stabil, wenn es eine σ-Algebra ist.*

§01.07 **Beweis** von Lemma §01.06. In der Vorlesung EWS. □

§01.08 **Lemma**. *Es sei $\mathscr{E} \subseteq 2^\Omega$ ein System von Teilmengen von Ω. Dann ist*

$$\sigma(\mathscr{E}) := \bigcap \{\mathscr{A} \mid \mathscr{A} \text{ ist } \sigma\text{-Algebra auf } \Omega \text{ und } \mathscr{E} \subseteq \mathscr{A}\} \qquad \text{und}$$

$$\delta(\mathscr{E}) := \bigcap \{\mathscr{D} \mid \mathscr{D} \text{ ist Dynkin-System auf } \Omega \text{ und } \mathscr{E} \subseteq \mathscr{D}\}$$

die kleinste σ-Algebra bzw. das kleinste Dynkin-System auf Ω mit $\mathscr{E} \subseteq \sigma(\mathscr{E})$ bzw. $\mathscr{E} \subseteq \delta(\mathscr{E})$. $\mathscr{E}$ heißt Erzeuger und $\sigma(\mathscr{E})$ die von $\mathscr{E}$ erzeugte σ-Algebra bzw. $\delta(\mathscr{E})$ das von $\mathscr{E}$ erzeugte Dynkin-System auf Ω.

§01.09 **Beweis** von Lemma §01.08. In der Vorlesung EWS. □

§01.10 **π-λ-Satz**. *Sei $\mathscr{E}$ $\cap$-stabil. Dann gilt $\sigma(\mathscr{E}) = \delta(\mathscr{E})$ und für jedes Dynkin-System $\mathscr{D}$ mit $\mathscr{E} \subseteq \mathscr{D}$ also $\sigma(\mathscr{E}) \subseteq \mathscr{D}$.*

§01.11 **Beweis** von Satz §01.10. In der Vorlesung EWS. □

§01.12 **Definition**. Sei $\mathscr{E} \subseteq 2^\Omega$ ein beliebiges System von Teilmengen und $A \subseteq \Omega$ eine nicht-leere Teilmenge. Das Mengensystem $\mathscr{E}\big|_A := \mathscr{E} \cap A := \{A \cap B \,|\, B \in \mathscr{E}\}$ heißt *Spur* oder *Einschränkung* von $\mathscr{E}$ auf A.

§01.13 **Bemerkung**. Ist $\mathscr{E}$ ein Semiring, (σ-)Ring oder (σ-)Algebra so ist $\mathscr{E}\big|_A$ ein Mengensystem auf A vom gleichen Typ. Für ein Dynkin-System gilt dies im Allgemeinen nicht. Desweiteren gilt $\sigma(\mathscr{E})\big|_A = \sigma(\mathscr{E}\big|_A)$. □

§01|02 Mengenfunktionen

§01.14 **Definition**. Sei $\mathscr{E} \subseteq 2^\Omega$. Eine Mengenfunktion $\mu : \mathscr{E} \to \overline{\mathbb{R}}^+ = [0, \infty]$ heißt

monoton, falls für je zwei Mengen $A, B \in \mathscr{E}$ mit $A \subseteq B$ gilt, dass $\mu(A) \leqslant \mu(B)$;

additiv, falls für je endlich viele paarweise disjunkte Mengen $(A_j)_{j \in [\![n]\!]}$ aus $\mathscr{E}$ mit $\biguplus_{j \in [\![n]\!]} A_j \in \mathscr{E}$

gilt, dass $\mu(\biguplus_{j \in [\![n]\!]} A_j) = \sum_{j \in [\![n]\!]} \mu(A_j)$;

σ-additiv, falls für je abzählbar viele paarweise disjunkte Mengen $(A_j)_{j \in \mathbb{N}}$ aus $\mathscr{E}$ mit $\biguplus_{j \in \mathbb{N}} A_j \in \mathscr{E}$

gilt, dass $\mu(\biguplus_{j \in \mathbb{N}} A_j) = \sum_{j \in \mathbb{N}} \mu(A_j)$;

subadditiv, falls für je endlich viele Mengen A und $(A_j)_{j \in [\![n]\!]}$ aus $\mathscr{E}$ mit $A \subseteq \bigcup_{j \in [\![n]\!]} A_j$ gilt, dass

$\mu(A) \leqslant \sum_{j \in [\![n]\!]} \mu(A_j)$;

σ-subadditiv, falls für je abzählbar viele Mengen A und $(A_j)_{j \in \mathbb{N}}$ aus $\mathscr{E}$ mit $A \subseteq \bigcup_{j \in \mathbb{N}} A_j$ gilt, dass

$\mu(A) \leqslant \sum_{j \in \mathbb{N}} \mu(A_j)$. □

§01.15 **Definition**. Sei $\mathscr{E} \subseteq 2^\Omega$ ein Semiring. Eine Mengefunktion $\mu : \mathscr{E} \to \overline{\mathbb{R}}^+$ mit $\mu(\emptyset) = 0$ heißt

Inhalt, falls μ additiv ist;

Prämaß, falls μ σ-additiv ist;

Maß, falls μ ein Prämaß ist und $\mathscr{E}$ eine σ-Algebra;

Wahrscheinlichkeitsmaß, falls μ ein Maß ist und $\mu(\Omega) = 1$.

Wir bezeichnen mit $\mathcal{M}(\mathscr{E})$ die Menge aller Prämaße auf $(\Omega, \mathscr{E})$. Ein Inhalt μ auf $\mathscr{E}$ heißt

endlich, falls $\mu(A) \in \mathbb{R}^+$ für jedes $A \in \mathscr{E}$;

σ-endlich, falls es Mengen $(\Omega_j)_{j \in \mathbb{N}}$ aus $\mathscr{E}$ gibt mit $\Omega = \bigcup_{j \in \mathbb{N}} \Omega_j$ und $\mu(\Omega_j) \in \mathbb{R}^+$ für jedes $j \in \mathbb{N}$.

Wir bezeichnen mit $\mathcal{M}_e(\mathscr{E})$ und $\mathcal{M}_\sigma(\mathscr{E})$ die Menge aller endlicher bzw. σ-endlicher Prämaße auf $(\Omega, \mathscr{E})$. Weiterhin, für eine σ-Algebra $\mathscr{A} \in 2^\Omega$ bezeichnet $\mathcal{W}(\mathscr{A})$ die Menge der Wahrscheinlichkeitsmaße auf $(\Omega, \mathscr{A})$.

§01.16 **Beispiel**.

(a) Für $A \in 2^\Omega$ bezeichnet $\mathbb{1}_A : \Omega \to \{0, 1\}$ mit $\mathbb{1}_A^{-1}(\{1\}) = A$ und $\mathbb{1}_A^{-1}(\{0\}) = A^c$ die *Indikatorfunktion* auf A. Für jede σ-Algebra $\mathscr{A} \subseteq 2^\Omega$ und jedes $\omega \in \Omega$ ist das *Einpunkt-*

oder *Diracmaß* $\boxed{\delta_\omega}: \mathscr{A} \to \{0,1\}$ mit $\delta_\omega(A) := \mathbb{1}_A(\omega)$ ein Wahrscheinlichkeitsmaß, also $\delta_\omega \in \mathcal{M}_1(\mathscr{A})$.

(b) Sei $\Omega \neq \emptyset$ abzählbar unendlich und $\mathscr{E} := \{A \in 2^\Omega : (|A| \wedge |A^c|) \in \mathbb{R}^+\}$, dann ist $\mathscr{E}$ eine Algebra. Sei $\mu : \mathscr{E} \to \{0, \infty\}$ für $A \in \mathscr{E}$ gegeben durch $\mu(A) = 0$ für $|A| \in \mathbb{R}^+$ und $\mu(A) = \infty$ für $|A^c| \in \mathbb{R}^+$. Dann ist μ ein Inhalt, das nicht σ-additiv ist, da $\mu(\Omega) = \infty$ aber $\sum_{\omega\in\Omega} \mu(\{\omega\}) = 0$. μ ist also kein Prämaß ist.

(c) Sei $\Omega \neq \emptyset$ abzählbar und $\mathbb{p} : \Omega \to \mathbb{R}^+$. Dann ist $\mu : 2^\Omega \to \overline{\mathbb{R}}^+$ mit $A \mapsto \mu(A) := \sum_{\omega\in\Omega} \mathbb{p}(\omega)\delta_\omega(A)$ ein σ-endliches Maß auf 2^Ω, also $\mu \in \mathcal{M}_\sigma(2^\Omega)$. Wir bezeichen $\mathbb{p}$ als *Zähldichte* von μ. Zur Erinnerung, erfüllt $\mathbb{p}$ zusätzlich $\sum_{\omega\in\Omega} \mathbb{p}(\omega) = 1$ dann ist $\mu \in \mathcal{W}(2^\Omega) = \mathcal{M}_1(2^\Omega)$ ein diskretes Wahrscheinlichkeitsmaß. Im Spezialfall $\mathbb{p}(\omega) = 1$ für jedes $\omega \in \Omega$ heißt $\boxed{\zeta_\Omega} := \mu$ das *Zählmaß* auf Ω.

§01.17 **Lemma**. *Sei $\mathscr{E}$ ein Semiring und μ ein Inhalt auf $\mathscr{E}$.*

(i) *Ist $\mathscr{E}$ ein Ring, so gilt $\mu(A \cup B) = \mu(A) + \mu(B \setminus A)$ und $\mu(B) = \mu(A \cap B) + \mu(B \setminus A)$, also $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$ für alle $A, B \in \mathscr{E}$.*

(ii) *μ ist monoton. Ist $\mathscr{E}$ ein Ring, so gilt $\mu(B) = \mu(A) + \mu(B \setminus A)$ für $A, B \in \mathscr{E}$ mit $A \subseteq B$.*

(iii) *μ ist subadditiv. Ist μ sogar σ-additiv, so ist μ auch σ-subadditiv.*

(iv) *Ist $\mathscr{E}$ ein Ring, so gilt für paarweise disjunkte Mengen $(A_j)_{j\in\mathbb{N}}$ aus $\mathscr{E}$ mit $\biguplus_{j\in\mathbb{N}} A_j \in \mathscr{E}$ stets*
$$\sum_{j\in[\![n]\!]} \mu(A_j) = \mu\Big(\biguplus_{j\in[\![n]\!]} A_j\Big) \leqslant \mu\Big(\biguplus_{j\in\mathbb{N}} A_j\Big) \text{ für jedes } n \in \mathbb{N}, \text{ also } \sum_{j\in\mathbb{N}} \mu(A_j) \leqslant \mu\Big(\biguplus_{j\in\mathbb{N}} A_j\Big).$$

(v) *Ist $\mathscr{E}$ ein Ring, so gilt für $n \in \mathbb{N}$ und $(A_j)_{j\in[\![n]\!]}$ aus $\mathscr{E}$ mit $\mu(\bigcup_{j\in[\![n]\!]} A_j) \in \mathbb{R}^+$ die Einschluss-Ausschlussformeln (Poincaré und Sylvester) $\mu\big(\bigcup_{i\in[\![n]\!]} A_i\big) = \sum_{\emptyset\neq\mathcal{I}\subseteq[\![n]\!]} (-1)^{|\mathcal{I}|-1} \mu\big(\bigcap_{i\in\mathcal{I}} A_i\big)$ und $\mu\big(\bigcap_{i\in[\![n]\!]} A_i\big) = \sum_{\emptyset\neq\mathcal{I}\subseteq[\![n]\!]} (-1)^{|\mathcal{I}|-1} \mu\big(\bigcup_{i\in\mathcal{I}} A_i\big)$.*

§01.18 **Beweis** von Lemma §01.17. Übung. □

§01.19 **Definition**. Ein Inhalt μ auf einem Ring $\mathscr{R} \subseteq 2^\Omega$ heißt

stetig von unten, falls für alle $A_n \uparrow A$ aus $\mathscr{R}$ gilt $\lim_{n\to\infty} \mu(A_n) = \mu(A)$.

stetig von oben, falls für alle $A_n \downarrow A$ aus $\mathscr{R}$ mit $\mu(A_1) \in \mathbb{R}^+$ gilt $\lim_{n\to\infty} \mu(A_n) = \mu(A)$.

$\emptyset$-stetig, falls für alle $A_n \downarrow \emptyset$ aus $\mathscr{R}$ mit $\mu(A_1) \in \mathbb{R}^+$ gilt $\lim_{n\to\infty} \mu(A_n) = 0 = \mu(\emptyset)$.

§01.20 **Lemma**. *Ein Inhalt μ auf einem Ring $\mathscr{R} \subseteq 2^\Omega$. Betrachte die Eigenschaften;* (i) *μ ist σ-additiv, also $\mu \in \mathcal{M}(\mathscr{R})$ ein Prämaß;* (ii) *μ ist σ-subadditiv;* (iii) *μ ist stetig von unten;* (iv) *μ ist $\emptyset$-stetig;* (v) *μ ist stetig von oben. Dann gelten die Implikationen* (i)$\Leftrightarrow$(ii)$\Leftrightarrow$(iii)$\Rightarrow$(iv)$\Leftrightarrow$(v). *Ist μ endlich, so gilt auch* (iv)$\Rightarrow$(iii). □

§01.21 **Beweis** von Lemma §01.20. In der Vorlesung. □

§01.22 **Beispiel** *(Beispiel §01.16* (b) *fortgesetzt)*. μ ist ein $\emptyset$-stetiger Inhalt, aber kein Prämaß. □

§01.23 **Definition**.

(a) Ein Paar $(\Omega, \mathscr{A})$ bestehend aus einer nichtleeren Menge Ω und einer σ-Algebra $\mathscr{A} \subseteq 2^\Omega$, heißt *messbarer Raum* oder *Messraum*. Die Mengen $A \in \mathscr{A}$ heißen *messbare Mengen*. Ist Ω höchstens abzählbar und $\mathscr{A} = 2^\Omega$, so wird $(\Omega, 2^\Omega)$ *diskret* genannt.

(b) Ein Tripel $\boxed{(\Omega, \mathscr{A}, \mu)}$ heißt *Maßraum*, wenn $(\Omega, \mathscr{A})$ ein Messraum ist und $\mu \in \mathcal{M}(\mathscr{A})$ ein Maß auf $\mathscr{A}$. □

§01|03 Fortsetzung von Maßen

§01.24 **Lemma** *(Eindeutigkeit)*. *Seien $(\Omega, \mathscr{A})$ ein Messraum, $\mathscr{E}$ ein $\cap$-stabiler Erzeuger von $\mathscr{A}$ und $\mu, \nu \in \mathcal{M}_\sigma(\mathscr{A})$ zwei σ-endliche Maße auf $\mathscr{A}$, die auf $\mathscr{E}$ übereinstimmen, d.h. $\mu(E) = \nu(E)$ für alle $E \in \mathscr{E}$. Es gebe weiterhin $(\Omega_n)_{n\in\mathbb{N}}$ in $\mathscr{E}$ mit $\bigcup_{n\in\mathbb{N}} \Omega_n = \Omega$ und $\mu(\Omega_n) \in \mathbb{R}^+$ für jedes $n \in \mathbb{N}$. Dann stimmen μ und ν auch auf $\mathscr{A}$ überein.*

Sind $\mu, \nu \in \mathcal{W}(\mathscr{A})$ Wahrscheinlichkeitsmaße, so gilt die Folgerung auch ohne die Existenz der Folge $(\Omega_n)_{n\in\mathbb{N}}$.

§01.25 **Beweis** von Lemma §01.24. In der Vorlesung. □

§01.26 **Bemerkung**. Mit anderen Worten, unter den Annahmen von Lemma §01.24 ist μ durch die Werte $\mu(E)$, $E \in \mathscr{E}$, eindeutig festgelegt. Die Eindeutigkeit ohne die Existenz der Folge $(\Omega_n)_{n\in\mathbb{N}}$ gilt im Allgemeinen nicht, falls $\mu \in \mathcal{M}_e(\mathscr{A})$ ein endliches Maß ist. In diesem Fall ist die Gesamtmasse $\mu(\Omega)$ im Allgemeinen nicht eindeutig festgelegt. Sei $\Omega = \{1, 2\}$, dann ist $\mathscr{E} = \{\{1\}\}$ ein $\cap$-stabiler Erzeuger von 2^Ω. Ein Wahrscheinlichkeitsmaß μ ist durch Angabe von $\mu(\{1\})$ eindeutig festgelegt. Andererseits sind $\mu = 0$ und $\nu = \delta_2$ zwei endliche Maße, die auf $\mathscr{E}$ übereinstimmen. □

§01.27 **Definition**. Eine Mengenfunktion $\mu^\star : 2^\Omega \to \overline{\mathbb{R}}^+$ heißt *äußeres Maß*, falls (äM1) $\mu^\star(\emptyset) = 0$, (äM2) $\mu^\star$ ist monoton, und (äM3) $\mu^\star$ ist σ-subadditiv. Eine Menge $A \in 2^\Omega$ heißt *$\mu^\star$-messbar*, falls $\mu^\star(A\cap B)+\mu^\star(A^c\cap B) = \mu^\star(B)$ für jedes $B \in 2^\Omega$. Setze $\sigma(\mu^\star) := \{A \in 2^\Omega : A \text{ ist } \mu^\star\text{-messbar}\}$. □

§01.28 **Bemerkung**. Wegen $\mu^\star(\emptyset) = 0$ ist stets $\Omega \in \sigma(\mu^\star)$. Da $\mu^\star$ subadditiv ist, gilt $A \in \sigma(\mu^\star)$ genau dann, wenn $\mu^\star(A \cap B) + \mu^\star(A^c \cap B) \leqslant \mu^\star(B)$ für jedes $B \in 2^\Omega$. □

§01.29 **Lemma**. *Sei $\mathscr{E} \subseteq 2^\Omega$ ein Mengensystem mit $\emptyset \in \mathscr{E}$ und $\mu : \mathscr{E} \to \overline{\mathbb{R}}^+$ eine Mengenfunktion mit $\mu(\emptyset) = 0$. Für $A \in 2^\Omega$ sei $\mathcal{U}(A) = \{\mathcal{F} \subseteq \mathscr{E} : \mathcal{F} \text{ ist abzählbar und } A \subseteq \bigcup_{F\in\mathcal{F}} F\}$ die Menge der abzählbaren[1] Überdeckungen $\mathcal{F}$ von A mit Mengen F aus $\mathscr{E}$. Setze*

$$\mu^\star : 2^\Omega \to \overline{\mathbb{R}}^+ \text{ mit } A \mapsto \mu^\star(A) := \inf\Big\{ \sum_{F\in\mathcal{F}} \mu(F) : \mathcal{F} \in \mathcal{U}(A)\Big\},$$

wobei $\inf\emptyset = \infty$. Dann ist $\mu^\star$ ein äußeres Maß. Ist μ zudem σ-subadditiv, so gilt $\mu^\star(E) = \mu(E)$ für alle $E \in \mathscr{E}$.

§01.30 **Beweis** von Lemma §01.29. In der Vorlesung. □

§01.31 **Lemma**. *Sei $\mu^\star$ ein äußeres Maß. Dann ist $\sigma(\mu^\star)$ eine σ-Algebra und die Einschränkung von $\mu^\star$ auf $\sigma(\mu^\star)$ ein Maß.*

§01.32 **Beweis** von Lemma §01.31. In der Vorlesung. □

§01.33 **Fortsetzungssatz für Maße**. *Sei $\mathscr{E}$ ein Semiring und $\mu : \mathscr{E} \to \overline{\mathbb{R}}^+$ eine additive, σ-subadditive, σ-endliche Mengenfunktion mit $\mu(\emptyset) = 0$.*

Dann existiert ein eindeutig bestimmtes, σ-endliches Maß $\widetilde{\mu} : \sigma(\mathscr{E}) \to \overline{\mathbb{R}}^+$ mit $\widetilde{\mu}(E) = \mu(E)$ für jedes $E \in \mathscr{E}$.

§01.34 **Beweis** von Satz §01.33. In der Vorlesung. □

§01.35 **Beispiel**.

(a) Es existiert ein eindeutig bestimmtes Maß λ^n auf $(\mathbb{R}^n, \mathscr{B}^n)$ mit der Eigenschaft $\lambda^n((a, b]) = \prod_{i\in[\![n]\!]}(b_i - a_i)$ für alle $a, b \in \mathbb{R}^n$ mit $a < b$. $\boxed{\lambda^n}$ heißt *Lebesgue-Maß* auf $(\mathbb{R}^n, \mathscr{B}^n)$ (vgl.

[1] eine höchstens abzählbar unendliche (d.h. endlich oder abzählbar unendliche) Menge

Vorlesung Analysis 3).

(b) Sei $\mathbb{F} : \mathbb{R} \to \mathbb{R}$ monoton wachsend und rechtsseitig stetig. Es existiert ein eindeutig bestimmtes Maß $\mu_{\mathbb{F}}$ auf $(\mathbb{R}, \mathscr{B})$ mit der Eigenschaft $\mu_{\mathbb{F}}((a, b]) = \mathbb{F}(b) - \mathbb{F}(a)$ für alle $a, b \in \mathbb{R}$ mit $a < b$. $\mu_{\mathbb{F}}$ heißt *Lebesgue-Stieltjes-Maß* auf $(\mathbb{R}, \mathscr{B})$ (Übung). Gilt weiterhin $\lim_{x\to\infty}(\mathbb{F}(x) - \mathbb{F}(-x)) = 1$, so ist $\mu_{\mathbb{F}}$ ein Wahrscheinlichkeitsmaß. □

§01.36 **Definition**. Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum.

(a) Eine Menge $N \in \mathscr{A}$ heißt *μ-Nullmenge*, oder kurz Nullmenge, falls $\mu(N) = 0$. Mit $\mathcal{N}_\mu$ bezeichnen wir das System aller Teilmengen von μ-Nullmengen.

(b) Eine Aussage gilt *μ-fast überall (μ-f.ü.)*, wenn es eine μ-Nullmenge $N \in \mathcal{N}_\mu$ gibt, so dass die Aussage für alle $\omega \in \Omega \setminus N = N^c$ gilt. Ist $A \in \mathscr{A}$, so sagen wir, eine Aussage gilt μ-fast überall auf A, falls $N \in \mathcal{N}_\mu$ gibt, so dass die Aussage für alle $\omega \in A \setminus N$ gilt. Ist $\mu = \mathbb{P} \in \mathcal{W}(\mathscr{A})$ ein Wahrscheinlichkleitsmaß, so sagen wir dann auch, dass die Aussage *$\mathbb{P}$-fast sicher ($\mathbb{P}$-f.s.)* gilt, beziehungsweise $\mathbb{P}$-fast sicher auf A.

(c) Der Maßraum $(\Omega, \mathscr{A}, \mu)$ heißt *vollständig*, falls $\mathcal{N}_\mu \subseteq \mathscr{A}$. □

§01.37 **Bemerkung**. Sei $(\Omega, \mathscr{A}, \mu)$ ein σ-endlicher Maßraum. Dann heißt $(\Omega, \sigma(\mu^\star), \mu^\star|_{\sigma(\mu^\star)})$ die Vervollständigung von $(\Omega, \mathscr{A}, \mu)$. Unter allen σ-Algbren $\mathscr{A}^\star$, die $\mathscr{A}$ enthalten, und Fortsetzungen $\mu^\star$ von μ auf $\mathscr{A}^\star$, derart dass $(\Omega, \mathscr{A}^\star, \mu^\star)$ vollständig ist, ist $\sigma(\mu^\star)$ die kleinste σ-Algebra. Ferner ist $\sigma(\mu^\star) = \sigma(\mathscr{A} \cup \mathcal{N}_\mu) = \{A \cup N : A \in \mathscr{A}, N \in \mathcal{N}_\mu\}$ und $\mu^\star(A \cup N) = \mu(A)$ für jedes $A \in \mathscr{A}$ und $N \in \mathcal{N}_\mu$. □

§01.38 **Definition**. Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum und $B \in \mathscr{A}$. Dann wird durch $\mu|_B(A) := \mu(A)$ für $A \in \mathscr{A}$ mit $A \subseteq B$ ein Maß auf der Spur $\mathscr{A}|_B$ von $\mathscr{A}$ über B (vgl. Beispiel A01.02 (g)) definiert. Dieses Maß wird *Einschränkung* von μ auf B genannt. □

§02 Integrationstheorie

§02|01 Das Integral

§02.01 **Erinnerung**. Seien $(\Omega, \mathscr{A}, \mu)$ ein Maßraum und $(\mathcal{S}, \mathscr{S})$ ein Messraum.

(i) Eine Funktion $f : \Omega \to \mathcal{S}$ heißt *$\mathscr{A}$-$\mathscr{S}$-messbar* (kurz *messbar*), falls $\sigma(f) := f^{-1}(\mathscr{S}) := \{f^{-1}(S) \mid S \in \mathscr{S}\} \subseteq \mathscr{A}$ gilt. Jede solche messbare Funktion wird auch ($(\mathcal{S}, \mathscr{S})$-wertige) *Zufallsvariable* genannt. $\sigma(f)$ wird die *von f erzeugte σ-Algebra* genannt und ist die kleinste Teil-σ-Algebra aus 2^Ω, so dass f messbar ist.

(ii) Eine messbare Funktion $f : (\Omega, \mathscr{A}) \to (\mathcal{S}, \mathscr{S})$ heißt

numerisch, kurz $f \in \overline{\mathscr{A}}$, falls $(\mathcal{S}, \mathscr{S}) = (\overline{\mathbb{R}}, \overline{\mathscr{B}})$;

positiv numerisch, kurz $f \in \overline{\mathscr{A}}^+$, falls $(\mathcal{S}, \mathscr{S}) = (\overline{\mathbb{R}}^+, \overline{\mathscr{B}}^+)$;

reell, kurz $f \in \mathscr{A}$, falls $(\mathcal{S}, \mathscr{S}) = (\mathbb{R}, \mathscr{B})$;

positiv reell, kurz $f \in \mathscr{A}^+$, falls $(\mathcal{S}, \mathscr{S}) = (\mathbb{R}^+, \mathscr{B}^+)$.

(iii) Eine reelle messbare Abbildung $f \in \mathscr{A}$, die nur endlich viele Werte annimmt, heißt *einfach* oder *elementar*. Ist $f \in \mathscr{A}$ einfach, so gibt es ein $n \in \mathbb{N}$, paarweise disjunkte, messbare Mengen $(A_i)_{i\in[\![n]\!]}$ aus $\mathscr{A}$ und reelle Zahlen $(a_i)_{i\in[\![n]\!]}$ aus $\mathbb{R}$ mit $f = \sum_{j\in[\![n]\!]} a_j \mathbb{1}_{A_j}$. Wir bezeichnen mit $\mathscr{A}_{\text{einf}}$ und $\mathscr{A}^+_{\text{einf}}$ die Menge der Elementarfunktionen bzw. positiven Elementarfunktionen auf $(\Omega, \mathscr{A})$. Sind $f = \sum_{j\in[\![n]\!]} a_j \mathbb{1}_{A_j}$ und $f = \sum_{j\in[\![m]\!]} b_j \mathbb{1}_{B_j}$ zwei Darstellungen von $f \in \mathscr{A}^+_{\text{einf}}$, so gilt $\sum_{j\in[\![n]\!]} a_j\mu(A_j) = \sum_{j\in[\![m]\!]} b_j\mu(B_j)$ (Nachrechnen!).

(iv) Ist $f \in \overline{\mathscr{A}}^+$ positiv numerisch, dann gibt es eine isotone Folge positiver Elementarfunktionen $(f_n)_{n\in\mathbb{N}}$ aus $\mathscr{A}^+_{\text{einf}}$ mit $f_n \uparrow f$ (vgl. Eigenschaft A02.06 (v)).

(v) Sind $f, g \in \overline{\mathscr{A}}$ numerische Funktionen, so schreiben wir $f \leqslant g$, falls $f(\omega) \leqslant g(\omega)$ für jedes $\omega \in \Omega$ gilt. Hingegen schreiben wir $f \leqslant g$ *μ-fast überall (μ-f.ü.)*, falls die schwächere Bedingung gilt, dass eine μ-Nullmenge N existiert mit $f(\omega) \leqslant g(\omega)$ für jedes $\omega \in N^c$. □

§02.02 **Satz**. *Für jedes Maß μ auf einem Messraum $(\Omega, \mathscr{A})$ heißt Integral bezüglich μ das eindeutig bestimmte Funktional $\mathbb{I}_\mu : \overline{\mathscr{A}}^+ \to \overline{\mathbb{R}}^+$, das die folgenden Bedingungen erfüllt:*

(I1) *Für alle $f, g \in \overline{\mathscr{A}}^+$ und $a, b \in \mathbb{R}^+$ gilt $\mathbb{I}_\mu(af + bg) = a\mathbb{I}_\mu(f) + b\mathbb{I}_\mu(g)$;* (linear)

(I2) *Für alle $(f_n)_{n\in\mathbb{N}} \uparrow f$ in $\overline{\mathscr{A}}^+$ gilt $\mathbb{I}_\mu(f_n) \uparrow \mathbb{I}_\mu(f)$;* (monoton konvergent)

(I3) *Für jedes $A \in \mathscr{A}$ gilt $\mathbb{I}_\mu(\mathbb{1}_A) = \mu(A)$.* (normiert)

Für jedes $f \in \overline{\mathscr{A}}^+$ heißt $\int f\, d\mu := \mathbb{I}_\mu(f)$ das Integral von f bezüglich μ. Für $A \in \mathscr{A}$ schreiben wir kurz $\int_A f\, d\mu := \int (f\mathbb{1}_A)\, d\mu$. □

§02.03 **Beweis** von Satz §02.02. Der Satz fasst die Hauptaussage dieses Abschnittes zusammen, der Beweis der Aussage erfolgt in mehreren Schritten. Wir zeigen zuerst in Satz §02.05 die Eindeutigkeitsaussage, und geben dann in Satz §02.09 ein Funktional $\mathbb{I}_\mu : \overline{\mathscr{A}}^+ \to \overline{\mathscr{B}}^+$ explizit an, für das wir die Bedingungen (I1)-(I3) nachweisen. Zusammenfassend zeigen wir damit dann in Satz §02.09 auch die Existenzaussage. □

§02.04 **Schreibweise**. Für $f \in \overline{\mathscr{A}}^+$ und $A \in \mathscr{A}$ schreiben wir auch $\mu(f) = \int f\, d\mu = \int_\Omega f(\omega)\mu(d\omega)$ sowie $\mu(f\mathbb{1}_A) = \int_A f\, d\mu = \int_A f(\omega)\mu(d\omega)$. □

§02.05 **Eindeutigkeitssatz**. *Das Integral ist eindeutig bestimmt.*

§02.06 **Beweis** von Satz §02.05. In der Vorlesung. □

Erinnerung §02.01 (iv) erlaubt es, die folgende Definition zu treffen, da der definierte Wert $\widetilde{\mathbb{I}}_\mu(f)$ nicht von der gewählten Darstellung von f abhängt.

§02.07 **Lemma**. *Die Abbildung $\widetilde{\mathbb{I}}_\mu : \mathscr{A}^+_{\text{einf}} \to \overline{\mathbb{R}}^+$ mit*

$$f = \sum_{j\in[\![n]\!]} a_j \mathbb{1}_{A_j} \mapsto \widetilde{\mathbb{I}}_\mu(f) := \sum_{j\in[\![n]\!]} a_j \mu(A_j).$$

ist normiert, positiv linear und monoton:

(i) *Für jedes $A \in \mathscr{A}$ gilt $\widetilde{\mathbb{I}}_\mu(\mathbb{1}_A) = \mu(A)$.* (normiert)

(ii) *Für alle $f, g \in \mathscr{A}^+_{\text{einf}}$ und $a, b \in \mathbb{R}^+$ gilt $\widetilde{\mathbb{I}}_\mu(af + bg) = a\widetilde{\mathbb{I}}_\mu(f) + b\widetilde{\mathbb{I}}_\mu(g)$;* (linear)

(iii) *Für alle $f, g \in \mathscr{A}^+_{\text{einf}}$ mit $f \leqslant g$ gilt $\widetilde{\mathbb{I}}_\mu(f) \leqslant \widetilde{\mathbb{I}}_\mu(g)$.* (monoton).

§02.08 **Beweis** von Lemma §02.07. Übung. □

§02.09 **Existenzsatz**. *Das Funktional $\mathbb{I}_\mu : \overline{\mathscr{A}}^+ \to \overline{\mathbb{R}}^+$ mit $f \mapsto \mathbb{I}_\mu(f) := \sup\left\{\widetilde{\mathbb{I}}_\mu(g) : g \in \mathscr{A}^+_{\text{einf}}, g \leqslant f\right\}$ ist ein Integral bzgl. μ, das heißt, es erfüllt die Bedingungen (I1)-(I3) aus Satz §02.02:*

(i) *Für jedes $A \in \mathscr{A}$ gilt $\mathbb{I}_\mu(\mathbb{1}_A) = \mu(A)$.* (normiert)

(ii) *Für alle $f, g \in \overline{\mathscr{A}}^+$ mit $f \leqslant g$ gilt $\mathbb{I}_\mu(f) \leqslant \mathbb{I}_\mu(g)$.* (monoton)

(iii) *Für alle $(f_n)_{n\in\mathbb{N}} \uparrow f$ in $\overline{\mathscr{A}}^+$ gilt $\mathbb{I}_\mu(f_n) \uparrow \mathbb{I}_\mu(f)$.* *(monoton konvergent)*

(iv) *Für alle $f, g \in \overline{\mathscr{A}}^+$ und $a, b \in \overline{\mathbb{R}}^+$ gilt $\mathbb{I}_\mu(af + bg) = a\mathbb{I}_\mu(f) + b\mathbb{I}_\mu(g)$,* (linear)
wobei wir die Konvention $\infty \cdot 0 = 0$ benutzen.

§02.10 **Beweis** von Satz §02.09. In der Vorlesung. □

§02.11 **Bemerkung**. Nach Lemma §02.07 (iii) gilt für $f \in \mathscr{A}^+_{\text{einf}}$ die Identität $\mathbb{I}_\mu(f) = \widetilde{\mathbb{I}}_\mu(f)$, sodass $\mathbb{I}_\mu$ eine Fortsetzung der Abbildung $\widetilde{\mathbb{I}}_\mu$ von $\mathscr{A}^+_{\text{einf}}$ auf die Menge der positiven numerischen Funktionen ist. □

§02.12 **Anmerkung**. Eine Partition $\mathcal{P} := \{A_i \,|\, i \in \mathcal{I}\} \subseteq \mathscr{A}$ von Ω nennen wir endlich, wenn die nicht-leere Indexmenge $\mathcal{I}$ endlich ist. Für jedes $A \in \mathcal{P}$ gilt also $\emptyset \neq A \in \mathscr{A}$. Setzen wir $\mathscr{P} := \{\mathcal{P} \subseteq \mathscr{A} \,|\, \mathcal{P} \text{ endliche Partition von } \Omega\}$ so erfüllt das Funktional $\mathbb{I}_\mu : \overline{\mathscr{A}}^+ \to \overline{\mathbb{R}}^+$ mit

$$f \mapsto \mathbb{I}_\mu(f) := \sup\Big\{ \sum_{A\in\mathcal{P}} \big(\inf_{\omega\in A} f(w)\big)\mu(A) : \mathcal{P} \in \mathscr{P} \Big\}.$$

ebenso die Bedingungen (I1)-(I3) aus Satz §02.02 und ist somit eine alternative aber äquivalente Definition des eindeutigen Integrals bzgl. μ. □

§02.13 **Schreibweise**. Für beliebige Maße $\mu, \nu \in \mathcal{M}(\mathscr{A})$ schreiben wir $\nu \leqslant \mu$, wenn $\nu(A) \leqslant \mu(A)$ für alle $A \in \mathscr{A}$ gilt. Offensichtlich, implizieren $\nu \leqslant \mu$ und $\mu \leqslant \nu$ gemeinsam $\mu = \nu$. □

§02.14 **Lemma** *(Eigenschaften)*. *Sei $(\Omega, \mathscr{A}, \mu)$ ein beliebiger Maßraum und $(f_n)_{n\in\mathbb{N}}$ eine Folge aus $\overline{\mathscr{A}}^+$.*

(i) *(Lemma von Fatou) Dann gilt $\mu\big(\liminf_{n\to\infty} f_n\big) = \int \big(\liminf_{n\to\infty} f_n\big)\, d\mu \leqslant \liminf_{n\to\infty} \int f_n \, d\mu$. Insbesondere für jede Folge $(A_n)_{n\in\mathbb{N}}$ von Mengen aus $\mathscr{A}$ gilt $\mu\big(\liminf_{n\to\infty} A_n\big) \leqslant \liminf_{n\to\infty} \mu\big(A_n\big)$. Ist $\mu \in \mathcal{M}_e(\mathscr{A})$ endlich, so gilt außerdem $\limsup_{n\to\infty} \mu\big(A_n\big) \leqslant \mu\big(\limsup_{n\to\infty} A_n\big)$.*

(ii) *Es gilt $\sum_{n\in\mathbb{N}} f_n \in \overline{\mathscr{A}}^+$ und $\mu\big(\sum_{n\in\mathbb{N}} f_n\big) = \int \big(\sum_{n\in\mathbb{N}} f_n\big)\, d\mu = \sum_{n\in\mathbb{N}} \mu(f_n)$.*

Seien weiterhin $f, g \in \overline{\mathscr{A}}^+$.

(iii) *Genau dann ist $f = 0$ μ-f.ü., wenn $\mu(f) = \int f \, d\mu = 0$ gilt. Ist $\mu(f) \in \mathbb{R}^+$, so gilt $f \in \mathbb{R}^+$ μ-f.ü. und die Einschränkung von μ auf $\{f \neq 0\}$ ist ein σ-endliches Maß.*

(iv) *Die Mengenfunktion $f\mu : \mathscr{A} \to \overline{\mathbb{R}}^+$ mit $A \mapsto f\mu(A) := \mu(\mathbb{1}_A f) = \int (\mathbb{1}_A f)\, d\mu$ ist ein Maß auf $(\Omega, \mathscr{A})$. Für alle $A \in \mathscr{A}$ mit $\mu(A) = 0$ gilt $f\mu(A) = 0$.*

(v) *Ist $f \leqslant g$ (bzw. $f = g$) μ-f.ü., so gilt $f\mu \leqslant g\mu$ (bzw. $f\mu = g\mu$). Ist (c1) f μ-integrierbar, oder (c2) $\mu \in \mathcal{M}_\sigma(\mathscr{A})$, oder (c3) $g\mu \in \mathcal{M}_\sigma(\mathscr{A})$ σ-endlich, so gilt auch die Umkehrung. Insbesondere, gilt $\mu(f) = \int f \, d\mu \leqslant \int g \, d\mu = \mu(g)$ (bzw. $\mu(f) = \mu(g)$).*

(vi) *Genau dann ist $\mu \in \mathcal{M}_\sigma(\mathscr{A})$ σ-endlich, wenn es $h \in \mathscr{A}^+_{\setminus 0}$ mit $\mu(h) \in \mathbb{R}^+$ (μ-integrierbar) gibt. Insbesondere, exisitiert zu jedem σ-endlichen $\mu \in \mathcal{M}_\sigma(\mathscr{A})$ ein $h \in \mathscr{A}^+_{\setminus 0}$, derart dass $h\mu \in \mathcal{M}_e(\mathscr{A})$ endlich ist und $h\mu$ besitzt die gleichen Nullmengen wie μ.*

(vii) *Für $f \in \mathscr{A}$ mit $f \geqslant 0$ μ-f.ü. gilt $\sum_{n\in\mathbb{N}} \mu\big(\{f \geqslant n\}\big) \leqslant \int f \, d\mu \leqslant \sum_{n\in\mathbb{N}_0} \mu\big(\{f > n\}\big)$ und $\int f \, d\mu = \int_0^\infty \mu\big(\{f \geqslant t\}\big) dt$.*

§02.15 **Beweis** von Lemma §02.14. In der Vorlesung. □

§02.16 **Definition**. Seien $(\Omega, \mathscr{A}, \mu)$ ein Maßraum und $\mathbb{f} \in \overline{\mathscr{A}}^+$. Wir sagen, dass das durch $\nu(A) := \int (\mathbb{1}_A \mathbb{f})\, d\mu$ für $A \in \mathscr{A}$ definierte Maß $\mathbb{f}\mu := \nu$ die *Dichte* $\mathrm{d}\nu/\mathrm{d}\mu := \frac{\mathrm{d}\nu}{\mathrm{d}\mu} := \mathbb{f}$ bezüglich μ besitzt.

§02.17 **Lemma** *(Eigenschaften)*. *Seien $\mu \in \mathcal{M}(\mathscr{A})$ ein Maß auf $(\Omega, \mathscr{A})$ und $\nu = \mathbb{f}\mu$ mit $\frac{\mathrm{d}\nu}{\mathrm{d}\mu} = \mathbb{f} \in \overline{\mathscr{A}}^+$.*

(i) *Für jede Funktion $g \in \overline{\mathscr{A}}^+$ gilt $\int g \, d\nu = \int (g\mathbb{f})\, d\mu = \mu(g\mathbb{f}) = \mathbb{f}\mu(g) = \nu(g)$.*

(ii) *Sei $\rho := \mathbb{q}\nu$ mit $\mathbb{q} \in \overline{\mathscr{A}}^+$. Dann gilt $\rho = \mathbb{q}\nu = \mathbb{q}(\mathbb{f}\mu) = (\mathbb{q}\mathbb{f})\mu$.*

(iii) *Die μ-Dichte $\mathrm{d}\nu/\mathrm{d}\mu$ von ν ist eindeutig bis auf Gleichheit μ-fast überall, wenn $\nu \in \mathcal{M}_\sigma(\mathscr{A})$ oder $\mu \in \mathcal{M}_\sigma(\mathscr{A})$ σ-endlich ist.*

(iv) *Ist $\nu \in \mathcal{M}_\sigma(\mathscr{A})$ σ-endlich, so ist $\frac{\mathrm{d}\nu}{\mathrm{d}\mu} = f \in \mathbb{R}^+$ μ-f.ü.. Die Umkehrung gilt, wenn $\mu \in \mathcal{M}_\sigma(\mathscr{A})$.*

§02.18 **Beweis** von Lemma §02.17. In der Vorlesung. □

§02.19 **Schreibweise**. Ist $f \in \overline{\mathscr{A}}$ eine numerische Funktion, so sind $f^+ := f \vee 0$, $f^- := (-f)^+$, $f^+ + f^- = |f| \in \overline{\mathscr{A}}^+$ positive numerische Funktionen (vgl. Eigenschaft A02.06). □

§02.20 **Definition**. Seien $(\Omega, \mathscr{A}, \mu)$ ein Maßraum und $f \in \overline{\mathscr{A}}$ eine numerische Funktion.

(a) Ist höchstens eines der beiden Integrale $\int f^+ \,d\mu$ und $\int f^- \,d\mu$ nicht endlich, dass heißt, $\mu(f^+) \wedge \mu(f^-) \in \mathbb{R}^+$, so definiert $\boxed{\int f \,d\mu} := \int f^+ \,d\mu - \int f^- \,d\mu$ das *Integral* von f bezüglich μ mit den üblichen Konventionen $\infty + x = \infty$ und $-\infty + x = -\infty$ für alle $x \in \mathbb{R}$. In diesem Fall wird f *quasiintegrierbar* genannt. Das Integral von f ist nicht definiert, wenn $\int f^+ \,d\mu = \infty = \int f^- \,d\mu$ gilt.

(b) Falls $\int |f| \,d\mu \in \mathbb{R}^+$, also falls $\int f^+ \,d\mu \in \mathbb{R}^+$ und $\int f^- \,d\mu \in \mathbb{R}^+$, gilt, dann heißt f *μ-integrierbar*. Die Menge aller μ-integrierbaren numerischen Funktionen bezeichnen wir mit $\boxed{\mathscr{L}_1} := \mathscr{L}_1(\mu) := \mathscr{L}_1(\mathscr{A}, \mu) := \{f \in \overline{\mathscr{A}} : \int |f| \,d\mu \in \mathbb{R}^+\}$.

(c) Für $p \in \mathbb{R}^+_{\setminus 0}$ definiere $\|f\|_{\mathscr{L}_p} := \big(\int (|f|^p) \,d\mu\big)^{1/p}$ und $\|f\|_{\mathscr{L}_\infty} := \inf\{x \in \mathbb{R}^+ : \mu(\{|f| > x\}) = 0\}$. Für $p \in \overline{\mathbb{R}}^+_{\setminus 0}$ heißt f *$\mathscr{L}_p$-integrierbar*, wenn $\|f\|_{\mathscr{L}_p} \in \mathbb{R}^+$. Die Menge aller $\mathscr{L}_p$-integrierbaren Funktionen bezeichnen wir mit $\boxed{\mathscr{L}_p} := \mathscr{L}_p(\mu) := \mathscr{L}_p(\mathscr{A}, \mu) := \{f \in \overline{\mathscr{A}} : \|f\|_{\mathscr{L}_p} \in \mathbb{R}^+\}$. Für $p \in [1, \infty]$ ist $\|\cdot\|_{\mathscr{L}_p}$ eine Pseudonorm auf dem Vektorraum $\mathscr{L}_p$.

(d) $\langle \cdot, \cdot \rangle_{\mathscr{L}_2} : \mathscr{L}_2(\mu) \times \mathscr{L}_2(\mu) \to \mathbb{R}$ mit $(f, g) \mapsto \langle f, g \rangle_{\mathscr{L}_2} := \int fg \,d\mu$ ist eine positive semidefinite symmetrische Bilinearform. □

§02.21 **Lemma** *(Eigenschaften)*. *Seien $f, g \in \mathscr{L}_1(\mathscr{A}, \mu)$.*

(i) *Für alle $a, b \in \mathbb{R}$ gilt $af + bg \in \mathscr{L}_1(\mu)$ und $\int (af + bg) \,d\mu = a \int f \,d\mu + b \int g \,d\mu$;* (linear)

(ii) *Sei $h \in \overline{\mathscr{A}}$. Ist $h = f$ μ-f.ü., dann gilt $h \in \mathscr{L}_1(\mu)$ und $\int h \,d\mu = \int f \,d\mu$.*
Ist $|h| \leqslant g$ μ-f.ü., dann gilt $h \in \mathscr{L}_1(\mu)$.

(iii) *Ist $f \leqslant g$ μ-f.ü., so ist $\int f \,d\mu \leqslant \int g \,d\mu$.* (monoton)
Insbesondere aus $f \in \overline{\mathbb{R}}^+$ μ-f.ü. folgt $\int f \,d\mu \in \mathbb{R}^+$. (positiv)

(iv) *Es gilt $|\int f \,d\mu| \leqslant \int |f| \,d\mu$.* (Dreiecksungleichung)

(v) *Es gilt $f = 0$ μ-f.ü. genau dann, wenn $\int_A f \,d\mu = 0$ für alle $A \in \mathscr{A}$ gilt.*

(vi) *Ist $\mu \in \mathcal{M}_e(\mathscr{A})$ endlich und $h \in \mathscr{A}$ beschränkt, also $\sup_{\omega \in \Omega} |h(\omega)| < \infty$, so gilt $h \in \mathscr{L}_1(\mu)$.*

(vii) *Für $\mu, \nu \in \mathcal{M}(\mathscr{A})$ ist $h \in \mathscr{L}_1(\mu + \nu)$ genau dann wenn $h \in \mathscr{L}_1(\mu) \cap \mathscr{L}_1(\nu)$ gilt. In diesem Fall ist $\int h \,d(\mu + \nu) = \int h \,d\mu + \int h \,d\nu$.*

(viii) *Sei $\nu = \mathfrak{f}\mu$ mit $\frac{\mathrm{d}\nu}{\mathrm{d}\mu} = \mathfrak{f} \in \overline{\mathscr{A}}^+$. Eine Funktion $g \in \overline{\mathscr{A}}$ ist genau dann ν-integrierbar, wenn $g\mathfrak{f}$ μ-integrierbar ist. In diesem Fall gilt $\nu(g) = \mu(g\mathfrak{f}) = \int (g\mathfrak{f}) \,d\mu = \int g \,d(\mathfrak{f}\mu) = \int g \,d\nu$.*

§02.22 **Beweis** von Lemma §02.21. In der Vorlesung. □

§02.23 **Korollar** *(Eigenschaften)*. *Seien nun $f, g \in \overline{\mathscr{A}}$.*

(i) *Sei $p \in \mathbb{R}^+_{\setminus 0}$. Genau dann ist $f \in \mathscr{L}_p(\mu)$, wenn $|f|^p \in \mathscr{L}_1(\mu)$. Es gilt $\mu(\{|f| > \|f\|_{\mathscr{L}_\infty}\}) = 0$.*

(ii) *Sei $p \in \overline{\mathbb{R}}^+_{\setminus 0}$. Genau dann ist $\|f\|_{\mathscr{L}_p} = 0$, wenn $f = 0$ μ-f.ü.. Für $a \in \mathbb{R}$ gilt $\|af\|_{\mathscr{L}_p} = |a| \|f\|_{\mathscr{L}_p}$. Ist $f \in \mathscr{L}_p(\mu)$ und $f = g$ μ-f.ü., so gilt $|f| < \infty$ μ-f.ü. sowie $\|f\|_{\mathscr{L}_p} = \|g\|_{\mathscr{L}_p}$.*

§02.24 **Beweis** von Korollar §02.23. Übung. □

§02.25 **Lemma** *(Eigenschaften)*. *Seien weiterhin $(\mathcal{X}, \mathscr{X})$ ein Messraum, $X : \Omega \to \mathcal{X}$ eine $\mathscr{A}$-$\mathscr{X}$-messbare Funktion und $\mu^X := \mu \circ X^{-1} \in \mathcal{M}(\mathscr{X})$ das Bildmaß von μ unter X auf $(\mathcal{X}, \mathscr{X})$. Für $h \in \mathscr{X}^+$ gilt dann $\int h(X)\, d\mu = \int h\, d\mu^X$. Genau dann ist $h \in \overline{\mathscr{X}}$ integrierbar bezüglich μ^X, wenn $h(X) \in \overline{\mathscr{A}}$ integrierbar bezüglich μ ist, und dann gilt ebenfalls $\int h(X)\, d\mu = \int h\, d\mu^X$.*

$$\begin{tikzcd}
(\Omega, \mathscr{A}) \arrow[r, "X"] \arrow[dr, "h(X) \in \mathscr{L}_1(\mu)"'] & (\mathcal{X}, \mathscr{X}) \arrow[d, "h \in \mathscr{L}_1(\mu^X)"] \\
 & (\overline{\mathbb{R}}, \overline{\mathscr{B}})
\end{tikzcd}$$

Ist speziell X eine Zufallsvariable auf $(\Omega, \mathscr{A}, \mathbb{P})$, so gilt

$$\int h(x)\mathbb{P}^X(dx) = \int h\, d\mathbb{P}^X = \mathbb{P}^X(h) = \mathbb{P}(h(X)) = \int h(X)\, d\mathbb{P} = \int h(X(\omega))\mathbb{P}(d\omega).$$

§02.26 **Beweis** von Lemma §02.25. In der Vorlesung. □

§02|02 Konvergenzarten

§02.27 **Definition**. Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum. Eine Folge $(f_n)_{n\in\mathbb{N}}$ aus $\mathscr{A}$ konvergiert gegen $f \in \overline{\mathscr{A}}$

μ-fast überall (μ-f.ü.), kurz $f_n \xrightarrow{\mu\text{-f.ü.}} f$, wenn $\limsup_{n\to\infty}|f_n - f| = 0$ μ-f.ü. gilt. Es also eine μ-Nullmenge $N \in \mathscr{A}$ gibt, sodass für jedes $\omega \in N^c := \Omega \setminus N$ gilt, dass $\lim_{n\to\infty}|f_n(\omega) - f(\omega)| = 0$.

μ-fast vollständig (μ-f.v.), kurz $f_n \xrightarrow{\mu\text{-f.v.}} f$, wenn für jedes $A \in \mathscr{A}$ mit $\mu(A) \in \mathbb{R}^+$ und für jedes $\varepsilon \in \mathbb{R}^+_{\setminus 0}$ gilt $\sum_{n\in\mathbb{N}} \mu\big(\{|f_n - f| > \varepsilon\} \cap A\big) \in \mathbb{R}^+$.

μ-stochastisch (oder dem Maße μ nach), kurz $f_n \xrightarrow{\mu} f$, wenn für jedes $A \in \mathscr{A}$ mit $\mu(A) \in \mathbb{R}^+$ und für jedes $\varepsilon \in \mathbb{R}^+_{\setminus 0}$ gilt $\lim_{n\to\infty} \mu\big(\{|f_n - f| > \varepsilon\} \cap A\big) = 0$.

im p-ten Mittel (oder in $\mathscr{L}_p(\mu)$), $p \in \overline{\mathbb{R}}^+_{\setminus 0}$, kurz $f_n \xrightarrow{\mathscr{L}_p(\mu)} f$, wenn $(f_n)_{n\in\mathbb{N}}$ und f aus $\mathscr{L}_p(\mu)$ sind und es gilt $\lim_{n\to\infty}\|f_n - f\|_{\mathscr{L}_p} = 0$.

Gelegentlich verwenden wir abkürzend $f_n \xrightarrow{\text{f.ü.}} f$ oder auch $f_n \xrightarrow{\mathscr{L}_p} f$ wenn das zu Grund liegende Maß μ aus dem Kontext hervorgeht. □

§02.28 **Bemerkung**. Konvergenz im p-ten Mittel als auch μ-fast überall legen ihren Grenzwert eindeutig fest bis auf Gleichheit μ-fast überall. Ist $\mu \in \mathcal{M}_\sigma(\mathscr{A})$ σ-endlich, so gilt dies auch für Konvergenz dem Maße μ nach. Ist $f_n \xrightarrow{\mu} f$ und $f_n \xrightarrow{\mu} g$, dann gilt (wegen $|f - g| \leqslant |f - f_n| + |g - f_n|$) für jedes $\varepsilon \in \mathbb{R}^+_{\setminus 0}$ und $A \in \mathscr{A}$ mit $\mu(A) \in \mathbb{R}^+$

$$\mu(\{|f - g| > \varepsilon\} \cap A) \leqslant \mu(\{|f - f_n| > \varepsilon/2\} \cap A) + \mu(\{|g - f_n| > \varepsilon/2\} \cap A) \xrightarrow{n\to\infty} 0.$$

Also ist $\mu(\{|f-g| > \varepsilon\} \cap A) = 0$. Damit gilt aber auch $\mu(\{f \neq g\} \cap A) = 0$ unter Verwendung von $\{f \neq g\} \cap A = \bigcup_{k\in\mathbb{N}}\{|f - g| > 1/k\} \cap A$. Wählen wir nun $A_n \uparrow \Omega$ mit $\mu(A_n) \in \mathbb{R}^+$ (da $\mu \in \mathcal{M}_\sigma(\mathscr{A})$), so folgt auch $f = g$ μ-f.ü.. Ist weiterhin $\mu \in \mathcal{M}_e(\mathscr{A})$ endlich, dann ist $f_n \xrightarrow{\mu} f$ äquivalent zu der Bedingung $\lim_{n\to\infty} \mu(\{|f_n - f| > \varepsilon\}) = 0$ für alle $\varepsilon \in \mathbb{R}^+_{\setminus 0}$. Ist dagegen $\mu \in \mathcal{M}_\sigma(\mathscr{A})$ σ-endlich, so gilt die Umkehrung nicht. In der Tat betrachten wir auf $(\mathbb{N}, 2^{\mathbb{N}})$ das σ-endliche Maß $\mu \in \mathcal{M}_\sigma(2^{\mathbb{N}})$ mit Zähldichte $\mathbb{p}(n) = 1/n$, $n \in \mathbb{N}$, bezüglich des Zählmaßes $\zeta_{\mathbb{N}}$ (vgl. Beispiel §01.16 (c)). Für $A_n := [n, \infty) \cap \mathbb{N}$, $n \in \mathbb{N}$ gilt $\mathbb{1}_{A_n} \xrightarrow{\mu} 0$, da für jedes $\varepsilon \in (0, 1)$ gilt $\{\mathbb{1}_{A_n} > \varepsilon\} = A_n$ und wegen $A_n \downarrow \emptyset$ (und Stetigkeit von oben) auch $\mu(A_n \cap A) \downarrow 0$ für jedes $A \in \mathscr{A}$ mit $\mu(A) \in \mathbb{R}^+$. Dagegen gilt $\mu(A_n) = \infty$ für alle $n \in \mathbb{N}$. □

§02.29 **Lemma**. *Sei $(\Omega, \mathscr{A}, \mu)$ ein beliebiger Maßraum.*

(i) *(Monotone Konvergenz) Sei $(f_n)_{n\in\mathbb{N}}$ aus $\mathscr{L}_1(\mu)$ und $f \in \overline{\mathscr{A}}$ mit $f_n \uparrow f$ μ-f.ü.. Dann gilt $\int f_n \, d\mu \uparrow \int f \, d\mu$.*

(ii) *(Dominierte Konvergenz) Sei $(f_n)_{n\in\mathbb{N}}$ aus $\overline{\mathscr{A}}$ μ-f.ü. konvergent und es gebe $g \in \mathscr{L}_1(\mu)$ mit $\sup_{n\in\mathbb{N}} |f_n| \leqslant g$ μ-f.ü.. Dann existiert ein $f \in \overline{\mathscr{A}}$ mit $\lim_{n\to\infty} f_n = f$ μ-f.ü., $(f_n)_{n\in\mathbb{N}}$ und f sind aus $\mathscr{L}_1(\mu)$ und es gilt $\lim_{n\to\infty} \int(|f - f_n|) \, d\mu = 0$ sowie $\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu$. Ist $g \in \mathscr{L}_p(\mu)$ für $p \in [1,\infty)$ dann sind auch $(f_n)_{n\in\mathbb{N}}$ und f aus $\mathscr{L}_p(\mu)$ und es gilt $\lim_{n\to\infty} \|f_n - f\|_{\mathscr{L}_p} = 0$ sowie $\lim_{n\to\infty} \|f_n\|_{\mathscr{L}_p} = \|f\|_{\mathscr{L}_p}$.*

(iii) *(Satz von Scheffé) Seien $(f_n)_{n\in\mathbb{N}}$ und f aus $\overline{\mathscr{A}}^+$ integrierbar und es gelte $f_n \xrightarrow{\mu\text{-f.ü.}} f$ sowie $\int f_n \, d\mu \xrightarrow{n\to\infty} \int f \, d\mu$. Dann gilt $f_n \xrightarrow{\mathscr{L}_1(\mu)} f$.*

(iv) *(Satz von Riesz) Seien $(f_n)_{n\in\mathbb{N}}$ und f aus $\mathscr{L}_p(\mu)$ mit $p \in [1,\infty)$ und es gelte $f_n \xrightarrow{\mu\text{-f.ü.}} f$. Dann gilt $\int |f_n|^p \, d\mu \xrightarrow{n\to\infty} \int |f|^p \, d\mu$ genau dann, wenn $f_n \xrightarrow{\mathscr{L}_p(\mu)} f$.*

(v) *Für f und $(f_n)_{n\in\mathbb{N}}$ aus $\mathscr{A}$ gilt $f_n \xrightarrow{\mu\text{-f.ü.}} f \Longrightarrow f_n \xrightarrow{\mu} f \Longleftarrow f_n \xrightarrow{\mathscr{L}_p(\mu)} f$. Ist $\mu \in \mathcal{M}_\sigma(\mathscr{A})$ σ-endlich, dann gilt auch $f_n \xrightarrow{\mu\text{-f.v.}} f \Longrightarrow f_n \xrightarrow{\mu\text{-f.ü.}} f$. Weiterhin gilt $f_n \xrightarrow{\mu} f$ genau dann, wenn zu jeder Teilfolge von $(f_n)_{n\in\mathbb{N}}$ eine gegen f μ-fast überall konvergente Teilfolge existiert.*

§02.30 **Beweis** von Korollar §02.23. In der Vorlesung. □

§02.31 **Erinnerung**. Eine Folge $(f_n)_{n\in\mathbb{N}}$ aus $\mathscr{L}_p(\mu)$ heißt $\mathscr{L}_p(\mu)$-Cauchy-Folge, wenn für jedes $\varepsilon \in \mathbb{R}^+_{\setminus 0}$ ein $n_o \in \mathbb{N}$ existiert, so dass für alle $m, n \in \mathbb{N} \cap (n_0, \infty)$ gilt $\|f_n - f_m\|_{\mathscr{L}_p} \leqslant \varepsilon$, kurz $\lim_{n,m\to\infty} \|f_n - f_m\|_{\mathscr{L}_p} = 0$. Wir halten fest, dass jede in $\mathscr{L}_p(\mu)$ konvergente Folge unter Verwendung der Minkowski Ungleichung (vgl. Lemma §02.50 (iii)) eine $\mathscr{L}_p(\mu)$-Cauchy-Folge ist. □

§02.32 **Lemma**. *Jede $\mathscr{L}_p(\mu)$-Cauchy-Folge $(f_n)_{n\in\mathbb{N}}$ konvergiert im p-ten Mittel gegen ein $f \in \mathscr{L}_p(\mu)$ für $p \in [1,\infty]$. Eine geeignete Teilfolge von $(f_n)_{n\in\mathbb{N}}$ konvergiert μ-f.ü. gegen f.*

§02.33 **Beweis** von Lemma §02.32. In der Vorlesung. □

§02.34 **Korollar**. *Eine $\mathscr{L}_p(\mu)$-Cauchy-Folge $(f_n)_{n\in\mathbb{N}}$ konvergiere μ-f.ü. gegen ein $f \in \mathscr{A}$ für $p \in [1,\infty]$. Dann liegt f in $\mathscr{L}_p(\mu)$ und die Folge $(f_n)_{n\in\mathbb{N}}$ konvergiert auch im p-ten Mittel gegen f.*

§02.35 **Beweis** von Korollar §02.34. In der Vorlesung. □

§02.36 **Vorbemerkung**. Seien $(\Omega, \mathscr{A}, \mu)$ ein beliebiger Maßraum, $p \in [1,\infty)$ und $f \in \overline{\mathscr{A}}$. Dann ist f μ-integrierbar genau dann, wenn für jedes $\varepsilon \in \mathbb{R}^+_{\setminus 0}$ ein $g \in \mathscr{L}_1(\mu) \cap \overline{\mathscr{A}}^+$ existiert mit $\mu(|f|\mathbb{1}_{\{|f|\geqslant g\}}) \leqslant \varepsilon$ oder äquivalant dazu $\inf\{\sup_{f\in\mathcal{F}} \mu(|f|\mathbb{1}_{\{|f|\geqslant g\}}) : g \in \mathscr{L}_1(\mu) \cap \overline{\mathscr{A}}^+\} = 0$. Ist $\mu(|f|) \in \mathbb{R}^+_{\setminus 0}$, so wähle $g = 2|f|$. Dann ist $\{|f| \geqslant g\} = \{f = 0\} \cup \{|f| = \infty\}$ und somit gemäß Korollar §02.23 (v) auch $\mu(|f|\mathbb{1}_{\{|f|\geqslant g\}}) = 0$. Umgekehrt gilt $\mu(|f|) = \mu(|f|\mathbb{1}_{\{|f|\geqslant g\}}) + \mu(|f|\mathbb{1}_{\{|f|<g\}}) \leqslant \varepsilon + \mu(g) \in \mathbb{R}^+_{\setminus 0}$, also auch $\mu(|f|) \in \mathbb{R}^+_{\setminus 0}$. □

§02.37 **Definition**. Eine Familie $\mathcal{F}$ aus $\mathscr{L}_1(\mu)$ heißt *gleichgradig μ-integrierbar*, wenn

$$\inf\Big\{\sup_{f\in\mathcal{F}} \mu\big(|f|\mathbb{1}_{\{|f|\geqslant g\}}\big) : g \in \mathscr{L}_1(\mu) \cap \overline{\mathscr{A}}^+\Big\} = 0.$$

Ist $\mu \in \mathcal{M}_e(\mathscr{A})$ endlich, so ist die gleichgradige μ-Intergrierbarkeit äquivalent zu

$$\inf\Big\{\sup_{f\in\mathcal{F}} \mu\big(|f|\mathbb{1}_{\{|f|\geqslant a\}}\big) : a \in \mathbb{R}^+\Big\} = 0.$$ □

§02.38 **Bemerkung**.

(i) Ist $\mathcal{F}$ gleichgradig μ-integrierbar und $\varepsilon \in \mathbb{R}^+_{\setminus 0}$, so existiert eine *$\varepsilon$-Schranke* $g \in \mathscr{L}_1(\mu) \cap \overline{\mathscr{A}}^+$ mit $\sup_{f\in\mathcal{F}} \mu\big(|f|\mathbb{1}_{\{|f|\geqslant g\}}\big) \leqslant \varepsilon$. Jede Funktion $h \in \mathscr{L}_1(\mu) \cap \overline{\mathscr{A}}^+$ mit $h \geqslant g$ ist dann auch eine ε-Schranke für $\mathcal{F}$.

(ii) Eine Familie $(f_i)_{i\in\mathcal{I}}$ aus $\overline{\mathscr{A}}$ heißt *gleichgradig μ-integrierbar*, wenn die Menge $\{f_i : i \in \mathcal{I}\}$ gleichgradig μ-integrierbar ist.

(iii) Sind $\mathcal{F}_i$, $i \in [\![n]\!]$, endlich viele gleichgradig μ-integrierbare Mengen aus $\overline{\mathscr{A}}$, dann ist auch $\mathcal{F} := \cap_{i\in[\![n]\!]}\mathcal{F}_i$ gleichgradig μ-integrierbar. In der Tat für $\varepsilon \in \mathbb{R}^+_{\setminus 0}$ und ε-Schranke g_i für $\mathcal{F}_i$, $i \in [\![n]\!]$, ist $g_1 \vee \cdots \vee g_n$ auch eine ε-Schranke für $\mathcal{F}$.

(iv) Ist $\mathcal{F} \subseteq \overline{\mathscr{A}}$ und $g \in \mathscr{L}_p(\mu) \cap \overline{\mathscr{A}}^+$ mit $|f| \leqslant g$ μ-f.ü. für alle $f \in \mathcal{F}$, dann ist $\mathcal{F}^p := \{|f|^p : f \in \mathcal{F}\}$ gleichgradig μ-integrierbar. Für $\varepsilon \in \mathbb{R}^+_{\setminus 0}$ ist jede ε-Schranke h für $\{g^p\}$ auch eine ε-Schranke für $\mathcal{F}^p$, da für alle $f \in \mathcal{F}$ gilt $\mu\big(|f|^p\mathbb{1}_{\{|f|^p>h\}}\big) \leqslant \mu\big(|g|^p\mathbb{1}_{\{|g|^p>h\}}\big) \leqslant \varepsilon$. □

§02.39 **Lemma**. *Sei $(\Omega, \mathscr{A}, \mu)$ ein beliebiger Maßraum.*

(i) *Eine $(f_i)_{i\in\mathcal{I}}$ aus $\mathscr{L}_1(\mu)$ mit endlicher Indexmenge $\mathcal{I}$ ist gleichgradig integrierbar.*

(ii) *Sind die Familien $(f_i)_{i\in\mathcal{I}}$ und $(g_j)_{j\in\mathcal{J}}$ aus $\mathscr{L}_1(\mu)$ gleichgradig integrierbar, dann sind auch $(f_i + g_j)_{i\in\mathcal{I},j\in\mathcal{J}}$, $(f_i - g_j)_{i\in\mathcal{I},j\in\mathcal{J}}$ sowie $(|f_i|)_{i\in\mathcal{I}}$ gleichgradig integrierbar.*

(iii) *Ist $\mathcal{F}$ gleichgradig integrierbar und existiert zu jedem $g \in \mathcal{G}$ ein $f \in \mathcal{F}$ mit $|g| \leqslant |f|$, so ist auch $\mathcal{G}$ gleichgradig integrierbar.*

(iv) *Sind $\mu \in \mathcal{M}_e(\mathscr{A})$ endlich, $p > 1$ und $\mathcal{F}$ beschränkt in $\mathscr{L}_p(\mu)$, das heißt $\sup_{f\in\mathcal{F}}\|f\|_{\mathscr{L}_p} \in \mathbb{R}^+$, dann ist $\mathcal{F}$ gleichgradig integrierbar.*

§02.40 **Beweis** von Lemma §01.08. Analog zur Vorlesung EWS / Übung. □

§02.41 **Satz**. *Sei $(\Omega, \mathscr{A}, \mu)$ ein beliebiger Maßraum. $\mathcal{F} \subseteq \overline{\mathscr{A}}$ ist genau dann gleichgradig μ-integrierbar, wenn die folgenden zwei Bedingungen gelten:*

(gI1) *$\mathcal{F}$ ist beschränkt in $\mathscr{L}_1(\mu)$, d.h. $\sup_{f\in\mathcal{F}} \mu(|f|) \in \mathbb{R}^+$;*

(gI2) $\forall\, \varepsilon \in \mathbb{R}^+_{\setminus 0} : \exists h \in \mathscr{L}_1(\mu) \cap \overline{\mathscr{A}}^+ : \exists \delta \in \mathbb{R}^+_{\setminus 0} : \forall A \in \mathscr{A} : \mu(h\mathbb{1}_A) \leqslant \delta \Rightarrow \sup_{f\in\mathcal{F}} \mu(|f|\mathbb{1}_A) \leqslant \varepsilon.$

§02.42 **Beweis** von Satz §02.41. In der Vorlesung. □

§02.43 **Satz**. *Sei $(\Omega, \mathscr{A}, \mu)$ ein beliebiger Maßraum. Sei $(f_n)_{n\in\mathbb{N}}$ reellwertig aus $\mathscr{L}_p(\mu)$, $p \in [1, \infty)$. Dann sind äquivalent:*

(i) *$(f_n)_{n\in\mathbb{N}}$ konvergiert in $\mathscr{L}_p(\mu)$;*

(ii) *$(|f_n|^p)_{n\in\mathbb{N}}$ ist gleichgradig μ-integrierbar und $(f_n)_{n\in\mathbb{N}}$ ist konvergent dem Maße μ nach.*

§02.44 **Beweis** von Satz §02.43. (i)⇒(ii) in der Vorlesung, für die Umkehrung siehe Bauer [1992, Satz 21.4, S.142] □

§02.45 **Bemerkung**. Der Satz §02.43 garantiert nur die Existenz einer p-fach μ-integrierbaren Funktion unter den möglichen stochastischen Limiten der Folge $(f_n)_{n\in\mathbb{N}}$. □

§02.46 **Korollar**. *Sei $\mu \in \mathcal{M}_\sigma(\mathscr{A})$ σ-endlich. Konvergiert eine Folge $(f_n)_{n\in\mathbb{N}}$ aus $\mathscr{L}_p(\mu)$, $p \in [1, \infty)$, μ-stochastisch gegen ein $f \in \mathscr{A}$, und ist die Folge $(|f_n|^p)_{n\in\mathbb{N}}$ gleichgradig μ-integrierbar, so liegt f auch in $\mathscr{L}_p(\mu)$ und die Folge $(f_n)_{n\in\mathbb{N}}$ konvergiert im p-ten Mittel gegen f.*

§02.47 **Beweis** von Korollar §02.46. In der Vorlesung. □

§02.48 **Zusammenfassung**. *Seien $(\Omega, \mathscr{A}, \mu)$ ein beliebiger Maßraum, $p \in [1, \infty]$ und $(f_n)_{n\in\mathbb{N}}$ aus $\mathscr{L}_p(\mu)$. Dann sind die folgenden Aussagen sind äquivalent:*

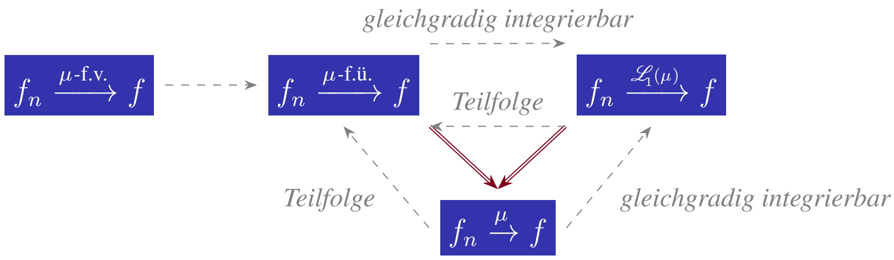
(i) *Es gibt ein* $f \in \mathscr{L}_p(\mu)$ *mit* $f_n \xrightarrow{\mathscr{L}_p(\mu)} f$.

(ii) $(f_n)_{n\in\mathbb{N}}$ *ist eine* $\mathscr{L}_p(\mu)$*-Cauchy Folge, also* $\lim_{n,m\to\infty} \|f_n - f_m\|_{\mathscr{L}_p} = 0$.

Ist $p < \infty$ *und* $\mu \in \mathcal{M}_\sigma(\mathscr{A})$ σ*-endlich, so sind (i) und (ii) zudem äquivalent zu*

(iii) $(|f_n|^p)_{n\in\mathbb{N}}$ *ist gleichgradig intergrierbar, und es gibt* $f \in \overline{\mathscr{A}}$ *mit* $f_n \xrightarrow{\mu} f$.

Die Limiten in (i) und (iii) stimmen überein.



Die Abbildung wurde auf der Grundlage von Klenke [2020, Abb.6.1, S.159] erstellt. □

§02|03 L_p-Räume

§02.49 **Bemerkung**. Zur Erinnerung für $p \in \overline{\mathbb{R}}^+_{\setminus 0}$ und $f, g \in \overline{\mathscr{A}}$ haben wir gezeigt, dass genau dann $\|f - g\|_{\mathscr{L}_p} = 0$ gilt, wenn $f = g$ μ-f.ü.. In diesem Fall sehen wir f und g als äquivalent an. Sei nun $\mathcal{N} = \{f \in \overline{\mathscr{A}} : f = 0\ \mu\text{-f.ü.}\}$, dann ist $\mathcal{N}$ ein Untervektorraum von $\mathscr{L}_p(\mu)$, so dass wir formal den Quotientenraum $\boxed{\mathrm{L}_p} := \mathrm{L}_p(\mu) := \mathrm{L}_p(\mathscr{A}, \mu) := \big\{[f] := f + \mathcal{N} : f \in \mathscr{L}_p(\mu)\big\}$ bilden können. Für $[f] \in \mathrm{L}_p(\mu)$ setzen wir $\|[f]\|_{\mathrm{L}_p} := \|f\|_{\mathscr{L}_p}$ für ein $f \in [f]$ und $\int [f]\, d\mu := \int f\, d\mu$, falls der Ausdruck für f definiert ist. Dabei hängt $\|[f]\|_{\mathrm{L}_p}$ nicht von der Wahl des Repräsentanten $f \in [f]$ ab. Analog für $[f], [g] \in \mathrm{L}_2(\mu)$ definieren wir $\langle [f], [g] \rangle_{\mathrm{L}_2} := \langle f, g \rangle_{\mathscr{L}_2}$, wobei $f \in [f]$ und $g \in [g]$. □

§02.50 **Lemma**. *Sei* $(\Omega, \mathscr{A}, \mu)$ *ein beliebiger Maßraum.*

(i) *(Hölder Ungleichung) Seien* $p, q \in [1, \infty]$ *mit* $\frac{1}{p} + \frac{1}{q} = 1$. *Dann gilt* $\|fg\|_{\mathscr{L}_1} \leqslant \|f\|_{\mathscr{L}_p} \|g\|_{\mathscr{L}_q}$.
(Cauchy-Schwarz Ungleichung) Insbesondere gilt $|\langle f, g \rangle_{\mathscr{L}_2}| \leqslant \|f\|_{\mathscr{L}_2} \|g\|_{\mathscr{L}_2}$.

(ii) *Seien* $\mu \in \mathcal{M}_e(\mathscr{A})$ *endlich,* $q \in \overline{\mathbb{R}}^+_{\setminus 0}$ *und* $p \in (0, q)$. *Dann gilt* $\mu(\Omega)^{1/q} \|f\|_{\mathscr{L}_p} \leqslant \mu(\Omega)^{1/p} \|f\|_{\mathscr{L}_q}$ *und somit* $\mathscr{L}_q \subseteq \mathscr{L}_p$.

(iii) *(Minkowski Ungleichung) Für* $p \in [1, \infty]$ *gilt* $\|f + g\|_{\mathscr{L}_p} \leqslant \|f\|_{\mathscr{L}_p} + \|g\|_{\mathscr{L}_p}$.

(iv) *(Fischer-Riesz) Für* $p \in [1, \infty]$ *ist* $(\mathrm{L}_p(\mu), \|\cdot\|_{\mathrm{L}_p})$ *ein Banachraum.*
Insbesondere ist $(\mathrm{L}_2(\mu), \langle \cdot, \cdot \rangle_{\mathrm{L}_2})$ *ein reeller Hilbertraum.*

§02.51 **Beweis** von Lemma §02.50. (i) und (iii) wie in der Vorlesung EWS oder Bauer [1992, Satz 14.1/14.2, S.85.86]. (ii) in der Vorlesung und für (iv) siehe Klenke [2020, Satz 7.18, S.171] □

§02.52 **Bemerkung**. Sei $(V, \langle \cdot, \cdot \rangle)$ ein Hilbertraum. Dann besagt der *Darstellungssatz von Riesz-Fréchet*, dass eine Abbildung $F : V \to \mathbb{R}$ genau dann stetig und linear ist, wenn es ein $f \in V$ gibt mit $F(x) = \langle f, x \rangle$ für alle $x \in V$. Dabei ist das Element $f \in V$ eindeutig bestimmt. Die Aussage des Satzes gilt auch für einen linearen Vektorraum mit vollständiger positiv semidefiniter symmetrischer Bilinearform (vgl. Klenke [2020] Abschnitt 7.3). □

§02.53 **Lemma**. *Die Abbildung* $F : \mathscr{L}_2(\mu) \to \mathbb{R}$ *ist genau dann stetig und linear, wenn es ein* $f \in \mathscr{L}_2(\mu)$ *gibt mit* $F(g) = \int gf\, d\mu$ *für alle* $g \in \mathscr{L}_2(\mu)$.

§02.54 **Beweis** von Lemma §02.50. siehe Klenke [2020, Satz 7.28, S.174] □

§03 Maße mit Dichten - Satz von Radon-Nikodym

§03.01 **Definition**. Seien $\nu, \mu \in \mathcal{M}(\mathscr{A})$ Maße auf $(\Omega, \mathscr{A})$.

$\nu \ll \mu$: ν heißt *(absolut-)stetig* bezüglich μ, kurz *μ-stetig*, oder *dominiert* durch μ, wenn jede μ-Nullmenge aus $\mathscr{A}$ auch eine ν-Nullmenge ist, also für jedes $A \in \mathscr{A}$ mit $\mu(A) = 0$ gilt auch $\nu(A) = 0$. Die Maße μ und ν heißen *äquivalent* (kurz $\mu \ll\gg \nu$), falls $\nu \ll \mu$ und $\mu \ll \nu$.

$\mu \perp \nu$: μ heißt *singulär* zu ν, kurz *ν-singulär*, wenn es eine μ-Nullmenge $N \in \mathscr{A}$ gibt mit $\nu = \mathbb{1}_N \nu$, also $\nu(A) = \nu(A \cap N)$ für alle $A \in \mathscr{A}$, oder äquivalent dazu $N^c := \Omega \setminus N$ ist eine ν-Nullmenge. □

§03.02 **Bemerkung**. Offenbar gilt genau dann $\mu \perp \nu$, wenn $\Omega_\mu, \Omega_\nu \in \mathscr{A}$ existieren mit $\Omega = \Omega_\mu \biguplus \Omega_\nu$ und $\mu(\Omega_\nu) = \nu(\Omega_\mu) = 0$, und somit genau dann wenn $\nu \perp \mu$. Daher werden Maße $\mu, \nu \in \mathcal{M}(\mathscr{A})$ mit $\mu \perp \nu$ auch zueinader singulär genannt. Die Forderung $\nu \perp \mu$ besagt, dass das Maß $\nu = \mathbb{1}_N \nu$ von einer μ-Nullmenge $N \in \mathscr{A}$ getragen wird. Aus $\nu \ll \mu$ und $\nu \perp \mu$ folgt $\nu(N) = 0$ daher $\nu = 0$. □

§03.03 **Lemma**. *Seien $\nu, \mu \in \mathcal{M}(\mathscr{A})$ Maße auf $(\Omega, \mathscr{A})$. ν heißt totalstetig bezüglich μ, falls es für jedes $\varepsilon \in \mathbb{R}^+_{\setminus 0}$ ein $\delta \in \mathbb{R}^+_{\setminus 0}$ gibt, so dass für jedes $A \in \mathscr{A}$ mit $\mu(A) \leqslant \delta$ gilt auch $\nu(A) \leqslant \varepsilon$. Ist ν totalstetig bezüglich μ, dann ist $\nu \ll \mu$. Ist $\nu \in \mathcal{M}_e(\mathscr{A})$ endlich, so gilt auch die Umkehrung.*

§03.04 **Beweis** von Lemma §03.03. In der Vorlesung. □

§03.05 **Lemma**. *Sind $\nu, \mu \in \mathcal{M}_e(\mathscr{A})$ endliche Maße mit $\nu \leqslant \mu$, so gibt es eine meßbare Funktion $h : \Omega \to [0, 1]$ mit $\nu = h\mu$.*

§03.06 **Beweis** von Lemma §03.05. In der Vorlesung. □

§03.07 **Satz von Radon-Nikodym**. *Seien $\mu \in \mathcal{M}_\sigma(\mathscr{A})$ ein σ-endliches Maß und $\nu \in \mathcal{M}(\mathscr{A})$ ein μ-stetiges Maß auf $(\Omega, \mathscr{A})$, also $\nu \ll \mu$. Dann hat ν eine Dichte $\mathbb{f} = \mathrm{d}\nu/\mathrm{d}\mu \in \overline{\mathscr{A}}^+$ bezüglich μ, d.h. es gilt $\nu = \mathbb{f}\mu$.*

§03.08 **Beweis** von Satz §03.07. In der Vorlesung. □

§03.09 **Bemerkung**. Es seien $\mu, \nu \in \mathcal{M}_\sigma(\mathscr{A})$ σ-endliche Maße und $\mathbb{f} = \mathrm{d}\nu/\mathrm{d}\mu \in \overline{\mathscr{A}}^+$ eine μ-Dichte von $\nu \ll \mu$. Dann folgen aus Lemma §02.17 direkt die üblichen *Kettenregeln*:

(i) Ist $g \in \overline{\mathscr{A}}$ quasiintegriebar, so gilt $\int_A g\, d\nu = \int_A g\mathbb{f}\, d\mu$ für alle $A \in \mathscr{A}$.

(ii) Ist $\rho \in \mathcal{M}_\sigma(\mathscr{A})$ ein σ-endliches Maß mit $\rho \ll \nu \ll \mu$ so gilt $\frac{\mathrm{d}\rho}{\mathrm{d}\mu} = \frac{\mathrm{d}\rho}{\mathrm{d}\nu}\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$ μ-f.ü..

(iii) Ist $h : \Omega \to [0, 1]$ messbar mit $h = \frac{\mathrm{d}\nu}{\mathrm{d}(\nu+\mu)}$ μ-f.ü., so gilt $\frac{\mathrm{d}\nu}{\mathrm{d}\mu} = \frac{h}{1-h}$ μ-f.ü.. □

§03.10 **Beispiel**.

(a) *Stetige Wahrscheinlichkeitsmaße* auf $(\mathbb{R}^k, \mathscr{B}^k)$ (vgl. Definition A01.04 (f)) sind somit gerade die bezüglich des Lebesgue-Maßes λ^k absolutstetigen Wahrscheinlichkeitsmaße mit entsprechender (Radon-Nikodym-) Dichte.

(b) Bezeichnen wir mit ζ_Ω das Zählmaß auf einer abzählbaren Menge Ω, so sind *diskrete Wahrscheinlichkeitsmaße* auf $(\Omega, 2^\Omega)$ (vgl. Definition A01.04 (e)) absolutstetig bezüglich ζ_Ω und ihre Zähldichte entspricht gerade der (Radon-Nikodym-) Dichte bezüglich ζ_Ω. Ist $\Omega \subseteq \mathbb{R}$ und $\mathbb{P}$ ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathscr{B})$ mit Zähldichte $\mathbb{p}$ wie in Beispiel A01.06 (b), dann ist $\mathbb{P}$ absolutstetig bezüglich des Zählmaßes ζ_Ω auf Ω aufgefasst als Maß auf $(\mathbb{R}, \mathscr{B})$ und $\mathbb{p}$ gerade die (Radon-Nikodym-) Dichte von $\mathbb{P}$ bezüglich ζ_Ω. □

§03.11 **Lebesgue'scher Zerlegungssatz**. *Seien $\mu, \nu \in \mathcal{M}_\sigma(\mathscr{A})$ σ-endliche Maße auf $(\Omega, \mathscr{A})$. Dann gibt es genau eine Zerlegung $\nu = \nu_a + \nu_s$ von ν in zwei Maße $\nu_a, \nu_s \in \mathcal{M}(\mathscr{A})$, derart dass $\nu_a \ll \mu$ der μ-stetige Anteil und $\nu_s \perp \mu$ der μ-singuläre Anteil von ν bezüglich μ ist. Dabei sind $\nu_a, \nu_s \in \mathcal{M}_\sigma(\mathscr{A})$ σ-endlich, und genau dann $\nu_a, \nu_s \in \mathcal{M}_e(\mathscr{A})$ endlich, wenn $\nu \in \mathcal{M}_e(\mathscr{A})$ endlich ist. ν_a besitzt eine μ-Dichte $\mathrm{d}\nu_a/\mathrm{d}\mu \in \overline{\mathscr{A}}^+$, die μ-fast überall rellwertig ist.*

§03.12 **Beweis** von Satz §03.11. In der Vorlesung. □

§03.13 **Bemerkung**. Wir können in Satz §03.11 die numerische Funktion $\mathbb{f} = \mathrm{d}\nu_a/\mathrm{d}\mu \in \overline{\mathscr{A}}^+$ durch die reelle Funktion $\tilde{\mathbb{f}} := \mathbb{f}\mathbb{1}_{\{\mathbb{f}\in\mathbb{R}^+\}} \in \mathscr{A}^+$ ersetzen, da $\mathbb{f} = \tilde{\mathbb{f}}$ μ-f.ü.. Mit anderen Worten $\tilde{\mathbb{f}} \in \mathscr{A}^+$ ist ebenfall eine Festlegung der Radon-Nikodym Dichte von ν_a bezüglich μ. Sind $\nu, \mu \in \mathcal{M}_\sigma(\mathscr{A})$ σ-endlich, so werden wir stets $\mathrm{d}\nu_a/\mathrm{d}\mu \in \mathscr{A}^+$ annehmen. Weiterhin definieren wir noch eine numerische Funktion $\mathrm{L} := \tilde{\mathbb{f}}\mathbb{1}_{N^c} + \infty\mathbb{1}_N \in \overline{\mathscr{A}}^+$ mit $\mu(N) = 0 = \nu_s(N^c)$, so ist $\{\mathrm{L} = \infty\} = N$ und die Lebesgue-Zerlegung schreibt sich in der Form $\nu = \mathrm{L}\mu + \mathbb{1}_{\{\mathrm{L}=\infty\}}\nu$, d.h. für alle $A \in \mathscr{A}$ gilt $\nu(A) = \int_A \mathrm{L}\, d\mu + \nu(A \cap \{\mathrm{L} = \infty\})$. □

§03.14 **Definition**. Seien $\nu, \mu \in \mathcal{M}_\sigma(\mathscr{A})$ σ-endliche Maße auf $(\Omega, \mathscr{A})$, wobei nicht notwendig $\nu \ll \mu$ gilt. Dann heißt jede positive numerische Funktion $\mathrm{L} \in \overline{\mathscr{A}}^+$ mit

$$\mu(\mathrm{L} = \infty) = 0 \text{ und } \nu = \mathrm{L}\mu + \mathbb{1}_{\{\mathrm{L}=\infty\}}\nu \tag{03.01}$$

ein *Dichtequotient (DQ)* von ν bezüglich μ. □

§03.15 **Lemma**. *Seien $\nu, \mu \in \mathcal{M}_\sigma(\mathscr{A})$ σ-endliche Maße. Dann ist der Dichtequotient $\mathrm{L} \in \overline{\mathscr{A}}^+$ von ν bezüglich μ eindeutig bis auf Gleichheit $\nu + \mu$-fast überall.*

§03.16 **Beweis** von Lemma §03.15. In der Vorlesung. □

Alternative Formulierung des Satzes von Radon-Nikodym

§03.17 **Definition**. Seien $(\Omega, \mathscr{A}, \mu)$ ein Maßraum und $\mathcal{F} \subseteq \overline{\mathscr{A}}$ eine Menge numerischer $\mathscr{A}$-messbarer Funktionen. Wir bezeichnen eine Funktion $g \in \overline{\mathscr{A}}$ als *μ-wesentliches Supremum* über $\mathcal{F}$, kurz $g = \mu\text{-ess}\sup_{f\in\mathcal{F}} f$, wenn (a) $f \leqslant g$ μ-f.ü. für alle $f \in \mathcal{F}$ gilt, und (b) erfüllt $h \in \overline{\mathscr{A}}$ für alle $f \in \mathcal{F}$: $f \leqslant h$ μ-f.ü., so gilt auch $g \leqslant h$ μ-f.ü.. □

§03.18 **Bemerkung**. Das μ-wesentliche Supremum läßt sich als Erweiterung des üblichen Begriffs des Supremums auffassen. Ist $\mathcal{F}$ abzählbar und $\mu \in \mathcal{M}_\sigma(\mathscr{A})$ σ-endlich, so erfüllt $g := \sup_{f\in\mathcal{F}} f$ die Bedingungen §03.17 (a) und (b), es gilt also $\sup_{f\in\mathcal{F}} f = \mu\text{-ess}\sup_{f\in\mathcal{F}} f$ μ-f.ü.. Ist dagegen $\mathcal{F}$ überabzählbar, etwa $\mathcal{F} = \{\mathbb{1}_{\{x\}}, x \in B\}$ mit überabzählbarem $B \in \mathscr{B}$ und $\lambda(B) \in \mathbb{R}^+_{\setminus 0}$, so ist das λ-wesentliche Supremum vom üblichen Supremum verschieden. In diesem Spezialfall gilt $\sup_{f\in\mathcal{F}} f = \mathbb{1}_B \neq 0 = \lambda\text{-ess}\sup_{f\in\mathcal{F}} f$. □

§03.19 **Lemma**. *Seien $\mu \in \mathcal{M}_\sigma(\mathscr{A})$ und $\mathcal{F} \subseteq \overline{\mathscr{A}}$ eine Menge numerischer Funktionen. Dann gilt:*

(i) *$g := \mu$-ess $\sup_{f\in\mathcal{F}} f$ existiert und ist μ-f.ü. eindeutig, d.h. mit $g \in \overline{\mathscr{A}}$ sind Lösungen von Definition §03.17 (a) und (b) genau diejenigen Funktionen $\tilde{g} \in \overline{\mathscr{A}}$ mit $\mu(\{g \neq \tilde{g}\}) = 0$.*

(ii) *Es gibt eine Folge $(f_n)_{n\in\mathbb{N}}$ aus $\mathcal{F}$ mit $g = \sup_{n\in\mathbb{N}} f_n$ μ-f.ü..*

(iii) *Ist $\mathcal{F}$ aufsteigend filtrierend (für alle $h, k \in \mathcal{F}$ existiert ein $f \in \mathcal{F}$ mit $f \geqslant h \vee k$), so gibt es eine isotone Folge $(f_n)_{n\in\mathbb{N}}$ aus $\mathcal{F}$ mit $f_n \uparrow g$ μ-f.ü..*

§03.20 **Beweis** von Lemma §03.19. Die Aussage wird in Witting [1985, Satz 1.102, S.105] gezeigt. □

§03.21 **Lemma**. *Seien $\mu, \nu \in \mathcal{M}_e(\mathscr{A})$ endliche und zueinander nicht singuläre Maße auf $(\Omega, \mathscr{A})$. Dann gibt es ein $\Omega_o \in \mathscr{A}$ mit $\mu(\Omega_o) \in \mathbb{R}^+_{\setminus 0}$ und ein $\varepsilon \in \mathbb{R}^+_{\setminus 0}$ mit $\varepsilon\mathbb{1}_{\Omega_o}\mu \leqslant \mathbb{1}_{\Omega_o}\nu$.*

§03.22 **Beweis** von Lemma §03.21. Die Aussage wird in Klenke [2020, Lemma 7.46, S.184] mit Hilfe der Hahn-Zerlegung für signierte Maße gezeigt. Einen alternativen Beweis der Aussage ist im Beweis von Bauer [1992, Satz 17.10, S.117] geführt, wobei dieser Bauer [1992, Lemma 17.9, S.114] benutzt. □

§03.23 **Lemma**. *Seien $\nu, \mu \in \mathcal{M}_e(\mathscr{A})$ endlich mit $\nu \leqslant \mu$. Setze $\mathcal{F} := \{f \in \overline{\mathscr{A}}^+ : f\mu \leqslant \nu\}$ und $g := \mu\text{-ess sup}_{f\in\mathcal{F}} f$. Dann gilt $\nu = g\mu$, d.h. g ist eine Festlegung der μ-Dichte von ν.*

§03.24 **Beweis** von Lemma §03.23. In der Vorlesung. □

§04 Maße auf Produkträumen

§04|01 Endliche Produktmaße

§04.01 **Erinnerung**. Seien $(\mathcal{S}_i, \mathscr{S}_i)$, $i \in \mathcal{I}$, messbare Räume mit beliebiger, nicht-leerer Indexmenge $\mathcal{I}$. Die Menge $\mathcal{S}_\mathcal{I} := \times_{i\in\mathcal{I}} \mathcal{S}_i$ aller Abbildungen $(s_i)_{i\in\mathcal{I}} : \mathcal{I} \to \cup_{i\in\mathcal{I}}\mathcal{S}_i$ sodass $s_i \in \mathcal{S}_i$ für alle $i \in \mathcal{I}$ gilt, heißt *Produktraum* oder kartesisches Produkt. Sind alle $\mathcal{S}_i$ gleich, etwa $\mathcal{S}_i = \mathcal{S}$, dann schreiben wir $\mathcal{S}^\mathcal{I} := \mathcal{S}_\mathcal{I}$, im Fall $n := |\mathcal{I}| < \infty$, auch nur kurz $\mathcal{S}^n := \mathcal{S}^\mathcal{I}$. Für jedes $\mathcal{J} \subseteq \mathcal{I}$ bezeichne $\Pi_{\mathcal{S}_\mathcal{J}} : \mathcal{S}_\mathcal{I} \to \mathcal{S}_\mathcal{J}$ mit $(s_i)_{i\in\mathcal{I}} \mapsto (s_j)_{j\in\mathcal{J}}$ die *kanonische Projektion* und speziell für $j \in \mathcal{I}$ die *Koordinantenabbildung* mit $\Pi_{\mathcal{S}_j} : \mathcal{S}_\mathcal{I} \to \mathcal{S}_j$ mit $(s_i)_{i\in\mathcal{I}} \mapsto s_j$, sodass $\times_{i\in\mathcal{I}} E_i = \bigcap_{i\in\mathcal{I}} \Pi_{\mathcal{S}_i}^{-1}(E_i)$ für $E_i \subseteq \mathcal{S}_i$, $i \in \mathcal{I}$. □

§04.02 **Definition**.

(a) Für eine Familie $(\mathscr{A}_i)_{i\in\mathcal{I}}$ von Teil-σ-Algebren von $\mathscr{A}$ mit beliebiger nicht-leerer Indexmenge $\mathcal{I}$ bezeichnet $\bigwedge_{i\in\mathcal{I}} \mathscr{A}_i := \bigcap_{i\in\mathcal{I}} \mathscr{A}_i$ und $\bigvee_{i\in\mathcal{I}} \mathscr{A}_i := \sigma(\bigcup_{i\in\mathcal{I}} \mathscr{A}_i)$ die *größte σ-Algebra*, die in allen $\mathscr{A}_i$, $i \in \mathcal{I}$, enthalten ist, bzw. die *kleinste σ-Algebra*, die alle $\mathscr{A}_i$, $i \in \mathcal{I}$, enthält.

(b) Die Produkt-σ-Algebra $\mathscr{S}_\mathcal{I} := \bigotimes_{i\in\mathcal{I}} \mathscr{S}_i$ auf dem Produktraum $\mathcal{S}_\mathcal{I}$ ist die kleinste σ-Algebra, sodass für jedes $i \in \mathcal{I}$ die Koordinantenabbildung $\Pi_{\mathcal{S}_i}$ $\mathscr{S}_\mathcal{I}$-$\mathscr{S}_i$-messbar ist, d.h.

$$\mathscr{S}_\mathcal{I} = \bigotimes_{i\in\mathcal{I}} \mathscr{S}_i := \bigvee_{i\in\mathcal{I}} \sigma(\Pi_{\mathcal{S}_i}) = \bigvee_{i\in\mathcal{I}} \Pi_{\mathcal{S}_i}^{-1}(\mathscr{S}_i).$$

Sind alle $(\mathcal{S}_i, \mathscr{S}_i)$ gleich, etwa $(\mathcal{S}_i, \mathscr{S}_i) = (\mathcal{S}, \mathscr{S})$, dann schreiben wir $\mathscr{S}^\mathcal{I} := \mathscr{S}_\mathcal{I}$, im Fall $n := |\mathcal{I}| < \infty$, auch nur $\mathscr{S}^n := \mathscr{S}^\mathcal{I}$.

§04.03 **Lemma**. *Für jedes $i \in [\![n]\!]$ sei $\mathscr{E}_i$ ein Erzeuger der σ-Algebra $\mathscr{S}_i$ auf $\mathcal{S}_i$, welcher eine Folge $(E_{ik})_{k\in\mathbb{N}}$ von Mengen mit $E_{ik} \uparrow \mathcal{S}_i$ enthält. Dann wid die Produkt-σ-Algebra $\mathscr{S}_{[\![n]\!]} = \bigotimes_{i\in[\![n]\!]} \mathscr{S}_i$ vom System aller Mengen $\{\times_{i\in[\![n]\!]} E_i : E_i \in \mathscr{E}_i, i \in [\![n]\!]\}$ erzeugt.*

§04.04 **Beweis** von Lemma §04.03. In der Vorlesung. □

§04.05 **Bemerkung**. Seien $n = 2$, $\mathscr{S}_1 = \{\emptyset, \mathcal{S}_1\}$, $\mathscr{E}_1 = \{\emptyset\}$ und $\mathscr{E}_2 = \mathscr{S}_2$ eine mindestens 4 Elemente enthaltende σ-Algebra in einer Menge $\mathcal{S}_2$. Dann ist das System aller Mengen $\{\emptyset \times E : E \in \mathscr{E}_2\}$ kein Erzeuger der Produkt-σ-Algebra $\mathscr{S}_1 \otimes \mathscr{S}_2$. Damit kann auf die einschränkende Vorraussetzung über die Erzeuger in Lemma §04.03 nicht ohne weiteres verzichtet werden. Andererseits wird die Produkt-σ-Algebra $\mathscr{S}_{[\![n]\!]} = \bigotimes_{i\in[\![n]\!]} \mathscr{S}_i$ nach Lemma §04.03 vom System aller Mengen $\{\times_{i\in[\![n]\!]} A_i : A_i \in \mathscr{A}_i, i \in [\![n]\!]\}$ erzeugt. □

§04.06 **Definition**. Sei $\mu_i \in \mathcal{M}(\mathscr{S}_i)$ ein Maß auf $(\mathcal{S}_i, \mathscr{S}_i)$ für jedes $i \in [\![n]\!]$. Ein Maß $\mu_{[\![n]\!]} \in \mathcal{M}(\mathscr{S}_{[\![n]\!]})$ auf $(\mathcal{S}_{[\![n]\!]}, \mathscr{S}_{[\![n]\!]})$ heißt *Produktmaß*, wenn für alle $E_i \in \mathscr{S}_i$, $i \in [\![n]\!]$ gilt $\mu_{[\![n]\!]}(\times_{i\in[\![n]\!]} E_i) = \mu_{[\![n]\!]}(\bigcap_{i\in[\![n]\!]} \Pi_{\mathcal{S}_i}^{-1}(E_i)) = \prod_{i\in[\![n]\!]} \mu_i(E_i)$. In dem Fall schreiben wir $\bigotimes_{i\in[\![n]\!]} \mu_i := \mu_{[\![n]\!]}$. Sind alle $\mu_i = \mu$ gleich, so schreiben wir $\mu^{\otimes n} := \mu_{[\![n]\!]}$. □

§04.07 **Lemma** *(Eindeutigkeit eines endlichen Produktmaßes). Jeder Erzeuger $\mathscr{E}_i$ von $\mathscr{S}_i$, $i \in [\![n]\!]$, sei $\cap$-stabil und enthalte eine Folge $(E_{ik})_{k\in\mathbb{N}}$ von Mengen mit $E_{ik} \uparrow \mathcal{S}_i$ und $\mu_i(E_{ik}) < \infty$ für jedes $k \in \mathbb{N}$. Dann gibt es höchstens ein Maß $\mu_{[\![n]\!]} \in \mathcal{M}(\mathscr{S}_{[\![n]\!]})$ auf $(\mathcal{S}_{[\![n]\!]}, \mathscr{S}_{[\![n]\!]})$ mit $\mu_{[\![n]\!]}\big(\bigtimes_{i\in[\![n]\!]} E_i\big) = \prod_{i\in[\![n]\!]} \mu_i(E_i)$ für alle $E_i \in \mathscr{E}_i$, $i \in [\![n]\!]$.*

§04.08 **Beweis** von Lemma §04.07. In der Vorlesung. □

§04.09 **Bemerkung**. Die Voraussetzungen von Lemma §04.07 haben zur Folge, dass für jedes $i \in [\![n]\!]$ das Maß $\mu_i \in \mathcal{M}_\sigma(\mathscr{S}_i)$ σ-endlich ist. □

§04.10 **Schreibweise**. Seien $(\mathcal{S}_i, \mathscr{S}_i)$, $i \in [\![2]\!]$, Messräume. Für jede Menge $A \subseteq \mathcal{S}_1 \times \mathcal{S}_2$ und jedes $s_1 \in \mathcal{S}_1$ bzw. $s_2 \in \mathcal{S}_2$ wird $A_{s_1} := \{s_2 \in \mathcal{S}_2 : (s_1, s_2) \in A\}$ und $A^{s_2} := \{s_1 \in \mathcal{S}_1 : (s_1, s_2) \in A\}$ der s_1- bzw. s_2-Schnitt von A genannt. □

§04.11 **Lemma**. *Für alle $A \in \mathscr{S}_1 \otimes \mathscr{S}_2$, $s_1 \in \mathcal{S}_1$ und $s_2 \in \mathcal{S}_2$ gilt $A_{s_1} \in \mathscr{S}_2$ und $A^{s_2} \in \mathscr{S}_1$.*

§04.12 **Beweis** von Lemma §04.11. In der Vorlesung. □

§04.13 **Bemerkung**. Sei nun $(\mathcal{S}_i, \mathscr{S}_i, \mu_i)$ ein beliebiger Maßraum, $i \in [\![2]\!]$. Für alle $A \in \mathscr{S}_1 \otimes \mathscr{S}_2$, $s_1 \in \mathcal{S}_1$ und $s_2 \in \mathcal{S}_2$ sind nach Lemma §04.11 dann $\mu_2(A_{s_1})$ und $\mu_1(A^{s_2})$ definiert. □

§04.14 **Lemma**. *Für $i \in [\![2]\!]$ sei $\mu_i \in \mathcal{M}_\sigma(\mathscr{S}_i)$ ein σ-endliches Maß auf $(\mathcal{S}_i, \mathscr{S}_i)$. Dann ist für jedes $A \in \mathscr{S}_1 \otimes \mathscr{S}_2$ die Abbildung $\mu_2(A_\bullet) : s_1 \mapsto \mu_2(A_{s_1})$ bzw. $\mu_1(A^\bullet) : s_2 \mapsto \mu_1(A^{s_2})$ auf $\mathcal{S}_1$ bzw. $\mathcal{S}_2$ definiert und aus $\overline{\mathscr{S}_1}^+$ bzw. $\overline{\mathscr{S}_2}^+$ (positiv numerisch).*

§04.15 **Beweis** von Lemma §04.11. In der Vorlesung. □

§04.16 **Satz** *(Existenz eines Produktmaßes). Sei jedes $\mu_i \in \mathcal{M}_\sigma(\mathscr{S}_i)$ ein σ-endliches Maß auf $(\mathcal{S}_i, \mathscr{S}_i)$, $i \in [\![2]\!]$. Dann existiert genau ein Produktmaß $\mu_{[\![2]\!]}$ auf $(\mathcal{S}_{[\![2]\!]}, \mathscr{S}_{[\![2]\!]})$. Dabei ist $\mu_{[\![2]\!]} \in \mathcal{M}_\sigma(\mathscr{S}_{[\![2]\!]})$ auch σ-endlich und für jedes $A \in \mathscr{S}_{[\![2]\!]}$ gilt $\mu_1\big(\mu_2(A_\bullet)\big) = \mu_{[\![2]\!]}(A) = \mu_2\big(\mu_1(A^\bullet)\big)$.*

§04.17 **Beweis** von Satz §04.16. In der Vorlesung. □

§04.18 **Bemerkung**. Die letzte Aussage kann problemlos auf endlich viele Faktoren ausgedehnt werden. Dabei ist zu beachten, dass die Produkte beliebig umgeklammert werden dürfen. Formal identifizieren wir die Produktmengen $\mathcal{S}_{[\![n-1]\!]} \times \mathcal{S}_n$ und $\mathcal{S}_{[\![n]\!]}$ wie üblich mit Hilfe der Bijektion $((s_i)_{i\in[\![n-1]\!]}, s_n) \mapsto (s_i)_{i\in[\![n]\!]}$. Die somit vereinbarte Gleichheit der Mengen impliziert dann sofort die Gleichheit der entsprechenden Produkte von σ-Algebren $\mathscr{S}_{[\![n-1]\!]} \otimes \mathscr{S}_n$ und $\mathscr{S}_{[\![n]\!]}$. Zusammenfassend erhält man allgemein die Assoziativität der Produktbildung $(\bigotimes_{i\in[\![m]\!]} \mathscr{S}_i) \otimes (\bigotimes_{i\in[\![n-m]\!]} \mathscr{S}_{m+i}) = \bigotimes_{i\in[\![n]\!]} \mathscr{S}_i$ für $m \in [\![n-1]\!]$. □

§04.19 **Korollar** *(Existenz eines Produktmaßes). Sei jedes $\mu_i \in \mathcal{M}_\sigma(\mathscr{S}_i)$ ein σ-endliches Maß auf $(\mathcal{S}_i, \mathscr{S}_i)$, $i \in [\![n]\!]$. Dann existiert genau ein Produktmaß $\mu_{[\![n]\!]}$ auf $(\mathcal{S}_{[\![n]\!]}, \mathscr{S}_{[\![n]\!]})$. Dabei ist $\mu_{[\![n]\!]} \in \mathcal{M}_\sigma(\mathscr{S}_{[\![n]\!]})$ auch σ-endlich.*

§04.20 **Beweis** von Korollar §04.19. In der Vorlesung. □

§04.21 **Bemerkung**. Für nicht notwendig σ-endliche Maße lässt sich zwar noch die Existenz, nicht mehr aber die Eindeutigkeit eines Produktmaßes beweisen. □

§04|02 Projektive Familie

§04.22 **Erinnerung**. Sind (Ω_i, τ_i), $i \in \mathcal{I}$, topologische Räume, so ist die Produkttopologie τ auf $\Omega_{\mathcal{I}}$ die gröbste Topologie, bezüglich der alle Koordinatenabbildungen Π_{Ω_i} stetig sind. □

§04.23 **Lemma**. *Sei $\mathcal{I}$ abzählbar, für jedes $i \in \mathcal{I}$ sei $\mathcal{S}_i$ separabler und vollständiger metrischer Raum (polnisch) mit Borel-σ-Algebra $\mathscr{B}_i := \mathscr{B}_{\mathcal{S}_i}$ und sei $\mathscr{B}_{\mathcal{S}_\mathcal{I}}$ die Borel-σ-Algebra bezüglich der Produkttopologie auf $\mathcal{S}_\mathcal{I} = \times_{i\in\mathcal{I}} \mathcal{S}_i$. Dann gilt $\mathscr{B}_{\mathcal{S}_\mathcal{I}} = \mathscr{B}_\mathcal{I} = \bigotimes_{i\in\mathcal{I}} \mathscr{B}_i$, also insbesondere $\mathscr{B}_{\mathbb{R}^n} = \mathscr{B}^n$.*

§04.24 **Beweis** von Lemma §04.03. siehe Klenke [2020, Satz 14.8, S.303] □

§04.25 **Definition**. Für jedes $\mathcal{J} \subseteq \mathcal{I}$ heißt $\mathcal{Z}_\mathcal{J} := \{\Pi_{\mathcal{S}_\mathcal{J}}^{-1}(E) : E \in \mathscr{S}_\mathcal{J}\} \subseteq \mathscr{S}_\mathcal{I}$ Menge der *Zylindermengen* mit Basis $\mathcal{J}$. Weiterhin heißt $\mathcal{Z}_\mathcal{J}^R := \{\Pi_{\mathcal{S}_\mathcal{J}}^{-1}(E_\mathcal{J}) : E_\mathcal{J} = \times_{j\in\mathcal{J}} E_j \in \mathscr{S}_\mathcal{J}\} \subseteq \mathscr{S}_\mathcal{I}$ Menge der *Rechteckzylinder* mit Basis $\mathcal{J}$. Für jedes $i \in \mathcal{I}$ sei weiterhin $\mathscr{E}_i \subseteq \mathscr{S}_i$. Dann schreiben wir $\mathcal{Z}_\mathcal{J}^{\mathscr{E},R} := \{\Pi_{\mathcal{S}_\mathcal{J}}^{-1}(E_\mathcal{J}) : E_\mathcal{J} = \times_{j\in\mathcal{J}} E_j, E_j \in \mathscr{E}_j, j \in \mathcal{J}\} \subseteq \mathscr{S}_\mathcal{I}$. Wir setzen weiterhin $\mathcal{Z} := \bigcup\{\mathcal{Z}_\mathcal{J} : \mathcal{J} \subseteq \mathcal{I} \text{ endlich}\}$ und definieren analog $\mathcal{Z}^R$ und $\mathcal{Z}^{\mathscr{E},R}$.

§04.26 **Bemerkung**. Jedes $\mathcal{Z}_\mathcal{J}$ ist eine σ-Algebra. $\mathcal{Z}$ ist eine Algebra und es gilt $\mathscr{S}_\mathcal{I} = \sigma(\mathcal{Z})$. Weiterhin, ist jedes $\mathscr{E}_i$ $\cap$-stabil, so ist $\mathcal{Z}^{\mathscr{E},R}$ auch $\cap$-stabil (Übung). □

§04.27 **Lemma**. *Für jedes $i \in \mathcal{I}$ sei $\mathscr{E}_i$ ein Erzeuger von $\mathscr{S}_i$.*

(i) *Für jedes abzählbare $\mathcal{J} \subseteq \mathcal{I}$ gilt $\mathscr{S}_\mathcal{J} = \sigma(\times_{j\in\mathcal{J}} E_j : E_j \in \mathscr{E}_j \cup \{\mathcal{S}_j\}, j \in \mathcal{J})$*

(ii) *Es gilt $\mathscr{S}_\mathcal{I} = \sigma(\mathcal{Z}^R) = \sigma(\mathcal{Z}^{\mathscr{E},R})$.*

(iii) *Seien (A1) $\mu \in \mathcal{M}_\sigma(\mathscr{S}_\mathcal{I})$ ein σ-endliches Maß auf $(\mathcal{S}_\mathcal{I}, \mathscr{S}_\mathcal{I})$, (A2) jedes $\mathscr{E}_i$ $\cap$-stabil und (A3) $\mathcal{Z}^{E,R}$ enthalte eine Folge $(E_n)_{n\in\mathbb{N}}$ von Mengen mit $E_n \uparrow \mathcal{S}_\mathcal{I}$ und $\mu(E_n) < \infty$ für jedes $n \in \mathbb{N}$. Dann ist μ durch die Angabe von $\mu(A)$ für jedes $A \in \mathcal{Z}^{\mathscr{E},R}$ eindeutig festgelegt.*

§04.28 **Beweis** von Lemma §04.27. Übung. □

§04.29 **Anmerkung**. Die Bedingung (A3) in Lemma §04.27 (iii) ist erfüllt, wenn $\mu \in \mathcal{M}_e(\mathscr{S}_\mathcal{I})$ endlich ist und für jedes $i \in \mathcal{I}$ ist $\mathcal{S}_i \in \mathscr{E}_i$ (vgl. Lemma §01.24). Insbesondere, ist ein Produktmaß $\mu_\mathcal{I} \in \mathcal{M}(\mathscr{S}_\mathcal{I})$ auf $(\mathcal{S}_\mathcal{I}, \mathscr{S}_\mathcal{I})$ eindeutig festgelegt, wenn es σ-endlich ist. □

§04.30 **Schreibweise**. Für $\mathcal{J} \subseteq \mathcal{K} \subseteq \mathcal{I}$ definiere die Koordinatenabbildung $\Pi_\mathcal{J}^\mathcal{K} \colon \mathcal{S}_\mathcal{K} \to \mathcal{S}_\mathcal{J}$ mit $(s_k)_{k\in\mathcal{K}} \mapsto (s_j)_{j\in\mathcal{J}}$, wobei offensichtlich $\Pi_\mathcal{J} := \Pi_{\mathcal{S}_\mathcal{J}} = \Pi_\mathcal{J}^\mathcal{I}$ gilt. □

§04.31 **Definition**. Für jede endliche Teilmenge $\mathcal{J} \subseteq \mathcal{I}$ sei $\mathbb{P}_\mathcal{J} \in \mathcal{W}(\mathscr{S}_\mathcal{J})$ ein Wahrscheinlichkeitsmaß auf $(\mathcal{S}_\mathcal{J}, \mathscr{S}_\mathcal{J})$. Die Familie $(\mathbb{P}_\mathcal{J}, \mathcal{J} \subseteq \mathcal{I} \text{ endlich})$ heißt *projektiv* oder *konsistent*, falls für alle endlichen $\mathcal{J} \subseteq \mathcal{K} \subseteq \mathcal{I}$ gilt $\mathbb{P}_\mathcal{J} = \mathbb{P}_\mathcal{K} \circ (\Pi_\mathcal{J}^\mathcal{K})^{-1}$. □

§04.32 **Kolmogorov'scher Erweiterungssatz**. *Sei $\mathcal{I}$ eine beliebige nicht-leere Indexmenge und für jedes $i \in \mathcal{I}$ sei $\mathcal{S}_i$ ein separabler und vollständiger metrischer Raum (polnisch) mit Borel-σ-Algebra $\mathscr{B}_i := \mathscr{B}_{\mathcal{S}_i}$. Sei $(\mathbb{P}_\mathcal{J} \in \mathcal{W}(\mathscr{B}_\mathcal{J}), \mathcal{J} \subseteq \mathcal{I} \text{ endlich})$ eine projektive Familie von I Wahrscheinlichkeitsmaßen. Dann gibt es ein eindeutig bestimmtes Wahrscheinlichkeitsmaß $\mathbb{P} \in \mathcal{W}(\mathscr{B}_\mathcal{I})$ auf $(\mathcal{S}_\mathcal{I}, \mathscr{B}_\mathcal{I})$ mit $\mathbb{P}_\mathcal{J} = \mathbb{P} \circ \Pi_{\mathcal{S}_\mathcal{J}}^{-1}$ für jedes endliche $\mathcal{J} \subseteq \mathcal{I}$.*

§04.33 **Beweis** von Satz §04.32. siehe Klenke [2020, Satz 14.39, S. 319] □

§04.34 **Definition**. Sei $\mathbb{P}_i \in \mathcal{W}(\mathscr{S}_i)$ ein Wahrscheinlichkeitsmaß auf $(\mathcal{S}_i, \mathscr{S}_i)$ für jedes $i \in \mathcal{I}$. Ein Wahrscheinlichkeitsmaß $\mathbb{P}_\mathcal{I} \in \mathcal{W}(\mathscr{S}_\mathcal{I})$ auf $(\mathcal{S}_\mathcal{I}, \mathscr{S}_\mathcal{I})$ heißt *Produktmaß*, wenn für alle endlichen $\mathcal{J} \subseteq \mathcal{I}$ und $E_j \in \mathscr{S}_j, j \in \mathcal{J}$ gilt $\mathbb{P}_\mathcal{I}\big(\bigcap_{j\in\mathcal{J}} \Pi_{\mathcal{S}_j}^{-1}(E_j)\big) = \prod_{j\in\mathcal{J}} \mathbb{P}_j(E_j)$. In dem Fall schreiben wir $\bigotimes_{i\in\mathcal{I}} \mathbb{P}_i := \mathbb{P}_\mathcal{I}$. Sind alle $\mathbb{P}_i = \mathbb{P}$ gleich, so schreiben wir $\mathbb{P}^{\otimes\mathcal{I}} := \mathbb{P}_\mathcal{I}$ und im Fall $n := |\mathcal{I}| \in \mathbb{N}$ auch $\mathbb{P}^{\otimes n} := \mathbb{P}_\mathcal{I}$. □

§04.35 **Bemerkung**. Sei $\mathcal{I}$ eine beliebige nicht-leere Indexmenge. Für jedes $i \in \mathcal{I}$ sei $\mathcal{S}_i$ ein separabler und vollständiger metrischer Raum (polnisch) mit Borel-σ-Algebra $\mathscr{B}_i := \mathscr{B}_{\mathcal{S}_i}$ und $\mathbb{P}_i \in \mathcal{W}(\mathscr{B}_i)$ ein Wahrscheinlichkeitsmaß auf $(\mathcal{S}_i, \mathscr{B}_i)$. Für jedes endliche $\mathcal{J} \subseteq \mathcal{I}$ sei $\mathbb{P}_\mathcal{J} := \bigotimes_{j\in\mathcal{J}} \mathbb{P}_j$ das Produktmaß der $\mathbb{P}_j$, $j \in \mathcal{J}$. Offenbar ist die Familie $(\mathbb{P}_\mathcal{J}, \mathcal{J} \subseteq \mathcal{I} \text{ endlich})$ *projektiv*. Unter Verwendung von Satz §04.32 existiert also genau ein Produktmaß $\mathbb{P} := \bigotimes_{i\in\mathcal{I}} \mathbb{P}_i \in \mathcal{W}(\mathscr{B}_\mathcal{I})$ auf $(\mathcal{S}_\mathcal{I}, \mathscr{B}_\mathcal{I})$. Unter $\mathbb{P}$ ist die Familie der Koordinatenabbildungen unabhängig, $\perp\!\!\!\perp_{i\in\mathcal{I}} \Pi_{\mathcal{S}_i}$ (siehe Abschnitt A03). □

§04|03 Integration bezüglich Produktmaßen

§04.36 **Schreibweise**. Bei gegebener Abbildung $h : \mathcal{S}_1 \times \mathcal{S}_2 \to \mathcal{S}_3$ bezeichnen wir für jedes $s_1 \in \mathcal{S}_1$ und $s_2 \in \mathcal{S}_2$ mit $h_{s_1} : \mathcal{S}_2 \to \mathcal{S}_3$ bzw. $h^{s_2} : \mathcal{S}_1 \to \mathcal{S}$ die Abbildung $s_2 \mapsto h_{s_1}(s_2) := h(s_1, s_2)$ bzw. $s_1 \mapsto h^{s_2}(s_1) := h(s_1, s_2)$. □

§04.37 **Lemma**. *Für beliebige Messräume $(\mathcal{S}_i, \mathscr{S}_i)$, $i \in [\![3]\!]$ und jede $\mathscr{S}_1 \otimes \mathscr{S}_2$-$\mathscr{S}_3$-messbare Abbildung $h : \mathcal{S}_1 \times \mathcal{S}_2 \to \mathcal{S}_3$ ist für jedes $s_1 \in \mathcal{S}_1$ und $s_2 \in \mathcal{S}_2$ die Funktion h_{s_1} $\mathscr{S}_2$-$\mathscr{S}_3$-messbar und die Funktion h^{s_2} $\mathscr{S}_1$-$\mathscr{S}_3$-messbar*

§04.38 **Beweis** von Lemma §04.37. In der Vorlesung. □

§04.39 **Satz von Tonelli**. *Seien $(\mathcal{S}_i, \mathscr{S}_i, \mu_i)$, $i \in [\![2]\!]$, σ-endliche Maßräume und sei $h \in \overline{\mathscr{S}_1 \otimes \mathscr{S}_2}^+$ positiv numerisch. Dann ist $\mu_1(h^\bullet) : \mathcal{S}_2 \to \overline{\mathbb{R}}^+$ mit $s_2 \mapsto \mu_1(h^{s_2})$ aus $\overline{\mathscr{S}_2}^+$ und $\mu_2(h_\bullet) : \mathcal{S}_1 \to \overline{\mathbb{R}}^+$ mit $s_1 \mapsto \mu_2(h_{s_1})$ aus $\overline{\mathscr{S}_1}^+$. Es gilt*

$$\begin{aligned}\mu_1 \otimes \mu_2(h) = \mu_2(\mu_1(h^\bullet)) = \int \mu_1(h^{s_2})\mu_2(ds_2) &= \int\int h(s_1, s_2)\mu_1(ds_1)\mu_2(ds_2)\\ &= \int \mu_2(h_{s_1})\mu_1(ds_1) = \mu_1(\mu_2(h_\bullet))\end{aligned}$$

§04.40 **Beweis** von Satz §04.39. In der Vorlesung. □

§04.41 **Definition**. Seien $(\Omega, \mathscr{A}, \mu)$ ein Maßraum, $(\mathcal{S}, \mathscr{S})$ ein Messraum und $N \in \mathscr{A}$ eine μ-Nullmenge. Eine Abbildung $h : N^c := \Omega \setminus N \to \mathcal{S}$ heißt *μ-fast überall definiert* sowie *$\mathscr{A}$-$\mathscr{S}$-messbar*, falls $h^{-1}(\mathscr{S}) \subseteq \mathscr{A}$ gilt.

§04.42 **Bemerkung**. Sind $h, g \in \overline{\mathscr{A}}$ μ-fast-überall endlich, so ist die Funktion $g - h$ μ-fast überall definiert und $\mathscr{A}$-$\overline{\mathscr{B}}$-messbar. Insbesondere gilt dies, wenn g und h μ-integrierbar sind. Ist f nun $\overline{\mathbb{R}}$-wertig, μ-fast-überall definiert mit μ-Nullmenge N und $\mathscr{A}$-$\overline{\mathscr{B}}$-messbar, so können wir $\tilde{f}(\omega) := 0$ für $\omega \in N$ und andernfalls $\tilde{f}(\omega) := f(\omega)$ definieren. Dann ist $\tilde{f} \in \overline{\mathscr{A}}$ numerisch. Ist $\tilde{f}$ weiterhin μ-integrierbar, so definieren wir für f das Integral $\mu(f) = \int f\, d\mu := \int \tilde{f}\, d\mu$. □

§04.43 **Korollar** *(Satz von Fubini)*. *Seien $(\mathcal{S}_i, \mathscr{S}_i, \mu_i)$, $i \in [\![2]\!]$, σ-endliche Maßräume und $h \in \mathscr{L}_1(\mu_1 \otimes \mu_2)$. Dann ist $\mu_2(h_\bullet) : s_1 \mapsto \mu_2(h_{s_1})$ μ_1-fast überall definiert und $\mathscr{S}_1$-$\overline{\mathscr{B}}$-messbar und $\mu_1(h^\bullet) : s_2 \mapsto \mu_1(h^{s_2})$ μ_2-fast überall definiert und $\mathscr{S}_2$-$\overline{\mathscr{B}}$-messbar. Es gilt*

$$\mu_2(\mu_1(h^\bullet)) = \int \mu_1(h^{s_2})\mu_2(ds_2) = \mu_1 \otimes \mu_2(h) = \int \mu_2(h_{s_1})\mu_1(ds_1) = \mu_1(\mu_2(h_\bullet)).$$

§04.44 **Beweis** von Korollar §04.43. In der Vorlesung. □

§04.45 **Bemerkung**. Die letzten Aussagen können wie in Bemerkung §04.18 problemlos auf endlich viele Faktoren ausgedehnt werden. □

§04.46 **Satz**. *Für jedes $i \in [\![n]\!]$ sei $(\mathcal{S}_i, \mathscr{S}_i, \mu_i)$ ein σ-endlicher Maßraum, $\mathfrak{f}_i \in \mathscr{S}_i^+$ und $\nu_i := \mathfrak{f}_i\mu_i$. Dann ist das Produktmaß $\nu_{[\![n]\!]} = \bigotimes_{i\in[\![n]\!]} \nu_i$ definiert, und absolut stetig bezüglich des Produktmaßes $\mu_{[\![n]\!]} = \bigotimes_{i\in[\![n]\!]} \mu_i$ mit Produktdichte $\prod_{i\in[\![n]\!]} \mathfrak{f}_i$, d.h. es gilt $\nu_{[\![n]\!]} = (\prod_{i\in[\![n]\!]} \mathfrak{f}_i)\mu_{[\![n]\!]}$.*

§04.47 **Beweis** von Lemma §04.49. In der Vorlesung. □

§04.48 **Erinnerung**. Seien nun $\nu = \mathbb{P}_0$ und $\mu = \mathbb{P}_1$ Wahrscheinlichkeitsmaße auf $(\mathcal{S}, \mathscr{S})$, wobei nicht notwendig $\mathbb{P}_0 \ll \mathbb{P}_1$ gilt. Dann heißt jede positive, numerische Funktion $\mathrm{L} \in \overline{\mathscr{S}}^+$ mit $\mathbb{P}_0 = \mathrm{L}\mathbb{P}_1 + \mathbb{1}_{\{\mathrm{L}=\infty\}}\mathbb{P}_0$ und $\mathbb{P}_1(\mathrm{L} \in \mathbb{R}^+) = 1$ Dichtequotient von $\mathbb{P}_0$ bezüglich $\mathbb{P}_1$ (vgl. Definition §03.14). Bezeichnet $\mu \in \mathcal{M}_\sigma(\mathscr{S})$ ein σ-endliches Maß mit $\mathbb{P}_i \ll \mu$, $i \in [\![2]\!]$, (z.Bsp. das endliche Maß $\mu = \mathbb{P}_0 + \mathbb{P}_1$) und bezeichnen $\mathfrak{f}_i \in \mathscr{S}^+$ μ-Dichten von $\mathbb{P}_i$, $i \in [\![2]\!]$, so ist $\mathrm{L}^\star := \frac{\mathfrak{f}_0}{\mathfrak{f}_1}\mathbb{1}_{\{\mathfrak{f}_1 \in \mathbb{R}^+_{\setminus 0}\}} + \infty \mathbb{1}_{\{\mathfrak{f}_1=0, \mathfrak{f}_0 \in \mathbb{R}^+_{\setminus 0}\}} \in \overline{\mathscr{S}}^+$ eine spezielle Festlegung des Dichtequotienten. In dem speziellen Fall $\mathbb{P}_0 \ll \mathbb{P}_1$ stimmt der Dichtequotient von $\mathbb{P}_0$ bezüglich $\mathbb{P}_1$ mit der $\mathbb{P}_1$-Dichte von $\mathbb{P}_0$ überein und ist $\mathbb{P}_1$-bestimmt. □

§04.49 **Lemma**. *Für jedes $i \in [\![n]\!]$ seien $\mathbb{P}_{0,i}, \mathbb{P}_{1,i} \in \mathcal{W}(\mathscr{S}_i)$ Wahrscheinlichkeitsmaße auf $(\mathcal{S}_i, \mathscr{S}_i)$ mit Dichtequotient L_i von $\mathbb{P}_{0,i}$ bezüglich $\mathbb{P}_{1,i}$. Dann ist das Produkt $\mathrm{L}(x) := \prod_{i\in[\![n]\!]} \mathrm{L}_i(x_i)$ für $x = (x_i)_{i\in[\![n]\!]}$ eine Festlegung des Dichtequotienten von $\mathbb{P}_0 := \bigotimes_{i\in[\![n]\!]} \mathbb{P}_{0,i}$ bezüglich $\mathbb{P}_1 := \bigotimes_{i\in[\![n]\!]} \mathbb{P}_{1,i}$.*

§04.50 **Beweis** von Lemma §04.49. In der Vorlesung. □

§04|04 Integration bezüglich Übergangskernen

§04.51 **Definition**. Seien $(\Omega_1, \mathscr{A}_1)$ und $(\Omega_2, \mathscr{A}_2)$ messbare Räume. Eine Abbildung $\kappa : \Omega_1 \times \mathscr{A}_2 \to \overline{\mathbb{R}}^+$ heißt (σ-)endlicher *Übergangskern* von $(\Omega_1, \mathscr{A}_1)$ nach $(\Omega_2, \mathscr{A}_2)$, falls sie die folgenden zwei Bedingungen erfüllt:

(Ük1) für alle $\omega_1 \in \Omega_1$ ist $\kappa_{\omega_1} : \mathscr{A}_2 \to \overline{\mathbb{R}}^+$ mit $A_2 \mapsto \kappa_{\omega_1}(A_2) := \kappa(\omega_1, A_2)$ ein (σ-)endliches Maß auf $(\Omega_2, \mathscr{A}_2)$, kurz $\kappa_{\omega_1} \in \mathcal{M}_e(\mathscr{A}_2)$ bzw. $\kappa_{\omega_1} \in \mathcal{M}_\sigma(\mathscr{A}_2)$;

(Ük2) für alle $A_2 \in \mathscr{A}_2$ ist $\kappa^{A_2} : \Omega_1 \to \overline{\mathbb{R}}^+$ mit $\omega_1 \mapsto \kappa^{A_2}(\omega_1) := \kappa(\omega_1, A_2)$ eine positive, numerische $\mathscr{A}_1$-messbare Abbildung, kurz $\kappa^{A_2} \in \overline{\mathscr{A}_i}^+$.

Ist für jedes $\omega_1 \in \Omega_1$ das Maß in (Ük1) ein Wahrscheinlichkeitmaß, $\kappa_{\omega_1} \in \mathcal{W}(\mathscr{A}_2)$, so heißt κ *Markovkern*. □

§04.52 **Bemerkung**. Es genügt die Bedingung (Ük2) nur für Mengen aus einem $\cap$-stabilen Erzeuger $\mathscr{E}$ von $\mathscr{A}_2$, der Ω_2 oder eine Folge $(E_n)_{n\in\mathbb{N}}$ von Mengen mit $E_n \uparrow \Omega_2$ enthält, zu fordern. Es ist dann nämlich $\mathscr{D} := \{A_2 \in \mathscr{A}_2 : \kappa^{A_2} \in \overline{\mathscr{A}_1}^+\}$ ein Dynkin-System (Übung) mit $\mathscr{E} \subseteq \mathscr{D} \subseteq \mathscr{A}_2$ und aus π-λ-Satz §01.10 folgt $\mathscr{D} = \sigma(\mathscr{E}) = \mathscr{A}_2$. □

§04.53 **Lemma**. *Sei κ ein endlicher Übergangskern von $(\Omega_1, \mathscr{A}_1)$ nach $(\Omega_2, \mathscr{A}_2)$, und sei $h \in \overline{\mathscr{A}_1 \otimes \mathscr{A}_2}^+$ positiv numerisch. Dann ist die Funktion $\kappa_\bullet(h_\bullet) : \Omega_1 \to \overline{\mathbb{R}}^+$ mit $\omega_1 \mapsto \kappa_{\omega_1}(h_{\omega_1}) = \int h_{\omega_1} d\kappa_{\omega_1}$ wohldefiniert und aus $\overline{\mathscr{A}_1}^+$.*

§04.54 **Beweis** von Lemma §04.53. In der Vorlesung. □

§04.55 **Schreibweise**. Für $\mathbb{1}_A \in (\mathscr{A}_1 \otimes \mathscr{A}_2)^+$, also $A \in \mathscr{A}_1 \otimes \mathscr{A}_2$, ist nach Lemma §04.53 die Funktion $\kappa_\bullet(A_\bullet) = \kappa_\bullet((\mathbb{1}_A)_\bullet) : \Omega_1 \to \overline{\mathbb{R}}^+$ mit $\omega_1 \mapsto \kappa_{\omega_1}(A_{\omega_1}) = \kappa_{\omega_1}((\mathbb{1}_A)_{\omega_1})$ wohldefiniert und aus $\overline{\mathscr{A}_1}^+$. □

§04.56 **Lemma**. *Sei $(\Omega_1, \mathscr{A}_1, \mu_1)$ ein endlicher Maßraum, $(\Omega_2, \mathscr{A}_2)$ ein Messraum und κ ein endlicher Übergangskern von $(\Omega_1, \mathscr{A}_1)$ nach $(\Omega_2, \mathscr{A}_2)$. Dann existiert ein eindeutig bestimmtes, σ-endliches Maß $\mu \odot \kappa \in \mathcal{M}_\sigma(\mathscr{A}_1 \otimes \mathscr{A}_2)$ auf dem Produktraum $(\Omega_1 \times \Omega_2, \mathscr{A}_1 \otimes \mathscr{A}_2)$ mit $\mu \odot \kappa(A) = \mu(\kappa_\bullet(A_\bullet))$ für $A \in \mathscr{A}_1 \otimes \mathscr{A}_2$, wobei für alle $A_1 \in \mathscr{A}_1$ und $A_2 \in \mathscr{A}_2$ gilt*

$$\mu \odot \kappa(A_1 \times A_2) = \mu(\mathbb{1}_{A_1}\kappa^{A_2}) = \int_{A_1} \kappa^{A_2}\, d\mu = \int_{A_1} \kappa(\omega_1, A_2)\mu(d\omega_1).$$

Ist κ ein Markovkern und μ ein Wahrscheinlichkeitsmaß, so ist $\mu \odot \kappa$ ein Wahrscheinlichkeitsmaß.

§04.57 **Beweis** von Lemma §04.56. In der Vorlesung. □

§04.58 **Satz von Tonelli/Fubini für Übergangskerne.** *Sei* $(\Omega_1, \mathscr{A}_1, \mu)$ *ein endlicher Maßraum,* $(\Omega_2, \mathscr{A}_2)$ *ein Messraum und* κ *ein endlicher Übergangskern von* $(\Omega_1, \mathscr{A}_1)$ *nach* $(\Omega_2, \mathscr{A}_2)$*. Ist* $h \in \overline{\mathscr{A}_1 \otimes \mathscr{A}_2}^+$ *oder* $h \in \mathscr{L}_1(\mu \odot \kappa)$ *dann gilt*

$$\mu \odot \kappa(h) = \mu(\kappa_{\bullet}(h_{\bullet})) = \int \kappa_{\omega_1}(h_{\omega_1})\mu(d\omega_1) = \int \Big(\int h_{\omega_1}\, d\kappa_{\omega_1} \Big)\mu(d\omega_1)$$
$$= \int \int h(\omega_1, \omega_2)\kappa(\omega_1, d\omega_2)\mu(d\omega_1).$$

§04.59 **Beweis** von Satz §04.58. In der Vorlesung. □

§04.60 **Schreibweise.** Betrachte einen Wahrscheinlichkeitsraum $(\Omega_1, \mathscr{A}_1, \mathbb{P})$, einen Messraum $(\Omega_2, \mathscr{A}_2)$, einen Markovkern κ von $(\Omega_1, \mathscr{A}_1)$ nach $(\Omega_2, \mathscr{A}_2)$ und mit Lemma §04.56 das eindeutig bestimmte Wahrscheinlichkeitsmaß $\mathbb{P} \odot \kappa$ auf - $(\Omega_1 \times \Omega_2, \mathscr{A}_1 \otimes \mathscr{A}_2)$. Dann bezeichnen wir mit

$$\kappa\mathbb{P}(A_2) := \mathbb{P}\big(\kappa^{A_2}\big) = \int \kappa^{A_2}\, d\mathbb{P} = \int \kappa(\omega_1, A_2)\mathbb{P}(d\omega_1), \quad \text{für alle } A_2 \in \mathscr{A}_2$$

die durch $\mathbb{P} \odot \kappa$ auf $(\Omega_2, \mathscr{A}_2)$ induzierte Randverteilung $\kappa\mathbb{P} \in \mathcal{W}(\mathscr{A}_2)$. □

Kapitel 2

Bedingte Erwartung

§05 Diskret- oder stetig-verteilte Zufallsvariablen

§05.01 **Erinnerung**. Sei $(\Omega, \mathscr{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Für $B \in \mathscr{A}$ mit $\mathbb{P}(B) \in \mathbb{R}^+_{\setminus 0}$ heißt $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$ die *bedingte Wahrscheinlichkeit* von A bei gegebenem B. □

§05.02 **Eigenschaft**.

(i) *Sei (S, X) eine diskret-verteilte Zufallsvariable mit Werten in $(\mathcal{S} \times \mathcal{X}, 2^{\mathcal{S} \times \mathcal{X}})$ und gemeinsamer Zähldichte $\mathbb{p}^{(S,X)} : \mathcal{S} \times \mathcal{X} \to [0,1]$ (bezüglich des Produktzählmaßes $\zeta_{\mathcal{S}} \otimes \zeta_{\mathcal{X}}$). Für jedes $s \in \mathcal{S}$ mit strikt-positiver Randzähldichte von S in s, also $\mathbb{p}^S(s) \in \mathbb{R}^+_{\setminus 0}$, ist die Abbildung*

$$\mathbb{p}^{X|S=s} : \mathcal{X} \to [0,1] \text{ mit } x \mapsto \mathbb{p}^{X|S=s}(x) := \frac{\mathbb{p}^{(S,X)}(x,s)}{\mathbb{p}^S(s)} \tag{05.01}$$

eine Zähldichte (bezüglich des Zählmaßes $\zeta_{\mathcal{X}}$).

(ii) *Sei (S, X) ein stetig-verteilter Zufallsvektor mit gemeinsamer Dichte $\mathbb{f}^{(S,X)} : \mathbb{R}^{m+n} \to \mathbb{R}^+$ (bezüglich des Lebesgue-Maßes λ^{m+n}). Für jedes $s \in \mathbb{R}^m$ mit strikt-positiver Randdichte von S in s, also $\mathbb{f}^S(s) \in \mathbb{R}^+_{\setminus 0}$, ist*

$$\mathbb{f}^{X|S=s} : \mathbb{R}^n \to \mathbb{R}^+ \text{ mit } x \mapsto \mathbb{f}^{X|S=s}(x) := \frac{\mathbb{f}^{(S,X)}(s,x)}{\mathbb{f}^S(s)} \tag{05.02}$$

eine Dichte (bezüglich des Lebesgue-Maßes λ^n). □

§05.03 **Bemerkung**. Da die Mengen $\{s \in \mathcal{S} : \mathbb{p}^S(s) = 0\}$ bzw. $\{s \in \mathbb{R}^m : \mathbb{f}^S(s) = 0\}$ Nullmengen bzgl. der Verteilung $\mathbb{P}^S$ von S sind, sind die Abbildungen $\mathbb{p}^{X|S=s}$ und $\mathbb{f}^{X|S=s}$ tatsächlich $\mathbb{P}^S$-fast überall, also nur außerhalb einer $\mathbb{P}^S$-Nullmenge, definiert. Wir können auf diesen Nullmengen beliebig Dichten, zum Beispiel die Randdichten $\mathbb{p}^{X|S=s} = \mathbb{p}^X$ und $\mathbb{f}^{X|S=s} = \mathbb{f}^X$, wählen. □

§05.04 **Definition**.

(a) Sei (S, X) eine diskret-verteilte Zufallsvariable mit Werten in $(\mathcal{S} \times \mathcal{X}, 2^{\mathcal{S} \times \mathcal{X}})$. Für jedes $s \in \mathcal{S}$ mit strikt-positiver Randzähldichte von S in s heißt die Verteilung $\boxed{\mathbb{P}^{X|S=s}}$ mit Zähldichte $\boxed{\mathbb{p}^{X|S=s}}$ in (05.01) *bedingte Verteilung* (mit *bedingter Zähldichte*) von X gegeben $S = s$. Für $h \in \overline{2^{\mathcal{X}}}^+$ bzw. $h \in \mathscr{L}_1(\mathbb{P}^{X|S=s})$ wird $\boxed{\mathbb{E}^{X|S=s}(h)} := \mathbb{P}^{X|S=s}(h) = \sum_{x \in \mathcal{X}} h(x)\mathbb{p}^{X|S=s}(x) = \int h\mathbb{p}^{X|S=s}\, d\zeta_{\mathcal{X}}$ *bedingter Erwartungswert* von $h(X)$ bei gegebenem $S = s$ genannt.

(b) Sei (S, X) ein stetig-verteilter Zufallsvektor mit Werten in $\mathbb{R}^{n+m}$. Für jedes $s \in \mathbb{R}^m$ mit strikt-positiver Randdichte von S in s heißt die Verteilung $\boxed{\mathbb{P}^{X|S=s}}$ mit Dichte $\boxed{\mathbb{f}^{X|S=s}}$ in (05.02) *bedingte Verteilung* bzw. *bedingte Dichte* von X gegeben $S = s$. Für $h \in \overline{\mathscr{B}^n}^+$ bzw. $h \in \mathscr{L}_1(\mathbb{P}^{X|S=s})$ wird $\boxed{\mathbb{E}^{X|S=s}(h)} := \mathbb{P}^{X|S=s}(h) = \int h(x)\mathbb{f}^{X|S=s}(x)\lambda^n(dx) = \int h\mathbb{f}^{X|S=s}\, d\lambda^n$ *bedingter Erwartungswert* von $h(X)$ gegeben $S = s$ genannt. □

§05.05 **Bemerkung**. Betrachten wir den stetigen Fall (für den diskreten Fall folgen die Aussagen analog). Da für $h \in \mathscr{L}_1(\mathbb{P}^X)$ auch außerhalb einer $\mathbb{P}^S$-Nullmenge $\int |h|\mathbb{f}_s^{(S,X)}\, d\lambda^n < \infty$ gilt, ist für jedes

s außerhalb einer $\mathbb{P}^S$-Nullmenge der bedingte Erwartungswert $\mathbb{E}^{X|S=s}(h) = \mathbb{P}^{X|S=s}(h)$ von $h(X)$ gegeben $S = s$, wie auch zuvor die bedingte Dichte, definiert, und auf der Nullmenge können wir diesen beliebig fortsetzen. Verschiedene Versionen (Festlegungen) unterscheiden sich damit nur auf einer Nullmenge. Wir sehen später, dass wir eine Festlegung derart wählen können, dass die Abbildung $s \mapsto \varphi(s) := \mathbb{E}^{X|S=s}(h) = \mathbb{P}^{X|S=s}(h)$ messbar ist. In diesem Fall gilt für alle $B \in \mathscr{B}^m$:

$$\begin{aligned}\mathbb{E}^S(\mathbb{1}_B\varphi) = \mathbb{P}^S(\mathbb{1}_B\varphi) = \int \mathbb{1}_B\varphi\, \mathbb{f}^S\, d\lambda^m = \int_{\mathbb{R}^m} \mathbb{1}_B(s) \int_{\mathbb{R}^n} h(x)\mathbb{f}^{(S,X)}(s,x)\lambda^N(dx)\lambda^m(ds)\\ = \int \mathbb{1}_B h\, \mathbb{f}^{(S,X)}\, d\lambda^{m+n} = \mathbb{P}^{(S,X)}(\mathbb{1}_B h) = \mathbb{E}^{(S,X)}(\mathbb{1}_B h)\end{aligned}$$

Die Messbarkeit und die letzte Gleichheit sind der Ausgangspunkt für den folgenden allgemeineren Zugang, der uns erlaubt, auch den diskreten sowie den stetigen Fall abzudecken. Die folgenden Resultate und Eigenschaften lassen sich dann auf die bisher vorgestellten bedingten Dichten, Verteilungen und Erwartungswerte übertragen. □

§05.06 **Beispiel.**

(a) Sei $X = |W_1 - W_2|$ und $S = W_1 + W_2$ der Absolutbetrag der Differenz bzw. die Summe der Augenzahlen von zwei unabhängigen fairen Würfeln (W_1, W_2). Dann ist (S, X) diskret-verteilt mit Werten in $[\![0,5]\!] \times [\![2,12]\!]$ und bedingten Zähldichten von X gegeben $S = s$:

		s											
	$\mathbb{p}^{X\|S=s}(x)$	2	3	4	5	6	7	8	9	10	11	12	$\mathbb{p}^X(x)$
	0	1	0	$\frac{1}{3}$	0	$\frac{1}{5}$	0	$\frac{1}{5}$	0	$\frac{1}{3}$	0	1	$\frac{6}{36}$
	1	0	1	0	$\frac{1}{2}$	0	$\frac{1}{3}$	0	$\frac{1}{2}$	0	1	0	$\frac{10}{36}$
x	2	0	0	$\frac{2}{3}$	0	$\frac{2}{5}$	0	$\frac{2}{5}$	0	$\frac{2}{3}$	0	0	$\frac{8}{36}$
	3	0	0	0	$\frac{1}{2}$	0	$\frac{1}{3}$	0	$\frac{1}{2}$	0	0	0	$\frac{6}{36}$
	4	0	0	0	0	$\frac{2}{5}$	0	$\frac{2}{5}$	0	0	0	0	$\frac{4}{36}$
	5	0	0	0	0	0	$\frac{1}{3}$	0	0	0	0	0	$\frac{2}{36}$
	$\mathbb{p}^S(s)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	1

(b) Seien X und S Zufallsvektoren mit

$$Y := \begin{pmatrix} X \\ S \end{pmatrix} \sim \mathrm{N}_{(\mu,\Sigma)} \text{ mit } \mu = \begin{pmatrix} \mu_X \\ \mu_S \end{pmatrix} \in \mathbb{R}^{n+m} \text{ und } \Sigma = \begin{pmatrix} \Sigma_X & \Sigma_{XS} \\ \Sigma_{SX} & \Sigma_S \end{pmatrix} > 0,$$

also $X \sim \mathrm{N}_{(\mu_X,\Sigma_X)}$, $S \sim \mathrm{N}_{(\mu_S,\Sigma_S)}$ und $\mathbb{C}\mathrm{ov}(X,S) = \Sigma_{XS} = \Sigma_{SX}^t$. Dann ist die bedingte Verteilung von X gegeben $S = s$ eine $\mathrm{N}_{(\mu_{X|S=s},\Sigma_{X|S=s})}$-Verteilung mit

$$\mu_{X|S=s} := \mu_X + \Sigma_{XS}\Sigma_S^{-1}(s - \mu_S) \in \mathbb{R}^m \text{ und } \Sigma_{X|S=s} := \Sigma_X - \Sigma_{XS}\Sigma_S^{-1}\Sigma_{SX} > 0.$$

Definitionsgemäß ist $\mu_{X|S=s}$ gerade der bedingte Erwartungswert von X gegeben $S = s$. Zum Beweis der Aussage benutzen wir, dass für $\Sigma > 0$ gilt $\Sigma_X > 0$, $\Sigma_S > 0$ und $E := \Sigma_X - \Sigma_{XS}\Sigma_S^{-1}\Sigma_{SX} > 0$ sowie mit $F := \Sigma_{XS}\Sigma_S^{-1}$ auch

$$\Sigma^{-1} = \begin{pmatrix} E^{-1} & -E^{-1}F \\ -FE^{-1} & \Sigma_S^{-1} + F^tE^{-1}F \end{pmatrix} > 0.$$

Für die gemeinsame Dichte $\mathbb{f}_{\mathrm{N}_{(\mu,\Sigma)}}$ von $Y = (X^t, S^t)^t$ und die Randdichte $\mathbb{f}_{\mathrm{N}_{(\mu_S,\Sigma_S)}}$ von S gilt mit $y = (x^t, s^t)$ und Normierungskonstante c

$$\frac{\mathbb{f}_{\mathrm{N}_{(\mu,\Sigma)}}(y)}{\mathbb{f}_{\mathrm{N}_{(\mu_S,\Sigma_S)}}(s)} = c \exp\big\{ -\tfrac{1}{2}(y-\mu)^t\Sigma^{-1}(y-\mu) + \tfrac{1}{2}(s-\mu_S)^t\Sigma^{-1}(s-\mu_S)\big\}$$
$$= c \exp\big\{ -\tfrac{1}{2}(x-(\mu_X + F(s-\mu_S)))^t E^{-1}(x-(\mu_X+F(s-\mu_S)))\big\}.$$

Da $\mu_{X|S=s} = \mu_X + F(s-\mu_S)$ und $\Sigma_{X|S=s} = E$ folgt die Behauptung. □

§06 Positive numerische Zufallsvariablen

Seien im Folgenden stets $(\Omega, \mathscr{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, $\mathbb{E}$ die Erwartung bzgl. $\mathbb{P}$ und $\mathscr{F} \subseteq \mathscr{A}$ eine Teil-σ-Algebra von $\mathscr{A}$.

§06.01 **Erinnerung**. Wir schreiben kurz $X \in \overline{\mathscr{A}}^+$, wenn X eine positive numerische Zufallsvariable auf $(\Omega, \mathscr{A})$, also $X : \Omega \to \overline{\mathbb{R}}^+$ eine $\mathscr{A}$-$\overline{\mathscr{B}}^+$-messbare Abbildung, ist. Insbesondere, ist $\overline{\mathscr{F}}^+ \subseteq \overline{\mathscr{A}}^+$ und für $Y \in \overline{\mathscr{F}}^+$ somit $\mathbb{E}(Y)$ definiert. □

§06.02 **Satz**. *Zu jedem $X \in \overline{\mathscr{A}}^+$ existiert eine Lösung $Y \in \overline{\mathscr{F}}^+$ der Radon-Nikodym-Gleichung $\mathbb{E}(\mathbb{1}_F Y) = \mathbb{E}(\mathbb{1}_F X)$ für jedes $F \in \mathscr{F}$, wobei Y bis auf Gleichheit $\mathbb{P}$-f.ü. eindeutig ist.*

§06.03 **Beweis** von Satz §06.02. In der Vorlesung. □

§06.04 **Definition**. Für $X \in \overline{\mathscr{A}}^+$ heißt jede Abbildung $Y : \Omega \to \overline{\mathbb{R}}^+$ *Festlegung (Version) des bedingten Erwartungswertes* von X gegeben $\mathscr{F}$, symbolisch $\boxed{\mathbb{E}(X|\mathscr{F})} := Y$, wenn sie die folgenden zwei Bedingungen erfüllt:

(bE1) Y ist $\mathscr{F}$-$\overline{\mathscr{B}}^+$-messbar, also $Y \in \overline{\mathscr{F}}^+$ und

(bE2) $\mathbb{E}(\mathbb{1}_F Y) = \mathbb{E}(\mathbb{1}_F X)$ für jedes $F \in \mathscr{F}$. (*Radon-Nikodym-Gleichung*)

Jede Abbildung $\boxed{\mathbb{E}(\bullet|\mathscr{F})} : \overline{\mathscr{A}}^+ \to \overline{\mathscr{F}}^+$ mit $X \mapsto \mathbb{E}(X|\mathscr{F})$ heißt *Festlegung der bedingten Erwartung* bzgl. $\mathbb{P}$ gegeben $\mathscr{F}$. Die von $\mathbb{E}(\bullet|\mathscr{F})$ implizierte Abbildung $\boxed{\mathbb{P}(\bullet|\mathscr{F})} : \mathscr{A} \to \overline{\mathscr{F}}^+$ mit $A \mapsto \mathbb{P}(A|\mathscr{F}) := \mathbb{E}(\mathbb{1}_A|\mathscr{F})$ wird Festlegung der bedingten Verteilung von $\mathbb{P}$ gegeben $\mathscr{F}$ genannt. Mit (bE2) ist jede Festlegung $\mathbb{P}(A|\mathscr{F})$ Lösung der Radon-Nikodym-Gleichung, d.h. $\mathbb{E}(\mathbb{1}_F \mathbb{P}(A|\mathscr{F})) = \int_F \mathbb{P}(A|\mathscr{F}) d\mathbb{P} = \mathbb{P}(F \cap A)$ für alle $F \in \mathscr{F}$. □

§06.05 **Bemerkung**. Nach Satz §06.02 unterscheiden sich für jedes $X \in \overline{\mathscr{A}}^+$ Festlegungen des bedingten Erwartungswertes von X gegeben $\mathscr{F}$ nur auf einer $\mathbb{P}$-Nullmenge. Diese Eigenschaft überträgt sich im Allgemeinen nicht auf Festlegungen der bedingten Erwartung bzgl. $\mathbb{P}$ gegeben $\mathscr{F}$, da wir zu jedem $X \in \overline{\mathscr{A}}^+$ eine Ausnahmemenge erhalten, und deren Vereinigung im Allgemeinen keine Nullmenge mehr ist. Betrachten wir eine Festlegung $\mathbb{P}(\bullet|\mathscr{F})$ der bedingten Verteilung von $\mathbb{P}$ gegeben $\mathscr{F}$, so ist eine wünschenswerte Eigenschaft, dass für jedes $\omega \in \Omega$ die entsprechende Abbildung $\mathbb{P}(\bullet|\mathscr{F})(\omega) : \mathscr{A} \to \overline{\mathbb{R}}^+$ mit $A \mapsto \mathbb{P}(A|\mathscr{F})(\omega)$ ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathscr{A})$ ist. □

§06.06 **Erinnerung** *(vgl. Definition §04.51)*. Seien $(\Omega_1, \mathscr{A}_1)$ und $(\Omega_2, \mathscr{A}_2)$ messbare Räume. Eine Abbildung $\kappa : \Omega_1 \times \mathscr{A}_2 \to [0,1]$ heißt *Markovkern* von $(\Omega_1, \mathscr{A}_1)$ nach $(\Omega_2, \mathscr{A}_2)$, falls sie die folgenden zwei Bedingungen erfüllt:

(Mk1) für alle $\omega_1 \in \Omega_1$ ist $\kappa_{\omega_1} : \mathscr{A}_2 \to [0,1]$ mit $A_2 \mapsto \kappa_{\omega_1}(A_2) := \kappa(\omega_1, A_2)$ ein Wahrscheinlichkeitsmaß auf $(\Omega_2, \mathscr{A}_2)$, kurz $\kappa_{\omega_1} \in \mathcal{W}(\mathscr{A}_2)$;

(Mk2) für alle $A_2 \in \mathscr{A}_2$ ist $\kappa^{A_2} : \Omega_1 \to [0,1]$ mit $\omega_1 \mapsto \kappa^{A_2}(\omega_1) := \kappa(\omega_1, A_2)$ eine $\mathscr{A}_1$-$\mathscr{B}_{[0,1]}$-messbare Abbildung, kurz $\kappa^{A_2} \in \overline{\mathscr{A}_1}^+$. □

§06.07 **Schreibweise** *(vgl. §04.60)*. Betrachte einen Wahrscheinlichkeitsraum $(\Omega_1, \mathscr{A}_1, \mathbb{P})$, einen messbaren Raum $(\Omega_2, \mathscr{A}_2)$ und einen Markovkern κ von $(\Omega_1, \mathscr{A}_1)$ nach $(\Omega_2, \mathscr{A}_2)$. Dann existiert nach Lemma §04.56 ein eindeutig bestimmtes Wahrscheinlichkeitsmaß $\mathbb{P} \odot \kappa$ auf $(\Omega_1 \times \Omega_2, \mathscr{A}_1 \otimes \mathscr{A}_2)$ mit

$$\mathbb{P} \odot \kappa(A_1 \times A_2) = \mathbb{P}(\mathbb{1}_{A_1} \kappa^{A_2}) = \int_{A_1} \kappa(\omega_1, A_2) \mathbb{P}(d\omega_1), \quad \text{für alle } A_1 \in \mathscr{A}_1, A_2 \in \mathscr{A}_2.$$

Ist $h \in \overline{\mathscr{A}_1 \otimes \mathscr{A}_2}^+$ oder $h \in \mathscr{L}_1(\mathbb{P} \odot \kappa)$ dann gilt

$$\mathbb{P} \odot \kappa(h) = \mathbb{P}(\kappa_\bullet(h_\bullet)) = \int_{\Omega_1} \int_{\Omega_2} h(\omega_1, \omega_2) \kappa(\omega_1, d\omega_2) \mathbb{P}(d\omega_1)$$

(vgl. Satz §04.58). Weiterhin bezeichnen wir mit $\kappa\mathbb{P}$ die durch $\kappa \odot \mathbb{P}$ auf $(\Omega_2, \mathscr{A}_2)$ induzierte Randverteilung, also $\kappa\mathbb{P}(A_2) = \mathbb{P}(\kappa^{A_2}) = \int_{\Omega_1} \kappa(\omega_1, A_2) \mathbb{P}(d\omega_1)$ für alle $A_2 \in \mathscr{A}_2$. □

§06.08 **Definition**.

(a) $\mathbb{P}(\bullet | \mathscr{F})$ heißt *reguläre Festlegung* der bedingten Verteilung von $\mathbb{P}$ gegeben $\mathscr{F}$, wenn die Abbildung $(\omega, A) \mapsto \mathbb{P}(A | \mathscr{F})(\omega)$ die Eigenschaften (Mk1) und (Mk2) erfüllt, also $\mathbb{P}(\bullet | \mathscr{F})$ ein Markovkern von $(\Omega, \mathscr{F})$ nach $(\Omega, \mathscr{A})$ ist.

(b) $\mathbb{E}(\bullet | \mathscr{F})$ heißt *reguläre Festlegung* der bedingten Erwartung bzgl. $\mathbb{P}$ gegeben $\mathscr{F}$, wenn die implizierte Festlegung $\mathbb{P}(\bullet | \mathscr{F})$ der bedingten Verteilung von $\mathbb{P}$ gegeben $\mathscr{F}$ regulär ist, und für jedes $\omega \in \Omega$ die Abbildung $X \mapsto \mathbb{E}(X | \mathscr{F})(\omega)$ die Erwartung bzgl. $\mathbb{P}(\bullet | \mathscr{F})(\omega)$ ist. □

§06.09 **Schreibweise**. Ist $\mathbb{E}(\bullet | \mathscr{F})$ eine reguläre Festlegung der bedingten Erwartung bzgl. $\mathbb{P}$ gegeben $\mathscr{F}$, und $\mathbb{P}(\bullet | \mathscr{F})$ die implizierte Festlegung der bedingten Verteilung bzgl. $\mathbb{P}$ gegeben $\mathscr{F}$, so verwenden wir für $X \in \overline{\mathscr{A}}^+$ wie bisher auch beide Schreibweisen $\mathbb{E}(X | \mathscr{F}) = \mathbb{P}(X | \mathscr{F})$. □

§06.10 **Satz**.

(i) *Zu jeder regulären Festlegung der bedingten Verteilung von $\mathbb{P}$ gegeben $\mathscr{F}$ existiert eine sie implizierende reguläre Festlegung der bedingten Erwartung bzgl. $\mathbb{P}$ gegeben $\mathscr{F}$.*

(ii) *Zu jedem Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\mathbb{R}^k, \mathscr{B}^k)$ und Teil-σ-Algebra $\mathscr{F} \subseteq \mathscr{B}^k$ von $\mathscr{B}^k$ existiert eine reguläre Festlegung der bedingten Verteilung von $\mathbb{P}$ gegeben $\mathscr{F}$.*

§06.11 **Beweis** von Satz §06.10. (i) unter Verwendung der Beweisstrategie durch Approximation mit einfachen Zufallsvariablen, (ii) Klenke [2020] Satz 8.37, S. 206 □

§06.12 **Bemerkung**. Bevor wir Eigenschaften der bedingten Erwartung festhalten, werden wir in einfachen Situationen konstruktiv Versionen der bedingten Erwartung bzw. Verteilung bestimmen. Dabei zeigt sich, dass auch eine reguläre Festlegung der bedingten Verteilung nicht eindeutig ist, aber in vielen für uns interessanten Situation bis auf $\mathbb{P}$-f.ü. Gleichheit eindeutig bestimmt ist. □

§06|01 Einfache Bedingungen

Wir betrachten zunächst eine Teil-σ-Algebra $\mathscr{F}$ von $\mathscr{A}$, die nur von einer einzelnen messbaren Menge erzeugt ist, dass heißt, $\mathscr{F} = \{\emptyset, B, B^c, \Omega\} = \sigma(\{B\})$ mit $B \in \mathscr{A}$.

§06.13 **Erinnerung**. Sei $(\Omega, \mathscr{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Zu $\mathbb{P}$ existiert nach Definition A04.01 (oder alternativ Satz §02.02) eine eindeutig bestimmte Erwartung $\mathbb{E} : \overline{\mathscr{A}}^+ \to \overline{\mathbb{R}}^+$, die (E1) linear und (E2) monoton konvergent ist mit (E3) $\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$ für alle $A \in \mathscr{A}$. □

§06.14 **Bedingte Erwartung**. *Sei $B \in \mathscr{A}$ mit $\mathbb{P}(B) \in \mathbb{R}^+_{\setminus 0}$. Dann ist*

$$\mathbb{P}(\bullet|B) : \mathscr{A} \to [0,1] \text{ mit } A \mapsto \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathscr{A})$. Entsprechend existiert zu $\mathbb{P}(\bullet|B)$ eine eindeutig bestimmte Erwartung, die wir mit $\mathbb{E}(\bullet|B)$ bezeichnen. Mit (E3) erfüllt diese

$$\forall A \in \mathscr{A} \ : \ \mathbb{E}(\mathbb{1}_A|B) = \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{E}(\mathbb{1}_A \mathbb{1}_B)}{\mathbb{P}(B)} = \mathbb{P}(\mathbb{1}_A|B).$$

Da für jedes $X \in \overline{\mathscr{A}}^+$ weiterhin gilt $X\mathbb{1}_B \in \overline{\mathscr{A}}^+$ erfüllt die Erwartung $\mathbb{E}(\bullet|B)$ auch

$$\forall X \in \overline{\mathscr{A}}^+ \ : \ \mathbb{E}(X|B) = \frac{\mathbb{E}(X\mathbb{1}_B)}{\mathbb{P}(B)} = \mathbb{P}(X|B).$$

Falls für $B^c = \Omega \setminus B \in \mathscr{A}$ auch $\mathbb{P}(B^c) \in \mathbb{R}^+_{\setminus 0}$ gilt, so können wir analog das Wahrscheinlichkeitsmaß $\mathbb{P}(\bullet|B^c)$ auf $(\Omega, \mathscr{A})$ und die entsprechende Erwartung $\mathbb{E}(\bullet|B^c)$ bzgl. $\mathbb{P}(\bullet|B^c)$ betrachten. Für jedes $X \in \overline{\mathscr{A}}^+$ ist dann

$$\mathbb{E}(X|\mathscr{F}) : \Omega \to \overline{\mathbb{R}}^+ \text{ mit } \omega \mapsto \mathbb{E}(X|\mathscr{F})(\omega) := \mathbb{E}(X|B)\mathbb{1}_B(\omega) + \mathbb{E}(X|B^c)\mathbb{1}_{B^c}(\omega)$$

messbar bzgl. der Teil-σ-Algebra $\mathscr{F} := \sigma(\{B\}) = \{\emptyset, B, B^c, \Omega\} \subseteq \mathscr{A}$, also eine einfache numerische Zufallsvariable auf $(\Omega, \mathscr{F})$, kurz $\mathbb{E}(X|\mathscr{F}) \in \overline{\mathscr{F}}^+$. Für jedes $\omega \in \Omega$ ist entweder $\omega \in B$ und $\mathbb{E}(X|\mathscr{F})(\omega) = \mathbb{E}(X|B)$ oder $\omega \in B^c$ sowie $\mathbb{E}(X|\mathscr{F})(\omega) = \mathbb{E}(X|B^c)$, so dass

$$\mathbb{E}(\bullet|\mathscr{F}) : \overline{\mathscr{A}}^+ \to \overline{\mathscr{F}}^+ \text{ mit } X \mapsto \mathbb{E}(X|\mathscr{F}) := \mathbb{E}(X|B)\mathbb{1}_B + \mathbb{E}(X|B^c)\mathbb{1}_{B^c}. \qquad (06.01)$$

Wir halten fest, dass gegeben $\omega \in B$, gilt $\mathbb{E}(\bullet|B) = \mathbb{E}(\bullet|\mathscr{F})(\omega)$. Für jedes $X \in \overline{\mathscr{A}}^+$ ist $\mathbb{E}(X|\mathscr{F})$ insbesondere Lösung der Radon-Nikodym-Gleichung (einfaches nachrechnen!)

$$\forall F \in \mathscr{F} \ : \ \mathbb{E}(\mathbb{1}_F \mathbb{E}(X|\mathscr{F})) = \mathbb{E}(\mathbb{1}_F X). \qquad (06.02)$$

Damit erfüllt die Abbildung $\mathbb{E}(\bullet|\mathscr{F})$ in (06.01) die Bedingungen (bE1) sowie (bE2), und sie ist somit eine Festlegung der bedingten Erwartung bzgl. $\mathbb{P}$ gegeben $\mathscr{F}$. □

§06.15 **Reguläre Festlegung**. *Wir halten weiterhin fest, dass die von der Festlegung $\mathbb{E}(\bullet|\mathscr{F})$ in (06.01) implizierte Festlegung der bedingten Verteilung*

$$\mathbb{P}(\bullet|\mathscr{F}) : \mathscr{A} \to \mathscr{F}^+ \text{ mit } A \mapsto \mathbb{P}(A|\mathscr{F}) := \mathbb{E}(\mathbb{1}_A|\mathscr{F})$$

ein Markovkern von $(\Omega, \mathscr{F})$ nach $(\Omega, \mathscr{A})$ ist. In der Tat erfüllt die Abbildung $\Omega \times \mathscr{A} \to \mathbb{R}^+$ mit

$$(\omega, A) \mapsto \mathbb{P}(A|\mathscr{F})(\omega) = \mathbb{P}(A|B)\mathbb{1}_B(\omega) + \mathbb{P}(A|B^c)\mathbb{1}_{B^c}(\omega),$$

die Bedingungen (Mk1) und (Mk2). Betrachte (Mk1). Für $\omega \in \Omega$ ist entweder $\omega \in B$, und somit $\mathbb{P}(\bullet|\mathscr{F})(\omega) = \mathbb{P}(\bullet|B)$, oder $\omega \in B^c$ und $\mathbb{P}(\bullet|\mathscr{F})(\omega) = \mathbb{P}(\bullet|B^c)$. In beiden Fällen ist $\mathbb{P}(\bullet|\mathscr{F})(\omega) \in \mathcal{W}(\mathscr{A})$ ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathscr{A})$. Wir halten fest, dass gegeben $\omega \in B$, gilt $\mathbb{P}(\bullet|B) = \mathbb{P}(\bullet|\mathscr{F})(\omega)$. Andererseits, gilt (Mk2). Da für jedes $A \in \mathscr{A}$ nach Konstruktion die Zufallsvariable $\mathbb{P}(A|\mathscr{F}) = \mathbb{E}(\mathbb{1}_A|\mathscr{F})$ die Bedingung (bE1) erfüllt, die Abbildung $\mathbb{P}(A|\mathscr{F}) : \Omega \to [0,1]$ mit $\omega \mapsto \mathbb{P}(A|\mathscr{F})(\omega)$ also $\mathscr{F}$-messbar ist. Folglich ist die Festlegung $\mathbb{P}(\bullet|\mathscr{F})$ der bedingten Verteilung von $\mathbb{P}$ gegeben $\mathscr{F}$ regulär. Andererseits ist per Konstruktion für jedes $\omega \in \Omega$ die Abbildung $\mathbb{E}(\bullet|\mathscr{F})(\omega)$ die Erwartung bzgl. $\mathbb{P}(\bullet|\mathscr{F})(\omega)$, so dass auch die Festlegung $\mathbb{E}(\bullet|\mathscr{F})$ in (06.01) regulär ist. □

§06.16 **Bedingte Verteilung gegeben eine Nullmenge** B. *Für $B \in \mathscr{A}$ mit $\mathbb{P}(B) = 0$ ist $\mathbb{P}(\bullet|B)$ und damit auch $\mathbb{E}(\bullet|\mathscr{F})$ sowie $\mathbb{P}(\bullet|\mathscr{F})$ bisher nicht definiert. Wir vereinbaren im Folgenden $\mathbb{P}(\bullet|B) := \mathbb{P}$ und somit $\mathbb{E}(\bullet|B) = \mathbb{E}$. Die Abbildung (06.01) wird dann zu*

$$\mathbb{E}(\bullet|\mathscr{F}) : \overline{\mathscr{A}}^+ \to \overline{\mathscr{F}}^+ \text{ mit } X \mapsto \mathbb{E}(X|\mathscr{F}) := \mathbb{E}(X)\mathbb{1}_B + \mathbb{E}(X|B^c)\mathbb{1}_{B^c}. \tag{06.03}$$

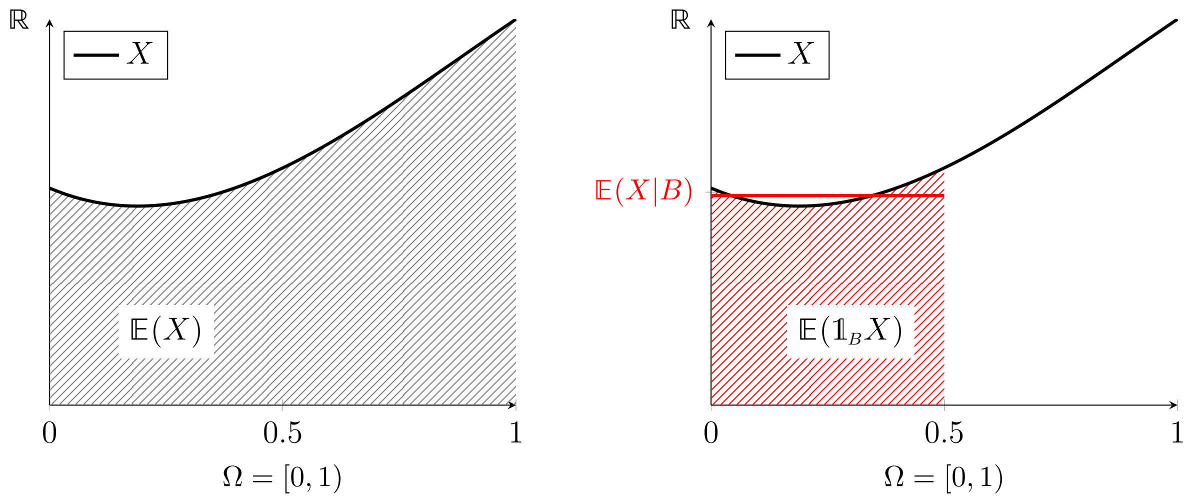
$\mathbb{E}(\bullet|\mathscr{F})$ erfüllt für alle $X \in \overline{\mathscr{A}}^+$ weiterhin die Bedingungen (bE1)-(bE2), d.h. $\mathbb{E}(X|\mathscr{F}) \in \overline{\mathscr{F}}^+$ sowie $\mathbb{E}(\mathbb{1}_F\mathbb{E}(X|\mathscr{F})) = \mathbb{E}(\mathbb{1}_F X)$ für alle $F \in \mathscr{F}$. Damit ist $\mathbb{E}(\bullet|\mathscr{F})$ in (06.03) eine Festlegung der bedingten Erwartung bzgl. $\mathbb{P}$ gegeben $\mathscr{F}$. Außerdem sind (Mk1) und (Mk2) immer noch von

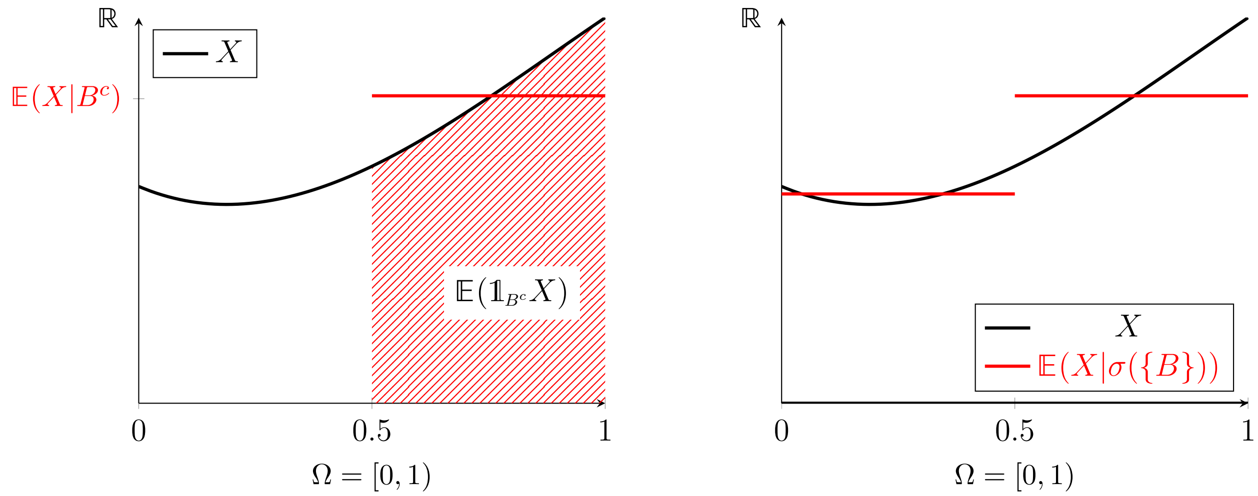
$$\mathbb{P}(\bullet|\mathscr{F}) : \mathscr{A} \to \overline{\mathscr{F}}^+ \text{ mit } A \mapsto \mathbb{P}(\bullet|\mathscr{F}) := \mathbb{P}(A)\mathbb{1}_B + \mathbb{P}(A|B^c)\mathbb{1}_{B^c} \tag{06.04}$$

erfüllt. Damit ist $\mathbb{P}(\bullet|\mathscr{F})$ weiterhin eine reguläre Festlegung der bedingten Verteilung von $\mathbb{P}$ gegeben $\mathscr{F}$. Per Konstruktion ist für jedes $\omega \in \Omega$ die Abbildung $\mathbb{E}(\bullet|\mathscr{F})(\omega)$ die Erwartung bzgl. $\mathbb{P}(\bullet|\mathscr{F})(\omega)$, so dass auch die Festlegung $\mathbb{E}(\bullet|\mathscr{F})$ in (06.03) regulär ist. □

§06.17 **Anmerkung**. Alle bisher getroffenen Aussagen für $\mathbb{E}(\bullet|\mathscr{F})$ und $\mathbb{P}(\bullet|\mathscr{F})$ gelten damit für beliebige Teil-σ-Algebren von $\mathscr{A}$ der Form $\mathscr{F} = \sigma(\{B\})$ mit $B \in \mathscr{A}$. Offensichtlich ist in der Situation $\mathbb{P}(B) = 0$ die Wahl $\mathbb{P}(\bullet|B) = \mathbb{P}$ willkürlich. Die Aussagen gelten genauso, wenn wir $\mathbb{P}(\bullet|B) = \tilde{\mathbb{P}}$ für ein beliebiges $\tilde{\mathbb{P}} \in \mathcal{W}(\mathscr{A})$ und mit der Erwartung $\tilde{\mathbb{E}}$ bzgl. $\tilde{\mathbb{P}}$ auch $\mathbb{E}(\bullet|B) = \tilde{\mathbb{E}}$ setzen. Es ist wichtig zu bemerken, dass in der Situation $\mathbb{P}(B) = 0$ eine andere Wahl $\mathbb{E}(\bullet|\mathscr{F})$ und $\mathbb{P}(\bullet|\mathscr{F})$ nur auf der Nullmenge B bzgl. $\mathbb{P}$ ändert, also $\mathbb{E}(\bullet|\mathscr{F})$ und $\mathbb{P}(\bullet|\mathscr{F})$ eindeutig sind bis auf Gleichheit $\mathbb{P}$-f.ü.. □

§06.18 **Skizze**. Betrachte den Wahrscheinlichkeitsraum $([0,1), \mathscr{B}_{[0,1)}, \mathrm{U}_{[0,1)})$. Die nächsten Graphiken stellen eine positive numerische Zufallsvariable $X : [0,1) \to \mathbb{R}$ sowie für $\mathscr{F} = \sigma(\{B\})$ mit $B = [0, 0.5)$ und $B^c = [0.5, 1)$ die vorgestellte reguläre Festlegung des bedingten Erwartungswertes von X gegeben $\mathscr{F}$ dar.





□

§06|02 Abzählbare Bedingungen

Wir betrachten nun eine Teil-σ-Algebra $\mathscr{F}$ von $\mathscr{A}$, die von einer abzählbaren und messbaren Partition $\{B_i, i \in \mathcal{I}\}$ von Ω erzeugt ist, dass heißt $\mathscr{F} = \sigma(\{B_i, i \in \mathcal{I}\})$.

§06.19 **Bedingte Erwartung**. *Für eine abzählbare und messbare Partition $\{B_i, i \in \mathcal{I}\}$ von Ω, wobei wir wie bisher $\mathbb{P}(\bullet|B_i) = \mathbb{P}$ für $\mathbb{P}(B_i) = 0$ setzen, sei für jedes $i \in \mathcal{I}$ weiterhin $\mathbb{E}(\bullet|B_i)$ die Erwartung bzgl. $\mathbb{P}(\bullet|B_i)$. Für $\mathscr{F} = \sigma(\{B_i, i \in \mathcal{I}\})$ erfüllt dann die Abbildung*

$$\mathbb{E}(\bullet|\mathscr{F}) : \overline{\mathscr{A}}^+ \to \overline{\mathscr{F}}^+ \text{ mit } X \mapsto \mathbb{E}(X|\mathscr{F}) := \sum_{i\in\mathcal{I}} \mathbb{E}(X|B_i)\mathbb{1}_{B_i} \tag{06.05}$$

die Bedingungen (bE1) und (bE2). Betrachte (bE1). Für $X \in \overline{\mathscr{A}}^+$ gilt $\mathbb{E}(X|\mathscr{F}) \in \overline{\mathscr{F}}^+$. In der Tat sind die Abbildungen $f : \Omega \to \mathcal{I}$ mit $\omega \mapsto f(\omega) = i \Leftrightarrow \omega \in B_i$ und $g : \mathcal{I} \to \overline{\mathbb{R}}^+$ mit $i \mapsto \mathbb{E}(X|B_i)$ $\mathscr{F}$-$2^{\mathcal{I}}$- bzw. $2^{\mathcal{I}}$-$\overline{\mathscr{B}}^+$-messbar, so dass $\mathbb{E}(X|\mathscr{F}) = g \circ f \in \overline{\mathscr{F}}^+$ eine $\mathscr{F}$-$\overline{\mathscr{B}}^+$-messbare Abbildung ist. Andererseits, ist $F \in \mathscr{F}$, so existiert $\mathcal{J} \subseteq \mathcal{I}$ mit $F = \biguplus_{j\in\mathcal{J}} B_j$ und (bE2) gilt, da

$$\begin{aligned}\mathbb{E}(\mathbb{1}_F\mathbb{E}(X|\mathscr{F})) &= \sum_{j\in\mathcal{J}} \mathbb{E}(\mathbb{1}_{B_j}\mathbb{E}(X|\mathscr{F})) = \sum_{j\in\mathcal{J}} \mathbb{E}(\mathbb{1}_{B_j}\mathbb{E}(X|B_j)) \\ &= \sum_{j\in\mathcal{J}} \mathbb{P}(B_j)\mathbb{E}(X|B_j) = \sum_{j\in\mathcal{J}} \mathbb{E}(\mathbb{1}_{B_j}X) = \mathbb{E}(\mathbb{1}_F X).\end{aligned}$$

Damit erfüllt die Abbildung $\mathbb{E}(\bullet|\mathscr{F})$ in (06.05) die Bedingungen (bE1) sowie (bE2), und sie ist somit eine Festlegung der bedingten Erwartung bzgl. $\mathbb{P}$ gegeben $\mathscr{F}$. Wir halten weiterhin fest, dass die von der Festlegung $\mathbb{E}(\bullet|\mathscr{F})$ in (06.05) implizierte Festlegung der bedingten Verteilung

$$\mathbb{P}(\bullet|\mathscr{F}) : \mathscr{A} \to \mathscr{F}^+ \text{ mit } A \mapsto \mathbb{P}(A|\mathscr{F}) := \mathbb{E}(\mathbb{1}_A|\mathscr{F}) = \sum_{i\in\mathcal{I}} \mathbb{P}(A|B_i)\mathbb{1}_{B_i} \tag{06.06}$$

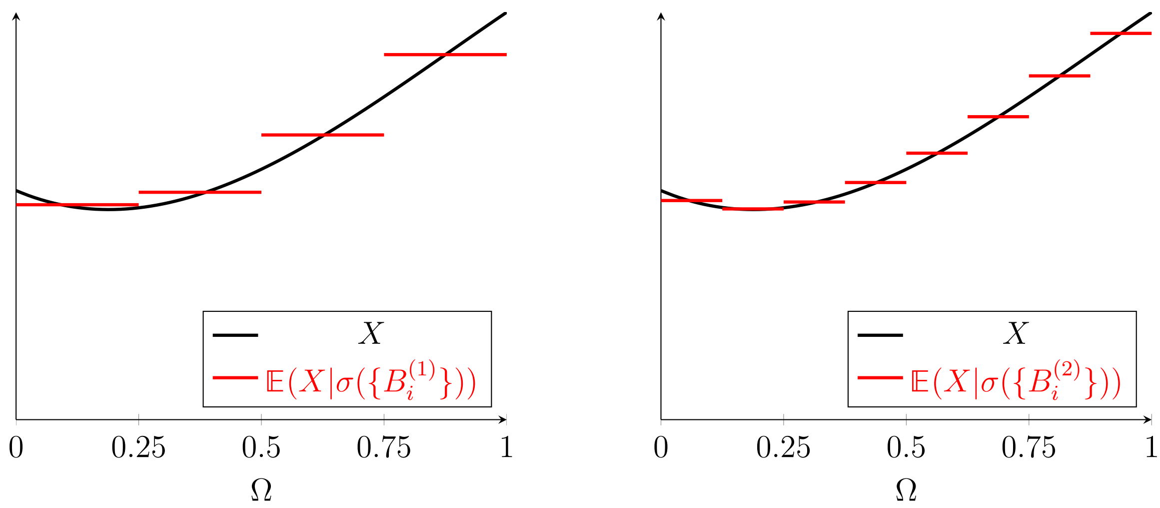
ein Markovkern von $(\Omega, \mathscr{F})$ nach $(\Omega, \mathscr{A})$ ist. In der Tat die Abbildung $\Omega \times \mathscr{A} \to [0, 1]$ mit

$$(\omega, A) \mapsto \mathbb{P}(A|\mathscr{F})(\omega) = \sum_{i\in\mathcal{I}} \mathbb{P}(A|B_i)\mathbb{1}_{B_i}(\omega),$$

erfüllt die Bedingungen (Mk1) und (Mk2). Betrachte (Mk1). Da für jedes $\omega \in \Omega$ genau ein $i \in \mathcal{I}$ mit $\omega \in B_i$ existiert, ist nach Konstruktion $\mathbb{P}(\bullet|\mathscr{F})(\omega) = \mathbb{P}(\bullet|B_i)$ ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathscr{A})$. Andererseits erfüllt für jedes $A \in \mathscr{A}$ nach Konstruktion die Zufallsvariable $\mathbb{P}(A|\mathscr{F}) = \mathbb{E}(\mathbb{1}_A|\mathscr{F})$ die Bedingung (bE1), die Abbildung $\mathbb{P}(A|\mathscr{F}) : \Omega \to [0, 1]$ mit

$\omega \mapsto \mathbb{P}(A|\mathscr{F})(\omega)$ ist also $\mathscr{F}$-messbar. Folglich ist die Festlegung $\mathbb{P}(\bullet|\mathscr{F})$ in (06.06) der bedingten Verteilung von $\mathbb{P}$ gegeben $\mathscr{F}$ regulär. Andererseits ist per Konstruktion für jedes $\omega \in \Omega$ die Abbildung $\mathbb{E}(\bullet|\mathscr{F})(\omega)$ die Erwartung bzgl. $\mathbb{P}(\bullet|\mathscr{F})(\omega)$, so dass auch die Festlegung $\mathbb{E}(\bullet|\mathscr{F})$ in (06.05) regulär ist. □

§06.20 **Skizze**. Betrachte wie in Skizze §06.18 den Wahrscheinlichkeitsraum $([0,1), \mathscr{B}_{[0,1)}, \mathrm{U}_{[0,1)})$. Die nächsten Graphiken stellen die vorgestellte reguläre Festlegung des bedingten Erwartungswertes von X gegeben $\mathscr{F} = \sigma(\{B_i^{(j)}\})$, $j \in [\![2]\!]$, für die zwei Partitionen $B_i^{(1)} := [\frac{i-1}{4}, \frac{i}{4})$,$i \in [\![4]\!]$ und $B_i^{(2)} := [\frac{i-1}{8}, \frac{i}{8})$, $i \in [\![8]\!]$ von $[0,1)$ dar.



□

§06|03 Induzierte bedingte Verteilung

§06.21 **Erinnerung**. Ist X eine Zufallsvariable auf $(\Omega, \mathscr{A}, \mathbb{P})$ mit Werten in einem messbaren Raum $(\mathcal{X}, \mathscr{X})$, so heißt $\mathbb{P}^X = \mathbb{P} \circ X^{-1} = \mathbb{P}(X^{-1}(\bullet))$ die von $\mathbb{P}$ auf $(\mathcal{X}, \mathscr{X})$ induzierte Verteilung. □

§06.22 **Definition**. Sei X eine Zufallsvariable auf $(\Omega, \mathscr{A}, \mathbb{P})$ mit Werten in einem messbaren Raum $(\mathcal{X}, \mathscr{X})$ und $\mathscr{F} \subseteq \mathscr{A}$ eine Teil-σ-Algebra.

(a) Zu jeder Festlegung $\mathbb{P}(\bullet|\mathscr{F})$ der bedingten Verteilung von $\mathbb{P}$ gegeben $\mathscr{F}$, heißt die von X induzierte Abbildung $\boxed{\mathbb{P}^{X|\mathscr{F}}} : \mathscr{X} \to \overline{\mathscr{F}}^+$ mit $B \mapsto \mathbb{P}^{X|\mathscr{F}}(B) := \mathbb{P}(X^{-1}(B)|\mathscr{F})$ *Festlegung der bedingten Verteilung von X gegeben $\mathscr{F}$*, wobei für alle $F \in \mathscr{F}$ und $B \in \mathscr{X}$ somit $\mathbb{E}(\mathbb{1}_{X^{-1}(B)}\mathbb{1}_F) = \mathbb{E}(\mathbb{P}^{X|\mathscr{F}}(B)\mathbb{1}_F)$ gilt. Ist weiterhin

$$(\omega, B) \mapsto \mathbb{P}^{X|\mathscr{F}}(B)(\omega)$$

ein Markovkern von $(\Omega, \mathscr{F})$ nach $(\mathcal{X}, \mathscr{X})$, so wird diese Festlegung *regulär* genannt.

(b) Zu jeder Festlegung $\mathbb{E}(\bullet|\mathscr{F})$ der bedingten Erwartung bzgl. $\mathbb{P}$ gegeben $\mathscr{F}$, heißt $\mathbb{E}^{X|\mathscr{F}}(h) := \mathbb{E}(h(X)|\mathscr{F})$ für $h \in \overline{\mathscr{X}}^+$ *Festlegung des bedingten Erwartungswertes von $h(X)$ gegeben $\mathscr{F}$* und die von X induzierte Abbildung $\boxed{\mathbb{E}^{X|\mathscr{F}}} : \overline{\mathscr{X}}^+ \to \overline{\mathscr{F}}^+$ mit $h \mapsto \mathbb{E}^{X|\mathscr{F}}(h)$ *Festlegung der bedingten Erwartung von X gegeben $\mathscr{F}$*, wobei für alle $F \in \mathscr{F}$ und $h \in \overline{\mathscr{X}}^+$ die Radon-Nikodym-Gleichung (bE2) gilt $\mathbb{E}(h(X)\mathbb{1}_F) = \mathbb{E}(\mathbb{E}(h(X)|\mathscr{F})\mathbb{1}_F) = \mathbb{E}(\mathbb{E}^{X|\mathscr{F}}(h)\mathbb{1}_F)$. Die Festlegung $\mathbb{E}^{X|\mathscr{F}}$ heißt *regulär*, falls die induzierte Festlegung $\mathbb{P}^{X|\mathscr{F}}$ regulär ist, und für jedes $\omega \in \Omega$, $\mathbb{E}^{X|\mathscr{F}}(\bullet)(\omega)$ die Erwartung bzgl. $\mathbb{P}^{X|\mathscr{F}}(\bullet)(\omega)$ ist. □

§06.23 **Bemerkung**. $\mathbb{P}^{X|\mathscr{F}}$ erfüllt definitionsgemäß die Bedingung (Mk2). Sei zusätzlich die Festlegung $\mathbb{P}(\bullet|\mathscr{F})$ regulär, so dass (Mk1) erfüllt ist, also für alle $\omega \in \Omega$ die Abbildung

$$A \mapsto \mathbb{P}(A|\mathscr{F})(\omega)$$

ein Wahrscheinlichkeitsmaß ist, kurz $\mathbb{P}(\bullet|\mathscr{F})(\omega) \in \mathcal{W}(\mathscr{A})$. Damit ist auch $\mathbb{P}^{X|\mathscr{F}}(\bullet)(\omega) = \mathbb{P}(X^{-1}(\bullet)|\mathscr{F})(\omega) \in \mathcal{W}(\mathscr{X})$ für jedes $\omega \in \Omega$. Die Festlegung $\mathbb{P}^{X|\mathscr{F}}$ erfüllt also auch (Mk1). Sie ist somit ein Markovkern, also regulär. Für eine reguläre Festlegung $\mathbb{E}(\bullet|\mathscr{F})$ ist definitionsgemäß die induzierte Festlegung $\mathbb{P}(\bullet|\mathscr{F})$ und somit auch $\mathbb{P}^{X|\mathscr{F}}$ regulär. Für jedes $\omega \in \Omega$ ist nach dem Transformationssatz (Eigenschaft A04.05 (viii)) die induzierte Festlegung $\mathbb{E}^{X|\mathscr{F}}(\bullet)(\omega)$ gerade die Erwartung bzgl. $\mathbb{P}^{X|\mathscr{F}}(\bullet)(\omega)$. □

§06|04 Bedingte Erwartung gegeben eine Zufallsvariable

§06.24 **Erinnerung**. Sei $\mathscr{F} = \sigma(S)$ für eine Zufallsvariable S auf $(\Omega, \mathscr{A}, \mathbb{P})$ mit Werten in einem messbaren Raum $(\mathcal{S}, \mathscr{S})$. Wir bezeichnen wie bisher mit $\mathbb{P}^S$ das Bildmaß und mit $\mathbb{E}^S$ die entsprechende Erwartung. Ist Y eine $\sigma(S)$-messbare, positive numerische Zufallsvariable, also $Y \in \overline{\sigma(S)}^+$, so existiert nach Eigenschaft A02.06 (iv) ein $\varphi \in \overline{\mathscr{S}}^+$ mit $Y = \varphi(S)$, also $Y(\omega) = \varphi(S(\omega))$, $\omega \in \Omega$. Die Abbildung φ ist durch Y nur auf $S(\Omega)$ eindeutig festgelegt, für $s \notin S(\Omega)$ kann sie beliebig (messbar) fortgesetzt werden. □

§06.25 **Definition**. Für $S : (\Omega, \mathscr{A}) \to (\mathcal{S}, \mathscr{S})$ sei $\mathbb{E}(X|\sigma(S)) \in \overline{\sigma(S)}^+$ eine Festlegung des bedingten Erwartungswertes von $X \in \overline{\mathscr{A}}^+$ gegeben $\sigma(S)$ und $\varphi \in \overline{\mathscr{S}}^+$ wie in §06.24, also $\mathbb{E}(X|\sigma(S))(\omega) = \varphi(S(\omega))$, $\omega \in \Omega$. $\boxed{\mathbb{E}(X|S)} := \varphi$ und $\mathbb{E}(X|S = s) := \varphi(s)$ für $s \in \mathcal{S}$, heißen *Festlegung des bedingten Erwartungswertes von X gegeben S bzw. gegeben $S = s$*. Jede Abbildung $\boxed{\mathbb{E}(\bullet|S)} : \overline{\mathscr{A}}^+ \to \overline{\mathscr{S}}^+$ mit $X \mapsto \mathbb{E}(X|S)$ heißt *Festlegung der bedingten Erwartung bzgl. $\mathbb{P}$ gegeben S*. Die von $\mathbb{E}(\bullet|S)$ implizierte Abbildung $\boxed{\mathbb{P}(\bullet|S)} : \mathscr{A} \to \overline{\mathscr{S}}^+$ mit $A \mapsto \mathbb{P}(A|S) := \mathbb{E}(\mathbb{1}_A|S)$ sowie $\mathbb{P}(\bullet|S = s)$: $\mathscr{A} \to \mathbb{R}^+$ mit $A \mapsto \mathbb{P}(A|S = s) := \mathbb{E}(\mathbb{1}_A|S = s)$ wird *Festlegung der bedingten Verteilung von $\mathbb{P}$ gegeben S bzw. gegeben $S = s$* genannt. Die entsprechende Radon-Nikodym-Gleichheit (bE2) lautet $\mathbb{E}^S(\mathbb{1}_B\mathbb{P}(A|S)) = \int_B \mathbb{P}(A|S)d\mathbb{P}^S = \mathbb{P}(S^{-1}(B) \cap A)$ für alle $B \in \mathscr{S}$ und $A \in \mathscr{A}$. Die Festlegung $\mathbb{P}(\bullet|S)$ heißt *regulär*, falls $(s, A) \mapsto \mathbb{P}(A|S = s)$ ein Markovkern von $(\mathcal{S}, \mathscr{S})$ nach $(\Omega, \mathscr{A})$ ist. Analog wird die Festlegung $\mathbb{E}(\bullet|S)$ *regulär* genannt, wenn die implizierte Festlegung $\mathbb{P}(\bullet|S)$ regulär ist, und für jedes $s \in \mathcal{S}$, $\mathbb{E}(\bullet|S = s)$ die Erwartung bzgl. $\mathbb{P}(\bullet|S = s)$ ist. □

§06.26 **Eigenschaft**. *Für $X \in \overline{\mathscr{A}}^+$ seien $\varphi(S)$ und $\tilde{\varphi}(S)$ mit $\varphi, \tilde{\varphi} \in \overline{\mathscr{S}}^+$ zwei Festlegungen des bedingten Erwartungswertes $\mathbb{E}(X|\sigma(S))$ von X gegeben $\sigma(S)$. Nach Satz §06.02 gilt $\varphi(S) = \tilde{\varphi}(S)$ $\mathbb{P}$-f.ü. und damit auch $\varphi = \tilde{\varphi}$ $\mathbb{P}^S$-f.ü.. Im Gegensatz zu $\mathbb{E}(X|\sigma(S))$ hängt φ nicht nur von $\sigma(S)$, sondern auch von der speziellen Wahl der Zufallsvariable S ab.* □

§06.27 **Bemerkung**. $\mathbb{P}(\bullet|S)$ erfüllt definitionsgemäß die Bedingung (Mk2). Sei zusätzlich die Festlegung $\mathbb{P}(\bullet|\sigma(S))$ regulär, so dass auch (Mk1) erfüllt ist, also für alle $\omega \in \Omega$ die Abbildung

$$A \mapsto \mathbb{P}(A|\sigma(S))(\omega) = \mathbb{P}(A|S = S(\omega))$$

ein Wahrscheinlichkeitsmaß ist, kurz $\mathbb{P}(\bullet|S = S(\omega)) \in \mathcal{W}(\mathscr{A})$. Damit gilt $\mathbb{P}(\bullet|S = s) \in \mathcal{W}(\mathscr{A})$ für jedes $s \in S(\Omega)$. Für jedes surjektive S erfüllt die Festlegung $\mathbb{P}(\bullet|S)$ also auch (Mk1). Sie ist somit ein Markovkern, also regulär. Ist S nicht surjektiv, aber $S(\Omega) \in \mathscr{S}$, so ist zum Beispiel die Festlegung $\mathbb{P}(\bullet|S)\mathbb{1}_{S(\Omega)} + \mathbb{P}\mathbb{1}_{\mathcal{S}\setminus S(\Omega)}$ ein Markovkern. Analog, ist die Festlegung $\mathbb{E}(\bullet|\sigma(S))$ regulär, also $\mathbb{P}(\bullet|\sigma(S))$ ist regulär und für jedes $\omega \in \Omega$ ist $\mathbb{E}(\bullet|\sigma(S))(\omega) =$

$\mathbb{E}\big(\bullet\,\big|S=S(\omega)\big)$ die Erwartung bzgl. $\mathbb{P}\big(A\big|\sigma(S)\big)(\omega)=\mathbb{P}\big(A\big|S=S(\omega)\big)$. Ist S surjektiv, so ist auch $\mathbb{E}\big(\bullet\,\big|S\big)$ regulär, falls $S(\Omega)=\mathcal{S}$ gilt. Ist S nicht surjektiv, aber $S(\Omega)\in\mathscr{S}$, so ist die Festlegung $\mathbb{E}\big(\bullet\,\big|S\big)\mathbb{1}_{S(\Omega)}+\mathbb{E}(\bullet)\mathbb{1}_{\mathcal{S}\setminus S(\Omega)}$ regulär. Andererseits, ist die Festlegung $\mathbb{P}\big(\bullet\,\big|S\big)$ regulär, also ein Markovkern von $(\mathcal{S},\mathscr{S})$ nach $(\Omega,\mathscr{A})$, dann ist auch $(\omega,A)\mapsto\mathbb{P}\big(A\big|S=S(\omega)\big)=:\mathbb{P}\big(A\big|\sigma(S)\big)(\omega)$ eine reguläre Festlegung der bedingten Verteilung von $\mathbb{P}$ gegeben $\sigma(S)$, also ein Markovkern $(\Omega,\sigma(S))$ nach $(\Omega,\mathscr{A})$. Offensichtlich gilt (Mk2), da die Hintereinanderausführung messbarer Abbildungen messbar ist. Weiterhin gilt für alle $\omega\in\Omega$ auch $\mathbb{P}\big(A\big|S=S(\omega)\big)\in\mathcal{W}(\mathscr{A})$, also (Mk1). □

§06.28 **Beispiel**. Sei nun $S:(\Omega,\mathscr{A})\to(\{0,1\},2^{\{0,1\}})$ eine einfache Zufallsvariable, also $\{S=1\}=S^{-1}(\{1\})\in\mathscr{A}$ und $\sigma(S)=\sigma(\{S^{-1}(\{1\})\})=:\mathscr{F}$. Eine Festlegung der bedingten Erwartung bzgl. $\mathbb{P}$ gegeben $\sigma(S)$ ist somit

$$\mathbb{E}\big(\bullet\,\big|\sigma(S)\big):\overline{\mathscr{A}}^+\to\overline{\sigma(S)}^+\text{ mit}$$
$$X\mapsto\mathbb{E}\big(X\big|\sigma(S)\big)=\mathbb{E}\big(X\big|S^{-1}(\{1\})\big)\mathbb{1}_{S^{-1}(\{1\})}+\mathbb{E}\big(X\big|S^{-1}(\{0\})\big)\mathbb{1}_{S^{-1}(\{0\})}=\varphi(S)$$
und $\varphi\in\overline{2^{\{0,1\}}}^+$ mit
$$s\mapsto\varphi(s)=\mathbb{E}\big(X\big|S^{-1}(\{1\})\big)\mathbb{1}_{\{1\}}(s)+\mathbb{E}\big(X\big|S^{-1}(\{0\})\big)\mathbb{1}_{\{0\}}(s).$$

Damit ist $\mathbb{E}\big(\bullet\,\big|S\big)=\mathbb{E}\big(\bullet\,\big|S^{-1}(\{1\})\big)\mathbb{1}_{\{1\}}+\mathbb{E}\big(\bullet\,\big|S^{-1}(\{0\})\big)\mathbb{1}_{\{0\}}$ und $\mathbb{E}\big(\bullet\,\big|S=s\big)=\mathbb{E}\big(\bullet\,\big|S^{-1}(\{1\})\big)\mathbb{1}_{\{1\}}(s)+\mathbb{E}\big(\bullet\,\big|S^{-1}(\{0\})\big)\mathbb{1}_{\{0\}}(s)$, $s\in\{0,1\}$ eine *Festlegung des bedingten Erwartungswertes von X gegeben S bzw. gegeben $S=s$*. Wir halten fest, dass insbesondere $\mathbb{E}\big(\bullet\,\big|S=1\big)=\mathbb{E}\big(\bullet\,\big|S^{-1}(\{1\})\big)$ und $\mathbb{E}\big(\bullet\,\big|S=0\big)=\mathbb{E}\big(\bullet\,\big|S^{-1}(\{0\})\big)$ gilt. □

§06|05 Induzierte bedingte Verteilung gegeben eine Zufallsvariable

§06.29 **Erinnerung**. Sei (S,X) eine Zufallsvariable auf $(\Omega,\mathscr{A},\mathbb{P})$ mit Werten in dem Produktraum $\mathcal{S}\times\mathcal{X}$ versehen mit der Produkt-σ-Algebra $\mathscr{S}\otimes\mathscr{X}$. Wir bezeichnen mit $\mathbb{P}^{(S,X)}$ die von (S,X) auf $(\mathcal{S}\times\mathcal{X},\mathscr{S}\otimes\mathscr{X})$ induzierte gemeinsame Verteilung. Seien wie bisher $\Pi_{\mathcal{X}}:\mathcal{S}\times\mathcal{X}\to\mathcal{X}$ und $\Pi_{\mathcal{S}}:\mathcal{S}\times\mathcal{X}\to\mathcal{S}$ mit $(s,x)\mapsto\Pi_{\mathcal{X}}(s,x):=x$ bzw. $(s,x)\mapsto\Pi_{\mathcal{S}}(s,x):=s$ die entsprechenden Koordinatenabbildungen. Für die Randverteilungen von X und S gilt dann $\mathbb{P}^X=\mathbb{P}\circ X^{-1}=\mathbb{P}\circ(\Pi_{\mathcal{X}}(S,X))^{-1}=\mathbb{P}^{(S,X)}\circ\Pi_{\mathcal{X}}^{-1}$ bzw. $\mathbb{P}^S=\mathbb{P}^{(X,S)}\circ\Pi_{\mathcal{S}}^{-1}$. Die Statistiken $\Pi_{\mathcal{X}}$ und $\Pi_{\mathcal{S}}$ sind offensichtlich surjektiv. Unser Ziel ist eine bedingte Verteilung von X gegeben S einzuführen unter Verwendung der Definitionen §06.22 und §06.25. Geben wir uns eine Festlegung $\mathbb{P}\big(\bullet\,\big|\sigma(S)\big)$ vor, so impliziert diese mit Definition §06.22 eine Festlegung $\mathbb{P}^{X|\sigma(S)}$ mit Werten in $\overline{\sigma(S)}^+$, welche weiterhin mit Definition §06.25 eine Festlegung $\mathbb{P}^{X|S}$ mit Werten in $\overline{\mathscr{S}}^+$ impliziert. Ist die Festlegung $\mathbb{P}\big(\bullet\,\big|\sigma(S)\big)$ regulär, so auch $\mathbb{P}^{X|\sigma(S)}$, aber wie wir in Bemerkung §06.27 gesehen haben, ist im Allgemeinen $\mathbb{P}^{X|S}$ nicht regulär. Andererseits, können wir uns eine Festlegung $\mathbb{P}^{(S,X)}\big(\bullet\,\big|\sigma(\Pi_{\mathcal{S}})\big)$ vorgeben, so impliziert diese mit Definition §06.22 eine Festlegung $\mathbb{P}^{X|\sigma(\Pi_{\mathcal{S}})}$ mit Werten in $\overline{\sigma(\Pi_{\mathcal{S}})}^+$, welche weiterhin mit Definition §06.25 eine Festlegung $\mathbb{P}^{X|S}$ mit Werten in $\overline{\mathscr{S}}^+$ impliziert. Da $\Pi_{\mathcal{S}}$ surjektiv ist, folgt die Regularität der Festlegung $\mathbb{P}^{X|S}$, falls wir mit einer regulären Festlegung $\mathbb{P}^{(S,X)}\big(\bullet\,\big|\sigma(\Pi_{\mathcal{S}})\big)$ angefangen haben. □

§06.30 **Definition**. Sei $\mathbb{P}^{(S,X)}$ die gemeinsame Verteilung von (S,X) auf $(\mathcal{S}\times\mathcal{X},\mathscr{S}\otimes\mathscr{X})$.

(a) Zu jeder Festlegung $\mathbb{P}^{(S,X)}\big(\bullet\,\big|\sigma(\Pi_{\mathcal{S}})\big)$ auf $(\mathcal{S}\times\mathcal{X},\mathscr{S}\otimes\mathscr{X})$ der bedingten Verteilung von $\mathbb{P}^{(S,X)}$ gegeben $\sigma(\Pi_{\mathcal{S}})$, heißt die Abbildung

$$\boxed{\mathbb{P}^{X|S}}:\mathscr{X}\to\overline{\mathscr{S}}^+\text{ mit }A\mapsto\mathbb{P}^{X|S}(A):=\varphi\text{ und}$$
$$\mathbb{P}^{X|\sigma(\Pi_{\mathcal{S}})}(A)=\mathbb{P}^{(X,S)}\big(\Pi_{\mathcal{X}}^{-1}(A)\big|\sigma(\Pi_{\mathcal{S}})\big)=\varphi(\Pi_{\mathcal{S}})$$

Festlegung der bedingten Verteilung von X gegeben S und $\boxed{\mathbb{P}^{X|S=s}(A)} = \varphi(s)$ *Festlegung der bedingten Wahrscheinlichkeit von $X \in A$ bei gegebenem $S = s$*. Die entsprechende Radon-Nikodym-Gleichheit (bE2) lautet $\mathbb{P}^S(\mathbb{P}^{X|S}(A)\mathbb{1}_B) = \int_B \mathbb{P}^{X|S}(A)\, d\mathbb{P}^S = \mathbb{P}(X^{-1}(A) \cap S^{-1}(B)) = \mathbb{P}^{(S,X)}(B \times A)$ für alle $A \in \mathscr{X}$ und $B \in \mathscr{S}$. Ist weiterhin $(s, A) \mapsto \mathbb{P}^{X|S=s}(A)$ ein Markovkern von $(\mathcal{S}, \mathscr{S})$ nach $(\mathcal{X}, \mathscr{X})$, so wird diese Festlegung *regulär* genannt, wobei dann auf Grund der Definition §06.04 (bE2) $\mathbb{P}^S \odot \mathbb{P}^{X|S} = \mathbb{P}^{(S,X)}$ gilt (vgl. Schreibweise §06.07). Besitzt in diesem Fall für ein $s \in \mathcal{S}$ die Verteilung $\mathbb{P}^{X|S=s}$ ein endliches erstes absolutes Moment, so schreiben wir auch $\boxed{\mathbb{E}^{X|S=s}(X)} := \mathbb{E}^{X|S=s}(\mathrm{id}_{\mathcal{X}}) = \int_{\mathcal{X}} x\mathbb{P}^{X|S=s}(dx)$.

(b) Zu jeder Festlegung $\mathbb{E}^{(S,X)}\big(\bullet\,\big|\sigma(\Pi_s)\big)$ der bedingten Erwartung bzgl. $\mathbb{P}^{(S,X)}$ gegeben $\sigma(\Pi_s)$, heißt die Abbildung

$$\boxed{\mathbb{E}^{X|S}} : \overline{\mathscr{X}}^+ \to \overline{\mathscr{S}}^+ \text{ mit } h \mapsto \mathbb{E}^{X|S}(h) := \varphi \text{ und}$$
$$\mathbb{E}^{X|\sigma(\Pi_s)}(h) = \mathbb{E}_{(S,X)}\big(h(\Pi_x)\big|\sigma(\Pi_s)\big) = \varphi(\Pi_s)$$

und $\boxed{\mathbb{E}^{X|S=s}(h)} = \varphi(s)$ *Festlegung der bedingten Erwartung von X gegeben S bzw. Festlegung des bedingten Erwartungswertes von $h(X)$ gegeben $S = s$*. Die Festlegung $\mathbb{E}^{X|S}$ heißt *regulär*, falls die induzierte Festlegung $\mathbb{P}^{X|S}$ regulär ist, und für jedes $s \in \mathcal{S}$, $\mathbb{E}^{X|S=s}$ die Erwartung bzgl. $\mathbb{P}^{X|S=s}$ ist. □

§07 Integrierbare Zufallsvariablen

§07.01 **Erinnerung**. Seien $(\Omega, \mathscr{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, $\mathscr{F} \subseteq \mathscr{A}$ eine Teil-σ-Algebra und $X \in \overline{\mathscr{A}}$ eine numerische Zufallsvariable. Mit Hilfe der Zerlegung $X = X^+ - X^-$ mit $X^+, X^- \in \overline{\mathscr{A}}^+$ haben wir in Definition A04.03 (b) für X mit $\mathbb{E}(|X|) \in \mathbb{R}^+$ also $\mathbb{E}(X^+) \in \mathbb{R}^+$ und $\mathbb{E}(X^-) \in \mathbb{R}^+$ den Erwartungswert $\mathbb{E}(X) := \mathbb{E}(X^+) - \mathbb{E}(X^-)$ definiert. Im Folgenden sei $\mathscr{L}_1(\mathscr{A}, \mathbb{P}) := \{X \in \overline{\mathscr{A}} : \mathbb{E}(|X|) \in \mathbb{R}^+\}$ und zu $\mathbb{P}$ bezeichne $\mathbb{E} : \mathscr{L}_1(\mathscr{A}, \mathbb{P}) \to \mathbb{R}$ die eindeutig bestimmte Erwartung. Da auch $\overline{\mathscr{F}} \subseteq \overline{\mathscr{A}}$ ist $\mathscr{L}_1(\mathscr{F}, \mathbb{P}) \subseteq \mathscr{L}_1(\mathscr{A}, \mathbb{P})$. Sei $X \in \mathscr{L}_1(\mathscr{A}, \mathbb{P})$, dann gilt $\mathbb{E}(X^+) \in \mathbb{R}^+$ und für jede Festlegung $\mathbb{E}\big(X^+\big|\mathscr{F}\big)$ gilt (bE1), $\mathbb{E}\big(X^+\big|\mathscr{F}\big) \in \overline{\mathscr{F}}^+$ und (bE2), $\mathbb{E}(\mathbb{1}_F \mathbb{E}\big(X^+\big|\mathscr{F}\big)) = \mathbb{E}(\mathbb{1}_F X^+)$ für alle $F \in \mathscr{F}$, also insbesondere mit $F = \Omega$ auch $\mathbb{E}(\mathbb{E}\big(X^+\big|\mathscr{F}\big)) = \mathbb{E}(X^+) \in \mathbb{R}^+$. Somit kann $\mathbb{E}\big(X^+\big|\mathscr{F}\big) \in \mathscr{F}^+$ und analog $\mathbb{E}\big(X^-\big|\mathscr{F}\big) \in \mathscr{F}^+$ gewählt werden. Damit erfüllt $Y := \mathbb{E}\big(X^+\big|\mathscr{F}\big) - \mathbb{E}\big(X^-\big|\mathscr{F}\big) \in \mathscr{F}$, als auch die Radon-Nikodym-Gleichung (bE2), wobei insbesondere $Y \in \mathscr{L}_1(\mathscr{F}, \mathbb{P})$. □

§07.02 **Definition**. Für $X \in \mathscr{L}_1(\mathscr{A}, \mathbb{P})$ heißt jede Abbildung $Y : \Omega \to \overline{\mathbb{R}}$ *Festlegung (Version) des bedingten Erwartungswertes* von X gegeben $\mathscr{F}$, symbolisch $\boxed{\mathbb{E}\big(X\big|\mathscr{F}\big)} := Y$, wenn sie die folgenden zwei Bedingungen erfüllt:

(bE1) Y ist $\mathscr{F}$-$\overline{\mathscr{B}}$-messbar, also $Y \in \overline{\mathscr{F}}$ und

(bE2) $\mathbb{E}(\mathbb{1}_F Y) = \mathbb{E}(\mathbb{1}_F X)$ für jedes $F \in \mathscr{F}$. (*Radon-Nikodym-Gleichung*)

Jede Abbildung $\mathbb{E}\big(\bullet\,\big|\mathscr{F}\big) : \mathscr{L}_1(\mathscr{A}, \mathbb{P}) \to \overline{\mathscr{F}}$ mit $X \mapsto \mathbb{E}\big(X\big|\mathscr{F}\big)$ heißt *Festlegung der bedingten Erwartung bzgl. $\mathbb{P}$ gegeben $\mathscr{F}$*. $\mathbb{E}\big(\bullet\,\big|\mathscr{F}\big)$ heißt *reguläre Festlegung* der bedingten Erwartung bzgl. $\mathbb{P}$ gegeben $\mathscr{F}$, wenn die implizierte Festlegung $\mathbb{P}\big(\bullet\,\big|\mathscr{F}\big)$ der bedingten Verteilung von $\mathbb{P}$ gegeben $\mathscr{F}$ regulär ist, und für jedes $\omega \in \Omega$ die Abbildung $\mathbb{E}\big(\bullet\,\big|\mathscr{F}\big)(\omega) : \mathscr{L}_1(\mathscr{A}, \mathbb{P}) \to \mathbb{R}$ die Erwartung bzgl. $\mathbb{P}\big(\bullet\,\big|\mathscr{F}\big)(\omega)$ ist. □

§07.03 **Bemerkung**.

(i) Sei X eine Zufallsvariable auf $(\Omega, \mathscr{A}, \mathbb{P})$ mit Werten in einem messbaren Raum $(\mathcal{X}, \mathscr{X})$ und $\mathscr{F} \subseteq \mathscr{A}$ eine Teil-σ-Algebra. Dann definieren wir analog zu Definition §06.22 (b) nun für $h \in \mathscr{L}_1(\mathscr{X}, \mathbb{P}^X)$ eine Festlegung $\mathbb{E}^{X|\mathscr{F}}(h) \in \overline{\mathscr{F}}$ des bedingten Erwartungswertes von $h(X)$ gegeben $\mathscr{F}$ sowie eine *(reguläre) Festlegung* $\boxed{\mathbb{E}^{X|\mathscr{F}}} : \mathscr{L}_1(\mathscr{X}, \mathbb{P}^X) \to \overline{\mathscr{F}}$ mit $h \mapsto \mathbb{E}^{X|\mathscr{F}}(h)$ *der bedingten Erwartung von X gegeben $\mathscr{F}$*.

(ii) Für $S : (\Omega, \mathscr{A}) \to (\mathcal{S}, \mathscr{S})$ definieren wir nun für $X \in \mathscr{L}_1(\mathscr{X}, \mathbb{P})$ wie in Definition §06.25 eine Festlegung $\mathbb{E}(X|S) \in \overline{\mathscr{S}}$ und $\mathbb{E}(X|S = s) \in \overline{\mathbb{R}}$ des bedingten Erwartungswertes von X gegeben S bzw. gegeben $S = s$ sowie eine *(reguläre) Festlegung* $\boxed{\mathbb{E}(\bullet|S)} : \mathscr{L}_1(\mathscr{A}, \mathbb{P}) \to \overline{\mathscr{S}}$ mit $X \mapsto \mathbb{E}(X|S)$ *der bedingten Erwartung bzgl. $\mathbb{P}$ gegeben S*.

(iii) Sei $(S, X) : (\Omega, \mathscr{A}) \to (\mathcal{S} \times \mathcal{X}, \mathscr{S} \otimes \mathscr{X})$ mit gemeinsamer Verteilung $\mathbb{P}^{(S,X)}$. Analog zu Definition §06.30 (b) definiere für $h \in \mathscr{L}_1(\mathscr{X}, \mathbb{P}^X)$ eine Festlegung $\mathbb{E}^{X|S}(h) \in \overline{\mathscr{S}}$ und $\mathbb{E}^{X|S=s}(h) \in \overline{\mathbb{R}}$ des bedingten Erwartungswertes von $h(X)$ gegeben S bzw. gegeben $S = s$ sowie eine *(reguläre) Festlegung* $\boxed{\mathbb{E}^{X|S}} : \mathscr{L}_1(\mathscr{X}, \mathbb{P}^X) \to \overline{\mathscr{S}}$ mit $h \mapsto \mathbb{E}^{X|S}(h)$ *der bedingten Erwartung von X gegeben S*. Ist analog zu Definition §06.30 (a) $\mathbb{P}^{X|S}$ *eine reguläre Festlegung der bedingten Verteilung von X gegeben S*, und besitzt für ein $s \in \mathcal{S}$ die Verteilung $\mathbb{P}^{X|S=s}$ zum Beispiel ein endliches erstes absolutes Moment, so schreiben wir auch $\boxed{\mathbb{E}^{X|S=s}(X)} := \mathbb{E}^{X|S=s}(\mathrm{id}_1) = \int_{\mathcal{X}} x \mathbb{P}^{X|S=s}(dx)$. □

§07.04 **Satz**. *Seien $X, Y \in \mathscr{L}_1(\mathscr{A}, \mathbb{P})$ und $\mathscr{F} \subseteq \mathscr{A}$ eine Teil-σ-Algebra. Für jede Festlegung der bedingten Erwartung bzgl. $\mathbb{P}$ gegeben $\mathscr{F}$ gelten die folgenden Aussagen* $\mathbb{P}$-f.s.,

(i) *Für alle $a, b \in \mathbb{R}$ gilt $\mathbb{E}(aX + bY|\mathscr{F}) = a\mathbb{E}(X|\mathscr{F}) + b\mathbb{E}(Y|\mathscr{F})$;* *(linear)*

(ii) *Für $X \leqslant Y$ gilt $\mathbb{E}(X|\mathscr{F}) \leqslant \mathbb{E}(Y|\mathscr{F})$;* *(monoton)*

(iii) $|\mathbb{E}(X|\mathscr{F})| \leqslant \mathbb{E}(|X||\mathscr{F})$; *(Dreiecksungleichung)*

(iv) *Für $S \in \overline{\mathscr{A}}$ mit $\mathbb{E}(|S||\mathscr{F}) \in \mathscr{B}^+$ gilt $\mathbb{P}(|S| = \infty) = 0$.* *(endlich)*

(v) *Für $\phi : \mathbb{R} \to \mathbb{R}$ konvex mit $\phi(X) \in \mathscr{L}_1(\mathscr{A}, \mathbb{P})$ gilt $\phi(\mathbb{E}(X|\mathscr{F})) \leqslant \mathbb{E}((\phi(X))|\mathscr{F})$.* *(Ungleichung von Jensen)*

(vi) *Für $X_n \uparrow X$ $\mathbb{P}$-f.ü. gilt $\sup_{n\in\mathbb{N}} \mathbb{E}(X_n|\mathscr{F}) = \mathbb{E}(X|\mathscr{F})$.* *(monotone Konvergenz)*

(vii) *Für $X_n \to X$ $\mathbb{P}$-f.ü. mit $|X_n| \leqslant Y$, $n \in \mathbb{N}$, gilt $\lim_{n\to\infty} \mathbb{E}(X_n|\mathscr{F}) = \mathbb{E}(X|\mathscr{F})$ $\mathbb{P}$-f.s. und in $\mathscr{L}_1(\mathscr{A}, \mathbb{P})$.* *(dominierte Konvergenz)*

Ist die Festlegung regulär, also $\mathbb{E}(\bullet|\mathscr{F})(\omega)$ eine Erwartung für alle $\omega \in \Omega$, so gelten die Aussagen (i)-(vii) für alle $\omega \in \Omega$.

§07.05 **Beweis** von Satz §07.04. In der Vorlesung, für die reguläre Festlegung folgt die Behauptung direkt aus den Eigenschaft A04.05. □

§07.06 **Satz**. *Seien $X, Y \in \mathscr{L}_1(\mathscr{A}, \mathbb{P})$ und $\mathscr{G} \subseteq \mathscr{F} \subseteq \mathscr{A}$ Teil-σ-Algebren. Für jede Festlegung der bedingten Erwartung gelten die folgenden Aussagen* $\mathbb{P}$-f.s.,

(i) *Für $\mathbb{E}(|XY|) < \infty$ und $Y \in \mathscr{F}$ gelten*

$$\mathbb{E}(XY|\mathscr{F}) = Y\mathbb{E}(X|\mathscr{F}) \text{ und } \mathbb{E}(Y|\mathscr{F}) = \mathbb{E}(Y|\sigma(Y)) = Y$$

(ii) $\mathbb{E}(\mathbb{E}(X|\mathscr{F})|\mathscr{G}) = \mathbb{E}(\mathbb{E}(X|\mathscr{G})|\mathscr{F}) = \mathbb{E}(X|\mathscr{G})$; *(Turmeigenschaft)*

(iii) *Sind $\sigma(X)$ und $\mathscr{F}$ unabhängig, so gilt $\mathbb{E}(X|\mathscr{F}) = \mathbb{E}(X)$;* *(unabhängig)*

(iv) *Für $\overline{\mathscr{T}} := \{A \in \mathscr{A} \,|\, \mathbb{P}(A) \in \{0, 1\}\}$ gilt $\mathbb{E}(X|\overline{\mathscr{T}}) = \mathbb{E}(X)$.*

(v) $\mathbb{E}(\mathbb{E}(X|\mathscr{F})) = \mathbb{E}(X)$. *(Satz von der totalen Wahrscheinlichkeit)*

Seien weiterhin S und T Zufallsvariablen von $(\Omega, \mathscr{A})$ nach $(\mathcal{S}, \mathscr{S})$ sowie $(\mathcal{T}, \mathscr{T})$.

(vi) *Für $h \in \mathscr{S}$ mit $\mathbb{E}(|h(S)X|) < \infty$ gilt $\mathbb{E}(h(S)X|S) = h(S)\mathbb{E}(X|S)$* $\mathbb{P}^S$-f.s.

(vii) *Sind (X,S) und (Y,T) unabhängig, so gilt* $\mathbb{E}\big(XY\big|(S,T)\big) = \mathbb{E}\big(X\big|S\big)\mathbb{E}\big(Y\big|T\big)$ $\mathbb{P}^{(S,T)}$-f.s..

§07.07 **Beweis** von Satz §07.06. In der Vorlesung. □

§07.08 **Satz**. *Es seien $\mathbb{P}_0, \mathbb{P}_1 \in \mathcal{W}(\mathscr{A})$ Wahrscheinlichkeitsmaße mit $\mathbb{P}_0 \ll \mathbb{P}_1$. Für jede Zufallsvariable $S : (\Omega, \mathscr{A}) \to (\mathcal{S}, \mathscr{S})$ gilt:*

(i) $\mathbb{P}_0^S \ll \mathbb{P}_1^S$.

(ii) $\dfrac{d\mathbb{P}_0^S}{d\mathbb{P}_1^S} = \mathbb{P}_1\left(\dfrac{d\mathbb{P}_0}{d\mathbb{P}_1}\,\middle|\, S\right)$ $\mathbb{P}_1^S$-f.s.

(iii) *Ist $h \in \overline{\mathscr{A}}$, für die $\mathbb{P}_0(h)$ existiert, so gilt* $\mathbb{P}_1\left(h\dfrac{d\mathbb{P}_0}{d\mathbb{P}_1}\,\middle|\, S\right) = \mathbb{P}_0\big(h\big|S\big)\dfrac{d\mathbb{P}_0^S}{d\mathbb{P}_1^S}$ $\mathbb{P}_1^S$-f.s.. *Insbesondere gilt* $\mathbb{P}_0\big(h\big|S\big) = \dfrac{\mathbb{P}_1\big(h\mathbb{f}\big|S\big)}{\mathbb{P}_1\big(\mathbb{f}\big|S\big)}$ $\mathbb{P}_0^S$-f.s. *mit* $\mathbb{f} := d\mathbb{P}_0/d\mathbb{P}_1$.

Bezeichnet $\widetilde{\mathbb{P}}_0 \in \mathcal{W}(\mathscr{A})$ ein weiteres Wahrscheinlichkeitsmaß mit $\widetilde{\mathbb{P}}_0 \ll \mathbb{P}_1$, L einen Dichtequotienten (DQ) von $\widetilde{\mathbb{P}}_0$ bezüglich $\mathbb{P}_0$, L^S einen DQ von $\widetilde{\mathbb{P}}_0^S$ bezüglich $\mathbb{P}_0^S$, sowie $\mathcal{N}$ bzw. $\mathcal{N}^S$ die singulären Bereiche von $\widetilde{\mathbb{P}}_0$ bezüglich $\mathbb{P}_0$ bzw. von $\widetilde{\mathbb{P}}_0^S$ bezüglich $\mathbb{P}_0^S$, so gilt mit $\varphi := \widetilde{\mathbb{P}}_0\big(N^c\big|S\big)$

(iv) $\widetilde{\mathbb{P}}_0(N) = \widetilde{\mathbb{P}}_0^S(\mathcal{N}^S) + \mathbb{P}_0(\mathrm{L}^S(S)(1-\varphi(S)))$,

(v) $\mathbb{P}_0\big(\mathrm{L}\big|S\big) = \mathrm{L}^S\varphi$ $\mathbb{P}_0^S$-f.s..

§07.09 **Beweis** von Satz §07.08. (i)-(iii) in der Vorlesung, (iv),(v) Witting [1985, Satz 1.121, S.121] □

§07|01 Reguläre Festlegungen

§07.10 **Erinnerung**. Sei (S,X) eine Zufallsvariable auf $(\Omega, \mathscr{A}, \mathbb{P})$ mit Werten in dem Produktraum $\mathcal{S} \times \mathcal{X}$ versehen mit der Produkt-σ-Algebra $\mathscr{S} \otimes \mathscr{X}$. Eine reguläre Festlegung der bedingten Verteilung $\mathbb{P}^{X|S}$ von X gegeben S, wenn sie existiert, ist aufgefasst als Abbildung $(s,A) \mapsto \mathbb{P}^{X|S=s}(A)$ ein Markovkern von $(\mathcal{S}, \mathscr{S})$ nach $(\mathcal{X}, \mathscr{X})$, der Lösung der Radon-Nikodym-Gleichung $\mathbb{P}^S \odot \mathbb{P}^{X|S} = \mathbb{P}^{(S,X)}$ ist, d.h. für jedes $A \in \mathscr{X}$ gilt

$$\mathbb{P}^{(S,X)}(B \times A) = \mathbb{P}(S^{-1}(B) \cap X^{-1}(A)) = \int_B \mathbb{P}^{X|S}(A)\, d\mathbb{P}^S = \mathbb{P}^S(\mathbb{1}_B \mathbb{P}^{X|S}(A)) \quad \forall\, B \in \mathscr{S}$$

Häufig ist es nützlich, neben $\mathbb{P}^{X|S}$ auch eine reguläre Festlegung der bedingten Verteilung $\mathbb{P}^{(S,X)|S}$ von (S,X) bei gegebenen S zu betrachten, also einen Markovkern $(s,D) \mapsto \mathbb{P}^{(S,X)|S=s}(D)$ von $(\mathcal{S}, \mathscr{S})$ nach $(\mathcal{S} \times \mathcal{X}, \mathscr{S} \otimes \mathscr{X})$, der für jedes $D \in \mathscr{S} \otimes \mathscr{X}$ Lösung der entsprechenden Radon-Nikodym-Gleichung

$$\begin{aligned}\mathbb{P}^{(S,X)}(D \cap (B \times \mathcal{X})) = \mathbb{P}((S,X)^{-1}(D) \cap S^{-1}(B)) &= \int_B \mathbb{P}^{(S,X)|S}(D)\, d\mathbb{P}^S \\ &= \mathbb{P}^S(\mathbb{P}^{(S,X)|S}(D)\mathbb{1}_B) \quad \forall\, B \in \mathscr{S}\end{aligned}$$

ist. Dann ergibt sich die bedingte Verteilung $\mathbb{P}^{X|S=s}$ aus der bedingten Verteilung $\mathbb{P}^{(S,X)|S=s}$ wieder als Randverteilung gemäß $\mathbb{P}^{X|S=s}(A) = \mathbb{P}^{(S,X)|S=s}(\mathcal{S} \times A)$, $A \in \mathscr{X}$. Wir bezeichnen weiterhin mit $D_s := \{x \in \mathcal{X} : (s,x) \in D\}$ bzw. $f_s(x) := f(s,x)$ den *s-Schnitt* von $D \subseteq \mathcal{S} \times \mathcal{X}$ bzw. $f : \mathcal{S} \times \mathcal{X} \to \mathbb{R}$, jeweils bei gegebenem $s \in \mathcal{S}$. □

§07.11 **Satz**. *Sei (S,X) eine Zufallsvariable auf $(\Omega, \mathscr{A}, \mathbb{P})$ mit Werten in $(\mathcal{S} \times \mathcal{X}, \mathscr{S} \otimes \mathscr{X})$. Es existiere eine reguläre Festlegung $\mathbb{P}^{X|S}$ der bedingten Verteilung von X bei gegebenen S.*

(i) *Dann ist $(s,D) \mapsto \mathbb{P}^{(S,X)|S=s}(D) := \mathbb{P}^{X|S=s}(D_s)$ eine reguläre Festlegung der bedingten Verteilung von (S,X) bei gegebenen S.*

(ii) *Für jede Funktion $f \in \mathscr{S} \otimes \mathscr{X}$, für die $\mathbb{P}^{(S,X)}(f) = \mathbb{E}(f(S,X))$ existiert, gilt*

$$\mathbb{P}^{(S,X)|S=s}(f) = \int f_s \, d\mathbb{P}^{X|S=s} = \mathbb{P}^{X|S=s}(f_s) \quad \textit{für } \mathbb{P}^S\text{-f.a. } s \in \mathcal{S},$$

sowie für alle $A \in \mathscr{X}$ und $B \in \mathscr{S}$ gilt

$$\mathbb{P}^{(S,X)}(f\mathbb{1}_{B\times A}) = \mathbb{P}^S\big(\mathbb{1}_B \mathbb{P}^{(S,X)|S}(\mathbb{1}_A f)\big) = \int_B \Big(\int_A f_s \, d\mathbb{P}^{X|S=s}\Big)\mathbb{P}^S(ds).$$

§07.12 **Beweis** von Satz §07.11. In der Vorlesung. □

§07.13 **Bemerkung**. Intuitiv liegt es nahe, $\mathbb{P}^{X|S}(f_S)$ für $\mathbb{P}^{(X,S)|S}(f)$ zu schreiben. Das ist jedoch nicht gerechtfertigt, da sich $\mathbb{P}^{X|S}(f_S)$ nicht als $\mathscr{S}$-messbare Lösung einer Radon-Nikodym-Gleichung gewinnen lässt, und somit nicht in der üblichen Weise als bedingter Erwartungswert formuliert werden kann (siehe Witting [1985, S.125] für eine weiterführende Diskussion). Andererseits kann zur vereinfachenden Berechnung im Integranden $f(S,X)$ durch $f(s,X)$ ersetzt werden, also die symbolische Schreibweise $S = s$ als ein Einsetzen verstanden werden. Die Rechtfertigung eines derartigen Einsetzens folgt daraus, dass $(s,D) \mapsto \mathbb{P}^{X|S=s}(D_s)$ unter den Voraussetzungen von Lemma §07.27 eine Version der regulären Festlegung $\mathbb{P}^{(S,X)|S}$ ist. Für $D = \{s\} \times \mathcal{X}$ gilt damit $\mathbb{P}^{(S,X)|S=s}(D) = \mathbb{P}^{X|S=s}(\mathcal{X}) = 1$ für $\mathbb{P}^S$-f.a. $s \in \mathcal{S}$. Die bedingte Verteilung $\mathbb{P}^{(S,X)|S}$ konzentriert sich also für $\mathbb{P}^S$-f.a. $s \in \mathcal{S}$ auf der Menge $\{S = s\}$. □

§07.14 **Beispiel**.

(a) Seien $X, S \in \mathscr{L}_1(\mathscr{A}, \mathbb{P})$ unabhängig. Dann gilt $\mathbb{P}$-f.s.

$$\mathbb{E}\big(X + S\big|\sigma(S)\big) = \mathbb{E}\big(X\big|\sigma(S)\big) + \mathbb{E}\big(S\big|\sigma(S)\big) = \mathbb{E}(X) + S.$$

Allgemeiner, sei $f \in \mathscr{S} \otimes \mathscr{X}$, für die $\mathbb{P}^{(S,X)}(f) = \mathbb{E}(f(S,X))$ existiert. Da $X \perp\!\!\!\perp S$ (siehe auch Satz §07.15) ist der Markovkern $(s,A) \mapsto \mathbb{P}^X(A)$ von $(\mathcal{S}, \mathscr{S})$ nach $(\mathcal{X}, \mathscr{X})$ eine reguläre Festlegung der bedingten Verteilung von X bei gegebenen S. Unter Verwendung von Satz §07.11 (ii) gilt damit $\mathbb{E}\big(f(S,X)\big|S = s\big) = \mathbb{P}^{(S,X)|S=s}(f) = \mathbb{P}^X(f_s) = \mathbb{E}\big(f(s,X)\big)$ für $\mathbb{P}^S$-f.a. $s \in \mathcal{S}$.

(b) Sei $(X_i)_{i\in[\![n]\!]}$ eine unabhängige Familie aus $\mathscr{L}_1(\mathscr{A}, \mathbb{P})$ mit $\mathbb{E}(X_i) = 0$, $i \in [\![n]\!]$. Für $m \in [\![n]\!]$ setze $\mathscr{F}_m := \sigma((X_i)_{i\in[\![m]\!]})$ und $S_m := \sum_{i\in[\![m]\!]} X_i$, dann gilt $\mathbb{P}$-f.s.

$$\mathbb{E}\big(S_n\big|\mathscr{F}_m\big) = \sum_{i\in[\![n]\!]} \mathbb{E}\big(X_i\big|\mathscr{F}_m\big) = \sum_{i\in[\![m]\!]} X_i + \sum_{i\in[\![m+1,n]\!]} \mathbb{E}(X_i) = S_m.$$

Da $\sigma(S_m) \subseteq \mathscr{F}_m$ folgt mit Satz §07.06 (ii) auch $\mathbb{P}$-f.s.

$$\mathbb{E}\big(S_n\big|\sigma(S_m)\big) = \mathbb{E}\big(\mathbb{E}\big(S_n\big|\mathscr{F}_m\big)\big|\sigma(S_m)\big) = \mathbb{E}\big(S_m\big|\sigma(S_m)\big) = S_m.$$ □

§07.15 **Satz**. *Sei (S,X) eine Zufallsvariable auf $(\Omega, \mathscr{A}, \mathbb{P})$ mit Werten in $(\mathcal{S} \times \mathcal{X}, \mathscr{S} \otimes \mathscr{X})$. Existiert eine reguläre Festlegung $\mathbb{P}^{X|S}$ der bedingten Verteilung von X bei gegebenen S, so sind die folgenden drei Aussagen äquivalent:*

(i) *X und S sind unabhängig, kurz $X \perp\!\!\!\perp S$, unter $\mathbb{P}$.*

(ii) *Es gibt eine Festlegung $\mathbb{P}^{X|S}$ die unabhängig von S ist.*

(iii) $\mathbb{P}^{X|S} = \mathbb{P}^X \quad \mathbb{P}^S$-f.ü..

In dem Fall $X \perp\!\!\!\perp S$ unter $\mathbb{P}$ gibt es für jedes $g : (\mathcal{S} \times \mathcal{X}, \mathscr{S} \otimes \mathscr{X}) \to (\mathcal{T}, \mathscr{T})$ eine $\mathbb{P}^S$-Nullmenge N derart, dass $\mathbb{P}^{g(S,X)|S=s}(C) = \mathbb{P}^{g(s,X)|S=s}(C)$ für alle $C \in \mathscr{T}$ und $s \in N^c$ gilt.

§07.16 **Beweis** von Satz §07.15. (i)⇔(ii)⇔(iii) Übung, Zusatz Witting [1985, Satz 1.129, S.130] □

§07|02 Bedingte Dichten

§07.17 **Vorbemerkung**. Eine explizite Darstellung einer regulären Festlegung der bedingten Verteilung $\mathbb{P}^{X|S}$ von X bei gegebenem S, und somit eine Möglichkeit des Nachweises der Existenz, können wir angeben, wenn die gemeinsame Verteilung eine Dichte bezüglich eines Produktmaßes σ-endlicher Maße besitzt. Wir bezeichnen weiterhin mit $D_s := \{x \in \mathcal{X} : (s,x) \in D\}$ bzw. $f_s(x) := f(s,x)$ den *s-Schnitt* von $D \subseteq \mathcal{S} \times \mathcal{X}$ bzw. $f : \mathcal{S} \times \mathcal{X} \to \overline{\mathbb{R}}$, jeweils bei gegebenem $s \in \mathcal{S}$. □

§07.18 **Definition**. Seien (S, X) eine Zufallsvariable auf $(\Omega, \mathscr{A}, \mathbb{P})$ mit Werten in $(\mathcal{S} \times \mathcal{X}, \mathscr{S} \otimes \mathscr{X})$ und $\mathbb{P}^{(S,X)} \ll \nu \otimes \mu$ mit $\nu \otimes \mu$-Dichte $\mathbb{f}^{(S,X)} \in (\mathscr{S} \otimes \mathscr{X})^+$ für σ-endliche Maße $\mu \in \mathcal{M}_\sigma(\mathscr{X})$ und $\nu \in \mathcal{M}_\sigma(\mathscr{S})$. Weiter bezeichne $\mathbb{f}^X \in \mathscr{X}^+$ und $\mathbb{f}^S \in \mathscr{S}^+$ Festlegungen der Randdichten (bzgl. μ bzw. ν), d.h. $\mathbb{f}^S(s) = \int (\mathbb{f}^{(S,X)})_s \, d\mu = \mu((\mathbb{f}^{(S,X)})_s)$ ν-f.ü. und $\mathbb{f}^X(x) = \nu((\mathbb{f}^{(S,X)})^x)$ μ-f.ü.. Die $\mathscr{S} \otimes \mathscr{X}$-$\mathbb{R}^+$- bzw. $\mathscr{X}$-$\mathbb{R}^+$-messbaren Abbildungen

$$\boxed{\mathbb{f}^{X|S}} := (\mathbb{f}^{(S,X)}/\mathbb{f}^S)\mathbb{1}_{\{\mathbb{f}^S \in \mathbb{R}^+_{\setminus 0}\}} + \mathbb{f}^X \mathbb{1}_{\{\mathbb{f}^S = 0\}} : \mathcal{S} \times \mathcal{X} \to \mathbb{R}^+ \text{ mit}$$

$$(s, x) \mapsto \mathbb{f}^{X|S=s}(x) := \frac{\mathbb{f}^{(S,X)}(s,x)}{\mathbb{f}^S(s)} \mathbb{1}_{\{\mathbb{f}^S(s) \in \mathbb{R}^+_{\setminus 0}\}} + \mathbb{f}^X(x) \mathbb{1}_{\{\mathbb{f}^S(s) = 0\}} \quad (07.01)$$

und $\mathbb{f}^{X|S=s} : \mathcal{X} \to \mathbb{R}^+$ heißen *reguläre Festlegungen der bedingten μ-Dichte* von X bei gegebenem S bzw. $S = s$. □

§07.19 **Satz**. *Unter den Annahmen von Definition §07.18 gilt*

(i) *Es gibt reguläre Festlegungen $\mathbb{P}^{X|S}$ und $\mathbb{P}^{(S,X)|S}$ von X bzw. (S, X) bei gegebenem S. Diese lassen sich explizit angeben und zwar bei festem $s \in \mathcal{S}$ mit $\mathbb{P}^{X|S=s} = \mathbb{f}^{X|S=s}\mu \in \mathcal{W}(\mathscr{X})$ und $\mathbb{P}^{(S,X)|S=s} \in \mathcal{W}(\mathscr{S} \otimes \mathscr{X})$ mit $D \mapsto \mathbb{P}^{(S,X)|S=s}(D) = \mu(\mathbb{1}_{D_s}\mathbb{f}^{X|S=s}) = \int_{D_s} \mathbb{f}^{X|S=s} \, d\mu$.*

(ii) *Für jede Funktion $h \in \overline{(\mathscr{S} \otimes \mathscr{X})}$, für die $\mathbb{P}^{(S,X)}(h) = \mathbb{E}(h(S,X))$ existiert, gilt*

$$\mathbb{P}^{(S,X)|S=s}(h) = \mu(h_s \mathbb{f}^{X|S=s}) = \mathbb{f}^{X|S=s}\mu(h_s) \quad \textit{für } \mathbb{P}^S\text{-f.a. } s \in \mathcal{S},$$

sowie für all $A \in \mathscr{X}$ und $B \in \mathscr{S}$ gilt

$$\mathbb{P}^{(S,X)}(h\mathbb{1}_{B \times A}) = \nu\big(\mathbb{1}_B \mathbb{f}^S \mathbb{P}^{(S,X)|S}(\mathbb{1}_A h)\big) = \int_B \mu(\mathbb{1}_A h_s \mathbb{f}^{X|S=s}) \, \mathbb{f}^S(s) \nu(ds).$$

§07.20 **Beweis** von Satz §07.19. In der Vorlesung. □

§07.21 **Bemerkung**. Die in der Definition §05.04 (a) und (b) eingeführten Begriffe können wir nun in den allgemeineren Ansatz einbetten. Dazu sei (S, X) eine Zufallsvariable mit gemeinsamer Verteilung $\mathbb{P}^{(S,X)}$ auf $(\mathcal{S} \times \mathcal{X}, \mathscr{S} \otimes \mathscr{X})$.

(i) Seien $\mathcal{S} \times \mathcal{X}$ abzählbar, also (S, X) *diskret-verteilt*, wie in Definition §05.04 (a). Für jedes $s \in S$ bezeichnen wir die Zähldichte $\mathbb{p}^{X|S=s} : 2^{\mathcal{X}} \to [0, 1]$ (bzgl. des Zählmaßes $\zeta_{\mathcal{X}}$) mit $x \mapsto \mathbb{p}^{X|S=s}(x) = \frac{\mathbb{p}^{(S,X)}(s,x)}{\mathbb{p}^S(s)} \mathbb{1}_{\{\mathbb{p}^S(s) \in \mathbb{R}^+_{\setminus 0}\}} + \mathbb{p}^X(x)\mathbb{1}_{\{\mathbb{p}^S(s) = 0\}}$ als *reguläre Festlegung der bedingten Zähldichte von X bei gegebenem $S = s$*. Dann ist $(s, A) \mapsto \mathbb{P}^{X|S=s}(A)$ ein Markovkern von $(\mathcal{S}, 2^{\mathcal{S}})$ nach $(\mathcal{X}, 2^{\mathcal{X}})$, der (bE1), also $\mathbb{P}^{X|S}(A) : 2^{\mathcal{S}} \to [0, 1]$, als auch die Bedingung (bE2) erfüllt. In der Tat, für jedes $A \in 2^{\mathcal{X}}$ und jedes $B \in 2^{\mathcal{S}}$ gilt

$$\begin{aligned}\mathbb{P}^{(S,X)}(\mathbb{1}_A(X)\mathbb{1}_B(S)) &= \sum_{s \in \mathcal{S}} \mathbb{1}_B(s) \sum_{x \in \mathcal{X}} \mathbb{1}_A(x) \mathbb{p}^{(S,X)}(x, s) \\ &= \sum_{s \in \mathcal{S}} \mathbb{1}_B(s) \mathbb{p}^S(s) \mathbb{P}^{X|S=s}(A) = \mathbb{P}^S(\mathbb{P}^{X|S}(A)\mathbb{1}_B).\end{aligned}$$

(ii) Seien (S,X) *stetig-verteilt* mit Werten in $(\mathbb{R}^{m+n},\mathscr{B}^{m+n})$, wie in Definition §05.04 (b). Für jedes $s\in\mathbb{R}^m$ bezeichnen wir die Dichte $\mathbb{f}^{X|S=s}\in(\mathscr{B}^n)^+$ (bzgl. des Lebesgue-Maßes λ^n) mit $x\mapsto \mathbb{f}^{X|S=s}(x)=\frac{\mathbb{f}^{(S,X)}(s,x)}{\mathbb{f}^S(s)}\mathbb{1}_{\{\mathbb{f}^S(s)\in\mathbb{R}^+_{\setminus 0}\}}+\mathbb{f}^X(x)\mathbb{1}_{\{\mathbb{f}^S(s)=0\}}$ als *reguläre Festlegung der bedingten Lebesgue-Dichte von X bei gegebenem $S=s$*. Dann ist $(s,A)\mapsto\mathbb{P}^{X|S=s}(A)$ ein Markovkern von $(\mathbb{R}^m,\mathscr{B}^m)$ nach $(\mathbb{R}^n,\mathscr{B}^n)$, der (bE1), also $\mathbb{P}^{X|S}(A)\in(\mathscr{B}^m)^+$, als auch die Bedingung (bE2) erfüllt. In der Tat, für jedes $A\in\mathscr{B}^n$ und jedes $B\in\mathscr{B}^m$ gilt

$$\mathbb{P}^{(S,x)}(\mathbb{1}_A(X)\mathbb{1}_B(S))=\int_{\mathbb{R}^m}\mathbb{1}_B(s)\int_{\mathbb{R}^n}\mathbb{1}_A(x)\mathbb{f}^{(S,X)}(s,x)dxds$$
$$=\int_{\mathbb{R}^m}\mathbb{1}_B(s)\mathbb{f}^S(s)\mathbb{P}^{X|S=s}(A)ds=\mathbb{E}^S(\mathbb{P}^{X|S}(A)\mathbb{1}_B).$$

□

§07.22 **Korollar**. *Unter den Annahmen von Definition §07.18 sind X und S genau dann unabhängig, kurz $X\perp\!\!\!\perp S$, unter $\mathbb{P}$, wenn $\mathbb{f}^{X|S}=\mathbb{f}^X$ $\mathbb{P}^S$-f.ü. gilt.*

§07.23 **Beweis** von Korollar §07.22. Die Behauptung folgt direkt aus den Sätze §07.11 und §07.19. □

§07.24 **Beispiel**.

(a) Unter Verwendung von Korollar §07.22 sind im Beispiel §05.06 (a) X und S somit nicht unabhängig, da zum Bsp. $\mathbb{p}^{X|S=2}\neq\mathbb{p}^X$ mit $\mathbb{P}^S(\{2\})=\mathbb{p}^S(2)=1/36>0$ gilt. Andererseits im Beispiel §05.06 (b) sind X und S unabhängig, d.h. $\mathbb{p}^{X|S}=\mathbb{p}^X$ gilt $\mathbb{P}^S$-f.s., genau dann, wenn $\Sigma_{XS}=0$ gilt, also wenn X und S unkorreliert sind.

(b) Für $a\in\mathbb{R}^+_{\setminus 0}$ seien $X\sim\mathrm{Exp}_a$ und $Y\sim\mathrm{Exp}_a$ unabhängig. Dann sind X und $S=X+Y$ abhängig, und es gilt $\mathbb{P}^{X|S=s}=\mathrm{U}_{[0,s]}$ für λ-f.a. $s\in\mathbb{R}^+$. Dagegen sind $W:=X/(X+Y)$ und S unabhängig, und es gilt $\mathbb{P}^{W|S}=\mathrm{U}_{[0,1]}$ λ-f.ü.. (Übung.) □

§07|03 Beste Vorhersage

§07.25 **Vorbemerkung**. Sei $(\mathbb{H},\langle\cdot,\cdot\rangle_{\mathbb{H}})$ ein Hilbertraum versehen mit der induzierten Norm $\|\cdot\|_{\mathbb{H}}$ und $\mathbb{U}$ ein abgeschlossener linearer Unterraum von $\mathbb{H}$. Für jedes Element $h\in\mathbb{H}$ existiert dann ein eindeutig bestimmtes Element $u_h\in\mathbb{U}$ mit $\|h-u_h\|_{\mathbb{H}}=\inf_{u\in\mathbb{U}}\|h-u\|_{\mathbb{H}}$. Die Abbildung $\Pi_{\mathbb{U}}:\mathbb{H}\to\mathbb{U}$ mit $h\mapsto\Pi_{\mathbb{U}}(h):=u_h$ heißt *orthogonale Projektion*. Bezeichnen wir mit $\mathbb{U}^\perp$ das orthogonale Komplement von $\mathbb{U}$ in $\mathbb{H}$, so gilt für alle $h,g\in\mathbb{H}$:

(i) $\Pi_{\mathbb{U}}\circ\Pi_{\mathbb{U}}=\Pi_{\mathbb{U}}$, (*Projektion*)

(ii) $h-\Pi_{\mathbb{U}}(h)\in\mathbb{U}^\perp$, (*orthogonal*)

(iii) $h=\Pi_{\mathbb{U}}(h)+(h-\Pi_{\mathbb{U}}(h))$ die eindeutig bestimmte Zerlegung von h in eine orthogonale direkte Summe von Elementen aus $\mathbb{U}$ und $\mathbb{U}^\perp$,

(iv) $\Pi_{\mathbb{U}}$ ist linear und selbst-adjungiert, also $\langle\Pi_{\mathbb{U}}(h),g\rangle_{\mathbb{H}}=\langle h,\Pi_{\mathbb{U}}(g)\rangle_{\mathbb{H}}$. □

§07.26 **Eigenschaft**. *Für jede Teil-σ-Algebra $\mathscr{F}\subseteq\mathscr{A}$ ist $\mathscr{L}_2(\mathscr{F},\mathbb{P})$ ein abgeschlossener linearer Unterraum von $\mathscr{L}_2(\mathscr{A},\mathbb{P})$.* □

§07.27 **Lemma**. *Sei $\mathscr{F}\subseteq\mathscr{A}$ eine Teil-σ-Algebra und $\mathbb{E}(\bullet|\mathscr{F})$ eine Festlegung der bedingten Erwartung. Dann ist $\mathbb{E}(\bullet|\mathscr{F}):\mathscr{L}_2(\mathscr{A},\mathbb{P})\to\mathscr{L}_2(\mathscr{F},\mathbb{P})$ eine orthogonale Projektion, also für alle $X\in\mathscr{L}_2(\mathscr{A},\mathbb{P})$ und $Y\in\mathscr{L}_2(\mathscr{F},\mathbb{P})$ gilt*

$$\|X-Y\|^2_{\mathscr{L}_2}=\mathbb{E}(|X-Y|^2)\geqslant\mathbb{E}(|X-\mathbb{E}(X|\mathscr{F})|^2)=\|X-\mathbb{E}(X|\mathscr{F})\|^2_{\mathscr{L}_2},$$

wobei Gleichheit genau dann gilt, wenn $Y=\mathbb{E}(X|\mathscr{F})$ $\mathbb{P}$-f.s..

§07.28 **Beweis** von Lemma §07.27. Übung □

§07.29 **Anmerkung**. Für jede Statistik $S:(\Omega,\mathscr{A})\to(\mathcal{S},\mathscr{S})$ und $\mathscr{F}=\sigma(S)$ ist $\mathbb{E}\big(X\big|\sigma(S)\big)=\varphi(S)$ nach Lemma §07.27 in $\mathscr{L}_2(\sigma(S),\mathbb{P})$ die beste Vorhersage von X, da für alle $h\in\mathscr{L}_2(\mathscr{S},\mathbb{P}^S)$ gilt $\mathbb{E}\big|X-\mathbb{E}\big(X\big|\sigma(S)\big)\big|^2\leqslant\mathbb{E}\big|X-h(S)\big|^2$. Nach Eigenschaft A04.12 ist $Z^*=\mathbb{E}(X)+\mathbb{C}\mathrm{ov}(X,S)(\mathbb{C}\mathrm{ov}(S))^+(S-\mathbb{E}(S))$ für Zufallsvektoren X und S in $\mathscr{L}_2$ die beste lineare Vorhersage von X durch S ist, also für alle A und b gilt $\mathbb{E}\big|X-Z^*\big|^2\leqslant\mathbb{E}\big|X-(AS+b)\big|^2$ (Definition A04.10). Offensichtlich ist eine beste Vorhersage, die auch linear ist, eine beste lineare Vorhersage, aber die Umkehrung gilt im Allgemeinen nicht. □

§07.30 **Beispiel**.

(a) Seien X und S wie in Beispiel §05.06 (b) gemeinsam normal-verteilt. Dann stimmt die beste Vorhersage $\mathbb{E}\big(X\big|\sigma(S)\big)=\mu_X+\Sigma_{XS}\Sigma_S^{-1}(S-\mu_S)$ mit der besten linearen Vorhersage in Eigenschaft A04.12 überein.

(b) Seien U und S unabhängig und identisch $\mathrm{N}_{(0,1)}$-verteilt, und $X:=S^2+S+U$. Dann ist für die beste Vorhersage $\mathbb{E}\big(X\big|\sigma(S)\big)=S^2+S$ (benutze Beispiel §07.14 (a)) der mittlere quadratische Vorhersagefehler $\mathbb{E}\big|X-\mathbb{E}\big(X\big|\sigma(S)\big)\big|^2=\mathbb{E}(U)^2=1$. Die beste lineare Vorhersage von X durch S ist in diesem Fall $Z^*=S+1$, da $\mathbb{C}\mathrm{ov}(X,S)=1$, $\mathbb{E}(X)=1$, $\mathbb{V}\mathrm{ar}(S)=1$ und $\mathbb{E}(S)=0$ ist. Damit gilt $\mathbb{E}\big|X-Z^*\big|^2=\mathbb{E}(S^4)+\mathbb{E}(U^2)-2\mathbb{E}(S^2)+1=3>1=\mathbb{E}\big|X-\mathbb{E}\big(X\big|\sigma(S)\big)\big|^2$. □

§07.31 **Lemma**. *Sei $p\in[1,\infty]$, $\mathscr{F}\subseteq\mathscr{A}$ eine Teil-σ-Algebra und $\mathbb{E}\big(\bullet\big|\mathscr{F}\big)$ eine Festlegung der bedingten Erwartung. Dann ist die lineare Abbildung $\mathbb{E}\big(\bullet\big|\mathscr{F}\big):\mathscr{L}_p(\mathscr{A},\mathbb{P})\to\mathscr{L}_p(\mathscr{F},\mathbb{P})$ eine Kontraktion, also $\|\mathbb{E}\big(X\big|\mathscr{F}\big)\|_{\mathscr{L}_p}\leqslant\|X\|_{\mathscr{L}_p}$, und damit beschränkt und stetig. Folglich für jede in $\mathscr{L}_p(\mathscr{A},\mathbb{P})$ konvergente Folge $(X_n)_{n\in\mathbb{N}}$ ist $(\mathbb{E}\big(X_n\big|\mathscr{F}\big))_{n\in\mathbb{N}}$ ebenso in $\mathscr{L}_p(\mathscr{F},\mathbb{P})$ konvergent.*

§07.32 **Beweis** von Lemma §07.31. In der Vorlesung. □

§07.33 **Erinnerung**. Eine Familie $(X_i)_{i\in\mathcal{I}}$ von Zufallsvariablen in $\mathscr{L}_1(\mathscr{A},\mathbb{P})$ mit beliebiger nicht-leerer Indexmenge $\mathcal{I}$, für die gilt

$$\inf\Big\{\sup_{i\in\mathcal{I}}\mathbb{P}\big(|X_i|\mathbb{1}_{\{|X_i|\geqslant a\}}\big):a\in\mathbb{R}^+\Big\}=0.$$

heißt *gleichgradig integrierbar* (Definition §02.37). □

§07.34 **Lemma**. *Für beliebige nicht-leere Indexmengen $\mathcal{I}$ und $\mathcal{J}$ seien $(X_i)_{i\in\mathcal{I}}$ eine gleichgradig integrierbare Familie von Zufallsvariablen in $\mathscr{L}_1(\mathscr{A},\mathbb{P})$ und $(\mathscr{F}_j)_{j\in\mathcal{J}}$ eine Familie von Teil-σ-Algebren von $\mathscr{A}$. Dann ist die Familie $\big(\mathbb{E}\big(X_i\big|\mathscr{F}_j\big)\big)_{i\in\mathcal{I},j\in\mathcal{J}}$ gleichgradig integrierbar in $\mathscr{L}_1(\mathscr{A},\mathbb{P})$. Insbesondere für $X\in\mathscr{L}_1(\mathscr{A},\mathbb{P})$ und $\mathcal{J}=\{\mathscr{F}:\mathscr{F}$ ist Teil-σ-Algebra von $\mathscr{A}\}$ ist die Familie $\big(\mathbb{E}\big(X\big|\mathscr{F}\big)\big)_{\mathscr{F}\in\mathcal{J}}$ in $\mathscr{L}_1(\mathscr{A},\mathbb{P})$ gleichgradig integrierbar.*

§07.35 **Beweis** von Lemma §07.34. Klenke [2020, Korollar 8.22 , S.199] □

§08 Bayes-Ansatz

§08.01 **Vorbemerkung**. Betrachten wir ein statistisches Experiment $(\mathcal{X},\mathscr{X},\mathbb{P}_\Theta)$, so gilt definitionsgemäß $\mathbb{P}_\theta\in\mathcal{W}(\mathscr{X})$ für jedes $\theta\in\Theta$. Nehmen wir zusätzlich an, dass $(\Theta,\mathscr{T})$ ein messbarer Raum ist, und dass für jedes $F\in\mathscr{X}$ die Abbildung $\mathbb{P}_\bullet(F):\Theta\to[0,1]$ mit $\theta\mapsto\mathbb{P}_\theta(F)$ $\mathscr{T}$-$\mathscr{B}_{[0,1]}$-messbar ist, so sind die beiden Bedingungen (Mk1) und (Mk2) (vgl. Erinnerung §06.06) eines Markovkernes von $(\Theta,\mathscr{T})$ nach $(\mathcal{X},\mathscr{X})$ erfüllt, den wir mit $\mathbb{P}_\bullet$ bezeichnen. □

§08.02 **Definition**. Es seien $(\mathcal{X}, \mathscr{X}, \mathbb{P}_{\Theta})$ ein statistisches Experiment, $(\Theta, \mathscr{T})$ ein messbarer Raum und $\mathbb{P}_{\bullet}$ mit $(\theta, F) \mapsto \mathbb{P}_{\theta}(F)$ ein Markovkern von $(\Theta, \mathscr{T})$ nach $(\mathcal{X}, \mathscr{X})$. Sei $\boldsymbol{\vartheta}$ eine $(\Theta, \mathscr{T})$-wertige Zufallsvariable, so dass die Parameter $\theta \in \Theta$ als Realisierung von $\boldsymbol{\vartheta}$ aufgefasst werden können. Die Wahrscheinlichkeitsverteilung $\mathbb{P}^{\boldsymbol{\vartheta}} \in \mathcal{W}(\mathscr{T})$ von $\boldsymbol{\vartheta}$ auf dem messbaren Raum $(\Theta, \mathscr{T})$ wird *a-priori Verteilung* des Parameters θ genannt und wir bezeichnen mit $\mathbb{E}^{\boldsymbol{\vartheta}}$ die Erwartung bezüglich $\mathbb{P}^{\boldsymbol{\vartheta}}$. □

§08.03 **Erinnerung**. Ein Markovkern $\mathbb{P}_{\bullet}$ von $(\Theta, \mathscr{T})$ nach $(\mathcal{X}, \mathscr{X})$ und eine a-priori Verteilung $\mathbb{P}^{\boldsymbol{\vartheta}} \in \mathcal{W}(\mathscr{T})$ bestimmen eindeutig (vgl. Lemma §04.56) die gemeinsame Verteilung $\mathbb{P}^{(\boldsymbol{\vartheta},X)} = \mathbb{P}^{\boldsymbol{\vartheta}} \odot \mathbb{P}_{\bullet} \in \mathcal{W}(\mathscr{T} \otimes \mathscr{X})$ auf dem messbaren Raum $(\Theta \times \mathcal{X}, \mathscr{T} \otimes \mathscr{X})$, wobei für alle $B \in \mathscr{T}$ und alle $A \in \mathscr{X}$

$$\mathbb{P}^{(\boldsymbol{\vartheta},X)}(B \times A) = \mathbb{P}^{\boldsymbol{\vartheta}} \odot \mathbb{P}_{\bullet}(B \times A) = \mathbb{P}^{\boldsymbol{\vartheta}}\big(\mathbb{1}_B \mathbb{P}_{\bullet}(A)\big) = \int_B \mathbb{P}_{\theta}(A)\mathbb{P}^{\boldsymbol{\vartheta}}(d\theta)$$

gilt. Die durch $\mathbb{P}^{(\boldsymbol{\vartheta},X)}$ auf $(\mathcal{X}, \mathscr{X})$ induzierte Randverteilung $\mathbb{P}^X := \mathbb{P}_{\bullet}\mathbb{P}^{\boldsymbol{\vartheta}}$ von X ist gegeben durch $\mathbb{P}^X(A) = \mathbb{P}_{\bullet}\mathbb{P}^{\boldsymbol{\vartheta}}(A) = \mathbb{P}^{\boldsymbol{\vartheta}}\big(\mathbb{P}_{\bullet}(A)\big) = \int \mathbb{P}_{\bullet}(A)\, d\mathbb{P}^{\boldsymbol{\vartheta}} = \int \mathbb{P}_{\theta}(A)\mathbb{P}^{\boldsymbol{\vartheta}}(d\theta)$ für alle $A \in \mathscr{X}$ (vgl. Schreibweise §06.07). □

§08.04 **Definition**. Seien $(\mathcal{X}, \mathscr{X}, \mathbb{P}_{\Theta})$ ein statistisches Experiment, $(\Theta, \mathscr{T})$ ein messbarer Raum, $(\boldsymbol{\vartheta}, X)$ eine $(\Theta \times \mathcal{X}, \mathscr{T} \otimes \mathscr{X})$-wertige Zufallsvariable, $\mathbb{P}^{\boldsymbol{\vartheta}}$ eine a-priori Verteilung von $\boldsymbol{\vartheta}$ auf $(\Theta, \mathscr{T})$, $\mathbb{P}_{\bullet}$ mit $(\theta, F) \mapsto \mathbb{P}_{\theta}(F)$ ein Markovkern von $(\Theta, \mathscr{T})$ nach $(\mathcal{X}, \mathscr{X})$. Die gemeinsame Verteilung $\mathbb{P}^{(\boldsymbol{\vartheta},X)} = \mathbb{P}^{\boldsymbol{\vartheta}} \odot \mathbb{P}_{\bullet}$ von $(\boldsymbol{\vartheta}, X)$ ist durch die reguläre bedingte Verteilung $\mathbb{P}^{X|\boldsymbol{\vartheta}} := \mathbb{P}_{\bullet}$ von X bei gegebenem $\boldsymbol{\vartheta}$ sowie $\mathbb{P}^{\boldsymbol{\vartheta}}$ festgelegt. Eine reguläre Festlegung der bedingten Verteilung $\mathbb{P}^{\boldsymbol{\vartheta}|X}$ von $\boldsymbol{\vartheta}$ bei gegebenem X, falls sie existiert, heißt *a-posteriori* Verteilung des zufälligen Parameters $\boldsymbol{\vartheta}$. Entsprechend wird eine reguläre Festlegung der bedingten Erwartung $\mathbb{E}^{\boldsymbol{\vartheta}|X}$ von $\boldsymbol{\vartheta}$ bei gegebenem X, falls sie existiert, *a-posteriori* Erwartung des zufälligen Parameters $\boldsymbol{\vartheta}$ genannt. Besitzt für ein $x \in \mathcal{X}$ die a-posteriori Verteilung $\mathbb{P}^{\boldsymbol{\vartheta}|X=x}$ ein endliches erstes absolutes Moment, so wird $\mathbb{E}^{\boldsymbol{\vartheta}|X=x}(\boldsymbol{\vartheta}) := \mathbb{E}^{\boldsymbol{\vartheta}|X=x}(\mathrm{id}_{\Theta}) = \mathbb{P}^{\boldsymbol{\vartheta}|X=x}(\mathrm{id}_{\Theta}) = \int_{\Theta} \theta \mathbb{P}^{\boldsymbol{\vartheta}|X=x}(d\theta)$ *a-posteriori* Erwartungswert des zufälligen Parameters $\boldsymbol{\vartheta}$ bei gegebenem $X = x$ genannt. Existiert ein endliches erstes absolutes Moment für alle $x \in \mathcal{X}$ so heißt die (messbare) Abbildung $\mathbb{E}^{\boldsymbol{\vartheta}|X}(\boldsymbol{\vartheta})$ mit $x \mapsto \mathbb{E}^{\boldsymbol{\vartheta}|X=x}(\boldsymbol{\vartheta})$ *a-posteriori* Erwartungswert des zufälligen Parameters $\boldsymbol{\vartheta}$ bei gegebenem X. □

§08.05 **Schreibweise**. Seien $(\mathcal{X}, \mathscr{X}, \mathbb{P}_{\Theta})$ ein statistisches Experiment, $\mathbb{P}_{\Theta}$ eine bezüglich eines σ-endlichen Maßes $\mu \in \mathcal{M}_{\sigma}(\mathscr{X})$ dominierte Verteilungsfamilie ($\mathbb{P}_{\theta} \ll \mu$ für alle $\theta \in \Theta$) mit entsprechenden μ-Dichten $(\mathbb{f}_{\theta})_{\theta \in \Theta}$ und $(\Theta, \mathscr{T})$ ein messbarer Raum. Fordern wir zusätzlich, dass $\mathbb{f}_{\bullet} : \Theta \times \mathcal{X} \to \mathbb{R}^+$ mit $(\theta, x) \mapsto \mathbb{f}_{\theta}(x)$ eine $(\mathscr{T} \otimes \mathscr{X})$-$\mathscr{B}^+$-messbare Funktion ist, dann ist mit Fubini $\mathbb{P}_{\bullet}$ mit $(\theta, F) \mapsto \mathbb{P}_{\theta}(F) = \mu(\mathbb{1}_F \mathbb{f}_{\theta}) = \int_F \mathbb{f}_{\theta}\, d\mu$ ein Markovkern von $(\Theta, \mathscr{T})$ nach $(\mathcal{X}, \mathscr{X})$. In dieser Situation bezeichnen wir $\mathbb{f}_{\bullet}$ als *(μ-) Dichte des Markovkernes* $\mathbb{P}_{\bullet}$ bezüglich des dominierenden Maßes μ. Betrachte eine $(\Theta \times \mathcal{X}, \mathscr{T} \otimes \mathscr{X})$-wertigen Zufallsvariable $(\boldsymbol{\vartheta}, X)$. Die a-priori Verteilung $\mathbb{P}^{\boldsymbol{\vartheta}}$ besitze eine ν-Dichte $\mathbb{f}^{\boldsymbol{\vartheta}}$ bezüglich eines dominierenden Maßes $\nu \in \mathcal{M}_{\sigma}(\mathscr{T})$. Die gemeinsame Verteilung $\mathbb{P}^{(\boldsymbol{\vartheta},X)} = \mathbb{P}^{\boldsymbol{\vartheta}} \odot \mathbb{P}^{X|\boldsymbol{\vartheta}}$, die durch die reguläre bedingte Verteilung $\mathbb{P}^{X|\boldsymbol{\vartheta}} := \mathbb{P}_{\bullet}$ und $\mathbb{P}^{\boldsymbol{\vartheta}}$ festgelegt sei, ist dann dominiert durch das Produktmaß $\nu \otimes \mu$, besitzt die *$\nu \otimes \mu$-Dichte* $\mathbb{f}^{\boldsymbol{\vartheta}}\mathbb{f}_{\bullet} : (\theta, x) \mapsto \mathbb{f}^{\boldsymbol{\vartheta}}(\theta)\mathbb{f}_{\theta}(x)$ und $\mathbb{f}^X := \widetilde{\mathbb{f}}^X \mathbb{1}_{\{\widetilde{\mathbb{f}}^X \in \mathbb{R}^+\}} \in \mathscr{X}^+$ mit $x \mapsto \widetilde{\mathbb{f}}^X(x) = \nu(\mathbb{f}_{\bullet}(x)\mathbb{f}^{\boldsymbol{\vartheta}}) = \int \mathbb{f}_{\bullet}(x)\mathbb{f}^{\boldsymbol{\vartheta}}\, d\nu = \int_{\Theta} \mathbb{f}_{\theta}(x)\mathbb{f}^{\boldsymbol{\vartheta}}(\theta)\nu(d\theta)$, ist eine *$\mu$-Dichte der Randverteilung* $\mathbb{P}^X = \mathbb{P}_{\bullet}\mathbb{P}^{\boldsymbol{\vartheta}}$ von X. Die Abbildung

$$\mathbb{f}^{\boldsymbol{\vartheta}|X} : \mathcal{X} \times \Theta \to \mathbb{R}^+ \text{ mit}$$

$$(x, \theta) \mapsto \mathbb{f}^{\boldsymbol{\vartheta}|X=x}(\theta) = \frac{\mathbb{f}_{\theta}(x)\mathbb{f}^{\boldsymbol{\vartheta}}(\theta)}{\mathbb{f}^X(x)} \mathbb{1}_{\{\mathbb{f}^X(x) \in \mathbb{R}^+_{\setminus 0}\}} + \mathbb{f}^{\boldsymbol{\vartheta}}(\theta)\mathbb{1}_{\{\mathbb{f}^X(x) = 0\}} \qquad (08.01)$$

ist dann eine $(\mathscr{X} \otimes \mathscr{T})$-$\mathscr{B}^+$-messbare Funktion und eine *ν-Dichte des Markovkernes* $\mathbb{P}^{\boldsymbol{\vartheta}|X}$ von $(\mathcal{X}, \mathscr{X})$ nach $(\Theta, \mathscr{T})$ mit $(x, B) \mapsto \mathbb{P}^{\boldsymbol{\vartheta}|X=x}(B) := \nu(\mathbb{1}_B \mathbb{f}^{\boldsymbol{\vartheta}|X=x})$ (vgl. Satz §07.19 (i)). □

§08.06 **Korollar**. *Seien $(\mathcal{X}, \mathscr{X}, \mathbb{P}_\Theta)$ ein statistisches Experiment, $(\Theta, \mathscr{T})$ ein messbarer Raum und $\mathbb{P}_\bullet$ ein Markovkern von $(\Theta, \mathscr{T})$ nach $(\mathcal{X}, \mathscr{X})$ mit Dichte $\mathbb{f}_\bullet$ bezüglich eines dominierenden Maßes μ. Betrachte eine $(\mathcal{X} \times \Theta, \mathscr{X} \otimes \mathscr{T})$-wertigen Zufallsvariable $(X, \boldsymbol{\vartheta})$. Die a-priori Verteilung $\mathbb{P}^{\boldsymbol{\vartheta}}$ besitze eine ν-Dichte $\mathbb{f}^{\boldsymbol{\vartheta}}$ bezüglich eines σ-endlichen Maßes $\nu \in \mathcal{M}_\sigma(\mathscr{T})$ und $\mathbb{P}^{(X,\boldsymbol{\vartheta})} := \mathbb{P}^{\boldsymbol{\vartheta}} \odot \mathbb{P}_\bullet$ sei die gemeinsame Verteilung von $(X, \boldsymbol{\vartheta})$. Der Markovkern $\mathbb{P}^{\boldsymbol{\vartheta}|X}$ von $(\mathcal{X}, \mathscr{X})$ nach $(\Theta, \mathscr{T})$ mit ν-Dichte $\mathbb{f}^{\boldsymbol{\vartheta}|X}$ gegeben in (08.01) ist dann eine reguläre Festlegung der bedingten Verteilung von $\boldsymbol{\vartheta}$ bei gegebenem X, also eine a-posteriori Verteilung.*

§08.07 **Beweis** von Korollar §08.06. Folgt direkt aus Satz §07.19 (i). □

Kapitel 3

Stochastische Prozesse und Stoppzeiten

§09 Stochastischer Prozess und Filtration

Im Folgenden sei $(\Omega, \mathscr{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, $(\mathcal{X}, \mathscr{X})$ ein Messraum und $\mathbb{T}$ eine nicht-leere Indexmenge.

§09.01 **Definition**. Ein Familie $(X_t)_{t\in\mathbb{T}}$ von Zufallsvariablen auf $(\Omega, \mathscr{A}, \mathbb{P})$ mit Werten in $(\mathcal{X}, \mathscr{X})$ heißt *stochastischer Prozess* mit Zeitbereich $\mathbb{T}$ und Zustandsraum $\mathcal{X}$. $(X_t)_{t\in\mathbb{T}}$ wird stochastischer Prozess in *diskreter* (oder *Zeitreihe*) bzw. *stetiger* Zeit genannt, wenn $\mathbb{T} \in \{\mathbb{N}, \mathbb{N}_0, \mathbb{Z}\}$ bzw. $(a, b) \subseteq \mathbb{T} \subseteq \mathbb{R}$ für $a, b \in \mathbb{R}$ mit $a < b$ ist. Für festes $\omega \in \Omega$ wird die Abbildung $t \mapsto X_t(\omega)$ *Pfad*, *Trajektorie* bzw. *Realisierung* von $(X_t)_{t\in\mathbb{T}}$ genannt. □

§09.02 **Beispiel**. Sei $(Y_j)_{j\in\mathbb{N}}$ eine Folge von unabhängigen und identisch-verteilten (u.i.v.) Rademacher-Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, \mathbb{P})$, d.h. Y_1 ist $\{-1, 1\}$-wertig mit $\mathbb{P}(Y_1 = -1) = 1/2 = \mathbb{P}(Y_1 = 1)$. Setze $X_0 := 0$ und $X_n := \sum_{j\in[\![n]\!]} Y_j$ für $n \in \mathbb{N}$. Der stochastischer Prozess in diskreter Zeit $(X_n)_{n\in\mathbb{N}_0}$ mit Zustandsraum $(\mathbb{Z}, 2^{\mathbb{Z}})$ wird *symmetrische einfache Irrfahrt* auf $\mathbb{Z}$ genannt. □

§09.03 **Bemerkung**. Da wir uns für das Verhalten der *zufälligen Funktion* $t \mapsto X_t$ und nicht einer einzelnen Zufallsvariable X_t interessieren, fassen wir $X = (X_t)_{t\in\mathbb{T}}$ als Zufallsvariable mit Werten im Produktraum $(\mathcal{X}^{\mathbb{T}}, \mathscr{X}^{\mathbb{T}})$ auf (vgl. Definition §04.02). Dann ist die Abbildung $X : \Omega \to \mathcal{X}^{\mathbb{T}}$ mit $\omega \mapsto X(\omega) = (X_t(\omega))_{t\in\mathbb{T}}$ $\mathscr{A}$-$\mathscr{X}^{\mathbb{T}}$-messbar. Wir bezeichnen wie bisher mit $\mathbb{P}^X := \mathbb{P} \circ X^{-1}$ die Verteilung von X, d.h. das Bildmaß von $\mathbb{P}$ unter X auf $(\mathcal{X}^{\mathbb{T}}, \mathscr{X}^{\mathbb{T}})$. Die uns im Folgenden interessierenden Fragen werden nur von dieser Verteilung abhängen, also ist es keine Einschränkung, wenn wir annehmen, dass $(\mathcal{X}^{\mathbb{T}}, \mathscr{X}^{\mathbb{T}}, \mathbb{P}^X)$ der zu Grunde liegende Wahrscheinlichkeitsraum ist und dass die Komponenten X_t des Prozesses gerade den Koordinatenabbildungen $\Pi_{\{t\}}$ entsprechen (vgl. Erinnerung §04.01 und Schreibweise §04.30). Falls X nämlich ein stochastischer Prozess auf einem abstrakten Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, \mathbb{P})$ ist, so bilden die Projektionen $\Pi_{\{t\}} : \mathscr{S}^{\mathbb{T}} \to \mathcal{X}$ einen auf $(\mathcal{X}^{\mathbb{T}}, \mathscr{X}^{\mathbb{T}}, \mathbb{P}^X)$ definierten Prozess, der die selbe Verteilung wie X besitzt. Die Beschreibung der Verteilung eines stochastischen Prozesses ist häufig kompliziert und im Allgemeine nicht durch eine explizite Formel gegeben. Deshalb, sind wir an ihrer Charakterisierung mit Hilfe der Verteilung der einzelnen Zufallsvariablen X_t interessiert. □

§09.04 **Definition**. Sei $X = (X_t)_{t\in\mathbb{T}}$ ein stochastischer Prozess mit Verteilung $\mathbb{P}^X$ auf $(\mathcal{X}^{\mathbb{T}}, \mathscr{X}^{\mathbb{T}})$. Für jedes endliches $\mathcal{J} \subseteq \mathbb{T}$ sei $\mathbb{P}^X_{\mathcal{J}} := \mathbb{P}^X \circ \Pi_{\mathcal{J}}^{-1}$ die gemeinsame Verteilung der Zufallsvariablen $(X_t)_{t\in\mathcal{J}} = \Pi_{\mathcal{J}} \circ X$. Die Familie von Wahrscheinlichkeitsmaßen $\{\mathbb{P}^X_{\mathcal{J}}, \mathcal{J} \subseteq \mathbb{T} \text{ endlich}\}$ heißt *Familie der endlichdimensionalen Verteilungen* von $\mathbb{P}^X$. □

§09.05 **Bemerkung**. Sei $\mathbb{P}^X$ die Verteilung eines stochastischen Prozesses X auf $(\mathcal{X}^{\mathbb{T}}, \mathscr{X}^{\mathbb{T}})$ dann ist die Familie der endlich dimensionalen Verteilungen $\{\mathbb{P}^X_{\mathcal{J}}, \mathcal{J} \subseteq \mathbb{T} \text{ endlich}\}$ eine projektive Familie (vgl. Definition §04.31). In der Tat für alle $\mathcal{J} \subseteq \mathcal{K} \subseteq \mathbb{T}$ endlich gilt $\mathbb{P}^X_{\mathcal{J}} = \mathbb{P}^X \circ \Pi_{\mathcal{J}}^{-1} = \mathbb{P}^X \circ (\Pi^{\mathcal{K}}_{\mathcal{J}} \circ \Pi_{\mathcal{K}})^{-1} = \mathbb{P}^X \circ \Pi_{\mathcal{K}}^{-1} \circ (\Pi^{\mathcal{K}}_{\mathcal{J}})^{-1} = \mathbb{P}^X_{\mathcal{K}} \circ (\Pi^{\mathcal{K}}_{\mathcal{J}})^{-1}$. Andererseits, unter Verwendung des *Kolmogorov'schen Erweiterungssatzes* §04.32 bestimmt die Familie der endlich dimensionalen Verteilungen eindeutig das Wahrscheinlichkeitsmaß $\mathbb{P}^X_{\mathcal{J}}$, wenn $\mathcal{X}$ ein polnischer Raum versehen mit der Borel-σ-Algebra $\mathscr{X} = \mathscr{B}_{\mathcal{X}}$ ist. □

§09.06 **Definition**. Stochastische Prozesse X und Y auf $(\Omega, \mathscr{A}, \mathbb{P})$ mit Zeitbereich $\mathbb{T}$ heißen

ununterscheidbar, wenn es $N \in \mathscr{A}$ gibt mit $\mathbb{P}(N) = 0$ und $\{X_t \neq Y_t\} \subseteq N$ für jedes $t \in \mathbb{T}$,

Modifikationen (oder *Versionen*) voneiander, falls für jedes $t \in \mathbb{T}$ gilt $X_t = Y_t$ $\mathbb{P}$-f.s.. □

§09.07 **Bemerkung**.

(i) Offenbar sind ununterscheidbare Prozesse auch Versionen voneinander. Die Umkehrung gilt im Allgemeinen nicht.

(ii) Sind X und Y Modifikationen voneinander, dann besitzen X und Y die selben endlichdimensionalen Verteilungen.

(iii) Seien X und Y Modifikationen voneinander. Ist der Zeitbereich $\mathbb{T}$ abzählbar, dann sind X und Y ununterscheidbar. In der Tat setzen wir $N_t := \{X_t \neq Y_t\}$ für $t \in \mathbb{T}$ und $N := \cup_{t\in\mathbb{N}} N_t$, wobei nach Voraussetzung gilt $\mathbb{P}(N_t) = 0$ für jedes $t \in \mathbb{T}$. Ist $\mathbb{T}$ abzählbar, so ist $N \in \mathscr{A}$ messbar und $\mathbb{P}(N) \leqslant \sum_{t\in\mathbb{T}} \mathbb{P}(N_t) = 0$. □

§09.08 **Definition**. Ein stochastischer Prozess X mit Zeitbereich $\mathbb{T} \subseteq \mathbb{R}$ und Zustandsraum $\mathcal{X}$ heißt

reellwertig (*numerisch*, *positiv reellwertig*, usw.), falls $\mathcal{X} = \mathbb{R}$ ($\mathcal{X} = \overline{\mathbb{R}}$, $\mathcal{X} = \mathbb{R}^+$, usw.),

Prozess mit *unabhängigen Zuwächsen*, wenn X reellwertig ist und für jedes $n \in \mathbb{N}$, $t_0 \in \mathbb{T}$ und alle $t_j \in \mathbb{T} \setminus (-\infty, t_{j-1}]$, $j \in [\![n]\!]$, gilt $\perp\!\!\!\perp_{j\in[\![n]\!]}(X_{t_j} - X_{t_{j-1}})$ (unabhängig),

Gauß'scher Prozess, wenn X reellwertig ist und für jedes $n \in \mathbb{N}$, und alle $t_j \in \mathbb{T}$, $j \in [\![n]\!]$, besitzt $(X_{t_j})_{j\in[\![n]\!]}$ eine multivariate Normalverteilung (vgl. Definition A05.02),

integrierbar (bzw. *$\mathscr{L}_p$-integrierbar*), wenn X numerisch ist und $\mathbb{P}(|X_t|) \in \mathbb{R}^+$ (bzw. $X_t \in \mathscr{L}_p(\mathbb{P})$) für jedes $t \in \mathbb{T}$ gilt.

Sei nun zusätzlich $\mathbb{T} \subseteq \mathbb{R}$ abgeschlossen unter Addition (z.Bsp $\mathbb{T} = \mathbb{N}$). Definiere für $s \in \mathbb{T}$ den *Shift* (-Operator) $\vartheta_s : \mathcal{X}^{\mathbb{T}} \to \mathcal{X}^{\mathbb{T}}$ mit $x = (x_t)_{t\in\mathbb{T}} \mapsto \vartheta_s(x) := (x_{t+s})_{t\in\mathbb{T}}$. Dann heißt X

stationär, falls $\mathbb{P}^X = \mathbb{P}^{\vartheta_s(X)}$ für jedes $s \in \mathbb{T}$,

Prozess mit *stationären Zuwächsen*, wenn X reellwertig ist und $\mathbb{P}^{\vartheta_t(X)_{s+r} - \vartheta_t(X)_r} = \mathbb{P}^{X_{s+r} - X_r}$ für alle $t, s, r \in \mathbb{T}$. (Ist $0 \in \mathbb{T}$, so genügt es $r = 0$ zu betrachten.) □

§09.09 **Beispiel**.

(a) Sind X_t, $t \in \mathbb{T}$, u.i.v. Zufallsvariablen, dann ist $(X_t)_{t\in\mathbb{T}}$ stationär.

(b) Sei $X = (X_z)_{z\in\mathbb{Z}}$ reellwertig und stationär, und seien $k \in \mathbb{N}$ und $c_i \in \mathbb{R}$, $i \in [\![0, k]\!]$. Dann definiert $Y_z := \sum_{i\in[\![0,k]\!]} c_i X_{z-i}$ einen stationären Prozess $Y = (Y_z)_{z\in\mathbb{Z}}$. Gilt zusätzlich $c_i \in \mathbb{R}^+$, $i \in [\![0, k]\!]$ und $\sum_{j\in[\![0,k]\!]} c_i = 1$, so wird Y das gleitende Mittel von X (mit Gewichten $(c_i)_{i\in[\![0,k]\!]}$) genannt. □

§10 Adaptierte stochastische Prozesse und Stoppzeiten

Im Folgenden sei der Zeitbereich $\mathbb{T} \subseteq \mathbb{R}$ eine Teilmenge der reellen Zahlen und $X = (X_t)_{t\in\mathbb{T}}$ ein stochastischer Prozess auf dem Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, \mathbb{P})$ mit Zustandsraum $(\mathcal{X}, \mathscr{X})$ und Bildwahrscheinlichkeitsmaß $\mathbb{P}^X$ auf $(\mathcal{X}^{\mathbb{T}}, \mathscr{X}^{\mathbb{T}})$.

§10.01 **Definition**. Eine Familie $\mathscr{F} = (\mathscr{F}_t)_{t\in\mathbb{T}}$ von Unter-σ-Algebren mit $\mathscr{F}_t \subseteq \mathscr{A}$ für jedes $t \in \mathbb{T}$, heißt *Filtration*, wenn $\mathscr{F}_s \subseteq \mathscr{F}_t$ für alle $s, t \in \mathbb{T}$ mit $s \leqslant t$. In diesem Fall wird $(\Omega, \mathscr{A}, \mathbb{P}, \mathscr{F})$ *filtrierter Wahrscheinlichkeitsraum* genannt. □

§10.02 **Definition**. Ein stochastischer Prozess $X = (X_t)_{t\in\mathbb{T}}$ heißt *adaptiert* an die Filtration $\mathscr{F} = (\mathscr{F}_t)_{t\in\mathbb{T}}$, wenn X_t bezüglich $\mathscr{F}_t$ messbar ist für jedes $t \in \mathbb{T}$. Gilt $\mathscr{F}_t = \sigma_t(X) := \sigma(X_s, s \in \mathbb{T}, s \leqslant t)$ für alle $t \in \mathbb{T}$, so nennen wir $\mathscr{F} = \sigma(X)$ die von X *erzeugte* oder *natürliche* Filtration. Wir definieren weiterhin $\mathscr{F}_{\sup\{\mathbb{T}\}} := \bigvee_{t\in\mathbb{T}} \mathscr{F}_t$ als auch $\sigma_{\sup\{\mathbb{T}\}}(X) := \sigma(X_s, s \in \mathbb{T})$. □

§10.03 **Bemerkung**. Ein stochastischer Prozess ist stets an seine erzeugte Filtration adaptiert. Die erzeugte Filtration ist die *kleinste* Filtration, an die ein Prozess adaptiert ist. Wir schreiben kurz, $\sigma(X) \subseteq \mathscr{F}$, wenn $\sigma_t(X) \subseteq \mathscr{F}_t$ für alle $t \in \mathbb{T}$ gilt. □

§10.04 **Beispiel**. Seien $Y = (Y_n)_{n\in\mathbb{N}}$ reelle Zufallsvariablen, $\sigma(Y) = (\sigma_n(Y) = \sigma(Y_j, j \in [\![n]\!]))_{n\in\mathbb{N}}$ die von Y erzeugte Filtration, sowie $X = (X_n := \sum_{j\in[\![n]\!]} Y_j)_{n\in\mathbb{N}}$. Dann ist X an $\sigma(Y)$ adaptiert, d.h. $\sigma(X) \subseteq \sigma(Y)$. Offenbar ist $(Y_j)_{j\in[\![n]\!]}$ auch messbar bezüglich $\sigma_n(X) = \sigma(X_j, j \in [\![n]\!])$, also auch $\sigma(Y) \subseteq \sigma(X)$, und daher gilt $\sigma(Y) = \sigma(X)$. Sei nun $Z_n := \sum_{j\in[\![n]\!]} \mathbb{1}_{\mathbb{R}^+}(Y_j)$. Dann ist Z_n an $\sigma(Y)$ adaptiert, jedoch ist im Allgemeinen $\sigma(Z) \subsetneqq \sigma(Y)$. □

§10.05 **Definition**. Sei $\mathbb{T} \in \{\mathbb{N}_0, \mathbb{N}\}$. Ein stochastischer Prozess $X = (X_n)_{n\in\mathbb{T}}$ heißt *vorhersagbar* (oder *previsibel*) bezüglich der Filtration $\mathscr{F} = (\mathscr{F}_n)_{n\in\mathbb{N}_0}$, wenn X_0 konstant ist (sofern $0 \in \mathbb{T}$ ist) und für jedes $n \in \mathbb{N}$ ist X_n $\mathscr{F}_{n-1}$-messbar. Ein numerischer Prozess X auf $(\Omega, \mathscr{A}, \mathbb{P}, \mathscr{F})$ wird *wachsend* genannt, wenn er $\mathscr{F}$-vorhersagbar ist und $X_n \leqslant X_{n+1}$ $\mathbb{P}$-f.s. gilt für jedes $n \in \mathbb{N}$. Ist $-X$ wachsend, so heißt X *fallend*. □

§10.06 **Definition**. Eine $\mathscr{A}$-$\overline{\mathscr{B}}$-messbare Zufallsvariable τ mit Werten in $\mathbb{T} \cup \{\sup\{\mathbb{T}\}\} =: \overline{\mathbb{T}} \subseteq \overline{\mathbb{R}}$ heißt *Stoppzeit bezüglich der Filtration* $\mathscr{F}$, kurz ($\mathscr{F}$-)Stoppzeit, wenn für jedes $t \in \mathbb{T}$ gilt, dass $\{\tau \leqslant t\} \in \mathscr{F}_t$, d.h. der Prozess $(\mathbb{1}_{\{\tau\leqslant t\}})_{t\in\mathbb{T}}$ ist an $\mathscr{F}$ adaptiert. □

§10.07 **Bemerkung**. Für alle $t \in \mathbb{T}$ gilt offensichtlich $\{\tau > t\} = \{\tau \leqslant t\}^c \in \mathscr{F}_t$. Auf dem Ereignis $\{\tau \leqslant t\}$ ist die Stoppzeit bis zum Zeitplunkt t schon eingetreten, und im Gegensatz dazu auf dem Ereignis $\{\tau > t\}$ wird sie erst noch eintreten. Da es $(t_n)_{n\in\mathbb{N}}$ aus $\mathbb{T}$ mit $t_n \uparrow \bar{t} := \sup\{\mathbb{T}\}$ gibt, gilt auch $\{\tau \in \mathbb{T}\} = \cup_{n\in\mathbb{N}}\{\tau \leqslant t_n\} \in \mathscr{F}_{\sup\{\mathbb{T}\}}$. Im Fall $\bar{t} \notin \mathbb{T}$ gilt somit auch $\{\tau = \bar{t}\} = \{\tau \in \mathbb{T}\}^c \in \mathscr{F}_{\sup\{\mathbb{T}\}}$. Insbesondere ist damit τ auch $\mathscr{F}_{\sup\{\mathbb{T}\}}$-$\overline{\mathscr{B}}$-messbar. □

§10.08 **Lemma**. *Sei $\mathbb{T}$ abzählbar. Dann ist τ genau dann eine $\mathscr{F}$-Stoppzeit, wenn $\{\tau = t\} \in \mathscr{F}_t$ für jedes $t \in \mathbb{T}$ gilt.*

§10.09 **Beweis** von Lemma §10.08. Übung. □

§10.10 **Beispiel**.

(a) Sei $t_o \in \overline{\mathbb{T}}$, dann ist $\tau = t_o$ (konstant) eine Stoppzeit. Da für alle $t \in \mathbb{T}$ entweder $\{\tau \leqslant t\} = \emptyset \in \mathscr{F}_t$ im Fall $t < t_o$ oder $\{\tau \leqslant t\} = \Omega \in \mathscr{F}_t$ im Fall $t_o \leqslant t$ gilt.

(b) Sei $X = (X_n)_{n\in\mathbb{N}_0}$ ein stochastischer Prozess mit Zustandsraum $(\mathcal{X}, \mathscr{X})$, der an die Filtration $\mathscr{F}$ adaptiert ist. Für $B \in \mathscr{X}$ betrachten wir den Zeitpunkt (Treffzeit / hitting time), zu dem X erstmals in B ist, d.h. für $\omega \in \Omega$

$$\tau_B(\omega) := \begin{cases} \min\{n \in \mathbb{N}_0 : X_n(\omega) \in B\}, & \text{falls } \omega \in \bigcup_{n\in\mathbb{N}_0} X_n^{-1}(B), \\ \infty, & \text{sonst.} \end{cases}$$

Dann ist τ_B eine Stoppzeit bezüglich $\mathscr{F}$, wobei $\tau_\emptyset = \infty$ und $\tau_{\mathcal{X}} = 0$. □

§10.11 **Lemma**. *Seien τ und σ Stoppzeiten. Dann gilt:*

(i) *$\tau \vee \sigma$ und $\tau \wedge \sigma$ sind Stoppzeiten.*

Sei zusätzlich $\mathbb{T} \subseteq \mathbb{R}^+$ abgeschlossen unter Addition.

(ii) *Dann ist auch $\tau + \sigma$ eine Stoppzeit.*

(iii) *Ist $s \in \mathbb{R}^+$, dann ist $\tau + s$ eine Stoppzeit, jedoch im Allgemeinen nicht $\tau - s$.*

§10.12 **Beweis** von Lemma §10.11. Übung. □

§10.13 **Bemerkung**. Die Eigenschaften (i) und (iii) erwarten wir von Stoppzeiten. Für (iii) ist zu beachten, dass $\tau - s$ um s Zeiteinheiten in die Zukunft blickt (denn, $\{\tau - s \leqslant t\} \in \mathscr{F}_{t+s}$), während $\tau + s$ um s Zeiteinheiten in die Vergangenheit schaut. Stoppzeiten ist aber nur der Blick in die Vergangenheit erlaubt. □

§10.14 **Beispiel** *(§10.10 (b) fortgesetzt)*. Für $B_1, B_2 \in \mathscr{X}$ betrachte die Treffzeiten τ_{B_1} and τ_{B_2}. Gilt $B_1 \subseteq B_2$, dann ist $\tau_{B_1} \geqslant \tau_{B_2}$. Insbesondere, gilt damit $\tau_{B_1} \wedge \tau_{B_2} \geqslant \tau_{B_1 \cup B_2}$ und $\tau_{B_1 \cap B_2} \geqslant \tau_{B_1} \vee \tau_{B_2}$. □

§10.15 **Definition**. Sei τ eine Stoppzeit auf $(\Omega, \mathscr{A})$ bezüglich der Filtration $\mathscr{F}$. Dann heißt

$$\mathscr{F}_\tau := \{A \in \mathscr{F}_{\sup\{\mathbb{T}\}} : A \cap \{\tau \leqslant t\} \in \mathscr{F}_t \text{ für jedes } t \in \mathbb{T}\}$$

die *σ-Algebra der τ-Vergangenheit*. □

§10.16 **Bemerkung**. Offensichtlich ist $\mathscr{F}_\tau \subseteq \mathscr{F}_{\sup\{\mathbb{T}\}}$ komplementstabil und σ-$\cup$-stabil, und es gilt $\Omega \in \mathscr{F}_\tau$, sodass $\mathscr{F}_\tau$ eine σ-Algebra ist. Da $\{\tau \leqslant \sup\{\mathbb{T}\}\} = \Omega$ gilt für alle $A \in \mathscr{F}_{\sup\{\mathbb{T}\}}$ automatisch $A \cap \{\tau \leqslant \sup\{\mathbb{T}\}\} = A \in \mathscr{F}_{\sup\{\mathbb{T}\}}$, und somit auch die nur scheinbar stärke Bedingung $\mathscr{F}_\tau = \{A \in \mathscr{F}_{\sup\{\mathbb{T}\}} : A \cap \{\tau \leqslant t\} \in \mathscr{F}_t \text{ für jedes } t \in \overline{\mathbb{T}}\}$. Weiterhin gilt $\{\tau \in \mathbb{T}\} \in \mathscr{F}_\tau$, da nämlich $\{\tau \in \mathbb{T}\} \in \mathscr{F}_{\sup\{\mathbb{T}\}}$ (vgl. Bemerkung §10.07) und $\{\tau \in \mathbb{T}\} \cap \{\tau \leqslant t\} = \{\tau \leqslant t\} \in \mathscr{F}_t$ für jedes $t \in \mathbb{T}$. Ist $\mathbb{T}$ insbesondere abzählbar, dann ist die σ-Algebra der τ-Vergangenheit auch gegeben durch $\mathscr{F}_\tau = \{A \in \mathscr{F}_{\sup\{\mathbb{T}\}} : A \cap \{\tau = t\} \in \mathscr{F}_t \text{ für jedes } t \in \mathbb{T}\}$. □

§10.17 **Beispiel**. Für eine konstante Stoppzeit $\tau = t_o$ mit $t_o \in \overline{\mathbb{T}}$ gilt $\mathscr{F}_\tau = \mathscr{F}_{t_o}$. In der Tat, es gilt $\mathscr{F}_\tau = \mathscr{F}_\tau \cap \{\tau \leqslant t_o\} \subseteq \mathscr{F}_{t_o}$. Andererseits, sei $A \in \mathscr{F}_{t_o}$ und $t \in \mathbb{T}$. Ist $t \geqslant t_o$, dann gilt $A \cap \{\tau \leqslant t\} = A \in \mathscr{F}_{t_o} \subseteq \mathscr{F}_t$ und für $t < t_o$ ist $A \cap \{\tau \leqslant t\} = \emptyset \in \mathscr{F}_t$. Also $A \in \mathscr{F}_\tau$ und somit auch $\mathscr{F}_{t_o} \subseteq \mathscr{F}_\tau$. □

§10.18 **Lemma**. *Sind τ und σ $\mathscr{F}$-Stoppzeiten, so gilt:* (i) $\mathscr{F}_\sigma \cap \{\sigma \leqslant \tau\} \subseteq \mathscr{F}_{\tau \wedge \sigma} = \mathscr{F}_\tau \cap \mathscr{F}_\sigma$, (ii) *$\tau$ ist $\mathscr{F}_\tau$-messbar,* (iii) $\mathscr{F}_\tau = \mathscr{F}_t$ *auf* $\{\tau = t\}$, *d.h.* $\mathscr{F}_\tau \cap \{\tau = t\} = \mathscr{F}_t \cap \{\tau = t\}$, *für alle* $t \in \overline{\mathbb{T}}$, *und* (iv) $\mathscr{F}_{\tau \vee \sigma} = \mathscr{F}_\tau \vee \mathscr{F}_\sigma$. *Insbesondere, folgt aus* (i), *dass* $\{\sigma \leqslant \tau\} \in \mathscr{F}_\sigma \cap \mathscr{F}_\tau$ *und* $\mathscr{F}_\sigma = \mathscr{F}_\tau$ *auf* $\{\sigma = \tau\}$, *sowie für* $\sigma \leqslant \tau$ *auch* $\mathscr{F}_\sigma \subseteq \mathscr{F}_\tau$.

§10.19 **Beweis** von Lemma §10.18. In der Vorlesung. □

§10.20 **Definition**. Für eine Stoppzeit τ und einen Prozess $X = (X_t)_{t \in \mathbb{T}}$ auf $(\Omega, \mathscr{A})$ definieren wir $X_\tau : \Omega \supset \{\tau \in \mathbb{T}\} \to \mathcal{X}$ mit $X_\tau(\omega) := X_{\tau(\omega)}(\omega)$ für alle $\omega \in \{\tau \in \mathbb{T}\}$ oder äquivalent dazu $X_\tau := X_t$ auf $\{\tau = t\}$ für alle $t \in \mathbb{T}$. □

§10.21 **Bemerkung**. Zur Erinnerung eine Funktion $h : \Omega \supset B \to \mathcal{X}$ mit $B \in \mathscr{A}$ heißt $\mathscr{A}$-$\mathscr{X}$ messbar, wenn $h^{-1}(\mathscr{X}) \subseteq \mathscr{A}$ gilt (vgl. Definition §04.41). Ist $\mathbb{T}$ nun abzählbar und τ eine $\mathscr{F}$-Stoppzeit, so ist $f : \Omega \supset \{\tau \in \mathbb{T}\} \to \mathcal{X}$ eine $\mathscr{F}_\tau$-$\mathscr{X}$-messbare Funktion, wenn für alle $t \in \mathbb{T}$, die Einschränkung von f auf $\{\tau = t\}$ eine $\mathscr{F}_t$-$\mathscr{X}$-messbare Funktion ist, d.h. $\{f \in B\} \cap \{\tau = t\} = f^{-1}(B) \cap \{\tau = t\} \in \mathscr{F}_t$ für jedes $B \in \mathscr{X}$. In der Tat gilt dann auch $f^{-1}(B) \cap \{\tau \leqslant t\} = \bigcup_{s \in \mathbb{T}: s \leqslant t} f^{-1}(B) \cap \{\tau = s\} \in \mathscr{F}_t$ für jedes $B \in \mathscr{X}$ und jedes $t \in \mathbb{T}$. Nach Definition §10.15 gilt somit $f^{-1}(\mathscr{X}) \subseteq \mathscr{F}_\tau$ oder mit anderen Worten f ist $\mathscr{F}_\tau$-$\mathscr{X}$-messbar. □

§10.22 **Lemma**. *Sei $\mathbb{T}$ abzählbar, X adaptiert und τ eine Stoppzeit. Dann ist X_τ $\mathscr{F}_\tau$-$\mathscr{X}$-messbar.*

§10.23 **Beweis** von Lemma §10.22. In der Vorlesung. □

§10.24 **Bemerkung**. Für überabzählbares $\mathbb{T}$ und festes ω ist die Abbildung $\mathbb{T} \to \mathcal{X}$ mit $t \mapsto X_t(\omega)$ im Allgemeinen nicht messbar, also ist auch die Zusammensetzung X_τ nicht immer messbar. Unter zusätzlichen Annahmen an die Regularität der Pfade $t \mapsto X_t(\omega)$ wie zum Beispiel Rechtsstetigkeit kann die Messbarkeit der Zusammensetzung X_τ gewährleistet werden (siehe Kallenberg [2002, Lemma 7.5, S.122]). □

§10.25 **Korollar**. *Sei $\mathbb{T}$ abzählbar, X adaptiert und $(\tau_t)_{t\in\mathbb{T}}$ eine isotone Familie endlichen Stoppzeiten, d.h. $\tau_t \leqslant \tau_s \in \mathbb{R}^+$ für alle $s, t \in \mathbb{T}$ mit $t \leqslant s$. Der Prozess $(X_{\tau_t})_{t\in\mathbb{T}}$ ist adaptiert an die Filtration $(\mathscr{F}_{\tau_t})_{t\in\mathbb{T}}$. Insbesondere ist $(X_{\tau\wedge t})_{t\in\mathbb{T}}$ adaptiert an beide Filtrationen $(\mathscr{F}_{\tau\wedge t})_{t\in\mathbb{T}}$ und $(\mathscr{F}_t)_{t\in\mathbb{T}}$.*

§10.26 **Beweis** von Korollar §10.25. In der Vorlesung. □

§10.27 **Definition**. Sei $\mathbb{T}$ abzählbar, $X = (X_t)_{t\in\mathbb{T}}$ $\mathscr{F}$-adaptiert und τ eine $\mathscr{F}$-Stoppzeit. Wer definieren den *gestoppten Prozess* $X^\tau = (X_t^\tau := X_{\tau\wedge t})_{t\in\mathbb{T}}$, der an beide Filtrationen $\mathscr{F} = (\mathscr{F}_t)_{t\in\mathbb{T}}$ und $\mathscr{F}^\tau = (\mathscr{F}_t^\tau := \mathscr{F}_{\tau\wedge t})_{t\in\mathbb{T}}$ adaptiert ist. □

Kapitel 4

Martingale

Im Folgenden seien stets $\mathbb{T} \subseteq \mathbb{R}$ eine nichtleere Indexmenge, $\mathscr{F} = (\mathscr{F}_t)_{t\in\mathbb{T}}$ eine Filtration und $(\Omega, \mathscr{A}, \mathbb{P}, \mathscr{F})$ ein filtrierter Wahrscheinlichkeitsraum. Für $Y \in \overline{\mathscr{A}}^+$ bzw. $Y \in \mathscr{L}_1(\mathbb{P})$ und Teil-σ-Algebra $\mathscr{S} \subseteq \mathscr{A}$ bezeichnet wie bisher $\mathbb{E}(Y|\mathscr{S})$ sowie auch $\mathbb{P}(Y|\mathscr{S})$ eine Festlegung des bedingten Erwartungswertes von Y bei gegebenem $\mathscr{S}$.

§11 Positive (Super-)Martingale

§11.01 **Definition.** Ein positiver (numerischer) adaptierter stochastischer Prozess $X = (X_t)_{t\in\mathbb{T}}$ auf einem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, \mathbb{P}, \mathscr{F})$ heißt (bezüglich $\mathscr{F}$)

positives Supermartingal, wenn $X_s \geqslant \mathbb{E}(X_t|\mathscr{F}_s)$ $\mathbb{P}$-f.s. für alle $s, t \in \mathbb{T}$ mit $t > s$,

positives Martingal, wenn $X_s = \mathbb{E}(X_t|\mathscr{F}_s)$ $\mathbb{P}$-f.s. für alle $s, t \in \mathbb{T}$ mit $t > s$.

Ein $\overline{\mathbb{R}}^d$-wertiger adaptierter stochastischer Prozess $X = ((X_t^1, \dots, X_t^d))_{t\in\mathbb{T}}$ auf $(\Omega, \mathscr{A}, \mathbb{P}, \mathscr{F})$ heißt *positives (Super-)Martingal* falls jeder Koordinatenprozess $X^k = (X_t^k)_{t\in\mathbb{T}}$ ein *positives (Super-)Martingal* ist. □

§11.02 **Bemerkung.**

(i) Ist $X = (X_t)_{t\in\mathbb{T}}$ ein $\mathscr{F}$-Supermartingal, dann gilt $\mathbb{E}(X_r|\mathscr{F}_s) \geqslant \mathbb{E}(X_t|\mathscr{F}_s)$ $\mathbb{P}$-f.s. für alle $s, r, t \in \mathbb{T}$ mit $s < r \leqslant t$, und somit ist die Abbildung $t \mapsto \mathbb{E}(X_t)$ monoton fallend, und im Fall eines Martingals konstant.

(ii) Ist $\mathbb{T} \in \{\mathbb{N}, \mathbb{N}_0, \mathbb{Z}\}$, dann genügt es, jeweils nur $t = s+1$ zu betrachten. In der Tat, für jedes $s \in \mathbb{T}$ gilt $\mathbb{E}(X_{s+1}|\mathscr{F}_s) \geqslant \mathbb{E}(\mathbb{E}(X_{s+2}|\mathscr{F}_{s+1})|\mathscr{F}_s) = \mathbb{E}(X_{s+2}|\mathscr{F}_s)$ $\mathbb{P}$-f.s. unter Verwendung der Turmeigenschaft der bedingten Erwartung (Satz §07.06 (ii)). Somit, wenn die definierende Ungleichung (beziehungsweise Gleichung) in einem Zeitschritt gilt, dann bei Induktion gilt diese auch für all anderen Zeitschritte.

(iii) Geben wir die Filtration $\mathscr{F}$ nicht explizit an, so nehmen wir stillschweigend an, dass $\mathscr{F} = \sigma(X)$ die von X erzeugte Filtration ist.

(iv) Sind $\mathscr{F}$ and $\mathscr{F}^\circ$ Filtrationen mit $\mathscr{F} \subseteq \mathscr{F}^\circ$, und ist X an $\mathscr{F}$ adaptiert und ein positives $\mathscr{F}^\circ$-(Super-)Martingal, dann ist X auch ein positives $\mathscr{F}$-(Super-)Martingal. In der Tat, für $s, t \in \mathbb{T}$ mit $s < t$ und den Fall, in dem X ein $\mathscr{F}^\circ$-Supermartingal ist, gilt $\mathbb{E}(X_t|\mathscr{F}_s) = \mathbb{E}(\mathbb{E}(X_t|\mathscr{F}_s^\circ)|\mathscr{F}_s) \leqslant \mathbb{E}(X_s|\mathscr{F}_s) = X_s$ $\mathbb{P}$-f.s.. Insbesondere ist ein positives $\mathscr{F}$-(Super-)Martingal X auch ein (Super-)Martingal bezüglich seiner eigenen natürlichen Filtration $\sigma(X)$. □

§11.03 **Beispiel.**

(a) Sei $Y = (Y_n)_{n\in\mathbb{N}}$ eine Folge unabhängiger positiver Zufallsvariablen aus $\overline{\mathscr{A}}^+$ mit $\mathbb{E}(Y_n) = 1$ für alle $n \in \mathbb{N}$. Weiterhin sei $\mathscr{F} = \sigma(Y)$ die von Y erzeugte Filtration, so dass $\sigma_n(Y)$ und Y_{n+1} unabhängig sind. Dann ist $(X_n := \prod_{i\in[\![n]\!]} Y_i)_{n\in\mathbb{N}}$ ein positives $\mathscr{F}$-Martingal, denn $\mathbb{E}(X_{n+1}|\mathscr{F}_n) = \mathbb{E}(X_n Y_{n+1}|\mathscr{F}_n) = X_n \mathbb{E}(Y_{n+1}|\mathscr{F}_n) = X_n \mathbb{E}(Y_{n+1}) = X_n$ $\mathbb{P}$-f.s..

(b) Sei $Z \in \overline{\mathscr{A}}^+$ eine positive Zufallsvariable. Dann ist $(X_n := \mathbb{E}(Z|\mathscr{F}_n))_{n\in\mathbb{N}}$ ein positives $\mathscr{F}$-Martingal, denn $\mathbb{E}(X_{n+1}|\mathscr{F}_n) = \mathbb{E}(\mathbb{E}(Z|\mathscr{F}_{n+1})|\mathscr{F}_n) = \mathbb{E}(Z|\mathscr{F}_n) = X_n$ $\mathbb{P}$-f.s.. □

§11.04 **Lemma**.

(i) *Seien X und Y positive (Super-)Martingale und $a, b \in \mathbb{R}^+$. Dann ist $aX + bY := (aX_t + bY_t)_{t\in\mathbb{T}}$ auch ein positives (Super-)Martingal.*

(ii) *Seien X und Y positive Supermartingale. Dann ist $Z := X \wedge Y := (X_t \wedge Y_t)_{t\in\mathbb{T}}$ ein positives Supermartingal.*

(iii) *Ist $(X_n)_{n\in\mathbb{N}}$ ein positives Supermartingal mit $\mathbb{E}(X_k) \geqslant \mathbb{E}(X_1)$ für ein $k \in \mathbb{N}$, dann ist $(X_n)_{n\in[\![k]\!]}$ ein positives Martingal. Existiert eine Teilfolge $(k_n)_{n\in\mathbb{N}}$ aus $\mathbb{N}$ mit $k_n \uparrow \infty$ und $\mathbb{E}(X_{k_n}) \geqslant \mathbb{E}(X_1)$, $n \in \mathbb{N}$, dann ist X ein positives Martingal.*

(iv) *Sind $(X_n)_{n\in\mathbb{N}}$ und $(Y_n)_{n\in\mathbb{N}}$ positive Supermartingale und τ eine Stoppzeit mit $X_\tau(\omega) \geqslant Y_\tau(\omega)$ für alle $\omega \in \{\tau \in \mathbb{N}\}$. Dann ist $Z := (Z_n := X_n \mathbb{1}_{\{n<\tau\}} + Y_n \mathbb{1}_{\{\tau\leqslant n\}})_{n\in\mathbb{N}}$ ein positives Supermartingal.*

§11.05 **Beweis** von Lemma §11.04. In der Vorlesung. □

§11.06 **Proposition** *(Maximalungleichung)*. *Sei $(X_n)_{n\in\mathbb{N}}$ ein positives Supermartingal. Dann ist $\sup_{n\in\mathbb{N}} X_n$ $\mathbb{P}$-f.s. endlich auf dem Ereignis $\{X_1 \in \mathbb{R}^+\}$ und für jedes $a \in \mathbb{R}^+_{\setminus 0}$ gilt*

$$\mathbb{P}\big(\sup_{n\in\mathbb{N}} X_n \geqslant a\big|\mathscr{F}_1\big) = \mathbb{E}\big(\mathbb{1}_{\{\sup_{n\in\mathbb{N}} X_n \geqslant a\}}\big|\mathscr{F}_1\big) \leqslant \min(X_1/a, 1) \quad \mathbb{P}\text{-f.s..}$$

§11.07 **Beweis** von Proposition §11.06. In der Vorlesung. □

§11.08 **Bemerkung**. Die letzte Aussage gilt weiterhin, wenn die Konstante $a \in \mathbb{R}^+_{\setminus 0}$ durch eine positive $\mathscr{F}_1$-messbare Zufallsvariable $A \in \overline{\mathscr{F}_1}^+$ ersetzt wird, d.h. auf dem Ereignis $\{A \in \overline{\mathbb{R}}^+_{\setminus 0}\}$ gilt ebenfalls $\mathbb{P}(\sup_{n\in\mathbb{N}} X_n \geqslant A|\mathscr{F}_1) \leqslant \min(X_1/A, 1)$ $\mathbb{P}$-f.s.. Insbesondere folgt damit:

(i) Für jede positive $\mathscr{F}_1$-messbare Zufallsvariable $A \in \overline{\mathscr{F}_1}^+$ mit $A \leqslant \sup_{n\in\mathbb{N}} X_n$ folgt $1 = \mathbb{P}(\sup_{n\in\mathbb{N}} X_n \geqslant A|\mathscr{F}_1) \leqslant \min\left(\frac{X_1}{A}, 1\right)$ $\mathbb{P}$-f.s. auf $\{A \in \overline{\mathbb{R}}^+_{\setminus 0}\}$ und, somit $A \leqslant X_1$ $\mathbb{P}$-f.s.. In anderen Worten, X_1 is die größte $\mathscr{F}_1$-messbare untere Schranke von $\sup_{n\in\mathbb{N}} X_n$.

(ii) Allgemeiner ist $\max_{n\in[\![k]\!]} X_n$, $k \in \mathbb{N}$, die größte $\mathscr{F}_k$-messbare untere Schranke von $\sup_{n\in\mathbb{N}} X_n$. In der Tat $(\max_{n\in[\![k]\!]} X_n, X_{k+1}, X_{k+2}, \dots)$ ist ein an die Filtration $(\mathscr{F}_{k+n})_{n\in\mathbb{N}_0}$ adaptiertes Supermartingal, und, unter Verwendung von Proposition §11.06 gilt für jede positive $\mathscr{F}_k$-messbare Zufallsvariable $A \in \overline{\mathscr{F}_k}^+$ mit $A \leqslant \sup_{n\in\mathbb{N}} X_n$, dass $A \leqslant \max_{n\in[\![k]\!]} X_n$ $\mathbb{P}$-f.s.. □

§11.09 **Definition**. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\overline{\mathbb{R}}$. Für $a, b \in \mathbb{R}$ mit $a < b$ definiere induktiv $\tau_0 := 1$, $\sigma_k := \inf\{n \in \mathbb{N} : n \geqslant \tau_{k-1} \wedge x_n \leqslant a\}$ und $\tau_k := \inf\{n \in \mathbb{N} : n \geqslant \sigma_k \wedge x_n \geqslant b\}$, $k \in \mathbb{N}$ (wobei $\inf \emptyset = \infty$). Die Anzahl der aufsteigenden Überquerungen des Intervalls $[a, b]$ durch die Folge $(x_n)_{n\in\mathbb{N}}$ bezeichnen wir mit $\beta_{a,b} := \sup\{k \in \mathbb{N}_0 : \tau_k \in \mathbb{N}\}$. □

§11.10 **Bemerkung**. Sei $X = (X_n)_{n\in\mathbb{N}}$ an die Filtration $\mathscr{F}$ adaptiert. Dann ist τ_k (und σ_k) definiert in Definition §11.09 eine $\mathscr{F}$-Stoppzeit für jedes $k \in \mathbb{N}$. In der Tat für jedes $n \in \mathbb{N}$ ist $\{\tau_k = n\}$ (und auch $\{\sigma_k = n\}$) ein $\mathscr{F}_n$-messbares Ereignis. Insbesondere gilt auch $\tau_k \leqslant \tau_{k+1}$, $k \in \mathbb{N}$. □

§11.11 **Lemma**. *Sei $X = (X_n)_{n\in\mathbb{N}}$ eine Folge numerischer Zufallsvariablen auf Ω und $a, b \in \mathbb{R}$ mit $a < b$. Für jedes $\omega \in \Omega$ bezeichne $\beta_{a,b}(\omega)$ die Anzahl der aufsteigenden Überquerungen des Intervalls $[a, b]$ durch die numerische Folge $(X_n(\omega))_{n\in\mathbb{N}}$. Dann ist $\beta_{a,b} : \Omega \to \overline{\mathbb{N}}_0 \subseteq \overline{\mathbb{R}}^+$ eine $\sigma_\infty(X)$-$\overline{\mathscr{B}}^+$-messbare Zufallsvariable.*

§11.12 **Beweis** von Lemma §11.11. Übung □

§11.13 **Bemerkung**. Wenn $\liminf_{n\to\infty} x_n < a < b < \limsup_{n\to\infty} x_n$ gilt, so ist offensichtlich $\beta_{a,b} = \infty$. Andererseits aus $\beta_{a,b} = \infty$ folgt $\liminf_{n\to\infty} x_n \leqslant a < b \leqslant \limsup_{n\to\infty} x_n$. In anderen Worten, die Folge $(x_n)_{n\in\mathbb{N}}$ aus $\overline{\mathbb{R}}$ ist genau dann konvergent, wenn $\beta_{a,b} < \infty$ für alle $a, b \in \mathbb{R}$ (oder in $\mathbb{Q}$) mit $a < b$ gilt. □

§11.14 **Lemma**. *Eine Folge $(X_n)_{n\in\mathbb{N}}$ numerischer Zufallsvariablen konvergiert genau dann $\mathbb{P}$-f.s., wenn die Anzahl der assoziierten aufsteigenden Überquerungen $\beta_{a,b}$ $\mathbb{P}$-f.s. endlich ist für alle $a, b \in \mathbb{Q}$ mit $a < b$.*

§11.15 **Beweis** von Lemma §11.14. Übung □

§11.16 **Lemma** *(Dubin'sche Ungleichung)*. *Seien $(X_n)_{n\in\mathbb{N}}$ ein positives $\mathscr{F}$-Supermartingal, $k \in \mathbb{N}$ und $a, b \in \mathbb{R}^+_{\setminus 0}$ mit $a < b$. Für die Anzahl der assoziierten aufsteigenden Überquerungen $\beta_{a,b}$ gilt*

$$\mathbb{P}\big(\beta_{a,b} \geqslant k \big| \mathscr{F}_1\big) \leqslant (a/b)^k \min\left(\frac{X_1}{a}, 1\right) \quad \mathbb{P}\text{-f.s..}$$

Insbesondere ist damit $\beta_{a,b}$ $\mathbb{P}$-f.s. endlich.

§11.17 **Beweis** von Lemma §11.16. In der Vorlesung. □

§11.18 **Bemerkung**. Sei $(X_z)_{z\in\mathbb{Z}}$ ein positives Supermartingal. Dann gilt auch $\mathbb{P}\big(\beta_{a,b} \geqslant k \big| \mathscr{F}_1\big) \leqslant (a/b)^k \min\left(\frac{\sup_{z\leqslant 1} X_z}{a}, 1\right)$ $\mathbb{P}$-f.s.. □

§11.19 **Satz**. *Jedes positive Supermartingal $(X_n)_{n\in\mathbb{N}}$ konvergiert $\mathbb{P}$-f.s., d.h. $X_n \xrightarrow{\mathbb{P}\text{-f.s.}} X_\infty$. Der $\mathbb{P}$-f.s. Grenzwert X_∞ ist insbesondere $\mathscr{F}_\infty$-messbar und erfüllt $X_n \geqslant \mathbb{E}\big(X_\infty \big| \mathscr{F}_n\big)$ $\mathbb{P}$-f.s. für alle $n \in \mathbb{N}$.*

§11.20 **Beweis** von Satz §11.19. In der Vorlesung. □

§11.21 **Bemerkung**.

(i) Da $X_n \geqslant \mathbb{E}\big(X_\infty \big| \mathscr{F}_n\big)$ $\mathbb{P}$-f.s. für alle $n \in \mathbb{N}$ gilt, ist X_∞ $\mathbb{P}$-f.s. endlich auf dem Komplement des Ereignisses $\bigcap_{n\in\mathbb{N}} \big\{X_n = \infty\big\}$. In der Tat, für jedes $n \in \mathbb{N}$ ist X_∞ $\mathbb{P}$-integrierbar auf dem Ereignis $\big\{\mathbb{E}\big(X_\infty \big| \mathscr{F}_n\big) \leqslant a\big\}$ für $a \in \mathbb{R}^+$. X_∞ ist damit $\mathbb{P}$-f.s. endlich auf dem Ereignis $\big\{\mathbb{E}\big(X_\infty \big| \mathscr{F}_n\big) \in \mathbb{R}^+\big\}$ und folglich auch auf dem Ereignis $\big\{X_n \in \mathbb{R}^+\big\}$.

(ii) Ist $(X_n)_{n\in\mathbb{N}}$ ein integrierbares positives Supermartingal, d.h. $X_n \in \mathscr{L}_1(\mathbb{P})$ für alle $n \in \mathbb{N}$, dann impliziert $X_n \geqslant \mathbb{E}\big(X_\infty \big| \mathscr{F}_n\big)$ auch $X_\infty \in \mathscr{L}_1(\mathbb{P})$. Im Allgemeinen konvergiert aber ein integrierbares positives Supermartingal nicht in $\mathscr{L}_1(\mathbb{P})$ gegen X_∞.

(iii) Ist $(X_n)_{n\in\mathbb{N}}$ ein positives Martingal, d.h. $X_n = \mathbb{E}\big(X_{n+1} \big| \mathscr{F}_n\big)$ $\mathbb{P}$-f.s. gilt für alle $n \in \mathbb{N}$, dann folgt aus Satz §11.19 $X_n \xrightarrow{\mathbb{P}\text{-f.s.}} X_\infty$ und $X_n \geqslant \mathbb{E}\big(X_\infty \big| \mathscr{F}_n\big)$ für alle $n \in \mathbb{N}$, wobei im Allgemeinen die Ungleichungen nicht Gleichungen werden. Die nächste Aussage charakterisiert eine Situation, in der dieses Phänomen nicht auftritt. □

§11.22 **Proposition**. *Sei $p \in [1, \infty)$. Für jedes $Z \in \mathscr{L}_p^+(\mathbb{P}) := \mathscr{L}_p(\mathbb{P}) \cap \overline{\mathscr{A}}^+$ ist der stochastische Prozess $(Z_n := \mathbb{E}\big(Z \big| \mathscr{F}_n\big))_{n\in\mathbb{N}}$ ein positives Martingal, das $\mathbb{P}$-f.s. und in $\mathscr{L}_p(\mathbb{P})$ gegen $Z_\infty := \mathbb{E}\big(Z \big| \mathscr{F}_\infty\big)$ konvergiert.*

§11.23 **Beweis** von Proposition §11.22. In der Vorlesung. □

§11.24 **Bemerkung**.

(i) Ein positives Martingal $(Z_n)_{n\in\mathbb{N}}$ wie in Proposition §11.22 und sein $\mathbb{P}$-f.s.-Grenzwert Z_∞ erfüllen die Gleichung $Z_n = \mathbb{E}\big(Z_\infty \big| \mathscr{F}_n\big)$ $\mathbb{P}$-f.s. für alle $n \in \mathbb{N}$, da $Z_n = \mathbb{E}\big(Z \big| \mathscr{F}_n\big) = \mathbb{E}\big(\mathbb{E}\big(Z \big| \mathscr{F}_\infty\big) \big| \mathscr{F}_n\big) = \mathbb{E}\big(Z_\infty \big| \mathscr{F}_n\big)$ $\mathbb{P}$-f.s. unter Verwendung der Turmeigenschaft (Satz §07.06 (ii)).

(ii) Sei $(X_n)_{n\in\mathbb{N}}$ ein positives Martingal, das in $\mathcal{L}_p(\mathbb{P})$ gegen $X_\infty \in \mathcal{L}_p(\mathbb{P})$ konvergiert. Die Martingalgleichung $X_n = \mathbb{E}\big(X_m \big| \mathscr{F}_n\big)$ $\mathbb{P}$-f.s. für $m \in \mathbb{N} \cap [n, \infty)$ und die Stetigkeit der bedingten Erwartung auf $\mathcal{L}_p(\mathbb{P})$ implizieren dann zusammen $X_n = \mathbb{E}\big(X_\infty \big| \mathscr{F}_n\big)$ $\mathbb{P}$-f.s. für alle $n \in \mathbb{N}$. Folglich impliziert Proposition §11.22, dass die Martingale der Form $(\mathbb{E}\big(Z\big|\mathscr{F}_n\big))_{n\in\mathbb{N}}$ mit $Z \in \mathcal{L}_p^+(\mathbb{P})$ genau die positiven in $\mathcal{L}_p(\mathbb{P})$ konvergenten Martingale in $\mathcal{L}_p(\mathbb{P})$ sind. Ein positives Martingal $(X_n)_{n\in\mathbb{N}}$ wird *abschließbar* (closable) in $\mathcal{L}_p(\mathbb{P})$ genannt, wenn $X \in \mathcal{L}_p^+(\mathbb{P})$ existiert mit $X_n = \mathbb{E}\big(X\big|\mathscr{F}_n\big)$ $\mathbb{P}$-f.s. für alle $n \in \mathbb{N}$.

(iii) Für $Z \in \mathcal{L}_p(\mathbb{P})$ gilt eine zu Proposition §11.22 entsprechende Aussage unter Verwendung der üblichen Zerlegung $Z = Z^+ - Z^-$. □

§11.25 **Korollar**. *Für jedes* $Z \in \overline{\mathscr{A}}^+$ *gilt* $\mathbb{E}\big(Z\big|\mathscr{F}_n\big) \xrightarrow{\mathbb{P}\text{-f.s.}} \mathbb{E}\big(Z\big|\mathscr{F}_\infty\big)$ *auf dem Komplement des Ereignisses* $\bigcap_{n\in\mathbb{N}}\{\mathbb{E}\big(Z\big|\mathscr{F}_n\big) = \infty\}$.

§11.26 **Beweis** von Korollar §11.25. Übung □

§11.27 **Bemerkung**. In Korollar §11.25 wird nicht Integrierbarkeit vorausgesetzt. In Neveu [1975, S. 31] wird eine $\mathscr{F}_\infty$-messbare Zufallsvariable Z konstruiert, die $\mathbb{P}$-f.s. endlich ist und $\mathbb{E}\big(Z\big|\mathscr{F}_n\big) = \infty$ $\mathbb{P}$-f.s. für alle $n \in \mathbb{N}$ erfüllt. In diesem Fall gilt $\mathbb{E}\big(Z\big|\mathscr{F}_n\big) \xrightarrow{\mathbb{P}\text{-f.s.}} \mathbb{E}\big(Z\big|\mathscr{F}_\infty\big) = Z$ nur auf einer Nullmenge. □

§11.28 **Erinnerung**. Seien $\mathbb{P}_0, \mathbb{P}_1 \in \mathcal{W}(\Omega, \mathscr{A})$ Wahrscheinlichkeitsmaße auf einem Messraum $(\Omega, \mathscr{A})$, wobei nicht notwendig $\mathbb{P}_0 \ll \mathbb{P}_1$ gilt. Dann heißt jede positive, numerische Funktion $\mathrm{L} \in \overline{\mathscr{A}}^+$ mit $\mathbb{P}_0 = \mathrm{L}\mathbb{P}_1 + \mathbb{1}_{\{\mathrm{L}=\infty\}}\mathbb{P}_0$ und $\mathbb{P}_1(\mathrm{L} \in \mathbb{R}^+) = 1$ Dichtequotient von $\mathbb{P}_0$ bezüglich $\mathbb{P}_1$ (vgl. Definition §03.14). In dem speziellen Fall $\mathbb{P}_0 \ll \mathbb{P}_1$ stimmt der Dichtequotient von $\mathbb{P}_0$ bezüglich $\mathbb{P}_1$ mit der $\mathbb{P}_1$-Dichte $\mathrm{d}\mathbb{P}_0/\mathrm{d}\mathbb{P}_1$ von $\mathbb{P}_0$ überein und ist $\mathbb{P}_1$-bestimmt. □

§11.29 **Eigenschaft**. *Seien* $\mathbb{P}_0, \mathbb{P}_1 \in \mathcal{W}(\Omega, \mathscr{A})$ *Wahrscheinlichkeitsmaße und* $\mathrm{L} \in \overline{\mathscr{A}}^+$ *ein Dichtequotient von* $\mathbb{P}_0$ *bezüglich* $\mathbb{P}_1$. *Dann gilt defintionsgemäß* $\mathbb{P}_1(\mathrm{L}) = \mathbb{P}_0(\mathrm{L} \in \mathbb{R}^+) \in [0, 1]$ *and* $\mathbb{P}_1(\mathrm{L} = \infty) = 0$, *und somit* $\mathbb{P}_0(\mathrm{L} = 0) = \mathrm{L}\mathbb{P}_1(\mathrm{L} = 0) + \mathbb{P}_0(\{\mathrm{L} = 0\} \cap \{\mathrm{L} = \infty\}) = 0$.

(i) $\mathbb{P}_0 \perp \mathbb{P}_1 \Leftrightarrow \exists A \in \mathscr{A} : \mathbb{P}_1(A) = 0$ *(somit* $\mathrm{L}\mathbb{P}_1(A) = 0$*) und* $\mathbb{P}_0(A) = 1$ *(somit* $\mathbb{P}_0(A \cap \{\mathrm{L} = \infty\}) = 1$*)* $\Leftrightarrow \mathbb{P}_0(\mathrm{L} = \infty) = 1 \Leftrightarrow \mathbb{P}_1(\mathrm{L}) = 0$;

(ii) $\mathbb{P}_0 \not\perp \mathbb{P}_1 \Leftrightarrow \forall A \in \mathscr{A} : \mathbb{P}_1(A) = 0$ *impliziert* $\mathbb{P}_0(A) \in [0, 1)$ *(insbesondere für* $A = \{\mathrm{L} = \infty\}$*)* $\Leftrightarrow \mathbb{P}_0(\mathrm{L} = \infty) \in [0, 1) \Leftrightarrow \mathbb{P}_1(\mathrm{L}) \in (0, 1]$;

(iii) $\mathbb{P}_0 \ll \mathbb{P}_1 \Leftrightarrow \forall A \in \mathscr{A} : \mathbb{P}_1(A) = 0$ *impliziert* $\mathbb{P}_0(A) = 0$ *(inbesondere für* $A = \{\mathrm{L} = \infty\}$*)* $\Leftrightarrow \mathbb{P}_0(\mathrm{L} = \infty) = 0 \Leftrightarrow \mathbb{P}_1(\mathrm{L}) = 1$. □

§11.30 **Lemma**. *Seien* $\mathbb{P}_0, \mathbb{P}_1 \in \mathcal{W}(\Omega, \mathscr{A})$ *Wahrscheinlichkeitsmaße und* $\mathscr{F} = (\mathscr{F}_n)_{n\in\mathbb{N}}$ *eine Filtration in* $\mathscr{A}$. *Für* $n \in \mathbb{N}$ *bezeichne* $\mathbb{P}_{0,n}, \mathbb{P}_{1,n} \in \mathcal{W}(\Omega, \mathscr{F}_n)$ *die Restriktionen von* $\mathbb{P}_0, \mathbb{P}_1$ *auf* $\mathscr{F}_n$.

(i) *Für* $n \in \mathbb{N}$ *sei* $\mathrm{L}_n \in \overline{\mathscr{F}_n}^+$ *ein Dichtequotient von* $\mathbb{P}_{0,n}$ *bezüglich* $\mathbb{P}_{1,n}$. *Dann ist* $(\mathrm{L}_n)_{n\in\mathbb{N}}$ *ein positives* $\mathscr{F}$*-Supermartingal auf dem filtrierten Wahrscheinlichkeitsraum* $(\Omega, \mathscr{A}, \mathbb{P}_1, \mathscr{F})$.

(ii) *Für alle* $n \in \mathbb{N}$ *sei* $\mathbb{P}_{0,n} \ll \mathbb{P}_{1,n}$ *und* $\frac{\mathrm{d}\mathbb{P}_{0,n}}{\mathrm{d}\mathbb{P}_{1,n}} \in \mathscr{F}_n^+$ *eine* $\mathbb{P}_{1,n}$*-Dichte von* $\mathbb{P}_{0,n}$, *wobei* $\mathrm{L}_n = \frac{\mathrm{d}\mathbb{P}_{0,n}}{\mathrm{d}\mathbb{P}_{1,n}}$ $\mathbb{P}_1$-f.s. *für alle* $n \in \mathbb{N}$ *gilt. Dann ist* $(\mathrm{L}_n)_{n\in\mathbb{N}}$ *ein positives* $\mathscr{F}$*-Martingal auf* $(\Omega, \mathscr{A}, \mathbb{P}_1, \mathscr{F})$.

(iii) *Sei* $\mathbb{P}_0 \ll \mathbb{P}_1$ *und* $\mathrm{L} = \frac{\mathrm{d}\mathbb{P}_0}{\mathrm{d}\mathbb{P}_1} \in \mathscr{A}^+$ *eine* $\mathbb{P}_1$*-Dichte von* $\mathbb{P}_0$. *Dann gilt* $\mathrm{L}_n = \mathbb{P}_1\big(\mathrm{L}\big|\mathscr{F}_n\big)$ $\mathbb{P}_1$-f.s. *für alle* $n \in \mathbb{N}$, *sodass das positive* $\mathscr{F}$*-Martingal* $(\mathrm{L}_n)_{n\in\mathbb{N}}$ *auf* $(\Omega, \mathscr{A}, \mathbb{P}_1, \mathscr{F})$ *abschliessbar ist.*

§11.31 **Beweis** von Lemma §11.30. In der Vorlesung. □

§11.32 **Bemerkung**. Die Folge $(\mathrm{L}_n)_{n\in\mathbb{N}}$ der Dichtequotienten in Lemma §11.30 (i)-(iii) konvergiert $\mathbb{P}_1$-f.s. nach Satz §11.19. Der $\mathbb{P}_1$-f.s.-Grenzwert L_∞ ist insbesondere $\mathscr{F}_\infty$-messbar und erfüllt $\mathrm{L}_n \geqslant \mathbb{P}_1\big(\mathrm{L}_\infty\big|\mathscr{F}_n\big)$ $\mathbb{P}_1$-f.s. für alle $n \in \mathbb{N}$. Unter Verwendung von Bemerkung §11.24 (ii) gilt für die Folge

$(\mathrm{L}_n)_{n\in\mathbb{N}}$ der Dichtequotienten in Lemma §11.30 (iii) zusätzlich $\mathrm{L}_n = \mathbb{P}_1\big(\mathrm{L}_\infty \big| \mathscr{F}_n\big)$ $\mathbb{P}_1$-f.s. für alle $n \in \mathbb{N}$. Nehmen wir zusätzlich $\mathscr{A} = \mathscr{F}_\infty$ an, so gilt $\mathrm{L}_\infty = \frac{\mathrm{d}\mathbb{P}_0}{\mathrm{d}\mathbb{P}_1}$ $\mathbb{P}_1$-f.s., und somit $\frac{\mathrm{d}\mathbb{P}_{0,n}}{\mathrm{d}\mathbb{P}_{1,n}} \xrightarrow{\mathbb{P}_1\text{-f.s.}} \frac{\mathrm{d}\mathbb{P}_0}{\mathrm{d}\mathbb{P}_1}$. □

§11.33 **Lemma**. *Seien $\mathbb{P}_0, \mathbb{P}_1 \in \mathcal{W}(\Omega, \mathscr{A})$ Wahrscheinlichkeitsmaße und $\mathscr{F} = (\mathscr{F}_n)_{n\in\mathbb{N}}$ eine Filtration mit $\mathscr{A} = \mathscr{F}_\infty$. Für $n \in \mathbb{N}$ bezeichne $\mathbb{P}_{0,n}, \mathbb{P}_{1,n} \in \mathcal{W}(\Omega, \mathscr{F}_n)$ die Restriktionen von $\mathbb{P}_0, \mathbb{P}_1$ auf $\mathscr{F}_n$. Für $n \in \mathbb{N}$ sei $\mathrm{L}_n \in \overline{\mathscr{F}_n}^+$ ein Dichtequotient von $\mathbb{P}_{0,n}$ bezüglich $\mathbb{P}_{1,n}$ und $\mathrm{L}_\infty \in \overline{\mathscr{F}_\infty}^+$ ihr $\mathbb{P}_1$-f.s.-Grenzwert. Schließlich sei $\mathrm{L} \in \overline{\mathscr{A}}^+$ ein Dichtequotient von $\mathbb{P}_0$ bezüglich $\mathbb{P}_1$. Dann konvergiert $\frac{1}{2}(\mathbb{P}_0 + \mathbb{P}_1)$-f.s. $(\mathrm{L}_n)_{n\in\mathbb{N}}$ gegen L, d.h. $\mathrm{L}_\infty = \mathrm{L}$ $\frac{1}{2}(\mathbb{P}_0 + \mathbb{P}_1)$-f.s..*

§11.34 **Beweis** von Lemma §11.33. In der Vorlesung. □

§11.35 **Satz** *(Kakutani Dichotomie)*. *Seien $(\mathcal{S}, \mathscr{B})$ und $(\mathcal{S}_n, \mathscr{B}_n)$, $n \in \mathbb{N}$, polnische Räume versehen mit der entsprechenden Borel-σ-Algebra. Dann gilt:*

(i) *Für $n \in \mathbb{N}$ seien $\mathbb{P}_{0,n}, \mathbb{P}_{1,n} \in \mathcal{W}(\mathcal{S}_n, \mathscr{B}_n)$ Wahrscheinlichkeitsmaße mit $\mathbb{P}_{0,n} \ll \mathbb{P}_{1,n}$. Bezeichne mit $\mathbb{P}_{0,\mathbb{N}} = \bigotimes_{n\in\mathbb{N}} \mathbb{P}_{0,n}$ und $\mathbb{P}_{1,\mathbb{N}} := \bigotimes_{n\in\mathbb{N}} \mathbb{P}_{1,n}$ das entsprechende Produktmaß auf dem Produktraum $(\mathcal{S}_\mathbb{N}, \mathscr{B}_\mathbb{N})$ (Definition §04.34). Dann gilt entweder $\mathbb{P}_{0,\mathbb{N}} \ll \mathbb{P}_{1,\mathbb{N}}$ oder $\mathbb{P}_{0,\mathbb{N}} \perp \mathbb{P}_{1,\mathbb{N}}$.*

(ii) *Seien $\mathbb{P}_0, \mathbb{P}_1 \in \mathcal{W}(\mathcal{S}, \mathscr{B})$ Wahrscheinlichkeitsmaße mit $\mathbb{P}_0 \ll\gg \mathbb{P}_1$ (äquivalent). Bezeichne mit $\mathbb{P}_0^{\otimes\mathbb{N}} := \bigotimes_{n\in\mathbb{N}} \mathbb{P}_0$ und $\mathbb{P}_1^{\otimes\mathbb{N}} := \bigotimes_{n\in\mathbb{N}} \mathbb{P}_1$ das entsprechende Produktmaß auf dem Produktraum $(\mathcal{X}^\mathbb{N}, \mathscr{B}^\mathbb{N})$ (Definition §04.34). Ist $\mathbb{P}_0 \neq \mathbb{P}_1$, dann gilt $\mathbb{P}_0^{\otimes\mathbb{N}} \perp \mathbb{P}_1^{\otimes\mathbb{N}}$.*

§11.36 **Beweis** von Satz §11.35. In der Vorlesung. □

§11.37 **Lemma** *(Optional Stopping)*. *Für jedes positive $\mathscr{F}$-(Super-)Martingal $(X_n)_{n\in\mathbb{N}}$ und für jede $\mathscr{F}$-Stoppzeit τ ist der gestoppte Prozess $X^\tau = (X_n^\tau := X_{\tau\wedge n})_{n\in\mathbb{N}}$ ein positives (Super-)Martingal.*

§11.38 **Beweis** von Lemma §11.37. Übung □

§11.39 **Satz** *(Optional Sampling)*. *Sei $(X_n)_{n\in\mathbb{N}}$ ein positives $\mathscr{F}$-Supermartingal mit $\mathbb{P}$-f.s.-Grenzwert X_∞. Für $\mathscr{F}$-Stoppzeiten τ, σ definiere $X_\tau \in \overline{\mathscr{F}_\tau}^+$ und $X_\sigma \in \overline{\mathscr{F}_\sigma}^+$ mit $X_\tau := X_n$ auf $\{\tau = n\}$ und $X_\sigma := X_n$ auf $\{\sigma = n\}$ für alle $n \in \overline{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$. Dann gilt $X_\tau \geqslant \mathbb{E}\big(X_\sigma \big| \mathscr{F}_\tau\big)$ $\mathbb{P}$-f.s. auf dem Ereignis $\{\tau \leqslant \sigma\}$.*

§11.40 **Beweis** von Satz §11.39. In der Vorlesung. □

§11.41 **Bemerkung**. Im Fall eines positiven Martingals $(X_n)_{n\in\mathbb{N}}$ wird in Satz §11.39 aus der Ungleichung $X_\tau \geqslant \mathbb{E}\big(X_\sigma \big| \mathscr{F}_\tau\big)$ im Allgemeinen keine Gleichung. □

§12 Integrierbare (Sub-, Super-)Martingale

§12.01 **Definition**. Ein ($\mathscr{L}_p$-)integrierbarer adaptierter stochastischer Prozess $X = (X_t)_{t\in\mathbb{T}}$ auf einem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, \mathbb{P}, \mathscr{F})$ heißt (bezüglich $\mathscr{F}$)

(($\mathscr{L}_p$-)integrierbares) Supermartingal, wenn $X_s \geqslant \mathbb{E}\big(X_t \big| \mathscr{F}_s\big)$ $\mathbb{P}$-f.s. für alle $s, t \in \mathbb{T}$ mit $t > s$,

(($\mathscr{L}_p$-)integrierbares) Submartingal, wenn $X_s \leqslant \mathbb{E}\big(X_t \big| \mathscr{F}_s\big)$ $\mathbb{P}$-f.s. für alle $s, t \in \mathbb{T}$ mit $t > s$,

(($\mathscr{L}_p$-)integrierbares) Martingal, wenn $X_s = \mathbb{E}\big(X_t \big| \mathscr{F}_s\big)$ $\mathbb{P}$-f.s. für alle $s, t \in \mathbb{T}$ mit $t > s$.

Ein $\mathbb{R}^d$-wertiger adaptierter stochastischer Prozess $X = ((X_t^1, \dots, X_t^d))_{t\in\mathbb{T}}$ auf $(\Omega, \mathscr{A}, \mathbb{P}, \mathscr{F})$ heißt *(($\mathscr{L}_p$-)integrierbares) (Sub-,Super-)Martingal* falls jeder Koordinatenprozess $X^k = (X_t^k)_{t\in\mathbb{T}}$ ein *(($\mathscr{L}_p$-)integrierbares) (Sub-,Super-)Martingal* ist. □

§12.02 **Lemma**.

(i) *Seien X und Y integrierbare Submartingale. Dann ist $X \vee Y := (X_t \vee Y_t)_{t\in\mathbb{T}}$ ein integrierbares Submartingal und daher insbesondere auch $X^+ = (X_t^+)_{t\in\mathbb{T}}$.*

(ii) *Sei X ein integrierbares Martingal und $f : \mathbb{R} \to \mathbb{R}$ eine konvexe Funktion. Ist* (A) *$\mathbb{P}(f(X_t)^+) \in \mathbb{R}^+$ für alle $t \in \mathbb{T}$ erfüllt, dann ist $(f(X_t))_{t\in\mathbb{T}}$ ein Submartingal.*

(iii) *Ist $\bar{t} := \sup\{\mathbb{T}\} \in \mathbb{T}$, so impliziert $\mathbb{P}(f(X_{\bar{t}})^+) \in \mathbb{R}^+$ schon* (A).

§12.03 **Beweis** von Lemma §12.02. In der Vorlesung. □

§12.04 **Bemerkung**. Sei X ein $\mathscr{L}_p$-integrierbares Martingal für $p \in [1, \infty]$. Da $x \mapsto |x|^p$ eine konvexe Funktion ist, ist $(|X_t|^p)_{t\in\mathbb{T}}$ ein Submartingal. □

§12.05 **Satz**. *Sei $X = (X_n)_{n\in\mathbb{N}}$ ein integrierbares Submartingal mit* (A) $\sup\{\mathbb{E}(X_n^+) : n \in \mathbb{N}\} \in \mathbb{R}^+$. *Dann existiert eine $\mathscr{F}_\infty$-messbare Zufallsvariable X_∞ mit $\mathbb{P}(|X_\infty|) \in \mathbb{R}^+$ und $X_n \xrightarrow{\mathbb{P}\text{-f.s.}} X_\infty$. Ist X ein integrierbares Martingal, dann sind* (A) *und* (B) $\sup\{\|X_n\|_{\mathscr{L}_1} : n \in \mathbb{N}\} \in \mathbb{R}^+$ *äquivalent.*

§12.06 **Beweis** von Satz §12.05. In der Vorlesung. □

§12.07 **Bemerkung**. Die Zerlegung $(X_n = M_n - A_n)_{n\in\mathbb{N}}$ eines integrierbaren Submartingals in ein positives integrierbares Martingal $(M_n)_{n\in\mathbb{N}}$ und ein positives integrierbares Supermartingal $(A_n)_{n\in\mathbb{N}}$, die wir im Beweis §12.06 gezeigt haben, wird *Krickeberg Zerlegung* genannt. □

§12.08 **Lemma**.

(i) *Sei $(X_n)_{n\in\mathbb{N}}$ ein integrierbares Martingal und sei τ eine beschränkte Stoppzeit, d.h. es gibt $K \in \mathbb{N}$ mit $\tau \leqslant K$. Dann gilt $X_\tau = \mathbb{E}(X_K | \mathscr{F}_\tau)$ und inbesondere $\mathbb{E}(X_\tau) = \mathbb{E}(X_1)$.*

(ii) *Sei allgemeiner $X = (X_n)_{n\in\mathbb{N}}$ ein integrierbarer adaptierter stochastischer Prozess. X ist genau dann integrierbares Martingal, wenn für jede beschränkte Stoppzeit τ gilt $\mathbb{E}(X_\tau) = \mathbb{E}(X_1)$.*

§12.09 **Beweis** von Lemma §12.08. In der Vorlesung. □

§12.10 **Definition**. Sei $(X_n)_{n\in\mathbb{N}_0}$ ein reellwertiger, $\mathscr{F}$-adaptierter Prozess und $(H_n)_{n\in\mathbb{N}}$ ein reellwertiger, $\mathscr{F}$-vorhersagbarer Prozess (Definition §10.05). Wir definieren den $\mathscr{F}$-adaptierten Prozess $H \bullet X = ((H \bullet X)_n)_{n\in\mathbb{N}_0}$ durch $(H \bullet X)_0 := 0$ und $(H \bullet X)_n := \sum_{k\in[\![n]\!]} H_k(X_k - X_{k-1})$ für $n \in \mathbb{N}$ und nennen $H \bullet X$ das *diskrete stochastische Integral* von H bezüglich X. Ist X ein Martingal, so wird $H \bullet X$ *Martingaltransformierte* von X genannt. □

§12.11 **Beispiel**. Sei $X = (X_n)_{n\in\mathbb{N}_0}$ ein (möglicherweise unfaires) Spiel, wobei $X_n - X_{n-1}$ den Spielgewinn pro Spiel in der n-ten Runde bezeichnet. Wir interpretieren H_n als die Anzahl Spielscheine, die für das n-te Spiel eingesetzt werden und verstehen H als eine Spielstrategie. Offenbar, muss der Wert von H_n zum Zeitpunkt $n-1$ bestimmt werden, also bevor das Ergebnis von X_n bekannt ist. In anderen Worten, H muss vorhersagbar sein. Sei X nun ein faires Spiel, also ein Martingal, und sei H *lokal beschränkt*, d.h. jedes H_n ist beschränkt. Wegen $\mathbb{E}(X_{n+1} - X_n | \mathscr{F}_n) = 0$ folgt $\mathbb{E}((H \bullet X)_{n+1} | \mathscr{F}_n) = \mathbb{E}((H \bullet X)_n + H_{n+1}(X_{n+1} - X_n) | \mathscr{F}_n) = (H \bullet X)_n + H_{n+1}\mathbb{E}(X_{n+1} - X_n | \mathscr{F}_n) = (H \bullet X)_n$, und somit ist die Martingaltransformierte $H \bullet X$ ebenfalls ein Martingal. Im folgenden Satz zeigen wir, dass auch die Umkehrung gilt, also X ein Martingal ist, wenn für hinreichend viele vorhersagbare Prozesse das diskrete stochastische Integral ein Martingal ist. □

§12.12 **Proposition**. *Sei $X = (X_n)_{n\in\mathbb{N}_0}$ ein reellwertiger, $\mathscr{F}$-adaptierter Prozess mit $X_0 \in \mathscr{L}_1(\mathbb{P})$.*

(i) *X ist genau dann ein integrierbares Martingal, wenn für jeden lokal beschränkten, vorhersagbaren Prozess H das diskrete stochastische Integral $H \bullet X$ ein integrierbares Martingal ist.*

(ii) *X is genau dann ein integrierbares Submartingal (Supermartingal), wenn $H \bullet X$ ein Submartingal (Supermartingal) ist für jeden lokal beschränkten, vorhersagbaren und positiven*

Prozess H.

§12.13 **Beweis** von Proposition §12.12. In der Vorlesung. □

§12.14 **Bemerkung**. Der vorangehende Satz sagt insbesondere, dass es keine (lokal beschränkte) Spielstrategie gibt, die aus einem Martingal (oder positiven Supermartingal im Fall einer Spielstrategie aus dem wahren Leben) ein Submartingal macht. Genau dies wird aber durch Aufforderungen zum so genannten „Systemlotto" und Ähnlichem nahe gelegt. □

§13 Reguläre integrierbare Martingale

§13.01 **Erinnerung**.

(i) Ein integrierbares $\mathscr{F}$-Martingal $(X_n)_{n\in\mathbb{N}}$ mit (B) $\sup\{\|X_n\|_{\mathscr{L}_1} : n \in \mathbb{N}\} \in \mathbb{R}^+$ konvergiert nach Satz §12.05 $\mathbb{P}$-f.s. gegen $X_\infty \in \mathscr{L}_1(\mathscr{F}_\infty, \mathbb{P})$.

(ii) Sei $(X_n)_{n\in\mathbb{N}}$ ein gleichgradig integrierbar (Definition §02.37) Prozess in $\mathscr{L}_1(\mathscr{A}, \mathbb{P})$, d.h. $\inf\big\{\sup_{n\in\mathbb{N}} \mathbb{P}\big(|X_n|\mathbb{1}_{\{|X_n|\geqslant a\}}\big) : a \in \mathbb{R}^+\big\} = 0$. Nach Satz §02.41 (gI1) gilt dann auch (B). Weiterhin konvergiert $(X_n)_{n\in\mathbb{N}}$ in $\mathscr{L}_1(\mathbb{P})$ nach Satz §02.43 genau dann, wenn $(X_n)_{n\in\mathbb{N}}$ gleichgradig integrierbar und stochastisch konvergent ist, wobei $\mathbb{P}$-f.s.-Konvergenz stochastische Konvergenz impliziert (Lemma §02.29 (v)). □

§13.02 **Satz**. *Für ein integrierbares Martingal $(X_n)_{n\in\mathbb{N}}$ auf einem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, \mathbb{P}, \mathscr{F})$ sind die folgenden vier Aussagen äquivalent:*

(i) *Die Folge $(X_n)_{n\in\mathbb{N}}$ konvergiert in $\mathscr{L}_1(\mathbb{P})$;*

(ii) *Es gilt (B) $\sup\{\|X_n\|_{\mathscr{L}_1} : n \in \mathbb{N}\} \in \mathbb{R}^+$. Der $\mathbb{P}$-f.s.-Grenzwert $X_\infty = \lim_{n\to\infty} X_n$, der nach Satz §12.05 in $\mathscr{L}_1(\mathbb{P})$ existiert, erfüllt $X_n = \mathbb{E}\big(X_\infty \big| \mathscr{F}_n\big)$ $\mathbb{P}$-f.s. für alle $n \in \mathbb{N}$;*

(iii) *Das Martingal $(X_n)_{n\in\mathbb{N}}$ ist abschliessbar, d.h. es gibt eine Zufallsvariable $X \in \mathscr{L}_1(\mathscr{A}, \mathbb{P})$ mit $X_n = \mathbb{E}\big(X \big| \mathscr{F}_n\big)$ $\mathbb{P}$-f.s. für alle $n \in \mathbb{N}$;*

(iv) *Die Folge $(X_n)_{n\in\mathbb{N}}$ ist gleichgradig integrierbar in $\mathscr{L}_1(\mathscr{A}, \mathbb{P})$.*

Das integrierbare Martingal $(X_n)_{n\in\mathbb{N}}$ heißt regulär *wenn eine der äquivalenten Aussagen* (i)-(iv) *erfüllt ist.*

§13.03 **Beweis** von Satz §13.20. In der Vorlesung. □

§13.04 **Bemerkung**. Für ein reguläres integrierbares $\mathscr{F}$-Martingal $(X_n)_{n\in\mathbb{N}}$ existiert nach Satz §13.20 also ein $X_\infty \in \mathscr{L}_1(\mathscr{F}_\infty, \mathbb{P})$ mit

$$X_n \xrightarrow{\mathbb{P}\text{-f.s.}} X_\infty \quad \text{und} \quad X_n \xrightarrow{\mathscr{L}_1(\mathbb{P})} X_\infty.$$

Ist weiterhin $X \in \mathscr{L}_1(\mathscr{A}, \mathbb{P})$ mit $X_n = \mathbb{E}\big(X \big| \mathscr{F}_n\big)$ $\mathbb{P}$-f.s. für alle $n \in \mathbb{N}$, so gilt $X_\infty = \mathbb{E}\big(X \big| \mathscr{F}_\infty\big)$ $\mathbb{P}$-f.s.. Im Allgemeinen ist nur die Zufallsvariable X_∞ $\mathbb{P}$-f.s. eindeutig, nicht aber die Zufallsvariable X. Für $\mathscr{F}_\infty = \mathscr{A}$ gilt jedoch $X_\infty = X$ $\mathbb{P}$-f.s.. □

§13.05 **Korollar** *(Optional sampling)*. *Sei $(X_n)_{n\in\mathbb{N}}$ ein reguläres integrierbares $\mathscr{F}$-Martingal mit $\mathbb{P}$-f.s.-Grenzwert X_∞.*

(i) *Für jede $\mathscr{F}$-Stoppzeit τ ist die Zufallsvariable $X_\tau \in \overline{\mathscr{F}_\tau}$ mit $X_\tau := X_n$ auf $\{\tau = n\}$ für alle $n \in \overline{\mathbb{N}}$ integrierbar.*

(ii) *Für jedes Paar von $\mathscr{F}$-Stoppzeiten τ, σ mit $\tau \leqslant \sigma$ $\mathbb{P}$-f.s. gilt auch $X_\tau = \mathbb{E}\big(X_\sigma \big| \mathscr{F}_\tau\big)$ $\mathbb{P}$-f.s..*

(iii) *Die Familie $\{X_\tau : \tau$ ist eine endliche $\mathscr{F}$-Stoppzeit$\}$ ist gleichgradig integrierbar.*

§13.06 **Beweis** von Korollar §13.05. In der Vorlesung. □

§13.07 **Bemerkung**. Für ein reguläres integrierbares $\mathscr{F}$-Martingal mit $\mathbb{P}$-f.s.-Grenzwert X_∞ und Zufallsvariable X_τ (bzw. X_σ), die definitionsgemäß gleich X_∞ auf $\{\tau = \infty\}$ (bzw. auf $\{\sigma = \infty\}$) ist. Da $\tau \wedge \sigma \leqslant \sigma$ Korollar §13.05 impliziert $X_{\tau\wedge\sigma} = \mathbb{E}(X_\sigma | \mathscr{F}_{\tau\wedge\sigma})$ $\mathbb{P}$-f.s.. Weiterhin gilt $\mathbb{E}(X_\sigma | \mathscr{F}_\tau) = \mathbb{E}(X_\sigma | \mathscr{F}_{\tau\wedge\sigma})$ $\mathbb{P}$-f.s., und damit für beliebige $\mathscr{F}$-Stoppzeiten τ, σ folgt $X_{\tau\wedge\sigma} = \mathbb{E}(X_\sigma | \mathscr{F}_\tau)$ $\mathbb{P}$-f.s.. In der Tat für alle $A \in \mathscr{F}_\tau$ gilt (unter Verwendung von Lemma §10.18 (i))

$$
\begin{aligned}
\mathbb{E}(\mathbb{E}(X_\sigma | \mathscr{F}_\tau)\mathbb{1}_A) = \mathbb{E}(X_\sigma \mathbb{1}_A) &= \mathbb{E}(X_\sigma \mathbb{1}_{\underbrace{A\cap\{\tau\leqslant\sigma\}}_{\in\mathscr{F}_{\tau\wedge\sigma}}}) + \mathbb{E}(X_\sigma \mathbb{1}_{A\cap\{\tau>\sigma\}}) \\
&= \mathbb{E}(\mathbb{E}(X_\sigma | \mathscr{F}_{\tau\wedge\sigma})\mathbb{1}_{A\cap\{\tau\leqslant\sigma\}}) + \mathbb{E}(X_{\tau\wedge\sigma}\mathbb{1}_{A\cap\{\tau>\sigma\}}) \\
&= \mathbb{E}\Big(\big(\mathbb{E}(X_\sigma | \mathscr{F}_{\tau\wedge\sigma})\mathbb{1}_{\{\tau\leqslant\sigma\}} + X_{\tau\wedge\sigma}\mathbb{1}_{\{\tau>\sigma\}}\big)\mathbb{1}_A\Big).
\end{aligned}
$$

Da (unter Verwendung von Lemma §10.18 (ii) und dem Zusatz $\{\tau \leqslant \sigma\} \in \mathscr{F}_{\tau\wedge\sigma} \subseteq \mathscr{F}_\tau$) $\mathbb{E}(X_\sigma | \mathscr{F}_{\tau\wedge\sigma})\mathbb{1}_{\{\tau\leqslant\sigma\}} + X_{\tau\wedge\sigma}\mathbb{1}_{\{\tau>\sigma\}}$ insbesondere $\mathscr{F}_{\tau\wedge\sigma}$-messbar und somit auch auch $\mathscr{F}_\tau$-messbar ist, gilt $\mathbb{E}(X_\sigma | \mathscr{F}_\tau) = \mathbb{E}(X_\sigma | \mathscr{F}_{\tau\wedge\sigma})\mathbb{1}_{\{\tau\leqslant\sigma\}} + X_{\tau\wedge\sigma}\mathbb{1}_{\{\tau>\sigma\}}$ $\mathbb{P}$-f.s. und damit auch $\mathbb{E}(X_\sigma | \mathscr{F}_\tau) = \mathbb{E}(\mathbb{E}(X_\sigma | \mathscr{F}_\tau) | \mathscr{F}_{\tau\wedge\sigma}) = \mathbb{E}(X_\sigma | \mathscr{F}_{\tau\wedge\sigma})$ $\mathbb{P}$-f.s. wegen $\mathscr{F}_{\tau\wedge\sigma} \subseteq \mathscr{F}_\tau$. □

§13.08 **Erinnerung**. Eine Folge $(X_n)_{n\in\mathbb{N}}$ ist gleichgradig integrierbar, wenn sie in $\mathscr{L}_p(\mathbb{P})$ mit $p > 1$ beschränkt ist, d.h. $\sup\{\|X_n\|_{\mathscr{L}_p} : n \in \mathbb{N}\} \in \mathbb{R}^+$ (Lemma §02.39 (iv)). □

§13.09 **Proposition**. *Jedes Martingal $(X_n)_{n\in\mathbb{N}}$ ist regulär, wenn es in $\mathscr{L}_p(\mathbb{P})$ mit $p > 1$ beschränkt ist, d.h. $\sup\{\|X_n\|_{\mathscr{L}_p} : n \in \mathbb{N}\} \in \mathbb{R}^+$. In diesem Fall konvergiert das Martingal in $\mathscr{L}_p(\mathbb{P})$ gegen den $\mathbb{P}$-f.s.-Grenzwert $X_\infty \in \overline{\mathscr{F}_\infty}$.*

§13.10 **Beweis** von Proposition §13.09. In der Vorlesung. □

§13.11 **Bemerkung**. Die letzte Aussage gilt im Allgemeinen nicht für $p = 1$. □

§13.12 **Lemma**. *Für jedes positive, integrierbare Submartingal $(X_n)_{n\in\mathbb{N}}$ und für alle $a \in \mathbb{R}^+_{\setminus 0}$ gilt*

$$
a\mathbb{P}\big(\sup_{m\in[\![n]\!]} X_m > a\big) \leqslant \mathbb{E}\big(X_n \mathbb{1}_{\{\sup_{m\in[\![n]\!]} X_m > a\}}\big) \quad \textit{für alle } n \in \mathbb{N}.
$$

§13.13 **Beweis** von Lemma §13.12. In der Vorlesung. □

§13.14 **Erinnerung**. Eine Folge $(X_n)_{n\in\mathbb{N}}$ ist gleichgradig integrierbar, wenn $\sup_{n\in\mathbb{N}} |X_n| \in \mathscr{L}_1(\mathbb{P})$ (Bemerkung §02.38 (iv)). □

§13.15 **Lemma**. *Für jedes in $\mathscr{L}_p(\mathscr{A}, \mathbb{P})$ mit $p > 1$ beschränkte Martingal $(X_n)_{n\in\mathbb{N}}$ ist die Zufallsvariable $\sup_{n\in\mathbb{N}} |X_n| \in \overline{\mathscr{A}}^+$ auch $\mathscr{L}_p$-integrierbar mit $\|\sup_{n\in\mathbb{N}} |X_n|\|_{\mathscr{L}_p} \leqslant \frac{p}{p-1} \sup_{n\in\mathbb{N}} \|X_n\|_{\mathscr{L}_p}$.*

§13.16 **Beweis** von Lemma §13.15. In der Vorlesung. □

§13.17 **Bemerkung**. Die letzte Aussage gilt im Allgemeinen nicht für $p = 1$. Erfüllt das Martingal $(X_n)_{n\in\mathbb{N}}$ die stärkere Bedingung $\sup_{n\in\mathbb{N}} \mathbb{E}(|X_n|(\log |X_n|)^+) \in \mathbb{R}^+$, dann ist die Zufallsvariable $\sup_{n\in\mathbb{N}} |X_n| \in \overline{\mathscr{A}}^+$ auch integrierbar (Neveu [1975, Proposition IV-2-10, S.70]) und das Martingal $(X_n)_{n\in\mathbb{N}}$ ist somit regulär. □

Die Begriffe der Filtration und des (Sub-, Super-)Martingals haben nirgends vorausgesetzt, dass die Indexmenge $\mathbb{T}$ (interpretiert als Zeit) eine $\mathbb{R}^+$-Teilmenge ist. Wir betrachten nun den Fall $\mathbb{T} = -\mathbb{N}_0$, sodass für eine Filtration $\mathscr{F} = (\mathscr{F}_z)_{z\in-\mathbb{N}_0}$ gilt $\mathscr{F}_{z-1} \subseteq \mathscr{F}_z$ für alle $z \in -\mathbb{N}_0$.

§13.18 **Definition**. Sei $\mathscr{F} = (\mathscr{F}_z)_{z\in-\mathbb{N}_0}$ eine Filtration und sei $(X_z)_{z\in-\mathbb{N}_0}$ ein integrierbares $\mathscr{F}$-(Sub-, Super-)Martingal, d.h., $X_z \in \mathscr{L}_1(\mathscr{F}_z, \mathbb{P})$ und $\mathbb{E}(X_z | \mathscr{F}_{z-1}) = X_{z-1}$ für alle $z \in \mathbb{T} = -\mathbb{N}_0$. Dann wird $X = (X_{-n})_{n\in\mathbb{N}_0}$ *(integrierbares) Rückwärts(sub-, super-)martingal* genannt. □

§13.19 **Bemerkung**. Ein integrierbares Rückwärtsmartingal ist stets gleichgradig integrierbar. Dies folgt aus Lemma §07.34 und $X_{-n} = \mathbb{E}(X_0 | \mathscr{F}_{-n})$ für alle $n \in \mathbb{N}_0$. □

§13.20 **Satz**. *Ein integrierbares Rückwärtsmartingal $(X_{-n})_{n\in\mathbb{N}_0}$ bezüglich einer Filtration $(\mathscr{F}_{-n})_{n\in\mathbb{N}_0}$ konvergiert* $\mathbb{P}$-f.s.. *Der* $\mathbb{P}$-f.s.*-Grenzwert $X_{-\infty}$ ist messbar bezüglich $\mathscr{F}_{-\infty} = \bigcap_{n\in\mathbb{N}_0} \mathscr{F}_{-n}$ und integrierbar. Weiterhin gilt $X_{-\infty} = \mathbb{E}(X_0 | \mathscr{F}_{-\infty})$* $\mathbb{P}$-f.s..

§13.21 **Beweis** von Satz §13.20. In der Vorlesung. □

§13.22 **Korollar** *(Kolmogorov'sche starkes Gesetz der Großen Zahlen). Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge unabhängiger und identisch-verteilter (u.i.v.) reellwertiger Zufallsvariablen in $\mathscr{L}_1(\Omega, \mathscr{A}, \mathbb{P})$, dann gilt*

$$n^{-1} \sum_{k\in[\![n]\!]} X_k \xrightarrow{n\to\infty} \mathbb{E}(X_1) \quad \mathbb{P}\text{-f.s. und in } \mathscr{L}_1(\mathbb{P}).$$

§13.23 **Beweis** von Korollar §13.22. In der Vorlesung. □

§14 Reguläre Stoppzeiten für integrierbare Martingale

§14.01 **Lemma**. *Sei $(X_n)_{n\in\mathbb{N}}$ ein integrierbares $\mathscr{F}$-(Sub-, Super-)Martingal. Für jede $\mathscr{F}$-Stoppzeit τ ist der gestoppte Prozess $X^\tau = (X_n^\tau := X_{\tau\wedge n})_{n\in\mathbb{N}}$ ein integrierbares $\mathscr{F}$-(Sub-, Super-)Martingal.*

§14.02 **Beweis** von Lemma §14.01. Übung □

§14.03 **Definition**. Eine $\mathscr{F}$-Stoppzeit τ heißt *regulär* für ein integriebares Martingal $(X_n)_{n\in\mathbb{N}}$, wenn der gestoppte Prozess $X^\tau = (X_n^\tau := X_{\tau\wedge n})_{n\in\mathbb{N}}$ regulär ist. □

§14.04 **Proposition**. *Für ein integrierbares Martingal $(X_n)_{n\in\mathbb{N}}$ auf einem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, \mathbb{P}, \mathscr{F})$ und eine $\mathscr{F}$-Stoppzeit τ sind die folgenden drei Aussagen äquivalent:*

(i) *τ ist regulär;*

(ii) (ii.c1) *$(X_n)_{n\in\mathbb{N}}$ konvergiert* $\mathbb{P}$-f.s. *auf dem Ereignis $\{\tau = \infty\}$ gegen $X_\infty \in \mathscr{F}_\infty$,* (ii.c2) *die Zufallsvariable $X_\tau \in \overline{\mathscr{F}_\tau}$ mit $X_\tau := X_n$ auf $\{\tau = n\}$ für alle $n \in \overline{\mathbb{N}}$ ist integrierbar; und* (ii.c3) *$X_{\tau\wedge n} = \mathbb{E}(X_\tau | \mathscr{F}_n)$* $\mathbb{P}$-f.s. *für alle $n \in \mathbb{N}$.*

(iii) (iii.c1) *$(X_n \mathbb{1}_{\{\tau>n\}})_{n\in\mathbb{N}}$ ist gleichgradig integrierbar und* (iii.c2) $\mathbb{E}(|X_\tau|\mathbb{1}_{\{\tau\in\mathbb{N}\}}) \in \mathbb{R}^+$.

§14.05 **Beweis** von Proposition §14.04. In der Vorlesung. □

§14.06 **Bemerkung**. Die Bedingungen (ii.c1)-(ii.c2) und (iii.c2) sind automatisch erfüllt, wenn für das Martingal $(X_n)_{n\in\mathbb{N}}$ gilt $\sup\{\mathbb{P}(|X_n|) : n \in \mathbb{N}\} \in \mathbb{R}^+$, insbesondere also im Fall eines positiven integrierbaren Martingals (da dann $\mathbb{P}(|X_n|) = \mathbb{P}(X_n) = \mathbb{P}(X_1)$). □

§14.07 **Satz** *(Optional sampling). Sei τ eine reguläre $\mathscr{F}$-Stoppzeit für ein integrierbares $\mathscr{F}$-Martingal $(X_n)_{n\in\mathbb{N}}$ mit* $\mathbb{P}$-f.s.*-Grenzwert $X_\infty \in \mathscr{F}_\infty$ auf $\{\tau = \infty\}$, und seien σ_1, σ_2 $\mathscr{F}$-Stoppzeiten mit $\sigma_1 \leqslant \sigma_2 \leqslant \tau$. Für $i \in [\![2]\!]$ definiere wie bisher $X_{\sigma_i} \in \overline{\mathscr{F}_{\sigma_i}}$ mit $X_{\sigma_i} := X_n$ auf $\{\sigma_i = n\}$ für alle $n \in \overline{\mathbb{N}}$. Dann sind X_{σ_1} und X_{σ_2} integrierbar und es gilt $X_{\sigma_1} = \mathbb{E}(X_{\sigma_2} | \mathscr{F}_{\sigma_1})$* $\mathbb{P}$-f.s..

§14.08 **Beweis** von Satz §14.07. In der Vorlesung. □

§14.09 **Korollar**. *Seien τ und σ zwei $\mathscr{F}$-Stoppzeiten mit $\tau \leqslant \sigma$* $\mathbb{P}$-f.s.. *Für ein integrierbares Martingal $(X_n)_{n\in\mathbb{N}}$ ist τ regulär, wenn σ regulär ist.*

§14.10 **Beweis** von Korollar §14.09. In der Vorlesung. □

§14.11 **Bemerkung**. Korollar §14.09 zeigt insbesondere, für ein reguläres integrierbares Martingal ist jede Stoppzeit regulär (wähle $\sigma = \infty$). Andererseits ist für ein integrierbares Martingal jede endliche konstante Stoppzeit regulär, und somit nach Korollar §14.09 auch jede beschränkte Stoppzeit. □

§14.12 **Korollar**. *Für ein integrierbares Martingal $(X_n)_{n\in\mathbb{N}}$ mit $\sup\{\mathbb{P}(|X_n|) : n \in \mathbb{N}\} \in \mathbb{R}^+$, insbesondere also für jedes positive, integrierbare Martingal, ist die Treffzeit τ_a mit $a \in \mathbb{R}^+$ (Beispiel §10.10 (b)) gegeben durch $\tau_a := \inf\{n \in \mathbb{N} : |X_n| > a\}$ (mit $\inf \emptyset = \infty$) regulär.*

§14.13 **Beweis** von Korollar §14.12. In der Vorlesung. □

§14.14 **Lemma**. *Sei $(X_n)_{n\in\mathbb{N}}$ ein integrierbares $\mathscr{F}$-Martingal. Eine $\mathscr{F}$-Stoppzeit τ ist genau dann regulär für $(X_n)_{n\in\mathbb{N}}$ und $(X_n)_{n\in\mathbb{N}}$ ist konvergent $\mathbb{P}$-f.s. gegen 0 auf $\{\tau = \infty\}$, wenn die folgenden zwei Bedingungen* (c1) $\mathbb{P}\big(|X_\tau|\mathbb{1}_{\{\tau\in\mathbb{N}\}}\big) \in \mathbb{R}^+$ *und* (c2) $\lim_{n\to\infty} \mathbb{P}\big(|X_n|\mathbb{1}_{\{\tau>n\}}\big) = 0$ *erfüllt sind.*

§14.15 **Beweis** von Lemma §14.14. In der Vorlesung. □

§14.16 **Beispiel** *(Wald'sche Gleichung)*. Sei $X = (X_n)_{n\in\mathbb{N}}$ eine Folge u.i.v. reellwertiger Zufallsvariablen definiert auf einem auf einem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, \mathbb{P}, \sigma(X))$, wobei $\sigma(X)$ die von X erzeugte Filtration ist (Definition §10.02). Ist weiterhin $X_1 \in \mathscr{L}_2(\mathbb{P})$, dann sind $S^\circ := (S_n - n\mathbb{P}(X_1))_{n\in\mathbb{N}}$ mit $S = (S_n := \sum_{i\in[\![n]\!]} X_i)_{n\in\mathbb{N}}$ und $V := ((S_n - n\mathbb{P}(X_1))^2 - n\,\mathbb{V}\mathrm{ar}(X_1))_{n\in\mathbb{N}}$ integrierbare Martingale, die nicht regulär sind, da sie $\mathbb{P}$-f.s. diverieren für $n \to \infty$. Andererseits, ist jede $\sigma(X)$)-Stoppzeit τ mit $\mathbb{P}(\tau) \in \mathbb{R}^+$ ist regulär für jedes der beiden Martingale S° und V. In diesem gelten die *Wald'schen Gleichungen* (a) $\mathbb{E}(S_\tau) = \mathbb{E}(\tau)\mathbb{E}(X_1)$ und (b) $\mathbb{E}\big(S_\tau - \tau\mathbb{E}(X_1)\big)^2 = \mathbb{E}(\tau)\,\mathbb{V}\mathrm{ar}(X_1)$. Ist weiterhin $\mathbb{E}(\tau^2) \in \mathbb{R}^+$, dann gilt auch (c) $\mathbb{V}\mathrm{ar}(S_\tau) = \mathbb{V}\mathrm{ar}(\tau)(\mathbb{E}X_1)^2 + \mathbb{E}(\tau)\,\mathbb{V}\mathrm{ar}(X_1)$. □

§15 Reguläre integrierbare Submartingale

§15.01 **Vorbemerkung**. Die Krickeberg-Zerlegung für Submartingale erlaubt es die Theorie für integrierbare Martingale ohne weitere Schwierigkeiten auf integrierbare Submartingale zu erweitern. □

§15.02 **Erinnerung**.

(i) Ein integrierbares $\mathscr{F}$-Submartingal $(X_n)_{n\in\mathbb{N}}$ mit (A) $\sup\{\mathbb{P}(X_n^+) : n \in \mathbb{N}\} \in \mathbb{R}^+$ ist nach Satz §12.05 $\mathbb{P}$-f.s.-konvergent gegen $X_\infty \in \mathscr{L}_1(\mathscr{F}_\infty, \mathbb{P})$.

(ii) Eine Folge $(X_n^+)_{n\in\mathbb{N}}$ ist gleichgradig integrierbar (Definition §02.37) in $\mathscr{L}_1(\mathscr{A}, \mathbb{P})$, d.h. $\inf\big\{\sup_{n\in\mathbb{N}} \mathbb{P}\big(X_n^+\mathbb{1}_{\{X_n^+\geqslant a\}}\big) : a \in \mathbb{R}^+\big\} = 0$, wenn $\sup_{n\in\mathbb{N}} X_n^+ \in \mathscr{L}_1(\mathbb{P})$ (Bemerkung §02.38 (iv)). □

§15.03 **Satz**. *Für ein integrierbares Submartingal $(X_n)_{n\in\mathbb{N}}$ auf einem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, \mathbb{P}, \mathscr{F})$ sind die folgenden vier Aussage äquivalent:*

(i) *Die Folge $(X_n^+)_{n\in\mathbb{N}}$ konvergiert in $\mathscr{L}_1(\mathbb{P})$;*

(ii) *Es gilt* (A) $\sup\{\mathbb{P}(X_n^+) : n \in \mathbb{N}\} \in \mathbb{R}^+$. *Der $\mathbb{P}$-f.s.-Grenzwert $X_\infty = \lim_{n\to\infty} X_n$, der nach Satz §12.05 in $\mathscr{L}_1(\mathbb{P})$ exisitiert, erfüllt $X_n \leqslant \mathbb{E}\big(X_\infty \big| \mathscr{F}_n\big)$ $\mathbb{P}$-f.s. für alle $n \in \mathbb{N}$;*

(iii) *Es gibt eine Zufallsvariable $Y \in \mathscr{L}_1(\mathscr{A}, \mathbb{P})$ mit $X_n \leqslant \mathbb{E}\big(Y \big| \mathscr{F}_n\big)$ für alle $n \in \mathbb{N}$;*

(iv) *Die Folge $(X_n^+)_{n\in\mathbb{N}}$ ist gleichgradig integrierbar in $\mathscr{L}_1(\mathscr{A}, \mathbb{P})$.*

Das integrierbare Submartingal $(X_n)_{n\in\mathbb{N}}$ heißt regulär *wenn eine der äquivalenten Aussagen* (i)-(iv) *erfüllt ist.*

§15.04 **Beweis** von Satz §15.03. In der Vorlesung. □

§15.05 **Bemerkung**. Für ein negatives integrierbares Submartingal (also für ein positive integrierbares Supermartingale mit gewechseltem Vorzeichen) sind die Aussagen in §15.03 offensichtlich erfüllt. Im Allgemeinen konvergiert das Submartingal aber nicht im Mittel, obwohl es immer $\mathbb{P}$-f.s. konvergiert. Insbesondere ist die Aussage (i) strikt schwächer als die Konvergenz im Mittel des Submartingals. Auf der anderen Seite für ein positives Submartingal folgt aus (i) natürlich schon die Konvergenz im Mittel des Submartingals. □

§15.06 **Korollar** *(Optional sampling). Sei $(X_n)_{n\in\mathbb{N}}$ ein reguläres integrierbares $\mathscr{F}$-Submartingal mit $\mathbb{P}$-f.s.-Grenzwert X_∞.*

(i) *Für jede $\mathscr{F}$-Stoppzeit τ ist die Zufallsvariable $X_\tau \in \overline{\mathscr{F}_\tau}$ mit $X_\tau := X_n$ auf $\{\tau = n\}$ für alle $n \in \overline{\mathbb{N}}$ integrierbar.*

(ii) *Für jedes Paar von $\mathscr{F}$-Stoppzeiten τ, σ mit $\tau \leqslant \sigma$ $\mathbb{P}$-f.s. gilt auch $X_\tau \leqslant \mathbb{E}\big(X_\sigma \big| \mathscr{F}_\tau\big)$ $\mathbb{P}$-f.s..*

§15.07 **Beweis** von Korollar §15.06. In der Vorlesung. □

§15.08 **Bemerkung**. Abschliessend, ist es ohne Schwierigkeiten möglich die Regularität einer Stoppzeit (vgl. Proposition §14.04 und Satz §14.07) auch für intergrierbare Submartingale einzuführen. Die einzigen Unterschiede in den Aussagen ist die Ersetzung des Wortes „Martingal" durch das Wort „Submartingal" und das Festhalten der Ungleichungen $X_{\tau\wedge n} \leqslant \mathbb{P}\big(X_\tau \big| \mathscr{F}_n\big)$ und $X_{\sigma_1} \leqslant \mathbb{P}\big(X_{\sigma_2} \big| \mathscr{F}_{\sigma_1}\big)$ anstelle der entsprechenden Gleichungen. □

§16 Doob-Zerlegung und quadratische Variation

§16.01 **Vorbemerkung**. Sei $X = (X_n)_{n\in\mathbb{N}_0}$ ein adaptierter, integrierbarer Prozess auf einem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, \mathbb{P}, \mathscr{F})$. Wir zerlegen X in die Summe aus einem vorhersagbaren Prozess $A = (A_n)_{n\in\mathbb{N}_0}$ (Definition §10.05), dem Trendanteil oder systematischen Anteil, und einem Martingal $M = (M_n)_{n\in\mathbb{N}_0}$ (Definition §12.01), dem Fehleranteil. Dazu definieren wir $M_0 := X_0$, $A_0 := 0$ und für $n \in \mathbb{N}$

$$M_n := X_0 + \sum_{k\in[\![n]\!]} \big(X_k - \mathbb{P}\big(X_k \big| \mathscr{F}_{k-1}\big)\big) \quad \text{und} \quad A_n := \sum_{k\in[\![n]\!]} \big(\mathbb{P}\big(X_k \big| \mathscr{F}_{k-1}\big) - X_{k-1}\big).$$

Offensichtlich gilt $X = (X_n = M_n + A_n)_{n\in\mathbb{N}_0} = M + A$. Per Konstruktion gilt für $n \in \mathbb{N}$

$$M_n - M_{n-1} = X_n - \mathbb{E}\big(X_n \big| \mathscr{F}_{n-1}\big) \quad \text{und} \quad A_n - A_{n-1} = \mathbb{E}\big(X_n \big| \mathscr{F}_{n-1}\big) - X_{n-1}$$

Folglich ist A vorhersagbar (mit $A_0 = 0$), und M ein Martingal, da $\mathbb{P}\big(M_n - M_{n-1} \big| \mathscr{F}_{n-1}\big) = \mathbb{P}\big(X_n - \mathbb{P}\big(X_n \big| \mathscr{F}_{n-1}\big) \big| \mathscr{F}_{n-1}\big) = 0$ $\mathbb{P}$-f.s.. □

§16.02 **Satz** *(Doob-Zerlegung). Sei $X = (X_n)_{n\in\mathbb{N}_0}$ ein adaptierter, integrierbarer Prozess auf einem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, \mathbb{P}, \mathscr{F})$. Dann existiert eine eindeutige Zerlegung $X = M + A$, wobei A vorhersagbar (mit $A_0 = 0$) und M ein Martingal ist. Diese Darstellung von X heißt Doob-Zerlegung. X ist genau dann ein Submartingal (bzw. Supermartingal), wenn A ein wachsender (bzw. fallender) Prozess ist (Definition §10.05).*

§16.03 **Beweis** von Satz §16.02. In der Vorlesung. □

§16.04 **Erinnerung**.

(i) Ein integrierbares $\mathscr{F}$-Submartingal $(X_n)_{n\in\mathbb{N}}$ mit (A) $\sup\{\mathbb{P}(X_n^+) : n \in \mathbb{N}\} \in \mathbb{R}^+$ ist nach Satz §12.05 $\mathbb{P}$-f.s.-konvergent gegen $X_\infty \in \mathscr{L}_1(\mathscr{F}_\infty, \mathbb{P})$.

(ii) Eine Folge $(X_n^+)_{n\in\mathbb{N}}$ ist gleichgradig integrierbar (Definition §02.37) in $\mathscr{L}_1(\mathscr{A}, \mathbb{P})$, d.h. $\inf\big\{\sup_{n\in\mathbb{N}} \mathbb{P}\big(X_n^+ \mathbb{1}_{\{X_n^+ \geqslant a\}}\big) : a \in \mathbb{R}^+\big\} = 0$, wenn $\sup_{n\in\mathbb{N}} X_n^+ \in \mathscr{L}_1(\mathbb{P})$ (Bemerkung §02.38 (iv)). □

§16.05 **Proposition**. *Sei $X = (X_n)_{n\in\mathbb{N}_0}$ ein integrierbares Submartingal auf einem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, \mathbb{P}, \mathscr{F})$ und sei $X = M + A$ seine Doob-Zerlegung.*

(i) *Die Bedingung* (A) $\sup\{\mathbb{P}(X_n^+) : n \in \mathbb{N}_0\} \in \mathbb{R}^+$ *gilt genau dann, wenn* (i.c1) $\sup\{\mathbb{E}(|M_n|) : n \in \mathbb{N}_0\} \in \mathbb{R}^+$ *und* (i.c2) $A_\infty \in \mathscr{L}_1(\mathbb{P})$ *gelten.*

(ii) *Das Submartingal X konvergiert in $\mathscr{L}_1(\mathbb{P})$ genau dann, wenn* (ii.c1) *M ein reguläres Martingal ist und* (i.c2) *gilt.*

(iii) *Sei τ eine reguläre Stoppzeit für das Martingal M. Dann ist die Zufallsvariable X_τ genau dann integrierbar, wenn A_τ integrierbar ist. In diesem Fall gilt $\mathbb{E}(X_\tau) = \mathbb{E}(M_0) + \mathbb{E}(A_\tau)$.*

§16.06 **Beweis** von Proposition §16.05. In der Vorlesung. □

§16.07 **Beispiel**. Sei $(X_n)_{n\in\mathbb{N}_0}$ ein quadratintegrierbares $\mathscr{F}$-Martingal, also $X_n \in \mathscr{L}_2(\mathscr{A}, \mathbb{P})$ für alle $n \in \mathbb{N}_0$. Dann ist $X^2 := (X_n^2)_{n\in\mathbb{N}_0}$ ein integrierbares Submartingal (Bemerkung §12.04) und es gilt weiterhin $\mathbb{E}\big(X_{i-1}X_i \big| \mathscr{F}_{i-1}\big) = X_{i-1}\mathbb{E}\big(X_i \big| \mathscr{F}_{i-1}\big) = X_{i-1}^2$ $\mathbb{P}$-f.s. für alle $i \in \mathbb{N}$. Sei $X^2 = M + A$ die Doob-Zerlegung von X^2, wobei $M = (X_n^2 - A_n)_{n\in\mathbb{N}_0}$ ein Martingal ist und für $n \in \mathbb{N}$,

$$A_n = \sum_{i\in[\![n]\!]} \big(\mathbb{E}\big(X_i^2 \big| \mathscr{F}_{i-1}\big) - X_{i-1}^2\big)$$
$$= \sum_{i\in[\![n]\!]} \big(\mathbb{E}\big((X_i - X_{i-1})^2 \big| \mathscr{F}_{i-1}\big) - 2X_{i-1}^2 + 2\mathbb{E}\big(X_{i-1}X_i \big| \mathscr{F}_{i-1}\big)\big)$$
$$= \sum_{i\in[\![n]\!]} \mathbb{E}\big((X_i - X_{i-1})^2 \big| \mathscr{F}_{i-1}\big) \quad \mathbb{P}\text{-f.s.}. \quad \square$$

§16.08 **Definition**. Sei $(X_n)_{n\in\mathbb{N}_0}$ ein quadratintegrierbares $\mathscr{F}$-Martingal. Der eindeutig bestimmte wachsende (vorhersagbare) Prozess A, mit dem $(X_n^2 - A_n)_{n\in\mathbb{N}_0}$ ein Martingal wird, heißt *quadratischer Variationsprozess* von X. Wir schreiben kurz $\langle X \rangle := (\langle X \rangle_n)_{n\in\mathbb{N}_0} := A$. □

§16.09 **Proposition**. *Sei X mit quadratischer Variationsprozess $\langle X \rangle$ wie in Definition §16.08. Dann ist für $n \in \mathbb{N}$, $\langle X \rangle_n = \sum_{i\in[\![n]\!]} \mathbb{E}\big((X_i - X_{i-1})^2 \big| \mathscr{F}_{i-1}\big)$ $\mathbb{P}$-f.s. und $\mathbb{E}\big(\langle X \rangle_n\big) = \mathbb{V}\mathrm{ar}(X_n - X_0)$.*

§16.10 **Beweis** von Proposition §16.09. Die Aussage ergibt sich sofort aus Beispiel §16.07. □

§16.11 **Beispiel**. Sei $X = (X_n)_{n\in\mathbb{N}}$ eine Folge unabhängiger, quadratintegrierbarer Zufallsvariablen definiert auf einem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, \mathbb{P}, \sigma(X))$, wobei $\sigma(X) = (\sigma_n(X))_{n\in\mathbb{N}_0}$ mit $\sigma_0(X) = \{\emptyset, \Omega\}$ und $\sigma_n(X) = \sigma(X_j, j \in [\![n]\!])$, $n \in \mathbb{N}$, die von X erzeugte Filtration ist (Beispiel §10.04).

(a) Wenn $\mathbb{E}(X_n) = 0$ für alle $n \in \mathbb{N}$ gilt, dann ist $S := (S_n)_{n\in\mathbb{N}_0}$ mit $S_0 := 0$ und $S_n := \sum_{i\in[\![n]\!]} X_i$, $n \in \mathbb{N}$, ein quadratintegrierbares $\sigma(X)$-Martingal mit quadratischem Variationsprozess $\langle S \rangle = (\langle S \rangle_n)_{n\in\mathbb{N}_0}$ gegeben durch $\langle S \rangle_0 = 0$ und $\langle S \rangle_n = \sum_{i\in[\![n]\!]} \mathbb{E}\big(X_i^2 \big| \sigma_{i-1}(X)\big) = \sum_{i\in[\![n]\!]} \mathbb{E}(X_i^2)$ $\mathbb{P}$-f.s. für $n \in \mathbb{N}$. Man beachte, dass es für diese einfache Darstellung von $\langle S \rangle$ nicht ausreicht, dass die Zufallsvariablen $(X_n)_{n\in\mathbb{N}}$ unkorrelliert sind.

(b) Wenn $\mathbb{E}(X_n) = 1$ für alle $n \in \mathbb{N}$ gilt, dann ist $P := (P_n)_{n\in\mathbb{N}_0}$ mit $P_0 := 1$ und $P_n := \prod_{i\in[\![n]\!]} X_i$, $n \in \mathbb{N}$, ein quadratintegrierbares $\sigma(X)$-Martingal mit quadratischem Variationsprozess $\langle P \rangle = (\langle P \rangle_n)_{n\in\mathbb{N}_0}$ gegeben durch $\langle P \rangle_0 = 0$ und $\langle P \rangle_n = \sum_{i\in[\![n]\!]} \mathbb{V}\mathrm{ar}(X_i)P_{i-1}^2$ $\mathbb{P}$-f.s. für $n \in \mathbb{N}$, da $\mathbb{E}\big((P_n - P_{n-1})^2 \big| \mathscr{F}_{n-1}\big) = \mathbb{E}\big((X_n - 1)^2 P_{n-1}^2 \big| \mathscr{F}_{n-1}\big) = \mathbb{V}\mathrm{ar}(X_n)P_{n-1}^2$ $\mathbb{P}$-f.s. für alle $n \in \mathbb{N}$ gilt. Damit ist der quadratische Variationsprozess ein echt zufälliger Prozess. □

§16.12 **Lemma**. *Sei $(X_n)_{n\in\mathbb{N}_0}$ ein quadratintegrierbares $\mathscr{F}$-Martingal mit quadratischer Variationsprozess $\langle X \rangle$, und sei τ eine $\mathscr{F}$-Stoppzeit. Dann ist der quadratischer Variationsprozess $\langle X^\tau \rangle$ des gestoppten Prozesses X^τ gegeben durch $\langle X^\tau \rangle = \langle X \rangle^\tau = (\langle X \rangle_{\tau\wedge n})_{n\in\mathbb{N}_0}$.*

§16.13 **Beweis** von Lemma §16.12. In der Vorlesung. □

§16.14 **Proposition**. *Sei $X = (X_n)_{n\in\mathbb{N}_0}$ ein quadratintegrierbares $\mathscr{F}$-Martingal mit $X_0 = 0$.*

(i) *Ist $\langle X\rangle_\infty := \sup\{\langle X\rangle_n : n \in \mathbb{N}\}$ integrierbar, dann konvergiert das Martingal X in $\mathscr{L}_2(\mathbb{P})$ und X ist somit regulär. Weiterhin gilt $\mathbb{E}\big(\sup_{n\in\mathbb{N}_0} X_n^2\big) \leqslant 4\mathbb{E}(\langle X\rangle_\infty) \in \mathbb{R}^+$.*

(ii) *Ist τ eine $\mathscr{F}$-Stoppzeit mit $\mathbb{E}\big(\langle X\rangle_\tau^{1/2}\big) \in \mathbb{R}^+$, so gilt $\mathbb{E}\big(\sup_{n\in[\![\tau]\!]} |X_n|\big) \leqslant 3\mathbb{E}\big(\langle X\rangle_\tau^{1/2}\big) \in \mathbb{R}^+$ und τ ist regulär für das Martingal X.*

(iii) *In jedem Fall konvergiert das Martingal X $\mathbb{P}$-f.s. gegen einen endlichen Grenzwert auf dem Ereignis $\big\{\langle X\rangle_\infty \in \mathbb{R}^+\big\}$.*

§16.15 **Beweis** von Proposition §16.14. In der Vorlesung. □

§16.16 **Korollar**. *Für ein quadratintegrierbares $\mathscr{F}$-Martingal $X = (X_n)_{n\in\mathbb{N}_0}$ mit quadratischem Variationsprozess $\langle X\rangle$ sind die folgenden vier Aussagen äquivalent:* (i) $\sup\{\mathbb{E}(X_n^2) : n \in \mathbb{N}_0\} \in \mathbb{R}^+$, (ii) $\mathbb{E}(\langle X\rangle_\infty) \in \mathbb{R}^+$, (iii) *$X$ konvergiert in $\mathscr{L}_2(\mathbb{P})$, und* (iv) *$X$ konvergiert $\mathbb{P}$-f.s. und in $\mathscr{L}_2(\mathbb{P})$.*

§16.17 **Beweis** von Korollar §16.16. In der Vorlesung. □

§16.18 **Proposition**. *Für ein quadratintegrierbares $\mathscr{F}$-Martingal $X = (X_n)_{n\in\mathbb{N}_0}$ mit quadratischem Variationsprozess $\langle X\rangle$ und $\alpha \in (1/2, \infty)$ gilt*

$$(X_n - X_0)/(\langle X\rangle_n)^\alpha \xrightarrow{n\to\infty} 0 \quad \mathbb{P}\text{-f.s.} \quad \textit{auf}\ \big\{\langle X\rangle_\infty = \infty\big\}.$$

§16.19 **Beweis** von Proposition §16.18. In der Vorlesung. □

§16.20 **Beispiel**. Wie in Beispiel §16.11 sei $X = (X_n)_{n\in\mathbb{N}}$ eine Folge unabhängiger, quadratintegrierbarer Zufallsvariablen definiert auf dem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, \mathbb{P}, \sigma(X))$. Analog zu Beispiel §16.11 (a) betrachte $S := (S_n)_{n\in\mathbb{N}_0}$ mit $S_0 := 0$ und $S_n := \sum_{i\in[\![n]\!]} (X_i - \mathbb{E}X_i)$, $n \in \mathbb{N}$, wobei $\langle S\rangle_n = \sum_{i\in[\![n]\!]} \mathbb{V}\mathrm{ar}(X_i)$ und nach Proposition §16.18 für jedes $\alpha \in (1/2, \infty)$ gilt $S_n/(\sum_{i\in[\![n]\!]} \mathbb{V}\mathrm{ar}(X_i))^\alpha \xrightarrow{\mathbb{P}\text{-f.s.}} 0$, falls $\sum_{i\in\mathbb{N}} \mathbb{V}\mathrm{ar}(X_i) = \infty$. Insbesondere, ist $X = (X_n)_{n\in\mathbb{N}}$ eine Folge unabhängiger, identisch-verteilter, quadratintegrierbarer Zufallsvariablen, dann gilt $S_n/n^\alpha \xrightarrow{\mathbb{P}\text{-f.s.}} 0$. Andererseits, ist $(a_n)_{n\in\mathbb{N}}$ eine wachsende, unbeschränkte Folge in $\mathbb{R}$, dann gilt für jede reellwertige Folge $(y_n)_{n\in\mathbb{N}}$ mit $\sum_{n\in\mathbb{N}} y_n/a_n \in \mathbb{R}^+$ auch $a_n^{-1} \sum_{i\in[\![n]\!]} y_i \xrightarrow{n\to\infty} 0$ (Kronecker's Lemma). Folglich, wenn $\sum_{i\in\mathbb{N}} \mathbb{V}\mathrm{ar}(X_i)/a_i^2 \in \mathbb{R}^+$, dann folgt aus Korollar §16.16, dass das Martingal $(\sum_{i\in[\![n]\!]} (X_i - \mathbb{E}(X_i))/a_i)_{n\in\mathbb{N}}$ $\mathbb{P}$-f.s.-konvergiert gegen einen endlichen Grenzwert und, somit nach Kronecker's Lemma $a_n^{-1} \sum_{i\in[\![n]\!]} (X_i - \mathbb{E}X_i) \xrightarrow{\mathbb{P}\text{-f.s.}} 0$. Im Fall, dass $(X_n)_{n\in\mathbb{N}}$ unabhängig und identisch-verteilt sind, schliessen wir damit $n^{-1} \sum_{i=1}^n (X_i - \mathbb{E}X_i) \xrightarrow{\mathbb{P}\text{-f.s.}} 0$. □

Kapitel 5

Markovketten

§17 Markovketten

§17.01 **Definition.** Seien $\mathbb{T} = \mathbb{N}_0$ (diskrete Zeit) oder $\mathbb{T} = \mathbb{R}^+$ (stetige Zeit) und $\mathcal{S}$ eine abzählbare nichtleere Menge (Zustandsraum) versehen mit der σ-Algebra $\mathscr{S} = 2^{\mathcal{S}}$. Ein stochastischer Prozess $(X_t)_{t\in\mathbb{T}}$ mit Zustandsraum $\mathcal{S}$ heißt *Markovkette* wenn für alle $n \in \mathbb{N}$, und alle $t_1 < t_2 < \ldots < t_n < t$ in $\mathbb{T}$ und alle $s_1, \ldots, s_n, s$ in $\mathcal{S}$ mit $\mathbb{P}(X_{t_1} = s_1, \ldots, X_{t_n} = s_n) \in (0,1]$ die *Markoveigenschaft* gilt: $\mathbb{P}\big(X_t = s \big| X_{t_1} = s_1, \ldots, X_{t_n} = s_n\big) = \mathbb{P}\big(X_t = s \big| X_{t_n} = s_n\big)$. Für eine Markovkette $(X_t)_{t\in\mathbb{T}}$ und $t_1 \leqslant t_2$ aus $\mathbb{T}$, $i, j \in \mathcal{S}$ bezeichnet $p_{ij}(t_1, t_2) := \mathbb{P}\big(X_{t_2} = j \big| X_{t_1} = i\big)$ die *Übergangswahrscheinlichkeit*, den Zustand j zum Zeitpunkt t_2 aus dem Zustand i zum Zeitpunkt t_1 zu erreichen (oder beliebig, wenn es nicht wohldefiniert ist). $P(t_1, t_2) := \big(p_{ij}(t_1, t_2)\big)_{i,j\in\mathcal{S}}$ wird dann *Übergangsmatrix* genannt. Die Übergangsmatrix und auch die Markovkette heiß *zeitlich homogen* (*time-homogeneous*), wenn $P(t_1, t_2) = P(0, t_2 - t_1) =: P(t_2 - t_1)$ für alle $t_1 \leqslant t_2$ und $\mathbb{T}$ gilt. □

§17.02 **Lemma.** *Die Übergangsmatrix genügt der Chapman-Kolmogorov Gleichung, d.h., für alle $t_1 \leqslant t_2 \leqslant t_3$ aus $\mathbb{T}$ gilt $P(t_1, t_3) = P(t_1, t_2)P(t_2, t_3)$ (Matrixmultiplikation). Im zeitlich homogenen Fall erhalten wir die Semigruppeneigenschaft $P(t_1 + t_2) = P(t_1)P(t_2)$ für alle $t_1, t_2 \in \mathbb{T}$, und insbesondere mit Einschrittübergangsmatrix $P := P(1)$ auch $P(n) = P^n$ für alle $n \in \mathbb{N}$.*

§17.03 **Beweis** von Lemma §17.02. In der Vorlesung. □

Im Folgenden bezeichnet $X = (X_n)_{n\in\mathbb{N}_0}$ eine zeitlich homogene Markovkette mit abzählbarem Zustandsraum $(\mathcal{S}, 2^{\mathcal{S}})$ und (Einschritt-)Übergangsmatrix $P = (P_{ij})_{i,j\in\mathcal{S}}$. Gegeben eine initiale (diskrete) Wahrscheinlichkeitsverteilung μ auf $(\mathcal{S}, 2^{\mathcal{S}})$ definiert

$$\mathbb{P}_{[\![0,n]\!]}(B_0 \times \cdots \times B_n) := \sum_{j_0\in B_0} \mu(\{j_0\}) \sum_{j_1\in B_1} P_{j_0,j_1} \cdots \sum_{j_n\in B_n} P_{j_{n-1},j_n}, \text{ für } B_0, B_1, \ldots B_n \subseteq \mathcal{S}$$

eine projektive Familie $\{\mathbb{P}_{\mathcal{J}}, \mathcal{J} \subseteq \mathbb{N}_0 \text{ endlich}\}$ von Wahrscheinlichkeitsmaßen auf dem Produktraum $(\mathcal{S}^{\mathbb{N}_0}, 2^{\mathcal{S}^{\mathbb{N}_0}})$, welche nach dem Kolmogorov'schen Erweiterungssatz §04.32 ein eindeutig bestimmtes Wahrscheinlichkeitsmaß $\mathbb{P}_\mu$ auf $(\mathcal{S}^{\mathbb{N}_0}, 2^{\mathcal{S}^{\mathbb{N}_0}})$ festlegt. Die Markovkette $X = (X_n)_{n\in\mathbb{N}_0}$ gegeben als kanonischer Koordinatenprozess, d.h. $X_n = \Pi_{\{n\}} : \mathcal{S}^{\mathbb{N}_0} \to \mathcal{S}$ mit $(j_m)_{m\in\mathbb{N}_0} \mapsto j_n$ (vgl. Schreibweise §04.30) besitzt dann als Bildwahrscheinlichkeitsmaß $\mathbb{P}_\mu$. Für $B_0, B_1, \ldots$ aus $2^{\mathcal{S}}$ gilt $\mathbb{P}_\mu(X_0 \in B_0, \ldots, X_n \in B_n) = \mathbb{P}_{[\![0,n]\!]}(B_0 \times \cdots \times B_n)$ und $\mathbb{P}_\mu(X_0 \in B_0) = \mu(B_0)$. Im Fall eines Einpunktmaßes $\mu = \delta_j$, $j \in \mathcal{S}$ (vgl. Beispiel §01.16) schreiben wir kurz $\mathbb{P}_j := \mathbb{P}_{\delta_j}$. Für jedes initiale Wahrscheinlichkeitsmaß μ und für jedes $A \in 2^{\mathcal{S}^{\mathbb{N}_0}}$ gilt $\mathbb{P}_\mu(A) = \sum_{j\in\mathcal{S}} \mu(\{j\})\mathbb{P}_j(A)$.

§17.04 **Definition.** Ein stochastischer Prozess $X = (X_n)_{n\in\mathbb{N}_0}$ mit Werten in einem abzählbaren Zustandsraum $(\mathcal{S}, 2^{\mathcal{S}})$ heißt zeitlich homogene Markovkette mit Familie $(\mathbb{P}_j)_{j\in\mathcal{S}}$ von Wahrscheinlichkeitsmaßen auf $(\mathcal{S}^{\mathbb{N}_0}, 2^{\mathcal{S}^{\mathbb{N}_0}})$, wenn

(a) Für jedes $j \in \mathcal{S}$ ist $(X_n)_{n\in\mathbb{N}_0}$ ein stochastischer Prozess auf dem Wahrscheinlichkeitsraum $(\mathcal{S}^{\mathbb{N}_0}, 2^{\mathcal{S}^{\otimes\mathbb{N}_0}}, \mathbb{P}_j)$ mit $\mathbb{P}_j(X_0 = j) = 1$.

(b) Die Abbildung $\kappa : \mathcal{S} \times 2^{\mathcal{S}^{\mathbb{N}_0}} \to [0,1]$ mit $(j, A) \mapsto \kappa(j, A) := \mathbb{P}_j(A)$ ist ein Markovkern (eine reguläre bedingte Wahrscheinlichkeitsverteilung). Für jedes $n \in \mathbb{N}_0$, ist die Abbildung $\kappa_n : \mathcal{S} \times 2^{\mathcal{S}} \to [0,1]$, $(j, B) \mapsto \kappa(j, \Pi_{\{n\}}^{-1}(B)) = \mathbb{P}_j(X_n \in B)$ ein Markovkern und die n-Schritt Übergangsmatrix $(P_{ij}^n)_{i,j\in\mathcal{S}}$ von X ist gegeben durch $P_{ij}^n = \kappa_n(i, \{j\}) = \mathbb{P}_i(X_n = j)$.

(c) $X = (X_n)_{n\in\mathbb{N}_0}$ besitzt bezüglich der natürlichen Filtration $\sigma(X) = (\sigma_n(X))_{n\in\mathbb{N}_0}$ mit $\sigma_0(X) = \{\emptyset, \Omega\}$ und $\sigma_n(X) = \sigma(X_j, j \in [\![n]\!])$, $n \in \mathbb{N}$, die zeitlich homogene Markoveigenschaft: für alle $i, j \in \mathcal{S}$ und $m, n \in \mathbb{N}_0$ gilt $\mathbb{P}_i\big(X_{n+m} = j \big| \sigma_m(X)\big) = \kappa_n(X_m, \{j\}) = \mathbb{P}_{X_m}(X_n = j) = P_{X_m j}^n$, $\mathbb{P}_i$-f.s..

Wir schreiben $\mathbb{E}_j$ für die Erwartung bezüglich $\mathbb{P}_j$, $\mathbb{P}_j^X = \mathbb{P}_j$ und $\mathbb{P}_j^{X|\mathscr{A}}$ für die induzierte Veteri- lung von X beziehungsweise induzierte reguläre bedingte Verteilung von X bei gegebenem $\mathscr{A}$ (vgl. Definition §06.22) sowie $\mathbb{E}_j\big(f(X)\big|\mathscr{A}\big)$ für den bedingten Erwartungswert von $f(X)$ bei gegebenem $\mathscr{A}$. Insbesondere benutzen wir die Schreibweise $\mathbb{P}_{X_m} = \kappa(X_m, \bullet)$, wobei wir X_m als Startwert einer zweiten Markovketten mit der gleichen Familie von Wahrscheinlichkeitsmaßen $(\mathbb{P}_j)_{j\in\mathcal{S}}$ verstehen. □

§17.05 **Bemerkung**. Die Existenz der Familie $(\kappa_n)_{n\in\mathbb{N}_0}$ von Markovkernen impliziert die Existenz des Markovkerns κ (z.Bsp. Klenke [2020], Satz 17.8, S.391). Damit, ist eine zeitlich homogene Markovkette ein stochastischer Prozess mit Markoveigenschaft, für die die Übergangswahrscheinlichkeiten zeitlich homogen sind. □

§17.06 **Erinnerung**. Für $n \in \mathbb{N}_0$ ist $\vartheta_n : \mathcal{S}^{\mathbb{N}_0} \to \mathcal{S}^{\mathbb{N}_0}$ mit $s = (s_m)_{m\in\mathbb{N}_0} \mapsto \vartheta_n(s) := (s_{m+n})_{m\in\mathbb{N}_0}$ der *Shift* (-Operator). □

§17.07 **Satz**. *Ein stochastischer Prozess $X = (X_n)_{n\in\mathbb{N}_0}$ ist genau dann eine zeitlich homogene Markovkette, wenn für jedes $n \in \mathbb{N}_0$ und $j \in \mathcal{S}$, $\mathbb{P}_j^{\vartheta_n(X)|\sigma_n(X)} = \mathbb{P}_{X_n}^X (= \mathbb{P}_{X_n})$ gilt und äquivalent dazu, wenn ein Markovkern $\kappa : \mathcal{S} \times 2^{\mathcal{S}^{\mathbb{N}_0}} \to [0,1]$ existiert, so dass für jede beschränkte Funktion $f : \mathcal{S}^{\mathbb{N}_0} \to \mathbb{R}$ und für jedes $n \in \mathbb{N}_0$ und $j \in \mathcal{S}$ gilt $\mathbb{E}_j\big(f(\vartheta_n(X))\big|\sigma_n(X)\big) = \mathbb{E}_{X_n}\big(f(X)\big) := \int_{\mathcal{S}^{\mathbb{N}_0}} f(s)\kappa(X_n, ds)$.* □

§17.08 **Beweis** von Satz §17.07. Klenke [2020, Satz 17.9, S.392 und Korollar 17.10, S.393] □

§17.09 **Definition**. Eine zeitlich homogene Markovkette $X = (X_n)_{n\in\mathbb{N}_0}$ mit Familie $(\mathbb{P}_j)_{j\in\mathcal{S}}$ von Wahrscheinlichkeitsmaßen besitzt die *starke Markoveigenschaft*, wenn für jede -f.s. endliche Stoppzeit τ, und jedes $j \in \mathcal{S}$, $\mathbb{P}_j^{\vartheta_\tau(X)|\sigma_\tau(X)} = \mathbb{P}_{X_\tau}^X := \kappa(X_\tau, \bullet)$ oder äquivalant dazu für jede beschränkte Funktion $f : \mathcal{S}^{\mathbb{N}_0} \to \mathbb{R}$ gilt $\mathbb{E}_j\big(f(\vartheta_\tau(X))\big|\sigma_\tau(X)\big) = \mathbb{E}_{X_\tau}\big(f(X)\big) := \int_{\mathcal{S}^{\mathbb{N}_0}} f(s)\kappa(X_\tau, ds)$. □

§17.10 **Lemma**. *Jede zeitlich homogene Markovkette $(X_n)_{n\in\mathbb{N}_0}$ besitzt die starke Markoveigenschaft.*

§17.11 **Beweis** von Lemma §17.10. In der Vorlesung. □

§18 Rekurrenz and Transienz

§18.01 **Definition**. Für $i, j \in \mathcal{S}$, $k \in \mathbb{N}$ definieren wir die *k-te Eintrittzeit von X in j* rekursiv durch $\tau_j^k := \inf\big\{n \in \mathbb{N} : n > \tau_j^{k-1} \wedge X_n = j\big\}$ (wobei $\inf \emptyset = \infty$) und $\tau_j^0 := 0$. Wir setzen $\tau_j := \tau_j^1$ und $\rho_{ij} := \mathbb{P}_i(\tau_j \in \mathbb{N})$. □

§18.02 **Bemerkung**. $\rho_{ij} = \mathbb{P}_i(\text{es gibt ein } k \in \mathbb{N} \text{ mit } X_k = j)$ ist die Wahrscheinlichkeit jemals von i zu j zu gelangen. Insbesondere, ist $\rho_{ij} \in (0,1]$ dann existiert ein $k \in \mathbb{N}$ mit $\mathbb{P}_i(X_k = j) = P_{ij}^k \in (0,1]$. Weiterhin ist ρ_{jj} die Rückkehrwahrscheinlichkeit (nach dem ersten Sprung) von j nach j, wobei $\tau_j \in \overline{\mathbb{N}}$ selbst bei Start in $X_0 = j$ gilt. □

§18.03 **Definition**. Ein Zustand $j \in \mathcal{S}$ heißt *rekurrent*, wenn $\rho_{jj} = 1$,

positiv rekurrent, wenn $\mathbb{E}_j(\tau_j) \in \mathbb{R}^+$,

nullrekurrent, wenn j rekurrent ist aber nicht positiv rekurrent,

transient, wenn $\rho_{jj} \in [0, 1)$, und

absorbierend, wenn $P_{jj} = 1$.

Die Markovkette X heißt (positiv-, null-) rekurrent, wenn jeder Zustand $j \in \mathcal{S}$ (positiv-, null-) rekurrent ist und transient, wenn jeder rekurrente Zustand absorbierend ist. □

§18.04 **Bemerkung**. Offenbar gilt: „absorbierend“ ⇒ „positiv rekurrent“ ⇒ „rekurrent“. □

§18.05 **Lemma**. *Für $k \in \mathbb{N}$ und $i, j \in \mathcal{S}$ gilt $\mathbb{P}_i(\tau_j^k \in \mathbb{N}) = \rho_{ij}\rho_{jj}^{k-1}$.*

§18.06 **Beweis** von Lemma §18.05. In der Vorlesung. □

§18.07 **Definition**. Für $i, j \in \mathcal{S}$ bezeichnet $N_j := \sum_{n\in\mathbb{N}_0} \mathbb{1}_{\{X_n=j\}}$ die *Gesamtanzahl der Besuche* von X im Zustand j und mit $G_{ij} = \mathbb{E}_i(N_j) = \sum_{n\in\mathbb{N}_0} \mathbb{P}_i(X_n = j) = \sum_{n\in\mathbb{N}_0} P_{ij}^n$ die *Greenfunktion* von X. □

§18.08 **Lemma**.

(i) *Ein Zustand $j \in \mathcal{S}$ ist genau dann rekurrent, wenn $G_{jj} = \infty$;*

(ii) *Ist $j \in \mathcal{S}$ ein transienter Zustand, dann gilt für alle $i \in \mathcal{S}$, $G_{ij} \in \mathbb{R}^+$ mit*

$$G_{ij} = \left\{ \begin{array}{ll} \frac{\rho_{ij}}{1-\rho_{jj}}, & \text{wenn } i \neq j \\ \frac{1}{1-\rho_{jj}}, & \text{sonst} \end{array} \right\} = \frac{\rho_{ij}}{1-\rho_{jj}} + \mathbb{1}_{\{i=j\}}.$$

§18.09 **Beweis** von Lemma §18.08. In der Vorlesung. □

§18.10 **Proposition**. *Ist eine Zustand $i \in \mathcal{S}$ rekurrent und $\rho_{ij} \in (0, 1]$ für ein $j \in \mathcal{S}$, dann ist der Zustand j auch rekurrent, und $\rho_{ij} = \rho_{ji} = 1$.*

§18.11 **Beweis** von Proposition §18.10. In der Vorlesung. □

§18.12 **Definition**. Eine Teilmenge $B \subseteq \mathcal{S}$ von Zuständen heißt

abgeschlossen, wenn $\rho_{ij} = 0$ für alle $i \in B$ and $j \in B^c = \mathcal{S}\backslash B$ gilt,

irreduzibel, wenn $\rho_{ij} \in (0, 1]$ für alle $i, j \in B$ gilt.

Ist der Zustandsraum $\mathcal{S}$ ist irreduzibel, dann heißt die Markovkette irreduzibel. □

§18.13 **Korollar**. *Eine irreduzible Markovkette ist entweder rekurrent oder transient. Wenn $|\mathcal{S}| \geqslant 2$, so gibt es keine absorbierenden Zustände.*

§18.14 **Beweis** von Korollar §18.13. Die Aussage folgt direkt aus Proposition §18.10. □

§18.15 **Proposition**. *Für eine irreduzible Markovkette mit endlichem Zustandsraum $\mathcal{S}$ sind alle Zustände rekurrent.*

§18.16 **Beweis** von Proposition §18.15. In der Vorlesung. □

§19 Invariante Verteilung

Im Folgenden sei $(X_n)_{n\in\mathbb{N}_0}$ eine Markovkette auf einem abzählbaren Zustandsraum $\mathcal{S}$ mit Übergangsmatrix $P = (P_{ij})_{i,j\in\mathcal{S}}$.

§19.01 **Definition**. Ist μ ein Maß auf $(\mathcal{S}, 2^{\mathcal{S}})$ und $f : \mathcal{S} \to \mathbb{R}$ eine Abbildung, dann schreiben wir $\mu P(\{j\}) = \sum_{i \in \mathcal{S}} \mu(\{i\}) P_{ij}$ und $Pf(i) = \sum_{j \in \mathcal{S}} P_{ij} f(j)$, falls die Summen konvergieren. □

§19.02 **Definition**.

(a) Ein σ-endliches Maß μ auf $(\mathcal{S}, 2^{\mathcal{S}})$ heißt *invariantes Maß*, falls $\mu P = \mu$. Ein Wahrscheinlichkeitsmaß, das ein invariantes Maß ist, wird *invariante Verteilung* genannt. Wir bezeichnen mit $\mathcal{I}$ die Menge der invarianten Verteilungen.

(b) Eine Funktion $f : \mathcal{S} \to \mathbb{R}$ heißt *subharmonisch* wenn Pf existiert and $f \leqslant Pf$ gilt. f heißt *superharmonisch*, fall $f \geqslant Pf$ gilt and *harmonisch*, falls $f = Pf$. □

§19.03 **Bemerkung**. Im Sinne der linearen Algebra ist ein invariantes Maß ein Links-Eigenvektor und eine harmonische Funktion ein Rechts-Eigenvektor, jeweils zum Eigenwert 1. □

§19.4 **Anmerkung**.

(i) Ist $\mathcal{S}$ endlich, dann kann jedes invariante Maß zu einer invariaten Verteilung normiert werden.

(ii) Jedes invariante Maß μ für P ist auch ein invariantes Maß für P^n für jedes $n \in \mathbb{N}$.

(iii) Ist P irreduzibel und $\mu \neq 0$ ein invariantes Maß für P, so gilt $\mu(\{j\}) \in \mathbb{R}^+_{\setminus 0}$ für jedes $j \in \mathcal{S}$. In der Tat, ist $i \in \mathcal{S}$ mit $\mu(\{j\}) \in \mathbb{R}^+_{\setminus 0}$ (da $\mu \neq 0$) und $j \in \mathcal{S}$ beliebig, so existiert ein $n \in \mathbb{N}$ mit $P^n_{ij} \in (0, 1]$ (da P irreduzibel und somit $\rho_{i,j} \in (0, 1]$) und es gilt $\mu(\{j\}) = \sum_{l \in \mathcal{S}} \mu(\{l\}) P^n_{lj} \geqslant \mu(\{i\}) P^n_{ij} \in \mathbb{R}^+_{\setminus 0}$.

(iv) Ist π eine invariante Verteilung, dann gilt für jedes $n \in \mathbb{N}_0$ und jedes $j \in \mathcal{S}$

$$\mathbb{P}_{\pi}(X_n = j) = \sum_{l \in \mathcal{S}} \pi(\{l\}) P^n_{lj} = \pi(\{j\}),$$

d.h. die mit Verteilung π gestartete Kette hat zu jedem Zeitpunkt die selbe Randverteilung π. □

§19.05 **Lemma**. *Ist f beschränkt und (sub-, super-) harmonisch, dann ist $((f(X_n)))_{n \in \mathbb{N}_0}$ ein (Sub-, Super-) Martingal bezüglich der natürlichen Filtration $\sigma(X)$.*

§19.06 **Beweis** von Lemma §19.05. In der Vorlesung. □

§19.07 **Satz**. *Seien P irreduzibel und rekurrent, $j \in \mathcal{S}$, $\tau_j := \inf\{n \in \mathbb{N} : X_n = j\}$ und*

$$\mu_j(\{i\}) := \mathbb{E}_j\Big(\sum_{n \in [\![\tau_j]\!]} \mathbb{1}_{\{X_n = j\}} \Big) \quad \textit{für alle } i \in \mathcal{S}.$$

Dann gelten die folgenden Aussagen

(i) $\mu_j = \mu_j P$

(ii) $\mu_j(\{i\}) \in \mathbb{R}^+_{\setminus 0}$ *für alle $i \in \mathcal{S}$,*

(iii) μ_j *ist das einzige invariante Maß mit Wert 1 in $\{j\}$.*

§19.08 **Beweis** von Satz §19.07. In der Vorlesung. □

§19.09 **Satz**. *Sei P irreduzibel, dann sind die folgenden drei Aussagen äquivalent:*

(i) *Es existiert eine invariante Verteilung $\pi \in \mathcal{W}(2^{\mathcal{S}})$.*

(ii) *Es gibt einen positiv rekurrenten Zustand in $\mathcal{S}$.*

(iii) *Alle Zustände in $\mathcal{S}$ sind positiv rekurrent.*

Sind diese Bedingungen erfüllt, so ist die invariante Verteilung π eindeutig bestimmt und durch $\pi(\{j\}) = \frac{1}{\mathbb{E}_j(\tau_j)}$ für alle $j \in \mathcal{S}$ gegeben.

§19.10 **Beweis** von Satz §19.09. In der Vorlesung. □

§19.11 **Bemerkung**. Eine irreduzible Markovkette gehört genau zu einem der folgenden drei Fälle:

transient: Die Kette ist transient, d.h. für alle $j \in \mathcal{S}$ gilt $\mathbb{P}_j(\tau_j \in \mathbb{N}) \in [0,1)$, und es gibt kein invariantes Maß.

nullrekurrent: Die Kette ist nullrekurrent, d.h. für alle $j \in \mathcal{S}$ gilt $\mathbb{P}_j(\tau_j \in \mathbb{N}) = 1$ und $\mathbb{E}_j(\tau_j) = \infty$, und es gibt ein invariantes Maß aber keine invariante Verteilung.

positiv rekurrent: Die Kette ist positiv rekurrent, d.h. für alle $j \in \mathcal{S}$ gilt $\mathbb{E}_j(\tau_j) = \mathbb{R}^+$, und es gibt eine invariante Verteilung.

Ist der Zustandsraum endlich, so tritt nur der positiv rekurrente Fall auf. □

Wir untersuchen, unter welchen Bedingungen eine positiv rekurrente Markovkette $(X_n)_{n\in\mathbb{N}_0}$ mit abzählbarem Zustandsraum $\mathcal{S}$ (und mit Übergangsmatrix P), die in einem beliebigen Punkt $i \in \mathcal{S}$ gestartet wird, bei divergierender Laufzeit gegen ihre invariante Verteilung $\pi \in \mathcal{W}(2^{\mathcal{S}})$ konvergiert, d.h. $\mathbb{P}_i(X_n = j)$ für jedes $j \in \mathcal{S}$ gegen $\pi(\{j\})$ für $n \to \infty$ konvergiert.

§19.12 **Schreibweise**. Den größten gemeinsamen Teiler aller $n \in M \subseteq \mathbb{N}$ bezeichnet $\mathrm{ggT}(M)$. □

§19.13 **Definition**. Für jedes $j \in \mathcal{S}$ heißt $d_j := \mathrm{ggT}(\{n \in \mathbb{N} : P^n_{ij} \in (0,1]\})$ die *Periode* des Zustandes j. Ist $d_j = 1$ für jedes $j \in \mathcal{S}$, so heißt X *aperiodisch*. □

§19.14 **Lemma**. *Ist P irreduzibel, dann gilt $d_j = d_i$ für alle $i,j \in \mathcal{S}$. Ist P zuätzlich aperiodisch, dann existiert für jede $i,j \in \mathcal{S}$ ein $n_{ij} \in \mathbb{N}$ so dass für alle $n \in \mathbb{N} \cap [n_{ij}, \infty)$ gilt $P^n_{ij} \in (0,1]$.*

§19.15 **Beweis** von Lemma §19.14. Klenke [2020, Lemma 18.3, S.433] □

Wir benutzen ein *Kopplungsargument*. Sei $(Y_n)_{n\in\mathbb{N}_0}$ einer weitere von $(X_n)_{n\in\mathbb{N}_0}$ unabhängige Markovkette mit der selben Übergangsmatrix P, die allerdings mit der invarianten Verteilung π gestartet wird. Damit ist die invarianten Verteilung π gerade die Randverteilung von Y_n zu jedem Zeitpunkt $n \in \mathbb{N}_0$.

§19.16 **Lemma**. *Seien $(X_n)_{n\in\mathbb{N}_0}$ und $(Y_n)_{n\in\mathbb{N}_0}$ zwei unabhängige Markovketten auf $\mathcal{S}$ mit irreduzibler, positiv rekurrenter, aperiodischer Übergangsmatrix P und invariante Verteilung $\pi \in \mathcal{W}(2^{\mathcal{S}})$. Die Übergangsmatrix $\widetilde{P} = (\widetilde{P}_{(x_1,y_1)(x_2,y_2)} := P_{x_1x_2}P_{y_1y_2})_{(x_1,y_1),(x_2,y_2)\in\mathcal{S}^2}$ der bivariaten Markovkette $(X_n, Y_n)_{n\in\mathbb{N}_0}$ ist dann irreduzibel, positiv rekurrent und aperiodisch. Das Produktmaß $\pi \otimes \pi$ ist die invariante Verteilung von $\widetilde{P}$. Insbesondere ist die Treffzeit $\tau = \inf\{n \in \mathbb{N}_0 : X_n = Y_n\}$ $\mathbb{P}_{\delta_\omega i \otimes \pi}$-f.s. endlich.*

§19.17 **Beweis** von Lemma §19.16. In der Vorlesung. □

§19.18 **Satz**. *Sei $(X_n)_{n\in\mathbb{N}_0}$ eine Markovkette auf $\mathcal{S}$ mit irreduzibler, positiv rekurrenter, aperiodischer Übergangsmatrix P und invarianter Verteilung $\pi \in \mathcal{W}(2^{\mathcal{S}})$. Für alle $i,j \in \mathcal{S}$ gilt dann*

$$\lim_{n\to\infty} \mathbb{P}_i(X_n = j) = \pi(\{j\}).$$

§19.19 **Beweis** von Satz §19.18. In der Vorlesung. □

Anhang A

Grundlagen der Wahrscheinlichkeitstheorie

In diesem Kapitel werden Begriffe, Notationen und Aussagen der Vorlesung Einführung in die Wahrscheinlichickeitstheorie und Statistik (EWS) wiederholt.

A01 Wahrscheinlichkeitsraum

A01.01 **Definition.**

(a) Eine nicht-leere Menge Ω aller möglichen Ausgänge eines zufälligen Experiments wird *Ergebnismenge* (*Grundmenge* oder *Stichprobenraum*) genannt. Ein möglicher Versuchsausgang ω des zufälligen Experiments, also ein Element von Ω, kurz $\omega \in \Omega$ heißt *Ergebnis* (*Stichprobe*).

(b) Ein *Ereignis* ist eine Teilmenge der Grundmenge Ω, also ein Element der Potenzmenge 2^Ω von Ω. Für einen Versuchsausgang $\omega \in \Omega$ ist ein Ereignis $A \in 2^\Omega$ eingetreten, wenn $\omega \in A$.

(c) Ein Paar $\boxed{(\Omega, \mathscr{A})}$ wird *messbarer Raum* genannt, wenn die *Menge der interessierenden Ereignisse* $\mathscr{A} \subseteq 2^\Omega$ eine *σ-Algebra* ist, also falls gilt: (i) $\Omega \in \mathscr{A}$; (ii) für alle $A \in \mathscr{A}$ gilt $A^c := \Omega \setminus A \in \mathscr{A}$; und (iii) für alle $A_n \in \mathscr{A}$ mit $n \in \mathbb{N}$ gilt $\left(\bigcup_{n\in\mathbb{N}} A_n\right) \in \mathscr{A}$. □

A01.02 **Beispiel.**

(a) Auf jeder nicht-leeren Ergebnismenge Ω existieren die *triviale σ-Algebra* $\{\emptyset, \Omega\}$ sowie die Potenzmenge 2^Ω als σ-Algebren.

(b) Sei $\mathscr{E} \subseteq 2^\Omega$ ein System von Teilmengen von Ω. Dann heißt die kleinste σ-Algebra auf Ω, die $\mathscr{E}$ enthält, $\boxed{\sigma(\mathscr{E})} := \bigcap\{\mathscr{A} \mid \mathscr{A} \text{ ist } \sigma\text{-Algebra auf } \Omega \text{ und } \mathscr{E} \subseteq \mathscr{A}\}$, die von $\mathscr{E}$ *erzeugte σ-Algebra* auf Ω. $\mathscr{E}$ heißt *Erzeuger* von $\sigma(\mathscr{E})$.

(c) Für jedes nicht-leere $A \subseteq \Omega$ gilt $\sigma(\{A\}) = \{\emptyset, \Omega, A, A^c\}$.

(d) Für eine höchstens abzählbar unendliche (d.h. endlich oder abzählbar unendliche), kurz *abzählbare*, Indexmenge $\mathcal{I}$ sei $\mathscr{E} = \{A_i \in 2^\Omega \setminus \{\emptyset\}, i \in \mathcal{I}, \text{ paarweise disjunkt und } \biguplus_{i\in\mathcal{I}} A_i = \Omega\}$ eine Partition von Ω. Dann ist $\sigma(\mathscr{E}) = \{\biguplus_{j\in\mathcal{J}} A_j \mid \mathcal{J} \in 2^{\mathcal{I}}\}$.

(e) Es sei $\mathcal{S}$ ein metrischer (oder topologischer) Raum und $\mathscr{O}$ das System der offenen Teilmengen von $\mathcal{S}$. Dann heißt die von den offen Mengen $\mathscr{O}$ erzeugte σ-Algebra $\mathscr{B}_\mathcal{S} := \sigma(\mathscr{O})$ die *Borel-σ-Algebra* über $\mathcal{S}$. Die Elemente von $\mathscr{B}_\mathcal{S}$ heißen *Borel-Mengen*.

(f) $\boxed{\mathscr{B}^n} := \mathscr{B}_{\mathbb{R}^n}$ bezeichnet die Borel-σ-Algebra über $\mathbb{R}^n$ versehen mit dem Euklidischen Abstand $d(x,y) = \|x-y\| := \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$ für $x = (x_i)_{i\in[\![n]\!]}$, $y = (y_i)_{i\in[\![n]\!]} \in \mathbb{R}^n$ mit $[\![n]\!] := [1, n] \cap \mathbb{N}$.

(g) Seien $(\Omega, \mathscr{A})$ ein messbarer Raum und $B \subseteq \Omega$ eine nicht-leere Teilmenge. Dann heißt die σ-Algebra über B, $\mathscr{A}\big|_B := \mathscr{A} \cap B := \{A \cap B \mid A \in \mathscr{A}\}$ *Spur* von $\mathscr{A}$ über B, wobei für $\mathscr{E} \in 2^\Omega$ gilt $\sigma(\mathscr{E})\big|_B = \sigma(\mathscr{E}\big|_B)$. □

A01.03 **Schreibweise.**

(i) Wir setzen $\mathbb{R}^+ := [0,\infty)$, $\mathbb{R}^+_{\setminus 0} := (0,\infty)$, $\mathbb{R}_{\setminus 0} := \mathbb{R}\setminus\{0\}$, $\overline{\mathbb{R}} := [-\infty,\infty]$, $\overline{\mathbb{R}}^+ := [0,\infty]$.

(ii) Für $a,b \in \mathbb{R}^n$ schreiben wir $a < b$, wenn $a_i < b_i$ für alle $i \in [\![n]\!]$ gilt. Für $a < b$, definieren wir den offenen *Quader* als das Kartesische Produkt $(a,b) := \times_{i=1}^n (a_i,b_i) := (a_1,b_1)\times(a_2,b_2)\times\cdots\times(a_n,b_n)$. Analog, sind $[a,b]$, $(a,b]$ sowie $[a,b)$ definiert. Weiterhin, sei $(-\infty,b) := \times_{i=1}^n(-\infty,b_i)$ und analog $(-\infty,b]$ definiert.

(iii) Wir bezeichnen mit $\overline{\mathscr{B}} := \mathscr{B}_{\overline{\mathbb{R}}}$ die Borel-σ-Algebra über der kompaktifizierten Zahlengerade $\overline{\mathbb{R}}$, wobei die Mengen $\{-\infty\}$, $\{\infty\}$ und $\mathbb{R}$ in $\overline{\mathbb{R}}$ abgeschlossen bzw. offen, also Borel-Mengen sind. Insbesondere, ist $\mathscr{B} := \mathscr{B}_{\mathbb{R}} = \overline{\mathscr{B}} \cap \mathbb{R}$ die Borel-σ-Algebra über $\mathbb{R}$. Weiterhin schreiben wir $\overline{\mathscr{B}}^+ := \overline{\mathscr{B}} \cap \overline{\mathbb{R}}^+$, $\mathscr{B}^+ := \mathscr{B}\cap\mathbb{R}^+$, $\mathscr{B}^+_{\setminus 0} := \mathscr{B}\cap\mathbb{R}^+_{\setminus 0}$ und $[0,1]^+_{\setminus 0} := \mathscr{B}\cap[0,1]$.

(iv) Ein Folge $(x_n)_{n\in\mathbb{N}}$ aus $\mathbb{R}$ heißt *monoton wachsend* (bzw. *fallend*), wenn $x_n \leqslant x_{n+1}$ (bzw. $x_{n+1} \leqslant x_n$) für alle $n \in \mathbb{N}$ gilt. Ist eine monotone wachsende (bzw. fallende) Folge konvergent, etwa $x = \lim_{n\to\infty} x_n$, so schreiben wir kurz $x_n \uparrow x$ (bzw. $x_n \downarrow x$).

(v) Sei $A_n \subseteq \Omega$ für $n \in \mathbb{N}$. Die Folge $(A_n)_{n\in\mathbb{N}}$ heißt *monoton wachsend* (bzw. *fallend*), wenn $A_n \subseteq A_{n+1}$ (bzw. $A_{n+1} \subseteq A_n$) für alle $n \in \mathbb{N}$ gilt. Weiterhin heißen $\boxed{A_\star} := \liminf_{n\to\infty} A_n := \bigcup_{n\geqslant 1}\bigcap_{m\geqslant n} A_m$ und $\boxed{A^\star} := \limsup_{n\to\infty} A_n := \bigcap_{n\geqslant 1}\bigcup_{m\geqslant n} A_m$ *Limes inferior* bzw. *Limes superior* der Folge $(A_n)_{n\in\mathbb{N}}$. Die Folge $(A_n)_{n\in\mathbb{N}}$ heißt *konvergent*, wenn $A_\star = A^\star =: A$ gilt. In diesem Fall schreiben wir kurz $\lim_{n\to\infty} A_n = A$. □

(vi) Jede monoton wachsende (bzw. fallende) Folge $(A_n)_{n\in\mathbb{N}}$ von Teilmengen von Ω ist konvergent mit $A := \lim_{n\to\infty} A_n = \bigcup_{n\in\mathbb{N}} A_n$ (bzw. $A := \lim_{n\to\infty} A_n = \bigcap_{n\in\mathbb{N}} A_n$). In diesem Fall schreiben wir kurz $A_n \uparrow A$ (bzw. $A_n \downarrow A$). □.

A01.04 **Definition.** Sei $(\Omega,\mathscr{A})$ ein messbarer Raum.

(a) Eine Abbildung $\boxed{\mathbb{P}} : \mathscr{A} \to [0,1]$ heißt *Wahrscheinlichkeitsmaß* (kurz *Verteilung*) auf $\mathscr{A}$, wenn sie folgenden Bedingungen genügt: (i) $\mathbb{P}(\Omega) = 1$ *(normiert)*; und (ii) $\mathbb{P}(\biguplus_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} \mathbb{P}(A_n)$ für jede Folge $(A_n)_{n\in\mathbb{N}}$ paarweise disjunkter Ereignisse aus $\mathscr{A}$, d.h. $A_n\cap A_m = \emptyset$ für alle $n,m\in\mathbb{N}$ mit $n \neq m$ *(σ-additiv)*.

(b) Wir bezeichnen mit $\mathcal{W} := \mathcal{W}(\mathscr{A})$ die Menge aller Wahrscheinlichkeitsmaße auf $\mathscr{A}$. Ein Tripel $\boxed{(\Omega,\mathscr{A},\mathbb{P})}$ bestehend aus einer Ergebnismenge Ω, einer σ-Algebra $\mathscr{A}$ interessierender Ereignisse sowie einem Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $\mathscr{A}$ wird *Wahrscheinlichkeitsraum* genannt.

(c) Für ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\Omega,\mathscr{A})$ heißt ein Element $\omega\in\Omega$ mit $\{\omega\}\in\mathscr{A}$ und $\mathbb{P}(\{\omega\}) > 0$ *Atom*. Die Wahrscheinlichkeit $\mathbb{P}(\{\omega\})$ wird *Masse* des Atoms ω genannt.

(d) Für ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\mathbb{R},\mathscr{B})$ wird $\boxed{\mathbb{F}} : \mathbb{R}\to[0,1]$ mit $x \mapsto \mathbb{F}(x) := \mathbb{P}((-\infty,x])$ die zugehörige *Verteilungsfunktion* genannt.

(e) Wenn Ω eine abzählbare Menge ist, dann wird ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $\mathscr{A} = 2^\Omega$ *diskret* genannt und $(\Omega,\mathscr{A},\mathbb{P})$ heißt *diskreter Wahrscheinlichkeitsraum*. Die Abbildung $\boxed{\mathbb{p}} : \Omega\to[0,1]$ mit $\omega\mapsto\mathbb{p}(\omega) := \mathbb{P}(\{\omega\})$ wird *Zähldichte* des Wahrscheinlichkeitsmaßes $\mathbb{P}$ genannt.

(f) Ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\mathbb{R}^n,\mathscr{B}^n)$ heißt *stetig*, wenn eine *Wahrscheinlichkeitsdichte* (kurz *Dichte*) $\boxed{\mathbb{f}}$ auf $\mathbb{R}^n$, also eine *Lebesgue-integrierbare Funktion* $\mathbb{f} : \mathbb{R}^n \to \mathbb{R}^+$ mit $\mathbb{P}(B) = \int_B \mathbb{f}(x)dx$ für alle $B\in\mathscr{B}^n$ existiert. $(\mathbb{R}^n,\mathscr{B}^n,\mathbb{P})$ wird dann *stetiger Wahrscheinlichkeitsraum* genannt. □

A01.05 **Eigenschaft.**

(i) *Ist* $(\mathbb{P}_n)_{n\in\mathbb{N}}$ *eine Folge von Wahrscheinlichkeitsmaße in* $\mathcal{W}$*, so ist auch jede Konvexkombination* $\sum_{n\in\mathbb{N}} w_n\mathbb{P}_n$ *mit* $w_n \geqslant 0$*,* $n \in \mathbb{N}$*, und* $\sum_{n\in\mathbb{N}} w_n = 1$ *in* $\mathcal{W}$*. Die Diracmaße bilden Extremalpunkte der konvexen Menge aller Wahrscheinlichkeitsmaße.*

(ii) *Ein Wahrscheinlichkeitsmaß* $\mathbb{P}$ *auf* $(\Omega, \mathscr{A})$ *besitzt abzählbar viele Atome.*

(iii) *Ist eine Abbildung* $\mathbb{F} : \mathbb{R} \to [0, 1]$ *monoton wachsend, rechtsstetig mit* $\lim_{x\to-\infty} \mathbb{F}(x) = 0$ *und* $\lim_{x\to\infty} \mathbb{F}(x) = 1$*, so existiert genau ein Wahrscheinlichkeitsmaß* $\mathbb{P}$ *auf* $(\mathbb{R}, \mathscr{B})$ *mit* $\mathbb{P}((x, y]) = \mathbb{F}(y) - \mathbb{F}(x)$ *für alle* $x, y \in \mathbb{R}$ *mit* $x < y$*. Insbesondere ist* $\mathbb{F}$ *die Verteilungsfunktion von* $\mathbb{P}$*.* □

A01.06 **Beispiel.**

(a) Sei $(\Omega, \mathscr{A})$ ein messbarer Raum. Für $A \subseteq \Omega$ bezeichnet $\boxed{\mathbb{1}_A} : \Omega \to \{0, 1\}$ mit $\mathbb{1}_A^{-1}(\{1\}) = A$ und $\mathbb{1}_A^{-1}(\{0\}) = A^c$ die *Indikatorfunktion* auf A. Für jedes $\omega \in \Omega$ ist das *Einpunkt-* oder *Diracmaß* $\boxed{\delta_\omega} : \mathscr{A} \to \{0, 1\}$ mit $\delta_\omega(A) := \mathbb{1}_A(\omega)$ für alle $A \in \mathscr{A}$ ein Wahrscheinlichkeitsmaß aus $\mathcal{W}(\mathscr{A})$.

(b) Sei $\emptyset \neq \Omega \subseteq \mathbb{R}$ abzählbar und $\mathbb{p}$ eine Zähldichte auf Ω. Das Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\mathbb{R}, \mathscr{B})$ mit $\mathbb{P}(B) := \sum_{\omega\in\Omega} \mathbb{p}(\omega)\delta_\omega$ für $B \in \mathscr{B}$ und die zugehörige Verteilungsfunktion $\mathbb{F} : \mathbb{R} \to [0, 1]$ mit $\mathbb{F}(x)= \mathbb{P}\big((-\infty, x]\big) = \sum_{\omega\in\Omega} \mathbb{p}(\omega)\delta_\omega\big((-\infty, x]\big) =: \sum_{\omega\leqslant x} \mathbb{p}(\omega)$ für $x \in \mathbb{R}$ werden *diskret* genannt.

(c) *Laplace-/Gleich-Verteilung*, kurz Lap_Ω, mit $|\Omega| < \infty$ und $\mathbb{p}_{\mathrm{Lap}_\Omega}(\omega) = \frac{1}{|\Omega|}$ für $\omega \in \Omega$.

(d) *Bernoulli-Schema*, kurz B_p^n, mit Erfolgswahrscheinlichkeit $p \in [0, 1]$, Länge $n \in \mathbb{N}$, $\Omega = \{0, 1\}^n$ und $\mathbb{p}_{\mathrm{B}_p^n}(\omega) = p^{\sum_{i=1}^n \omega_i}(1 - p)^{n-\sum_{i=1}^n \omega_i}$ für $\omega = (\omega_i)_{i\in[\![n]\!]} \in \{0, 1\}^n$.
Bernoulli-Verteilung. Im Fall $n = 1$, schreiben wir kurz B_p, also $\mathbb{p}_{\mathrm{B}_p}(\omega) = p^\omega(1 - p)^{1-\omega}$ für $\omega \in \{0, 1\}$.

(e) *Binomial-Verteilung*, kurz $\mathrm{Bin}_{(n,p)}$, mit Erfolgswahrscheinlichkeit $p \in [0, 1]$, Länge $n \in \mathbb{N}$, $\Omega := [\![0, n]\!]$ und $\mathbb{p}_{\mathrm{Bin}_{(n,p)}}(\omega) = \binom{n}{\omega}p^\omega(1 - p)^{n-\omega}$ für $\omega \in [\![0, n]\!]$.

(f) *Poissonverteilung*, kurz Poi_λ, mit Parameter $\lambda \in \mathbb{R}^+_{\setminus 0}$, $\Omega = \mathbb{N}_0$ und $\mathbb{p}_{\mathrm{Poi}_\lambda}(\omega) = \exp(-\lambda)\frac{\lambda^\omega}{\omega!}$ für $\omega \in \mathbb{N}_0$.

(g) *Gleich-/Uniformverteilung*, kurz U_G, auf $G \in \mathscr{B}$ mit Lebesgue-Maß $\lambda(G) \in \mathbb{R}^+_{\setminus 0}$ und $\mathbb{f}_{\mathrm{U}_G}(x) = \frac{1}{\lambda(G)}\mathbb{1}_G(x)$ für $x \in \mathbb{R}$.

(h) *Exponentialverteilung*, kurz Exp_λ, mit $\lambda \in \mathbb{R}^+_{\setminus 0}$ und $\mathbb{f}_{\mathrm{Exp}_\lambda}(x) = \lambda\exp(-\lambda x)\mathbb{1}_{\mathbb{R}^+}(x)$ für $x \in \mathbb{R}$.

(i) *Normalverteilung*, kurz $\mathrm{N}_{(\mu,\sigma^2)}$, mit Parametern $\mu \in \mathbb{R}$ und $\sigma^2 \in \mathbb{R}^+_{\setminus 0}$ sowie $\mathbb{f}_{\mathrm{N}_{(\mu,\sigma^2)}}(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{1}{2\sigma^2}(x - \mu)^2)$ für $x \in \mathbb{R}$. Weiterhin vereinbaren wir $\mathrm{N}_{(\mu,0)} := \delta_\mu$. □

A01.07 **Schreibweise.** Sei $((\mathcal{S}_i, \mathscr{S}_i))_{i\in\mathcal{I}}$ eine Familie messbarer Räume mit beliebiger, nicht-leerer Indexmenge $\mathcal{I}$. Die Menge $\mathcal{S}_\mathcal{I} := \bigtimes_{i\in\mathcal{I}} \mathcal{S}_i$ aller Abbildungen $(s_i)_{i\in\mathcal{I}} : \mathcal{I} \to \cup_{i\in\mathcal{I}}\mathcal{S}_i$ sodass $s_i \in \mathcal{S}_i$ für alle $i \in \mathcal{I}$ ist, heißt *Produktraum* oder kartesisches Produkt. Sind alle $\mathcal{S}_i$ gleich, etwa $\mathcal{S}_i = \mathcal{S}$, dann schreiben wir $\mathcal{S}^\mathcal{I} := \mathcal{S}_\mathcal{I}$, im Fall $n := |\mathcal{I}| < \infty$, auch nur kurz $\mathcal{S}^n := \mathcal{S}^\mathcal{I}$. Für jedes $j \in \mathcal{I}$ bezeichne $\Pi_{\mathcal{S}_j}: \mathcal{S}_\mathcal{I} \to \mathcal{S}_j$ mit $(s_i)_{i\in\mathcal{I}} \mapsto s_j$ die Koordinantenabbildung. □

A01.08 **Definition.**

(a) Für eine Familie $(\mathscr{A}_i)_{i\in\mathcal{I}}$ von Teil-σ-Algebren von $\mathscr{A}$ mit beliebiger nicht-leerer Indexmenge $\mathcal{I}$ bezeichnet $\bigwedge_{i\in\mathcal{I}}\mathscr{A}_i := \bigcap_{i\in\mathcal{I}}\mathscr{A}_i$ und $\bigvee_{i\in\mathcal{I}}\mathscr{A}_i := \sigma(\bigcup_{i\in\mathcal{I}}\mathscr{A}_i)$ die *größte σ-Algebra*, die in allen $\mathscr{A}_i$, $i\in\mathcal{I}$, enthalten ist, bzw. die *kleinste σ-Algebra*, die alle $\mathscr{A}_i$, $i\in\mathcal{I}$, enthält.

(b) Die Produkt-σ-Algebra $\mathscr{S}_\mathcal{I} := \bigotimes_{i\in\mathcal{I}}\mathscr{S}_i$ auf dem Produktraum $\mathcal{S}_\mathcal{I}$ ist die kleinste σ-Algebra, sodass für jedes $i\in\mathcal{I}$ die Koordinantenabbildung $\Pi_{\mathcal{S}_i}$ $\mathscr{S}_\mathcal{I}$-$\mathscr{S}_i$-messbar ist, d.h. $\mathscr{S}_\mathcal{I} = \bigotimes_{i\in\mathcal{I}}\mathscr{S}_i := \bigvee_{i\in\mathcal{I}}\sigma(\Pi_{\mathcal{S}_i}) = \bigvee_{j\in\mathcal{I}}\Pi_{\mathcal{S}_j}^{-1}(\mathscr{S}_j)$. Sind alle $(\mathcal{S}_i,\mathscr{S}_i)$ gleich, etwa $(\mathcal{S}_i,\mathscr{S}_i) = (\mathcal{S},\mathscr{S})$, dann schreiben wir $\mathscr{S}^\mathcal{I} := \mathscr{S}_\mathcal{I}$, im Fall $n := |\mathcal{I}| < \infty$, auch nur $\mathscr{S}^n := \mathscr{S}^\mathcal{I}$.

(c) Ist für jedes $i\in\mathcal{I}$ weiterhin $\mathbb{P}_i$ ein Wahrscheinlichkeitsmaß auf $(\mathcal{S}_i,\mathscr{S}_i)$, dann heißt ein Wahrscheinlichkeitsmaß $\mathbb{P}_\mathcal{I}$ auf $\bigotimes_{i\in\mathcal{I}}\mathscr{S}_i$ *Produktmaß*, wenn für alle endlichen $\mathcal{J}\subseteq\mathcal{I}$ und $S_j\in\mathscr{S}_j$, $j\in\mathcal{J}$ gilt $\mathbb{P}_\mathcal{I}\big(\bigcap_{j\in\mathcal{J}}\Pi_{\mathcal{S}_j}^{-1}(S_j)\big) = \prod_{j\in\mathcal{J}}\mathbb{P}_j(S_j)$. In dem Fall schreiben wir $\mathbb{P}_\mathcal{I} = \bigotimes_{i\in\mathcal{I}}\mathbb{P}_i$. Sind alle $\mathbb{P}_i$ gleich, etwa $\mathbb{P}_i = \mathbb{P}$, dann schreiben wir kurz $\mathbb{P}^\mathcal{I} := \mathbb{P}_\mathcal{I}$ und im Fall $n := |\mathcal{I}| < \infty$ auch $\mathbb{P}^n := \mathbb{P}_\mathcal{I}$. □

A01.09 **Eigenschaft.**

(i) *Sei $\mathcal{I}$ abzählbar, für jedes $i\in\mathcal{I}$ sei $\mathcal{S}_i$ separabler und vollständiger metrischer Raum (polnisch) mit Borel-σ-Algebra $\mathscr{B}_i$ und sei $\mathscr{B}_{\mathcal{S}_\mathcal{I}}$ die Borel-σ-Algebra bzgl. der Produkttopologie auf $\mathcal{S}_\mathcal{I} = \bigtimes_{i\in\mathcal{I}}\mathcal{S}_i$. Dann gilt $\mathscr{B}_{\mathcal{S}_\mathcal{I}} = \mathscr{B}_\mathcal{I} = \bigotimes_{i\in\mathcal{I}}\mathscr{B}_i$, also insbesondere $\mathscr{B}_{\mathbb{R}^n} = \mathscr{B}^n$ (vgl.* **?** *Satz 14.8).*

(ii) *Seien $\mathbb{P}_1,\dots,\mathbb{P}_n$ diskrete Wahrscheinlichkeitsmaße auf Ω mit Zähldichten $\mathbb{p}_1,\dots,\mathbb{p}_n$. Die Produktzähldichte $\prod_{i=1}^n\mathbb{p}_i(\omega_i)$ für $(\omega_i)_{i\in[\![n]\!]}\in\Omega^n$ ist die Zähldichte des Produktmaßes $\mathbb{P}_{[\![n]\!]} = \bigotimes_{i\in[\![n]\!]}\mathbb{P}_i$ auf $(\Omega^n, 2^{\Omega^n})$.*

(iii) *Seien $\mathbb{P}_1,\dots,\mathbb{P}_n$ stetige Wahrscheinlichkeitsmaße auf $\mathbb{R}$ mit Wahrscheinlichkeitsdichten $\mathbb{f}_1,\dots,\mathbb{f}_n$. Die Produktdichte $\prod_{i=1}^n\mathbb{f}_i(x_i)$ für $(x_i)_{i\in[\![n]\!]}\in\mathbb{R}^n$ ist Dichte des Produktmaßes $\mathbb{P}_{[\![n]\!]} = \bigotimes_{i\in[\![n]\!]}\mathbb{P}_i$ auf $(\mathbb{R}^n,\mathscr{B}^n)$.* □

A02 Zufallsvariablen

Im Folgenden seien $(\Omega,\mathscr{A},\mathbb{P})$ ein Wahrscheinlichkeitsraum, $(\mathcal{S},\mathscr{S})$ ein messbarer Raum und $((\mathcal{S}_i,\mathscr{S}_i))_{i\in\mathcal{I}}$ eine Familie messbarer Räume mit beliebiger nicht-leerer Indexmenge $\mathcal{I}$.

A02.01 **Definition.**

(a) Eine Funktion $X:\Omega\to\mathcal{S}$ heißt *$\mathscr{A}$-$\mathscr{S}$-messbar* (kurz *messbar*), falls $\sigma(X) := X^{-1}(\mathscr{S}) := \{X^{-1}(S)\,|\,S\in\mathscr{S}\}\subseteq\mathscr{A}$ gilt. Jede solche messbare Funktion wird ($(\mathcal{S},\mathscr{S})$-wertige) *Zufallsvariable* genannt. $\sigma(X)$ wird die *von X erzeugte σ-Algebra* genannt und ist die kleinste Teil-σ-Algebra aus 2^Ω, so dass f messbar ist.

(b) Das Bildmaß $\mathbb{P}^X := \mathbb{P}\circ X^{-1}$ von $\mathbb{P}$ unter X auf $(\mathcal{S},\mathscr{S})$ wird die Verteilung von X, kurz $X\sim\mathbb{P}^X$, genannt. Mit der Verteilungsfunktion $\mathbb{F}^X$ (Dichte $\mathbb{f}^X$, Zähldichte $\mathbb{p}^X$) von X werden wir stets die zu $\mathbb{P}^X$ gehörigen Größen bezeichnen. X heißt *diskret-verteilte Zufallsvariable*, wenn $\mathbb{P}^X$ ein diskretes Wahrscheinlichkeitsmaß auf $(\mathcal{S},\mathscr{S})$ ist. Eine reelle Zufallsvariable (Zufallsvektor) X mit stetigem Bildmaß $\mathbb{P}^X$ wird *stetig-verteilt* genannt.

(c) $(\mathcal{S},\mathscr{S})$-wertige Zufallsvariablen X_i, $i\in\mathcal{I}$ heißen *identisch verteilt* (kurz *i.v.*), wenn für alle $i\in\mathcal{I}$ gilt $\mathbb{P}^{X_i} = \mathbb{P}$ für ein Wahrscheinlichkeitsmaß $\mathbb{P}\in\mathcal{W}(\mathscr{S})$.

(d) Das Bildmaß $\mathbb{P}^X := \mathbb{P}\circ X^{-1}$ unter $X := (X_i)_{i\in\mathcal{I}}$ auf $(\mathcal{S}_\mathcal{I},\mathscr{S}_\mathcal{I})$ heißt *gemeinsame Verteilung* der Familie $(X_i)_{i\in\mathcal{I}}$. Für jedes $i\in\mathcal{I}$ wird das Bildmaß $\mathbb{P}^{X_i} := \mathbb{P}\circ X_i^{-1} = \mathbb{P}\circ(\Pi_i\circ X)^{-1} = \mathbb{P}^X\circ\Pi_i^{-1}$ *Randverteilung* (*marginale Verteilung*) von X_i bzgl. $\mathbb{P}^{X_i}$ genannt. □

A02.02 **Eigenschaft**. *Sei $X : \Omega \to \mathcal{S}$ eine Abbildung.*

(i) *Für jedes Teilmengensystem $\mathscr{E}$ aus $2^{\mathcal{S}}$ gilt $\sigma\big(X^{-1}(\mathscr{E})\big) = X^{-1}\big(\sigma(\mathscr{E})\big)$.*

(ii) *Falls $X^{-1}(\mathscr{E}) \subseteq \mathscr{A}$ für einen Erzeuger $\mathscr{E}$ von $\mathscr{S}$, also $\mathscr{S} = \sigma(\mathscr{E})$, gilt, so ist X eine $\mathscr{A}$-$\mathscr{S}$-messbare Funktion, also eine $(\mathcal{S}, \mathscr{S})$-wertige Zufallsvariable.*

(iii) *Für Abbildungen $X_i : \Omega \to \mathcal{S}_i$, $i \in \mathcal{I}$ ist $\sigma((X_i)_{i\in\mathcal{I}}) := \bigvee_{i\in\mathcal{I}} \sigma(X_i) = \sigma(\cup_{i\in\mathcal{I}} X_i^{-1}(\mathscr{S}_i))$ die kleinste Teil-σ-Algebra, so dass alle X_i, $i \in \mathcal{I}$, messbar sind.*

(iv) *Jede stetige Funktion $g : \mathcal{S} \to \mathcal{T}$ zwischen metrischen Räumen S und T ist $\mathscr{B}_{\mathcal{S}}$-$\mathscr{B}_{\mathcal{T}}$-messbar, kurz* Borel-messbar. □

A02.03 **Definition**. Sei $X : (\Omega, \mathscr{A}) \to (\mathcal{S}, \mathscr{S})$ eine Zufallsvariable.

(a) Falls $(\mathcal{S}, \mathscr{S}) = (\overline{\mathbb{R}}, \overline{\mathscr{B}})$, so heißt X *numerische Zufallsvariable*, kurz $X \in \overline{\mathscr{A}}$.
Falls $(\mathcal{S}, \mathscr{S}) = (\overline{\mathbb{R}}^+, \overline{\mathscr{B}}^+)$, so heißt X *positive numerische Zufallsvariable*, kurz $X \in \overline{\mathscr{A}}^+$.

(b) Falls $(\mathcal{S}, \mathscr{S}) = (\mathbb{R}, \mathscr{B})$, so heißt X *reelle Zufallsvariable*, kurz $X \in \mathscr{A}$.
Falls $(\mathcal{S}, \mathscr{S}) = (\mathbb{R}^+, \mathscr{B}^+)$, so heißt X *positive reelle Zufallsvariable*, kurz $X \in \mathscr{A}^+$.

(c) Falls $(\mathcal{S}, \mathscr{S}) = (\mathbb{R}^n, \mathscr{B}^n)$, so heißt X *Zufallsvektor*, kurz $X \in \mathscr{A}^n$. □

A02.04 **Beispiel**.

(a) Seien $(\Omega, \mathscr{A})$ ein messbarer Raum und $\mathbb{1}_A$ mit $A \subseteq \Omega$ die Indikatorfunktion. Dann gilt $\mathbb{1}_A \in \mathscr{A}$, d.h. $\mathbb{1}_A$ ist genau dann eine reelle Zufallsvariable, wenn $A \in \mathscr{A}$ gilt.

(b) Für jedes $n \in \mathbb{N}$ ist die identische Abbildung $\mathrm{id}_{\mathbb{R}^n} : \mathbb{R}^n \to \mathbb{R}^n$ mit $x \mapsto \mathrm{id}_{\mathbb{R}^n}(x) := x$ ist ein Zufallsvektor auf $(\mathbb{R}^n, \mathscr{B}^n)$, also $\mathrm{id}_{\mathbb{R}^n} \in \mathscr{B}^n$.

(c) Eine reelle Zufallsvariable $X \in \mathscr{A}$ heißt *einfach* oder *elementar*, falls sie nur endlich viele reelle Werte annimmt, d.h. für $X(\Omega) = \{x_i, i \in [\![n]\!]\} \subseteq \mathbb{R}$ mit $n \in \mathbb{N}$ besitzt X eine Darstellung der Form

$$X = \sum_{i\in[\![n]\!]} x_i \mathbb{1}_{A_i} \quad \text{mit } A_i := X^{-1}(\{x_i\}) = \big\{X = x_i\big\} \in \mathscr{A}.$$

(d) Sei $(\mathbb{R}^n, \mathscr{B}^n, \mathbb{P})$ ein stetiger Wahrscheinlichkeitsraum mit einer Wahrscheinlichkeitsdichte $\mathfrak{f} : \mathbb{R}^n \to \mathbb{R}^+$. Dann ist $\mathfrak{f}$ eine positive reelle Zufallsvariable. □

A02.05 **Schreibweise**.

(i) Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge numerischer (bzw. reeller) Zufallsvariablen auf $(\Omega, \mathscr{A})$. Die Folge $(X_n)_{n\in\mathbb{N}}$ heißt *(punktweise) monoton wachsend* (kurz *isoton*) bzw. *fallend* (kurz *antiton*), wenn $X_n(\omega) \leqslant X_{n+1}(\omega)$ bzw. $X_{n+1}(\omega) \leqslant X_n(\omega)$ für alle $\omega \in \Omega$ und für alle $n \in \mathbb{N}$ gilt. Für jedes $\omega \in \Omega$ definiere $[\liminf_{n\to\infty} X_n](\omega) := \sup_{n\geqslant 1} \inf_{m\geqslant n} X_m(\omega)$ und $[\limsup_{n\to\infty} X_n](\omega) := \inf_{n\geqslant 1} \sup_{m\geqslant n} X_m(\omega)$. Dann heißen $\boxed{X_\star} := \liminf_{n\to\infty} X_n$ und $\boxed{X^\star} := \limsup_{n\to\infty} X_n$ *Limes inferior* bzw. *Limes superior* der Folge $(X_n)_{n\in\mathbb{N}}$. Die Folge $(X_n)_{n\in\mathbb{N}}$ heißt *konvergent*, wenn $X_\star = X^\star =: X$ gilt, d.h. der punktweise Grenzwert existiert überall. In diesem Fall schreiben wir $\lim_{n\to\infty} X_n = X$.

(ii) Jede *monoton wachsende* (bzw. *fallende*) Folge $(X_n)_{n\in\mathbb{N}}$ von numerischen oder reellen Zufallsvariablen ist konvergent mit $X := \lim_{n\to\infty} X_n = \sup_{n\in\mathbb{N}} X_n$ (bzw. $X := \lim_{n\to\infty} X_n = \inf_{n\in\mathbb{N}} X_n$). In diesem Fall schreiben wir kurz $X_n \uparrow X$ (bzw. $X_n \downarrow X$).

(iii) Sei $(A_n)_{n\in\mathbb{N}}$ Folge von Teilmengen aus Ω. Für $A_\star := \liminf_{n\to\infty} A_n := \bigcup_{n\geqslant 1} \bigcap_{k\geqslant n} A_k$ und $A^\star := \limsup_{n\to\infty} A_n := \bigcap_{n\geqslant 1} \bigcup_{k\geqslant n} A_k$ gilt dann $\mathbb{1}_{A_\star} = (\mathbb{1}_{A_n})_\star = \liminf_{n\to\infty} \mathbb{1}_{A_n}$ und $\mathbb{1}_{A^\star} = (\mathbb{1}_{A_n})^\star = \limsup_{n\to\infty} \mathbb{1}_{A_n}$. □

A02.06 **Eigenschaft.**

(i) *Für Zufallsvariablen* $X, Y \in \overline{\mathscr{A}}$ *und* $a \in \mathbb{R}$ *gelten:* $aX \in \overline{\mathscr{A}}$ *(mit der Konvention* $0 \times \infty = 0$*);* $X \vee Y := \max(X,Y), X \wedge Y := \min(X,Y) \in \overline{\mathscr{A}}$ *und insbesondere* $X^+ := X \vee 0,\ X^- := (-X)^+ \in \overline{\mathscr{A}}^+$ *und* $|X| \in \overline{\mathscr{A}}^+$*;* $\{X < Y\}, \{X \leqslant Y\}, \{X = Y\} \in \mathscr{A}$*. Für* $x \in \overline{\mathbb{R}}^+$ *sei* $\lfloor x \rfloor := \sup\{z \in \mathbb{Z} : z \leqslant x\}$*. Dann gilt* $\lfloor X \rfloor \in \overline{\mathscr{A}}^+$*.*

(ii) *Es seien* $X_1, \dots, X_n \in \mathscr{A}$ *und* $h : \mathbb{R}^n \to \mathbb{R}^m$ *Borel-messbar. Dann ist* $h(X_1, \dots, X_n) \in \mathscr{A}^m$*. Insbesondere gilt also* $(X_1, \dots, X_n) \in \mathscr{A}^n$ *und* $X_1 + X_2$*,* $X_1 - X_2$*,* $X_1 X_2 \in \mathscr{A}$ *sowie, falls überall wohldefiniert,* $X_1/X_2 \in \mathscr{A}$*.*

(iii) *Sei* $(X_n)_{n\in\mathbb{N}}$ *aus* $\overline{\mathscr{A}}$*. Dann gilt* $\sup_{n\in\mathbb{N}} X_n \in \overline{\mathscr{A}}$*,* $\inf_{n\in\mathbb{N}} X_n \in \overline{\mathscr{A}}$*,* $X_\star = \liminf_{n\to\infty} X_n \in \overline{\mathscr{A}}$ *und* $X^\star = \limsup_{n\to\infty} X_n \in \overline{\mathscr{A}}$*. Falls* $X := \lim_{n\to\infty} X_n$ *existiert, dann ist* $X \in \overline{\mathscr{A}}$*.*

(iv) *Sei* $S : (\Omega, \mathscr{A}) \to (\mathcal{S}, \mathscr{S})$ *messbar und* $Y : \Omega \to \overline{\mathbb{R}}$*. Dann sind äquivalent:* (a) *Y ist messbar bzgl.* $\sigma(S)$*, kurz* $Y \in \overline{\sigma(S)}$*;* (b) *Es existiert* $h : (\mathcal{S}, \mathscr{S}) \to (\overline{\mathbb{R}}, \overline{\mathscr{B}})$ *messbar, kurz* $h \in \overline{\mathscr{S}}$*, mit* $Y = h(X)$*. Falls Y reell, beschränkt oder positiv ist, so erbt h diese Eigenschaft.*

$$\begin{tikzcd}
(\Omega,\mathscr{A}) \arrow[r, "S"] \arrow[dr, "Y=h(S)\in\overline{\sigma(S)}"'] & (\mathcal{S},\mathscr{S}) \arrow[d, "h\in\overline{\mathscr{S}}"] \\
& (\overline{\mathbb{R}},\overline{\mathscr{B}})
\end{tikzcd}$$

(v) *Für jedes* $X \in \overline{\mathscr{A}}^+$ *ist* $(X_n)_{n\in\mathbb{N}}$ *mit* $X_n := (2^{-n}\lfloor 2^n X\rfloor) \wedge n$ *für* $n \in \mathbb{N}$ *eine Folge einfacher Zufallsvariablen aus* $\mathscr{A}^+$*, derart dass gilt* (a) $X_n \uparrow X$*;* (b) $X_n \leqslant X \wedge n$*;* (c) *Für jedes* $c \in \mathbb{R}^+$ *gilt* $\lim_{n\to\infty} X_n = X$ *gleichmäßig auf* $\{X \leqslant c\}$*.* □

A03 Unabhängigkeit

A03.01 **Definition.**

(a) Zwei Ereignisse $A, B \in \mathscr{A}$ heißen *(stochastisch) unabhängig* (unter $\mathbb{P}$), kurz $A \perp\!\!\!\perp B$, wenn $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ gilt.

(b) Eine Familie $(A_i)_{i\in\mathcal{I}}$ von Ereignissen aus $\mathscr{A}$ mit beliebiger nicht-leerer Indexmenge $\mathcal{I}$ heißt *(stochastisch) unabhängig* (unter $\mathbb{P}$), kurz $\perp\!\!\!\perp_{i\in\mathcal{I}} A_i$, wenn für jede endliche Teilmenge $\mathcal{J} \subseteq \mathcal{I}$ gilt $\mathbb{P}\big(\bigcap_{j\in\mathcal{J}} A_j\big) = \prod_{j\in\mathcal{J}} \mathbb{P}(A_j)$.

(c) Eine Familie $(\mathscr{E}_i)_{i\in\mathcal{I}}$ von Teilmengensystemen aus $\mathscr{A}$ mit beliebiger nicht-leerer Indexmenge $\mathcal{I}$, d.h. $\mathscr{E}_i \subseteq \mathscr{A}$ für alle $i \in \mathcal{I}$, heißt *(stochastisch) unabhängig* (unter $\mathbb{P}$), kurz $\perp\!\!\!\perp_{i\in\mathcal{I}} \mathscr{E}_i$, wenn $\perp\!\!\!\perp_{i\in\mathcal{I}} A_i$ für alle $A_i \in \mathscr{E}_i$ und $i \in \mathcal{I}$ gilt.

(d) Sei $(\mathscr{A}_n)_{n\in\mathbb{N}}$ eine Folge von Teil-σ-Algebren aus $\mathscr{A}$, dann heißt $\mathscr{A}_\infty := \bigwedge_{n\geqslant 1} \bigvee_{m\geqslant n} \mathscr{A}_m$ die *asymptotische (terminale) σ-Algebra*. Ein Ereignis $A \in \mathscr{A}_\infty$ wird *asymptotisch (terminal)* bzgl. $(\mathscr{A}_n)_{n\in\mathbb{N}}$ genannt.

(e) Die Familie $(X_i)_{i\in\mathcal{I}}$ von Zufallsvariablen heißt *unabhängig*, kurz $\perp\!\!\!\perp_{i\in\mathcal{I}} X_i$, wenn die Familie $(\sigma(X_i))_{i\in\mathcal{I}}$ von Teil-σ-Algebren von $\mathscr{A}$ unabhängig ist

(f) Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von $(\mathcal{S}_n, \mathscr{S}_n)$-wertigen Zufallsvariablen auf $(\Omega, \mathscr{A}, \mathbb{P})$. Dann heißt $\mathscr{A}_X := \bigwedge_{n\geqslant 1} \bigvee_{m\geqslant n} \sigma(X_m)$ die *asymptotische σ-Algebra*. Ein Ereignis $A \in \mathscr{A}_X$ wird *asymptotisch* bzgl. $(X_n)_{n\in\mathbb{N}}$ genannt, d.h. A hängt für alle $n \geqslant 1$ nur von $(X_m)_{m\geqslant n}$ ab. □

A03.02 **Beispiel.**

(a) Ist $(A_i)_{i\in\mathcal{I}}$ eine Familie von Ereignissen, so gilt $\sigma(\{A_i\}) = \sigma(X_i)$ für die Bernoulli-Zufallsvariable $X_i := \mathbb{1}_{A_i}$, $i \in \mathcal{I}$. Weiterhin gilt $\perp\!\!\!\perp_{i\in\mathcal{I}} A_i$ genau dann, wenn $\perp\!\!\!\perp_{i\in\mathcal{I}} X_i$.

(b) Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge von Ereignissen aus $\mathscr{A}$, dann sind $A_\star = \liminf_{n\to\infty} A_n$ und $A^\star = \limsup_{n\to\infty} A_n$ asymptotische Ereignisse bzgl. $(\sigma(\{A_n\}))_{n\in\mathbb{N}}$.

(c) Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge von *unabhängigen* Ereignissen aus $\mathscr{A}$. Da $A_\star = \liminf_{n\to\infty} A_n$ und $A^\star = \limsup_{n\to\infty} A_n$ asymptotische Ereignisse bzgl. der Folge unabhängiger σ-Algebren $(\sigma(A_n))_{n\in\mathbb{N}}$ sind, folgt aus §A03.03 (iii) $\mathbb{P}(A_\star) \in \{0,1\}$ und $\mathbb{P}(A^\star) \in \{0,1\}$.

(d) Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge numerischer Zufallsvariablen, so sind die Abbildungen $X_\star := \liminf_{n\to\infty} X_n$ und $X^\star := \limsup_{n\to\infty} X_n$ messbar bzgl. $\mathscr{A}_X$.

(e) Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge *unabhängiger* numerischer Zufallsvariablen auf $(\Omega, \mathscr{A}, \mathbb{P})$, so sind $X_\star := \liminf_{n\to\infty} X_n$ und $X^\star := \limsup_{n\to\infty} X_n$ fast sicher konstant, das heißt, es gibt $x_\star, x^\star \in \overline{\mathbb{R}}$ mit $\mathbb{P}(X_\star = x_\star) = 1$ und $\mathbb{P}(X^\star = x^\star) = 1$. □

A03.03 **Eigenschaft.**

(i) *(Satz von Borel-Cantelli) Für $(A_n)_{n\in\mathbb{N}}$ aus $\mathscr{A}$ gilt: (a) $\mathbb{P}\big(\limsup_{n\to\infty} A_n\big) = 0$, falls $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) < \infty$; (b) Ist zusätzlich $\perp\!\!\!\perp_{n\in\mathbb{N}} A_n$, so auch $\mathbb{P}\big(\limsup_{n\to\infty} A_n\big) = 1$, falls $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = \infty$.*

(ii) *Für jede unabhängige Familie $(\mathscr{A}_i)_{i\in\mathcal{I}}$ von Teil-σ-Algebren aus $\mathscr{A}$ mit beliebiger nicht-leerer Indexmenge $\mathcal{I}$, also $\perp\!\!\!\perp_{i\in\mathcal{I}} \mathscr{A}_i$, und jede Partition $\{\mathcal{J}_k : k \in \mathcal{K}\}$ von $\mathcal{I}$ ist die Familie $(\bigvee_{i\in\mathcal{I}_k} \mathscr{A}_i)_{k\in\mathcal{K}}$ von Teil-σ-Algebren aus $\mathscr{A}$ unabhängig, also $\perp\!\!\!\perp_{k\in\mathcal{K}} \bigvee_{i\in\mathcal{J}_k} \mathscr{A}_i$.*

(iii) *(0-1-Gesetz von Kolmogorov) Sei $(\mathscr{A}_n)_{n\in\mathbb{N}}$ eine Folge von unabhängigen Teil-σ-Algebren aus $\mathscr{A}$. Dann ist die Wahrscheinlichkeit für jedes bzgl. $(\mathscr{A}_n)_{n\in\mathbb{N}}$ asymptotischen Ereignisses entweder 0 oder 1, also $\mathscr{A}_\infty \subseteq \overline{\mathscr{T}} := \big\{A \in \mathscr{A} \,|\, \mathbb{P}(A) \in \{0,1\}\big\}$. $\overline{\mathscr{T}}$ wird $\mathbb{P}$-triviale σ-Algebra genannt.*

(iv) *Seien $(\mathcal{T}_i, \mathscr{T}_i)_{i\in\mathcal{I}}$ eine Familie messbarer Räume und für jedes $i \in \mathcal{I}$, $h_i : \mathcal{S}_i \to \mathcal{T}_i$ eine $\mathscr{S}_i$-$\mathscr{T}_i$-messbare Funktion. Ist $\perp\!\!\!\perp_{i\in\mathcal{I}} X_i$, so gilt auch $\perp\!\!\!\perp_{i\in\mathcal{I}} h_i(X_i)$. Für jede Partition $\{\mathcal{J}_k : k \in \mathcal{K}\}$ von $\mathcal{I}$ ist $((h_i(X_i))_{i\in\mathcal{J}_k})_{k\in\mathcal{K}}$ eine Familie von unabhängigen Zufallsvariablen, also $\perp\!\!\!\perp_{k\in\mathcal{K}} (h_i(X_i))_{i\in\mathcal{J}_k}$.*

(v) *Für jede Familie $(\mathcal{S}_i, \mathscr{S}_i, \mathbb{P}_i)_{i\in\mathcal{I}}$ von Wahrscheinlichkeitsräumen mit beliebiger nicht-leerer Indexmenge $\mathcal{I}$ existiert eine Familie $(X_i)_{i\in\mathcal{I}}$ unabhängiger $(\mathcal{S}_i, \mathscr{S}_i)$-wertiger Zufallsvariablen definiert auf einem gemeinsamen Wahrscheinlichkeitsraum mit Randverteilung $\mathbb{P}_i$ und dem Produktmaß $\bigotimes_{i\in\mathcal{I}} \mathbb{P}_i$ als gemeinsamer Verteilung auf dem Produktraum $(\mathcal{S}_\mathcal{I}, \mathscr{S}_\mathcal{I})$.*

(vi) *(0-1-Gesetz von Kolmogorov) Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von unabhängigen Zufallsvariablen auf $(\Omega, \mathscr{A}, \mathbb{P})$. Dann ist die Wahrscheinlichkeit für jedes bzgl. $(X_n)_{n\in\mathbb{N}}$ asymptotischen Ereignisses entweder 0 oder 1, also $\mathscr{A}_X \subseteq \overline{\mathscr{T}}$.* □

A04 Erwartung

A04.01 **Definition.** Für jedes Wahrscheinlichkeitsmaß $\mathbb{P}$ auf einem messbaren Raum $(\Omega, \mathscr{A})$ heißt das eindeutig bestimmte Funktional $\boxed{\mathbb{E}} : \overline{\mathscr{A}}^+ \to \overline{\mathbb{R}}^+$, das die folgenden Bedingungen erfüllt:

(E1) Für alle $X, Y \in \overline{\mathscr{A}}^+$ und $a, b \in \overline{\mathbb{R}}^+$ gilt $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$; *(linear)*

(E2) Für alle $(X_n)_{n\in\mathbb{N}} \uparrow X$ in $\overline{\mathscr{A}}^+$ gilt $\mathbb{E}(X_n) \uparrow \mathbb{E}(X)$; *($\sigma$-stetig)*

(E3) Für jedes $A \in \mathscr{A}$ gilt $\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$. *(normiert)*

Erwartung bzgl. $\mathbb{P}$ und für jedes $X \in \overline{\mathscr{A}}^+$ heißt $\boxed{\mathbb{E}(X)}$ der *Erwartungswert* von X, wobei für $\mathscr{P} := \{\mathcal{P} \subseteq \mathscr{A} \,|\, \mathcal{P} \text{ endliche Partition von } \Omega\}$ gilt $\mathbb{E}(X) = \sup_{\mathcal{P}\in\mathscr{P}} \Big\{ \sum_{A\in\mathcal{P}:} \big(\inf_{\omega\in A} X(\omega)\big)\mathbb{P}(A)\Big\}$. □

A04.02 **Eigenschaft.**

(i) *Für $X \in \overline{\mathscr{A}}^+$ ist $\mathbb{P}(X = 0) = 1$ genau dann wenn $\mathbb{E}(X) = 0$. Insbesondere für $\mathbb{E}(X) < \infty$ gilt $\mathbb{P}(X = +\infty) = 0$.*

(ii) *Für $X, Y \in \overline{\mathscr{A}}^+$ gilt $\mathbb{P}(X \leqslant Y) = 1$ genau dann, wenn $\mathbb{E}(X\mathbb{1}_A) \leqslant \mathbb{E}(Y\mathbb{1}_A)$ für alle $A \in \mathscr{A}$ gilt. Insbesondere, aus $\mathbb{P}(X \leqslant Y) = 1$ folgt $\mathbb{E}(X) \leqslant \mathbb{E}(Y)$. Weiterhin gilt $\mathbb{P}(X = Y) = 1$ genau dann, wenn $\mathbb{E}(X\mathbb{1}_A) = \mathbb{E}(Y\mathbb{1}_A)$ für alle $A \in \mathscr{A}$ gilt.*

(iii) *(Lemma von Fatou) Für $(X_n)_{n\in\mathbb{N}}$ aus $\overline{\mathscr{A}}^+$ gilt $\mathbb{E}\big(\liminf_{n\to\infty} X_n\big) \leqslant \liminf_{n\to\infty} \mathbb{E}(X_n)$.*

(iv) *Eine Familie $(\mathscr{A}_i)_{i\in\mathcal{I}}$ von Teil-σ-Algebren aus $\mathscr{A}$ mit beliebiger nicht-leerer Indexmenge $\mathcal{I}$ ist unabhängig, also $\perp\!\!\!\perp_{i\in\mathcal{I}}\mathscr{A}_i$, genau dann, wenn für jede Familie $(X_i)_{i\in\mathcal{I}}$ positiver numerischer Zufallsvariablen mit $X_i \in \overline{\mathscr{A}}_i^+$, $i \in \mathcal{I}$, und jede endliche nicht-leere Teilmenge $\mathcal{J} \subseteq \mathcal{I}$ gilt $\mathbb{E}(\prod_{j\in\mathcal{J}} X_j) = \prod_{j\in\mathcal{J}} \mathbb{E}(X_j)$.* □

A04.03 **Definition.** Sei $X \in \overline{\mathscr{A}}$ eine numerische Zufallsvariable.

(a) Ist höchstens einer der beiden Erwartungswerte $\mathbb{E}(X^+)$ und $\mathbb{E}(X^-)$ nicht endlich, dass heißt, $\mathbb{E}(X^+) \wedge \mathbb{E}(X^-) < \infty$, so definiert $\boxed{\mathbb{E}(X)} := \mathbb{E}(X^+) - \mathbb{E}(X^-)$ den *Erwartungswert* von X mit den üblichen Konventionen $\infty + x = \infty$ und $-\infty + x = -\infty$ für alle $x \in \mathbb{R}$. Der Erwartungswert von X ist nicht definiert, wenn $\mathbb{E}(X^+) = \mathbb{E}(X^-) = \infty$ gilt.

(b) Falls $\mathbb{E}(|X|) < \infty$, also falls $\mathbb{E}(X^+) < \infty$ und $\mathbb{E}(X^-) < \infty$, gilt, dann heißt X *integrierbar*. Die Menge aller integrierbaren numerischen Zufallsvariablen bezeichnen wir mit $\boxed{\mathscr{L}_1} := \mathscr{L}_1(\mathbb{P}) := \mathscr{L}_1(\mathscr{A}, \mathbb{P}) := \{X \in \overline{\mathscr{A}} : \mathbb{E}(|X|) < \infty\}$.

(c) Für $p \in \mathbb{R}^+_{\setminus 0}$ definiere $\|X\|_{\mathscr{L}_p} := \big(\mathbb{E}(|X|^p)\big)^{1/p}$ und $\|X\|_{\mathscr{L}_\infty} := \inf\{x \in \mathbb{R}^+ : \mathbb{P}\big(|X| > x\big) = 0\}$. Für $p \in \overline{\mathbb{R}}^+_{\setminus 0}$ heißt X *$\mathscr{L}_p$-integrierbar*, wenn $\|X\|_{\mathscr{L}_p} < \infty$. Die Menge aller $\mathscr{L}_p$-integrierbaren Zufallsvariablen bezeichnen wir mit $\boxed{\mathscr{L}_p} := \mathscr{L}_p(\mathbb{P}) := \mathscr{L}_p(\mathscr{A}, \mathbb{P}) := \{X \in \overline{\mathscr{A}} : \|X\|_{\mathscr{L}_p} < \infty\}$.

(d) Für $X \in \mathscr{L}_p$ und $p \in \mathbb{N}$ heißt $\mathbb{E}(X^p)$ das *p-te Moment* von X; für $X \in \mathscr{L}_p$ und $p \in \mathbb{R}^+_{\setminus 0}$ heißt $\mathbb{E}(|X|^p)$ das *p-te absolute Moment* von X.

(e) Für Zufallsvariablen $X, Y \in \mathscr{L}_2$ bezeichnet $\mathbb{C}\mathrm{ov}(X, Y) := \mathbb{E}(\{X - \mathbb{E}(X)\}\{Y - \mathbb{E}(Y)\}) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$ die *Kovarianz* zwischen X und Y. Mit $\mathbb{V}\mathrm{ar}(X) := \mathbb{C}\mathrm{ov}(X, X) = \mathbb{E}(\{X - \mathbb{E}(X)\}^2) = \mathbb{E}(X^2) - \{\mathbb{E}(X)\}^2$ bzw. $\mathrm{std}(X) := \sqrt{\mathbb{V}\mathrm{ar}(X)}$ bezeichnet man die *Varianz* bzw. *Standardabweichung* (standard deviation) von X.

(f) Für $X, Y \in \mathscr{L}_2$ mit $\mathrm{std}(X), \mathrm{std}(Y) \in \mathbb{R}^+_{\setminus 0}$ heißt $\rho(X, Y) := \frac{\mathbb{C}\mathrm{ov}(X,Y)}{\mathrm{std}(X)\,\mathrm{std}(Y)} \in [-1, 1]$ *Korrelation* zwischen X und Y. Falls $\mathbb{C}\mathrm{ov}(X, Y) = 0$ gilt, heißen X und Y *unkorreliert*. □

A04.04 **Schreibweise.** Für $k \in \mathbb{N}$ bezeichnen wir mit $\boxed{\mathcal{W}_k} := \mathcal{W}_k(\mathscr{B})$ die Menge aller Wahrscheinlichkeitsmaße auf $(\mathbb{R}, \mathscr{B})$ mit endlichem k-ten absolutem Moment, also für alle $\mathbb{P} \in \mathcal{W}_k$ gilt $\mathrm{id}_{\mathbb{R}} \in \mathscr{L}_k(\mathbb{P})$ (vgl. Beispiel A02.04 (b)). Weiterhin, bezeichnet $\mathbb{E}$ die Erwartung bzgl. $\mathbb{P}$, so schreiben wir für $Y \sim \mathbb{P}$ zum Beispiel auch $\boxed{\mathbb{E}_{\mathbb{P}}(Y)} := \mathbb{E}(\mathrm{id}_{\mathbb{R}}) = \int_{\mathbb{R}} y\mathbb{P}(dy)$. □

A04.05 **Eigenschaft.** *Für $p \in [1, \infty]$ ist $\mathscr{L}_p$ ein Vektorraum. Für das Funktional $\mathbb{E} : \mathscr{L}_p \to \mathbb{R}$ mit $X \mapsto \mathbb{E}(X)$ gilt:*

(i) *Für alle $X, Y \in \mathscr{L}_p$ und $a, b \in \mathbb{R}$ gilt $aX + bY \in \mathscr{L}_p$ und $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$, falls $X \geqslant 0$ so auch $\mathbb{E}(X) \geqslant 0$; $|\mathbb{E}(X)| \leqslant \mathbb{E}(|X|)$; falls $X \leqslant Y$ so auch $\mathbb{E}(X) \leqslant \mathbb{E}(Y)$;*

(ii) *Seien $X \in \mathscr{L}_1$ und $Y \in \overline{\mathscr{A}}$ mit $\mathbb{P}(X = Y) = 1$, dann gilt $Y \in \mathscr{L}_1$ und $\mathbb{E}(X) = \mathbb{E}(Y)$.*

(iii) *Falls $X, Y \in \mathscr{L}_1$ unabhängig sind, so gilt $XY \in \mathscr{L}_1$ und $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. Somit sind unabhängige Zufallsvariablen unkorreliert. Die Umkehrung gilt nicht. Für unkorrellierte Zufallsvariablen $X, Y \in \mathscr{L}_2$ gilt $\mathbb{V}\mathrm{ar}(X+Y) = \mathbb{V}\mathrm{ar}(X) + \mathbb{V}\mathrm{ar}(Y)$.*

(iv) *Für $X \in \mathscr{L}_1$ ist $\mathbb{P}(X = 0) = 1$ genau dann, wenn $\mathbb{E}(X\mathbb{1}_A) = 0$ für alle $A \in \mathscr{A}$ gilt.*

(v) *(Ungleichung von Jensen) Seien $X \in \mathscr{L}_1$ und $\phi : \mathbb{R} \to \mathbb{R}$ eine konvexe Funktion mit $\phi(X) \in \mathscr{L}_1$. Dann gilt $\phi\big(\mathbb{E}(X)\big) \leqslant \mathbb{E}\big(\phi(X)\big)$.*

(vi) *(Monotone Konvergenz) Sei $(X_n)_{n\in\mathbb{N}}$ aus $\mathscr{L}_1$ mit $X_n \uparrow X$ (bzw. $X_n \downarrow X$) für ein $X \in \overline{\mathscr{A}}$ und $\sup_{n\in\mathbb{N}} |\mathbb{E}(X_n)| < \infty$. Dann gilt $X \in \mathscr{L}_1$ und $\mathbb{E}(X_n) \uparrow \mathbb{E}(X)$ (bzw. $\mathbb{E}(X_n) \downarrow \mathbb{E}(X)$).*

(vii) *(Dominierte Konvergenz) Sei $(X_n)_{n\in\mathbb{N}}$ aus $\mathscr{L}_1$ mit $\lim_{n\to\infty} X_n = X$ für ein $X \in \overline{\mathscr{A}}$ und existiert $Y \in \mathscr{L}_1$ mit $\sup_{n\in\mathbb{N}} |X_n| \leqslant Y$, also $\sup_{n\in\mathbb{N}} |X_n| \in \mathscr{L}_1$, dann gilt $X \in \mathscr{L}_1$, $\lim_{n\to\infty} \mathbb{E}(|X - X_n|) = 0$ und $\lim_{n\to\infty} \mathbb{E}(X_n) = \mathbb{E}(X)$.*

(viii) *Sei X eine $(\mathcal{S}, \mathscr{S})$-wertige Zufallsvariable auf $(\Omega, \mathscr{A}, \mathbb{P})$, sei $\mathbb{P}^X = \mathbb{P} \circ X^{-1}$ die Verteilung von X auf $(\mathcal{S}, \mathscr{S})$ und sei $h \in \overline{\mathscr{S}}$ eine numerische Zufallsvariable auf $(\mathcal{S}, \mathscr{S})$. Für $h \geqslant 0$, also $h \in \overline{\mathscr{S}}^+$, gilt $\mathbb{E}_{\mathbb{P}}(h(X)) = \mathbb{E}_{\mathbb{P}^X}(h)$. Weiterhin gilt $h(X) \in \mathscr{L}_1(\Omega, \mathscr{A}, \mathbb{P})$ genau dann, wenn $h \in \mathscr{L}_1(\mathcal{S}, \mathscr{S}, \mathbb{P}^X)$.*

$$\begin{tikzcd}
(\Omega, \mathscr{A}) \arrow[r, "X"] \arrow[dr, "h(X) \in \mathscr{L}_1(\mathbb{P})"'] & (\mathcal{S}, \mathscr{S}) \arrow[d, "h \in \mathscr{L}_1(\mathbb{P}^X)"] \\
 & (\overline{\mathbb{R}}, \overline{\mathscr{B}})
\end{tikzcd}$$

(ix) *Für $p \in \mathbb{R}^+_{\setminus 0}$ gilt $X \in \mathscr{L}_p$ genau dann, wenn $|X|^p \in \mathscr{L}_1$ gilt. Für $p = \infty$ gilt $\mathbb{P}\big(|X| > \|X\|_{\mathscr{L}_\infty}\big) = 0$.*

(x) *Für $p, q \in \overline{\mathbb{R}}^+_{\setminus 0}$ mit $p \leqslant q$ und $X \in \mathscr{L}_q$ gilt $\|X\|_{\mathscr{L}_p} \leqslant \|X\|_{\mathscr{L}_q}$ und somit $\mathscr{L}_q \subseteq \mathscr{L}_p$.*

(xi) *Sei $p \in \overline{\mathbb{R}}^+_{\setminus 0}$. Für $X \in \overline{\mathscr{A}}$ gilt $\|X\|_{\mathscr{L}_p} = 0$ genau dann, wenn $\mathbb{P}(X = 0) = 1$. Für $a \in \mathbb{R}$ gilt $\|aX\|_{\mathscr{L}_p} = |a|\|X\|_{\mathscr{L}_p}$. Für $X \in \mathscr{L}_p$ gilt $\mathbb{P}(|X| < \infty) = 1$. Für $Y \in \overline{\mathscr{A}}$ mit $\mathbb{P}(X = Y) = 1$ gilt $\|X\|_{\mathscr{L}_p} = \|Y\|_{\mathscr{L}_p}$.*

(xii) *(Hölder Ungleichung) Seien $X, Y \in \overline{\mathscr{A}}$ und $p, q \in [1, \infty]$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt: $\mathbb{E}(|XY|) \leqslant \|X\|_{\mathscr{L}_p}\|Y\|_{\mathscr{L}_q}$.*
(Cauchy-Schwarz Ungleichung) Für $X, Y \in \mathscr{L}_2$ gilt $XY \in \mathscr{L}_1$ und $|\mathbb{E}(XY)|^2 \leqslant \|X\|^2_{\mathscr{L}_2}\|Y\|^2_{\mathscr{L}_2} = \mathbb{E}(|X|^2)\mathbb{E}(|Y|^2)$.

(xiii) *(Minkowski Ungleichung) Für $X, Y \in \mathscr{L}_p$ mit $p \in [1, \infty]$ gilt $X + Y \in \mathscr{L}_p$ und $\|X + Y\|_{\mathscr{L}_p} \leqslant \|X\|_{\mathscr{L}_p} + \|Y\|_{\mathscr{L}_p}$. Insbesondere ist $\|\cdot\|_{\mathscr{L}_p}$ für $p \in [1, \infty]$ eine Pseudonorm auf $\mathscr{L}_p$, das heißt für $X, Y \in \mathscr{L}_p$ und $a \in \mathbb{R}$ gilt:* (a) *$\|X\|_{\mathscr{L}_p} \geqslant 0$ für alle X und $\|X\|_{\mathscr{L}_p} = 0$, falls $X = 0$* $\mathbb{P}$-f.s.*;* (b) *$\|aX\|_{\mathscr{L}_p} = |a|\|X\|_{\mathscr{L}_p}$ und* (c) *$\|X + Y\|_{\mathscr{L}_p} \leqslant \|X\|_{\mathscr{L}_p} + \|Y\|_{\mathscr{L}_p}$.*

(xiv) *(Markov Ungleichung) Für $Y \in \overline{\mathscr{A}}^+$ und $p, \varepsilon \in \mathbb{R}^+_{\setminus 0}$ gilt $\varepsilon^p \mathbb{1}_{\{Y > \varepsilon\}} \leqslant Y^p$, sodass $\mathbb{P}(Y > \varepsilon) \leqslant \varepsilon^{-p}\mathbb{E}(Y^p)$.*
(Tschebischeff Ungleichung) Für $X \in \mathscr{L}_2$ und $\varepsilon \in \mathbb{R}^+_{\setminus 0}$ gilt $\mathbb{P}(|X - \mathbb{E}(X)| > \varepsilon) \leqslant \varepsilon^{-2}\,\mathbb{V}\mathrm{ar}(X)$. □

A04.06 **Schreibweise.** Im Folgenden fassen wir Vektoren als Spaltenvektoren auf, dass heißt $a = (a_1 \cdots a_n)^t \in \mathbb{R}^n$. Wir bezeichnen mit $\langle \cdot, \cdot \rangle$ das Standardskalarprodukt auf $\mathbb{R}^n$, dass heißt, $\langle a, b \rangle = \sum_{i=1}^n a_i b_i = b^t a$ für alle $a, b \in \mathbb{R}^n$. Für $p \in \overline{\mathbb{R}}^+_{\setminus 0}$ sei $\|\cdot\|_p$ die übliche $\mathscr{L}_p$-Norm auf $\mathbb{R}^n$, dass heißt, für alle $a \in \mathbb{R}^n$ ist $\|a\|_p = (\sum_{i=1}^n |a_i|^p)^{1/p}$ für $p \in \mathbb{R}^+_{\setminus 0}$ sowie $\|a\|_\infty = \max_{i\in[\![n]\!]} |a_i|$. Für die vom Standardskalarprodukt $\langle\cdot,\cdot\rangle$ auf $\mathbb{R}^n$ induzierte Norm schreiben wir kurz $\|\cdot\| := \|\cdot\|_2$, also $\|a\| = \sqrt{\langle a, a\rangle} = \sqrt{a^t a} = (\sum_{i=1}^n a_i^2)^{1/2}$ für alle $a \in \mathbb{R}^n$. Für eine symmetrische Matrix $\Sigma = (\Sigma_{ij}) \in \mathbb{R}^{(p,p)}$ schreiben wir $\Sigma \geqslant 0$ oder $\Sigma \in \mathbb{R}^{(p,p)}_{\geqslant}$, falls sie positiv semi-definit ist, sowie $\Sigma > 0$ oder $\Sigma \in \mathbb{R}^{(p,p)}_{>}$, falls sie strikt positiv definit ist. □

A04.07 **Definition**. Sei $X = (X_i)_{i\in[\![n]\!]} \in \overline{\mathscr{A}}^n$ ein numerischer Zufallsvektor (aufgefasst als Spaltenvektor) auf $(\Omega, \mathscr{A}, \mathbb{P})$, also $X_i \in \overline{\mathscr{A}}$, $i \in [\![n]\!]$, sowie $\|X\|_p \in \overline{\mathscr{A}}$ für $p \in \overline{\mathbb{R}}^+_{\setminus 0}$.

(a) Für $p \in \overline{\mathbb{R}}^+_{\setminus 0}$ heißt X $\mathscr{L}_p$-integrierbar, falls $\|X\|_p \in \mathscr{L}_p(\mathscr{A}, \mathbb{P})$ oder dazu äquivalent $X_i \in \mathscr{L}_p$ für alle $i \in [\![n]\!]$ gilt. Wir definieren $\|X\|_{\mathscr{L}_p} := \|\|X\|_p\|_{\mathscr{L}_p}$ sowie $\mathscr{L}_p := \mathscr{L}_p(\mathbb{P}) := \mathscr{L}_p(\mathscr{A}, \mathbb{P}) := \{X \in \overline{\mathscr{A}}^n : \|X\|_{\mathscr{L}_p} < \infty\}$.

(b) Für $X \in \mathscr{L}_1$ heißt $\mathbb{E}(X) := \big(\mathbb{E}(X_i)\big)_{i\in[\![n]\!]} \in \mathbb{R}^n$ *Erwartungswertvektor* von X.

(c) Falls $X \in \mathscr{L}_2$, dann heißt

$$\begin{aligned}\mathbb{C}\mathrm{ov}(X) &:= \big(\mathbb{C}\mathrm{ov}(X_i, X_j)\big)_{i,j\in[\![n]\!]} := \Big(\mathbb{E}\big(\{X_i - \mathbb{E}(X_i)\}\{X_j - \mathbb{E}(X_j)\}\big)\Big)_{i,j\in[\![n]\!]}\\ &=: \mathbb{E}\big((X - \mathbb{E}(X))(X - \mathbb{E}(X))^t\big) = \mathbb{E}(XX^t) - \mathbb{E}(X)(\mathbb{E}(X))^t \in \mathbb{R}^{(n,n)}\end{aligned}$$

Kovarianzmatrix von X. □

A04.08 **Schreibweise**. Für $k \in \mathbb{N}$ bezeichen wir mit $\mathcal{W}_k := \mathcal{W}_k(\mathscr{B}^n)$ die Menge aller Wahrscheinlichkeitsmaße auf $(\mathbb{R}^n, \mathscr{B}^n)$ mit endlichem k-ten absolutem Moment, also für alle $\mathbb{P} \in \mathcal{W}_k(\mathscr{B}^n)$ gilt $\mathrm{id}_{\mathbb{R}^n} \in \mathscr{L}_k(\mathscr{B}^n, \mathbb{P})$ (vgl. Beispiel A02.04 (b)). Weiterhin, bezeichnet $\mathbb{E}$ die Erwartung bzgl. $\mathbb{P}$, so schreiben wir für $Y \sim \mathbb{P}$ auch $\mathbb{E}_{\mathbb{P}}(Y) := \mathbb{E}(\mathrm{id}_{\mathbb{R}^n}) = \big(\int_{\mathbb{R}^n} y_i \mathbb{P}(dy)\big)_{i\in[\![n]\!]}$. □

A04.09 **Eigenschaft**.

(i) *(Ungleichung von Jensen) Es seien $\mathcal{E} \subseteq \mathbb{R}^n$ konvex und $X = (X_i)_{i\in[\![n]\!]} \in \overline{\mathscr{A}}^n$ und in $\mathscr{L}_1(\mathbb{P})$ mit $\mathbb{P}(X \in \mathcal{E}) = 1$, so dass $\mathbb{E}(X) = \big(\mathbb{E}(X_i)\big)_{i\in[\![n]\!]} \in \mathcal{E}$. Ist $\psi : \mathcal{E} \to \mathbb{R}$ eine konvexe Funktion dann gilt $\mathbb{E}(\psi(X)) \geqslant \psi(\mathbb{E}(X))$.*

(ii) *Für die Kovarianzmatrix $\Sigma := \mathbb{C}\mathrm{ov}(X)$ eines $\mathbb{R}^n$-wertigen Zufallsvektors $X \in \mathscr{L}_2$ gilt $\mathbb{C}\mathrm{ov}(\langle a, X\rangle, \langle X, b\rangle) = \sum_{i=1}^n \sum_{j=1}^n a_i b_j \mathbb{C}\mathrm{ov}(X_i, X_j) = b^t \Sigma a = \langle \Sigma a, b\rangle$ für alle $a, b \in \mathbb{R}^n$. Insbesondere gilt $\Sigma_{ij} = \mathbb{C}\mathrm{ov}(X_i, X_j) = \mathbb{C}\mathrm{ov}(X_j, X_i) = \Sigma_{ji}$ für alle $i, j \in [\![n]\!]$ und $\langle \Sigma c, c\rangle = \mathbb{V}\mathrm{ar}(\langle X, c\rangle) \geqslant 0$ für alle $c \in \mathbb{R}^n$, also $\Sigma \in \mathbb{R}^{(n,n)}_{\geqslant}$.*

(iii) *Sei $X \in \mathscr{A}^m$ in $\mathscr{L}_2$. Für alle $b \in \mathbb{R}^n$ und $A \in \mathbb{R}^{(n,m)}$ ist dann $Y = AX + b \in \mathscr{A}^n$ in $\mathscr{L}_2$. Bezeichnen wir weiterhin mit $\mu := \mathbb{E}(X) \in \mathbb{R}^m$ und $\Sigma := \mathbb{C}\mathrm{ov}(X) \in \mathbb{R}^{(m,m)}$ den Erwartungswertvektor und die Kovarianzmatrix von X, dann gilt $\mathbb{E}(Y) = A\mu + b$ und $\mathbb{C}\mathrm{ov}(Y) = A\Sigma A^t$.* □

A04.10 **Definition**. Für $X \in \mathscr{A}^k, Y \in \mathscr{A}^n$ in $\mathscr{L}_2$ heißt $Z^* = A^* X + b^*$ mit $A^* \in \mathbb{R}^{(n,k)}$ und $b^* \in \mathbb{R}^n$ eine *lineare Vorhersage* von Y durch X. Z^* wird *beste lineare Vorhersage* von Y durch X genannt, wenn $\mathbb{E}\|Y - Z^*\|^2 \leqslant \mathbb{E}\|Y - (AX + b)\|^2$ für alle $A \in \mathbb{R}^{(n,k)}$ und $b \in \mathbb{R}^n$ gilt. □

A04.11 **Bemerkung**. Für eine Matrix $A \in \mathbb{R}^{(n,k)}$ heißt eine Matrix $A^+ \in \mathbb{R}^{(k,n)}$ *Moore-Penrose Inverse* von A, wenn $AA^+A = A$, $A^+AA^+ = A^+$ und AA^+ sowie A^+A symmetrisch sind. Die Moore-Penrose Inverse ist eindeutig festgelegt. Ist $B \in \mathbb{R}^n$ symmetrisch, so dass $U^t BU = \mathrm{Diag}(\lambda)$ für eine orthogonale Matrix U und eine Diagonalmatrix $\mathrm{Diag}(\lambda)$ mit reellen Diagonaleinträgen $\lambda = (\lambda_i)_{i\in[\![n]\!]}$. Setzen wir $\lambda^+ = (\lambda_i^+)_{i\in[\![n]\!]}$ mit $\lambda_i^+ = \lambda_i^{-1}\mathbb{1}_{\mathbb{R}^+_{\setminus 0}}(\lambda_i)$, dass heißt, für $\lambda_i \in \mathbb{R}^+_{\setminus 0}$ ist $\lambda_i^+ = \lambda_i^{-1}$ und ansonsten $\lambda_i^+ = 0$. Dann ist $B^+ = U\,\mathrm{Diag}(\lambda^+)U^t$ die Moore-Penrose Inverse von B. Allgemein ist $A^+ = (A^tA)^+A^t \in \mathbb{R}^{(k,n)}$ die Moore-Penrose Inverse von $A \in \mathbb{R}^{(n,k)}$. □

A04.12 **Eigenschaft**. *Für $X \in \mathscr{A}^k$, $Y \in \mathscr{A}^n$ in $\mathscr{L}_2$ mit $\mathbb{C}\mathrm{ov}(Y,X) := \mathbb{E}((Y-\mathbb{E}(Y))(X-\mathbb{E}(X))^t)$ ist $Z^* = \mathbb{E}(Y) + \mathbb{C}\mathrm{ov}(Y,X)\,\mathbb{C}\mathrm{ov}(X)^+(X-\mathbb{E}(X))$ die beste lineare Vorhersage von Y durch X. Der Fehler $\varepsilon := Y - Z^*$ und AX für beliebiges $A \in \mathbb{R}^{(n,k)}$ sind unkorreliert, also $\mathbb{C}\mathrm{ov}(\varepsilon, AX) = 0$. Es gilt $\mathbb{C}\mathrm{ov}(\varepsilon) = \mathbb{C}\mathrm{ov}(Y) - \mathbb{C}\mathrm{ov}(Y,X)\,\mathbb{C}\mathrm{ov}(X)^+\,\mathbb{C}\mathrm{ov}(X,Y)$ und $\mathbb{E}(\varepsilon) = 0$.* □

A05 Multivariate Normalverteilung

Nicht degenerierte multivariate Normalverteilungen können direkt über ihre Dichte definiert werden. Eine Normalverteilung heißt degeneriert, falls ihre Kovarianzmatrix nicht strikt positiv definit ist (nicht vollen Rang hat). In der Vorlesung werden wir auch Zufallsvariablen mit degenerierten Normalverteilungen betrachten. Beispiele für solche Zufallsvariablen sind Projektionen von nicht degenerierten normalverteilten Zufallsvariablen auf lineare Teilräume.

A05.01 **Satz von Cramér-Wold**. *Die Verteilung eines $\mathbb{R}^n$-wertigen Zufallsvektors X ist eindeutig festgelegt durch die Verteilungen der linearen Formen $\langle X, c\rangle$ für alle $c \in \mathbb{R}^n$.* □

A05.02 **Definition**. Ein $\mathbb{R}^n$-wertiger Zufallsvektor X besitzt eine *multivariate Normalverteilung* $\mathrm{N}_{(\mu,\Sigma)}$ mit $\mu \in \mathbb{R}^n$ und positiv semi-definiter Matrix $\Sigma \in \mathbb{R}^{(n,n)}$, falls für alle $c \in \mathbb{R}^n$ die reelle Zufallsvariable $\langle X, c\rangle$ eine $\mathrm{N}_{(\langle\mu,c\rangle,\langle\Sigma c,c\rangle)}$-Verteilung besitzt. Das Produktmaß $\mathrm{N}_{(0,E_n)} = \bigotimes_{i=1}^n \mathrm{N}_{(0,1)} = \mathrm{N}_{(0,1)}^n$ heißt insbesondere *(n-dimensionale) Standardnormalverteilung*, wobei E_n die n-dimensionale Einheitsmatrix ist. □

A05.03 **Vorbemerkung.**. Für eine Matrix $A \in \mathbb{R}^{(n,m)}$ mit Spaltenvektoren $a_{\bullet 1}, \dots, a_{\bullet m}$ bezeichnet $\mathrm{Bild}(A) = \langle a_{\bullet 1}, \dots, a_{\bullet m}\rangle \subseteq \mathbb{R}^n$ die lineare Hülle der Spaltenvektoren, also das Bild der linearen Abbildung $\mathbb{R}^m \to \mathbb{R}^n$ mit $x \mapsto Ax$. Für einen linearen Unterraum $\mathcal{U} \subseteq \mathbb{R}^n$ bezeichnet $\mathbb{R}^n = \mathcal{U} \oplus \mathcal{U}^\perp$ die direkte orthogonale Summe, dass heißt, $\mathcal{U}$ und $\mathcal{U}^\perp$ sind orthogonal, also für alle $u \in \mathcal{U}$ und $v \in \mathcal{U}^\perp$ gilt $\langle u, v\rangle = 0$, und jedes Element $x \in \mathbb{R}^n$ hat eine eindeutige Darstellung $x = u + v$ mit $u \in \mathcal{U}$ und $v \in \mathcal{U}^\perp$. Wir bezeichnen mit $\Pi_{\mathcal{U}}$ die Darstellungsmatrix der orthogonalen Projektion von $\mathbb{R}^n$ in $\mathcal{U}$, also $\mathcal{U} \oplus \mathcal{U}^\perp \to \mathcal{U}$ mit $x = u + v \mapsto u = \Pi_{\mathcal{U}} x$. Eine Matrix $U \in \mathbb{R}^{(n,m)}$ heißt partielle Isometrie, falls $UU^t = \Pi_{\mathrm{Bild}(U)}$ und $U^tU = \Pi_{\mathrm{Bild}(U^t)}$. □

A05.04 **Eigenschaft**. *Seien $Z \sim \mathrm{N}_{(0,E_m)}$ und $Y \sim \mathrm{N}_{(0,E_k)}$, dann gelten die folgenden Aussagen:*

(i) *Falls $A \in \mathbb{R}^{(n,m)}$ und $B \in \mathbb{R}^{(n,k)}$ mit $AA^t = BB^t$ gilt, dann sind die $\mathbb{R}^n$-wertigen Zufallsvektoren AZ und BY identisch verteilt.*

(ii) *Falls $U \in \mathbb{R}^{(n,m)}$ eine partielle Isometrie ist, dann gilt $UZ \sim \mathrm{N}_{(0,\Pi_{\mathrm{Bild}(U)})}$.*

(iii) *Falls $A \in \mathbb{R}^{(m,n)}$ und $B \in \mathbb{R}^{(m,k)}$ mit $A^tB = 0$. Dann sind $\Pi_{\mathrm{Bild}(A)}Z \sim \mathrm{N}_{(0,\Pi_{\mathrm{Bild}(A)})}$ und $\Pi_{\mathrm{Bild}(B)}Z \sim \mathrm{N}_{(0,\Pi_{\mathrm{Bild}(B)})}$ unabhängig.*

Sei $X \sim \mathrm{N}_{(\mu,\Sigma)}$ mit $\mu \in \mathbb{R}^n$ und $\Sigma \in \mathbb{R}^{(n,n)}_{\geqslant}$, dann gelten die folgenden Aussagen:

(iv) *Für alle $i \in [\![n]\!]$ gilt $X_i \sim \mathrm{N}_{(\mu_i,\Sigma_{ii})}$.*

(v) *Für alle $i, j \in [\![n]\!]$ mit $i \neq j$ sind die Koordinaten X_i und X_j von X genau dann unabhängig, wenn $\Sigma_{ij} = 0$ gilt.*

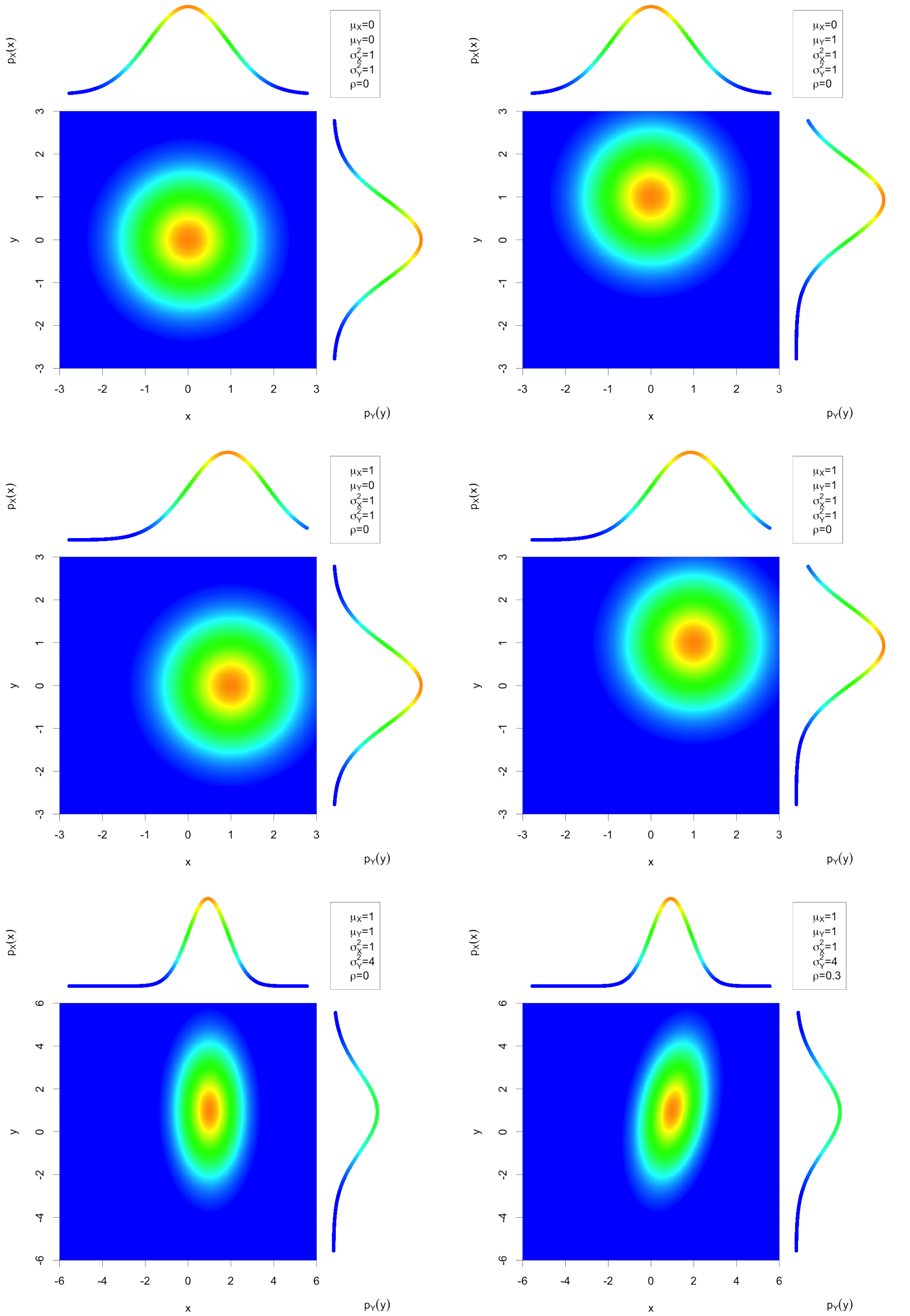
(vi) *Für $A \in \mathbb{R}^{(m,n)}$ und $b \in \mathbb{R}^m$ gilt $Y = AX + b \sim \mathrm{N}_{(A\mu+b, A\Sigma A^t)}$.*

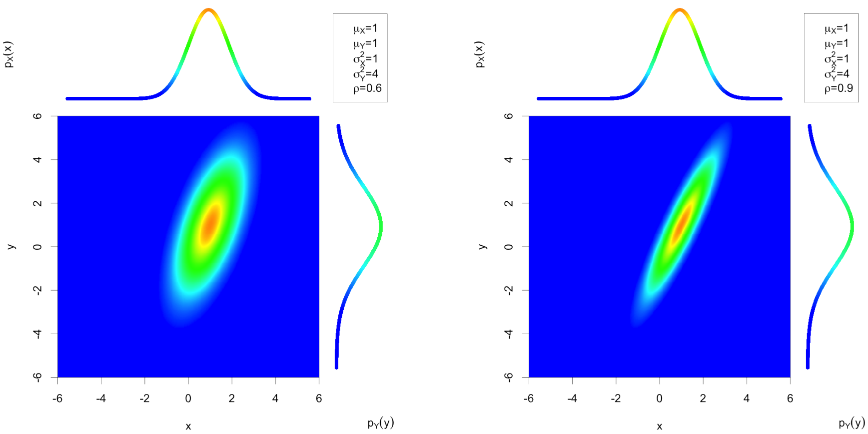
(vii) *Ist Σ positiv definit, dann ist X stetig verteilt mit Lebesgue-Dichte $f(x) = (2\pi)^{-n/2}(\det\Sigma)^{-1/2}\exp\left\{-\tfrac{1}{2}\langle\Sigma^{-1}(x-\mu),(x-\mu)\rangle\right\}$, $x \in \mathbb{R}^n$.* □

A05.05 **Beispiel**. Der stetig-verteilte Zufallsvektor (X, Y) wird *bivariat normalverteilt* mit Parametern $\mu_X, \mu_Y \in \mathbb{R}$, $\sigma_X, \sigma_Y \in \mathbb{R}^+_{\backslash 0}$ und $\rho \in (-1, 1)$ genannt, wenn die gemeinsame Dichte durch

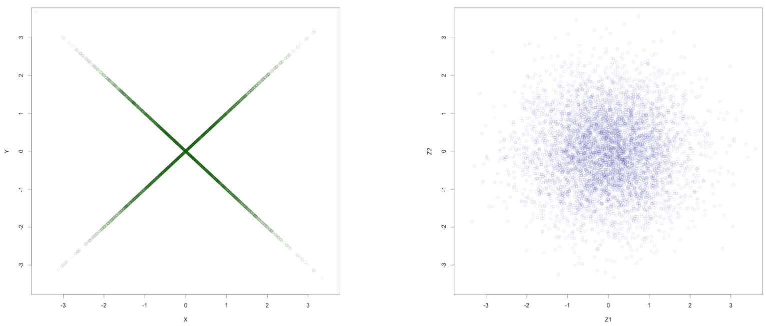
$$\mathbb{f}^{(}_{(X,Y)}x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\Big(-\frac{(x-\mu_X)^2}{2(1-\rho^2)\sigma_X^2}\Big) \exp\Big(\frac{2\rho(x-\mu_X)(y-\mu_Y)}{2(1-\rho^2)\sigma_X\sigma_Y}\Big) \exp\Big(-\frac{(y-\mu_Y)^2}{2(1-\rho^2)\sigma_Y^2}\Big)$$

gegeben ist, wobei $\mu_X = \mathbb{E}(X)$, $\mu_Y = \mathbb{E}(Y)$, $\sigma_X^2 = \mathbb{V}\mathrm{ar}(X)$, $\sigma_Y^2 = \mathbb{V}\mathrm{ar}(Y)$ und $\rho = \mathbb{C}\mathrm{orr}(X, Y) = \frac{\mathbb{C}\mathrm{ov}(X,Y)}{\sigma_X\sigma_Y}$. Die nächsten Graphiken stellen die gemeinsame und die marginalen Dichten für verschiedene Werte der Parameter dar.





X und Y sind genau dann unabhängig, wenn sie unkorreliert sind, also $\rho = 0$ gilt (vgl. §A05.04 (v)). *Achtung*, es ist natürlich möglich, dass $X \sim \mathrm{N}_{(\mu_X,\sigma_X^2)}$ und $Y \sim \mathrm{N}_{(\mu_Y,\sigma_Y^2)}$ unkorreliert sind, aber der Vektor (X, Y) nicht bivariat normalverteilt ist. Betrachte dazu zwei unabhängige Zufallsvariablen X und V, wobei $X \sim \mathrm{N}_{(0,1)}$ und V ist eine Rademacher-Zufallsvariable, d.h. $V \in \{-1, 1\}$ mit $P(V = -1) = 1/2 = P(V = 1)$. Es ist nun leicht zu zeigen, dass die Zufallsvariablen $Y := VX$ und X unkorreliert sind und dass $Y \sim \mathrm{N}_{(0,1)}$ (Nachrechnen!). Die Zufallsvariablen X und Y sind somit standardnormalverteilt und unkorreliert, aber ihre gemeinsame Verteilung ist keine Normalverteilung (warum?). Die nächsten Graphiken zeigen 5000 Realisierungen von (X, Y) (in grün) und zum Vergleich 5000 Realisierungen einer bivariaten Standardnormalverteilung.



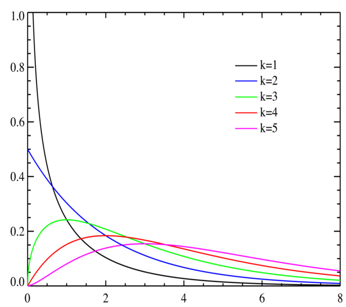
□

Im Folgenden sind $(Z_i)_{i\in[\![0,m+k]\!]}$ unabhängige und identisch $\mathrm{N}_{(0,1)}$-verteilte Zufallsvariablen, also $(Z_i)_{i\in[\![0,m+k]\!]} \sim \mathrm{N}_{(0,1)}^{1+m+k}$.

A05.06 **χ^2-Verteilung.** Die Verteilung der Zufallsvariable

$$Q := \sum_{i=1}^{k} Z_i^2$$

heißt *(zentrale) χ^2-Verteilung* mit k *Freiheitsgraden* , kurz $Q \sim \chi_k^2$. Für $\alpha \in (0, 1)$ bezeichnen wir weiterhin den Wert $\chi_{k,\alpha}^2 \in \mathbb{R}_{\setminus 0}^{+}$ als α-Quantil einer (zentralen) χ^2-Verteilung mit k Freiheitsgraden, falls $\mathbb{P}(Q \leqslant \chi_{k,\alpha}^2) = \alpha$ gilt.



χ_k^2-Dichtefunktionen

Für $\delta \in \mathbb{R}$ heißt die Verteilung der Zufallsvariable

$$Q := (Z_1 + \delta)^2 + \sum_{i=2}^{k} Z_i^2$$

nicht-zentrale χ_k^2-Verteilung mit k *Freiheitsgraden* und *Nichtzentralitätsparameter* δ^2, kurz

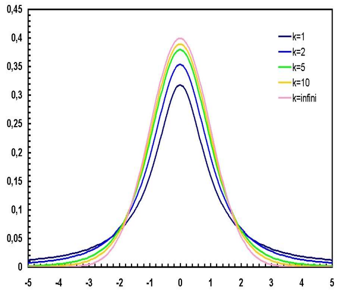
$Q \sim \chi^2_k(\delta^2)$. $\chi^2_{k,\alpha}(\delta^2) \in \mathbb{R}^+_{\setminus 0}$ bezeichnet das α-Quantil einer nicht-zentralen χ^2-Verteilung mit k Freiheitsgraden und Nichtzentralitätsparameter δ^2, d.h. $\mathbb{P}\big(Q \leqslant \chi^2_{k,\alpha}(\delta^2)\big) = \alpha$. □

A05.07 **Eigenschaft.** *Sei $Q \sim \chi^2_k$ und $W \sim \chi^2_k(\delta^2)$, dann gilt $\mathbb{E}(Q) = k$, $\mathbb{V}\mathrm{ar}(Q) = 2k$ und $\mathbb{E}(W) = \delta^2 + k$. Für $Z \sim \mathrm{N}^m_{(0,1)}$, $v \in \mathbb{R}^m$ und $A \in \mathbb{R}^{m\times p}$ mit $\mathrm{rg}(A) = p$ gelten außerdem: (i) $\|\Pi_{\mathrm{Bild}(A)} Z\|^2 \sim \chi^2_p$ und (ii) $\|Z + v\|^2 \sim \chi^2_m(\|v\|^2)$.* □

A05.08 **(Student-) t-Verteilung.** Die Verteilung der Zufallsvariable

$$T := \frac{Z_0}{\sqrt{\frac{1}{k}\sum_{i=1}^{k} Z_i^2}}$$

heißt *(Student-) t-Verteilung* mit k *Freiheitsgraden*, kurz $T \sim \mathrm{t}_k$, und $\mathrm{t}_{k,\alpha} \in \mathbb{R}$ bezeichnet das α-Quantil einer (Student-) t-Verteilung mit k-Freiheitsgraden, d.h. $\mathbb{P}(T \leqslant \mathrm{t}_{k,\alpha}) = \alpha$.



t_k-Dichtefunktionen

A05.09 **Eigenschaft.**

(i) *Die (Student-) t_1-Verteilung mit einem ($k = 1$) Freiheitsgrad entspricht gerade der Cauchy-Verteilung.*

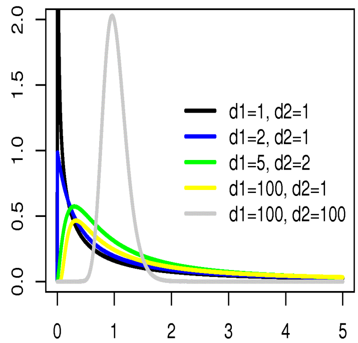
(ii) *Für jedes $k \in \mathbb{N}$ besitzt die t_k-Verteilung endliche Momente nur bis zur Ordnung $p < k$ (sie ist heavy-tailed). Insbesondere, ist $T \sim \mathrm{t}_k$ so gilt $\mathbb{E}(T) = 0$ für $k > 1$, sowie $\mathbb{V}\mathrm{ar}(T) = k/(k-2)$ für $k > 2$.*

(iii) *Für $(X_i)_{i\in[\![n]\!]} \sim \mathrm{N}^n_{(\mu,\sigma^2)}$, $\overline{X}_n := \sum_{i=1}^n X_i$ und $\widehat{S}^{(2)}_n := \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$ sind $\frac{\sqrt{n}}{\sigma}(\overline{X}_n - \mu) \sim \mathrm{N}_{(0,1)}$ und $\frac{(n-1)}{\sigma^2}\widehat{S}^{(2)}_n \sim \chi^2_{n-1}$ unabhängig, so dass $\widehat{T}_n = \frac{\sqrt{n}}{\widehat{S}_n}(\overline{X}_n - \mu) \sim \mathrm{t}_{n-1}$ mit $\widehat{S}_n := \sqrt{\widehat{S}^{(2)}_n}$ gilt.* □

A05.10 **(Fisher-) F-Verteilung.** Die Verteilung der Zufallsvariable

$$F := \frac{\frac{1}{m}\sum_{i=1}^{m} Z_i^2}{\frac{1}{k}\sum_{i=m+1}^{m+k} Z_i^2}$$

heißt *zentrale (Fisher-) F-Verteilung* mit m und k *Freiheitsgraden*, kurz $F \sim \mathrm{F}_{m,k}$. $\mathrm{F}_{m,k,\alpha}$ bezeichnet das α-Quantil einer zentralen Fisher-$\mathrm{F}_{m,k}$-Verteilung mit m und k Freiheitsgraden, d.h. $P(F \leqslant \mathrm{F}_{m,k,\alpha}) = \alpha$.



$\mathrm{F}_{d1,d2}$-Dichtefunktionen

Für $\delta \in \mathbb{R}$ heißt die Verteilung der Zufallsvariable

$$F := \frac{\frac{1}{m}\{(Z_1 + \delta)^2 + \sum_{i=2}^{m} Z_i^2\}}{\frac{1}{k}\sum_{i=m+1}^{m+k} Z_i^2}$$

nicht-zentrale (Fisher-) F-Verteilung mit m und k *Freiheitsgraden* und *Nichtzentralitätsparameter* δ^2, kurz $F \sim \mathrm{F}_{m,k}(\delta^2)$. $\mathrm{F}_{m,k,\alpha}(\delta^2) \in \mathbb{R}^+$ bezeichnet das α-Quantil einer nicht-zentralen

$\mathrm{F}_{n,k}(\delta^2)$-Verteilung mit m und k Freiheitsgraden und Nichtzentralitätsparameter δ^2, das heißt $P\big(F \leqslant \mathrm{F}_{m,k,\alpha}(\delta^2)\big) = \alpha$. □

A05.11 **Eigenschaft.**

(i) *Sei $F \sim \mathrm{F}_{n,k}$ mit $k > 1$, dann ist F^{-1} eine $\mathrm{F}_{k,m}$-verteilte Zufallsvariable. Für $T \sim \mathrm{t}_k$ ist T^2 eine $\mathrm{F}_{1,k}$-verteilte Zufallsvariable.*

(ii) *Sei $F_k \sim \mathrm{F}_{m,k}$, $k \in \mathbb{N}$, dann konvergiert die Folge von Zufallsvariablen $(mF_k)_{k\in\mathbb{N}}$ für $k \to \infty$ in Verteilung gegen ein χ^2_m-verteilte Zufallsvariable.* □

A06 Grenzwertsätze

A06.01 **Schreibweise.** Im Folgenden bezeichnen wir mit $\mathcal{C}_b := \mathcal{C}_b(\mathbb{R}^k)$ die Menge aller beschränkten, stetigen, reellen Funktionen auf $\mathbb{R}^k$. Für $h \in \mathcal{C}_b$ ist somit $\|h\|_\infty := \sup_{x\in\mathbb{R}^k} |h(x)| < \infty$, so dass für jedes Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\mathbb{R}^k, \mathscr{B}^k)$, kurz $\mathbb{P} \in \mathcal{W}(\mathscr{B}^k)$, gilt $\mathcal{C}_b \subseteq \mathscr{L}_\infty(\mathscr{B}^k, \mathbb{P})$. □

A06.02 **Definition.**

(a) Eine Folge $(X_n)_{n\in\mathbb{N}}$ reeller Zufallsvariablen in $\mathscr{A}$ konvergiert

- *$\mathbb{P}$-fast sicher* ($\mathbb{P}$-f.s.) gegen die numerische Zufallsvariable $X \in \overline{\mathscr{A}}$, kurz $\boxed{X_n \xrightarrow{\mathbb{P}\text{-f.s.}} X}$, wenn $\limsup_{n\to\infty}|X_n - X| = 0$ $\mathbb{P}$-f.s., dass heißt $\mathbb{P}(\limsup_{n\to\infty}|X_n - X| = 0) = 1$ gilt.

- *$\mathbb{P}$-fast vollständig* ($\mathbb{P}$-f.v.) gegen die numerische Zufallsvariable $X \in \overline{\mathscr{A}}$, kurz $\boxed{X_n \xrightarrow{\mathbb{P}\text{-f.v.}} X}$, wenn für alle $\varepsilon \in \mathbb{R}^+_{\setminus 0}$ gilt $\sum_{n\in\mathbb{N}} \mathbb{P}\big(|X_n - X| > \varepsilon\big) < \infty$.

- *stochastisch* gegen die numerische Zufallsvariable $X \in \overline{\mathscr{A}}$, kurz $\boxed{X_n \xrightarrow{\mathbb{P}} X}$, wenn für alle $\varepsilon \in \mathbb{R}^+_{\setminus 0}$ gilt $\lim_{n\to\infty} \mathbb{P}\big(|X_n - X| > \varepsilon\big) = 0$.

(b) Eine Folge $(X_n)_{n\in\mathbb{N}}$ in $\mathscr{L}_p$, $p \in \overline{\mathbb{R}}^+_{\setminus 0}$, *konvergiert in $\mathscr{L}_p$* gegen $X \in \mathscr{L}_p$, kurz $\boxed{X_n \xrightarrow{\mathscr{L}_p} X}$, wenn gilt $\lim_{n\to\infty}\|X_n - X\|_{\mathscr{L}_p} = 0$.

(c) Eine Folge $(\mathbb{P}_n)_{n\in\mathbb{N}}$ von Wahrscheinlichkeitsmaßen auf $(\mathbb{R}^k, \mathscr{B}^k)$ *konvergiert schwach* gegen ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\mathbb{R}^k, \mathscr{B}^k)$, wenn $\lim_{n\to\infty} \mathbb{E}_{\mathbb{P}_n}(h) = \mathbb{E}_{\mathbb{P}}(h)$ für alle $h \in \mathcal{C}_b$ gilt. Wir schreiben kurz $\boxed{\mathbb{P}_n \xrightarrow{w} \mathbb{P}}$ oder $\mathbb{P} = w\text{-}\lim_{n\to\infty} \mathbb{P}_n$.

(d) Eine Folge $(X_n)_{n\in\mathbb{N}}$ von $\mathbb{R}^k$-wertigen Zufallsvektoren *konvergiert in Verteilung* gegen einen $\mathbb{R}^k$-wertigen Zufallsvektor X, kurz $\boxed{X_n \xrightarrow{D} X}$, wenn $\lim_{n\to\infty} \mathbb{E}\big(h(X_n)\big) = \mathbb{E}\big(h(X)\big)$ für alle $h \in \mathcal{C}_b$, also $\mathbb{P}^X = w\text{-}\lim_{n\to\infty} \mathbb{P}^{X_n}$ gilt.

(e) Für eine Folge $(X_n)_{n\in\mathbb{N}}$ definieren wir *Konvergenz in Verteilung* mit einem Wahrscheinlichkeitsmaß $\mathbb{P}$ als Grenzwert, kurz $\boxed{X_n \xrightarrow{D} \mathbb{P}}$, allgemein durch $\mathbb{P}^{X_n} \xrightarrow{w} \mathbb{P}$. □

A06.03 **Eigenschaft.** *Es seien X, X_n und Y_n, $n \in \mathbb{N}$, reelle Zufallsvariablen.*

(i) *Sei $h : \mathbb{R} \to \mathbb{R}$ eine stetige Funktion und $(X_n)_{n\in\mathbb{N}}$ konvergiere $\mathbb{P}$-f.s. (stochastisch bzw. in Verteilung) gegen X. Dann konvergiert $(h(X_n))_{n\in\mathbb{N}}$ auch gegen $h(X)$ $\mathbb{P}$-f.s. (stochastisch bzw. in Verteilung).*

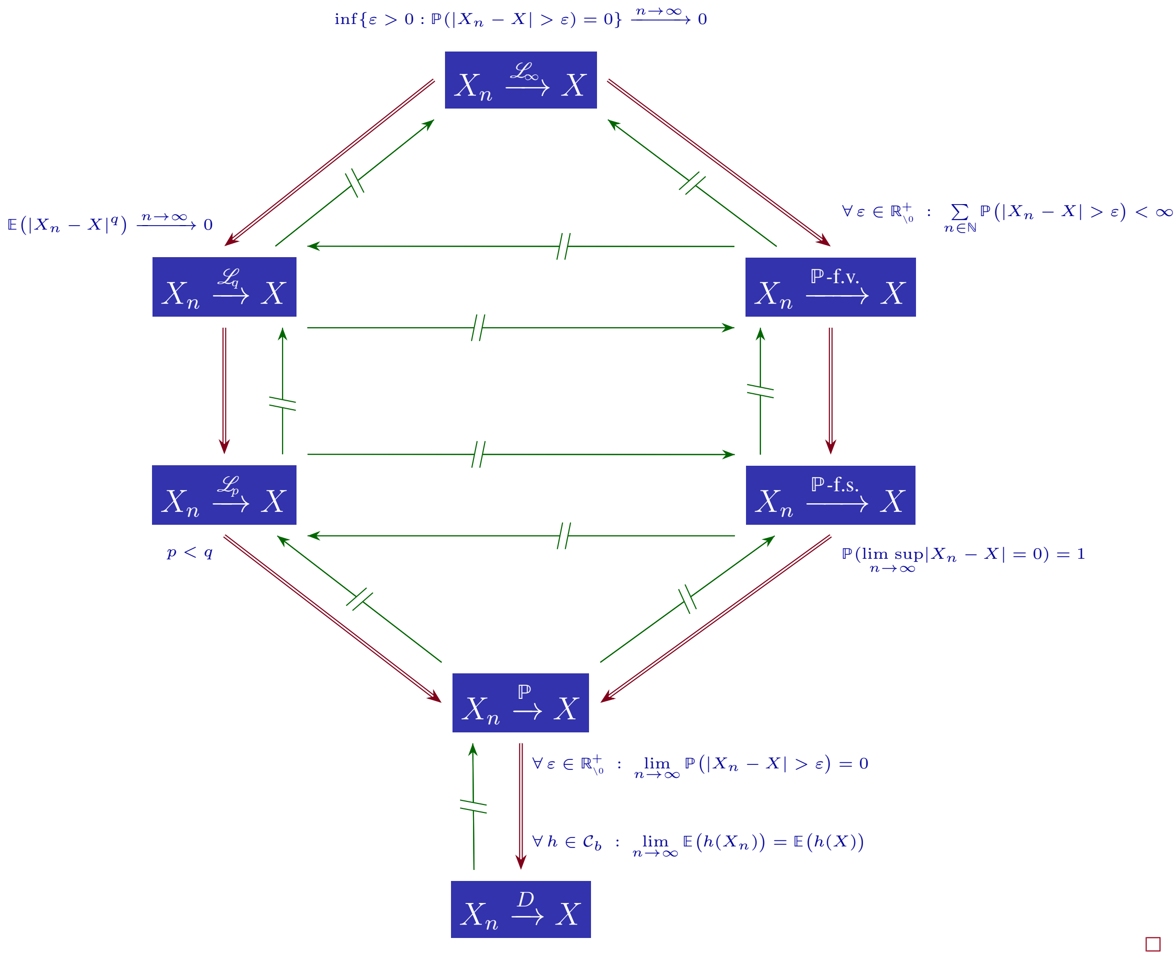
(ii) *Konvergiert $(Y_n)_{n\in\mathbb{N}}$ in Verteilung gegen X, also $Y_n \xrightarrow{D} X$, und konvergiert $(X_n - Y_n)_{n\in\mathbb{N}}$ stochastisch gegen Null, also $|X_n - Y_n| \xrightarrow{\mathbb{P}} 0$, dann konvergiert $(X_n)_{n\in\mathbb{N}}$ in Verteilung auch gegen X, also $X_n \xrightarrow{D} X$. Falls $X_n \xrightarrow{D} X$ und $Y_n \xrightarrow{\mathbb{P}} a$, so gilt auch $X_n + Y_n \xrightarrow{D} X + a$ sowie $Y_nX_n \xrightarrow{D} aX$.*

(iii) *$(X_n)_{n\in\mathbb{N}}$ konvergiert in Verteilung gegen eine Konstante $a \in \mathbb{R}$, also $X_n \xrightarrow{D} a$, genau dann, wenn sie stochastisch gegen a konvergiert, also $X_n \xrightarrow{\mathbb{P}} a$.*

(iv) *(Delta Methode) Seien $x \in \mathbb{R}$ und $a_n \uparrow \infty$ mit $a_n(X_n - x) \xrightarrow{D} X$. Dann gilt $X_n \xrightarrow{\mathbb{P}} x$. Ist weiterhin $f \in \mathscr{B}$ differenzierbar in x, so gilt $a_n(f(X_n) - f(x)) - a_n(X_n - x)f'(x) \xrightarrow{\mathbb{P}} 0$ und somit auch $a_n(f(X_n) - f(x)) \xrightarrow{D} f'(x)Y$.*

(v) *(Cramér-Wold device) Für eine Folge $\mathbb{R}^d$-wertiger Zufallsvektoren $(X_n)_{n\in\mathbb{N}}$ sind äquivalent: (a) Es gibt einen Zufallsvektor X mit $X_n \xrightarrow{D} X$. (b) Für jedes $v \in \mathbb{R}^d$ existiert ein X^v mit $\langle v, X_n\rangle \xrightarrow{D} X^v$. Falls (a) und (b) gelten, so sind X^v und $\langle v, X\rangle$ identisch verteilt.*

(vi) *Gegenbeispiele zeigen, dass die Umkehrungen (in grün) der folgenden direkten Implikationen (in rot) nicht gelten.*



□

A06.04 **Definition**. Eine Familie $(X_{n,j})_{j\in[\![n]\!],n\in\mathbb{N}}$ reeller Zufallsvariablen in $\mathscr{L}_2$ heißt *standardisiertes Dreiecksschema*, wenn (i) $(X_{n,j})_{j\in[\![n]\!]}$ sind unabhängig; sowie (ii) $\mathbb{E}(X_{n,j}) = 0$, $j \in [\![n]\!]$, und $\sum_{j=1}^n \mathbb{V}\mathrm{ar}(X_{n,j}) = 1$ für jedes $n \in \mathbb{N}$ gilt.
Ein standardisiertes Dreiecksschema $(X_{n,j})_{j\in[\![n]\!],n\in\mathbb{N}}$ erfüllt

(a) die *Lindeberg-Bedingung*, wenn $\lim_{n\to\infty} \sum_{j=1}^n \mathbb{E}\big(X_{n,j}^2 \mathbb{1}_{\{|X_{n,j}|\geqslant\delta\}}\big) = 0$ für jedes $\delta \in \mathbb{R}^+_{\setminus 0}$ gilt;

(b) die *Lyapunov-Bedingung*, wenn $\lim_{n\to\infty} \sum_{j=1}^n \mathbb{E}\big(|X_{n,j}|^{2+\delta}\big) = 0$ für ein $\delta \in \mathbb{R}^+_{\setminus 0}$ gilt. □

A06.05 **Eigenschaft**. *Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge unabhängiger reeller Zufallsvariablen.*

(i) *(Starkes Gesetz der großen Zahlen) Seien X_n, $n \in \mathbb{N}$, identisch-verteilt. $X_1 \in \mathscr{L}_1$ gilt genau* *dann, wenn* $\lim_{n\to\infty} \overline{X}_n = \mathbb{E}(X_1)$ $\mathbb{P}$-f.s. *(und dann auch in* $\mathscr{L}_1$*).*

(ii) *(Lévy's Äquivalenzsatz) Die Folge der Partialsummen* $(S_n)_{n\in\mathbb{N}}$ *mit* $S_n := \sum_{i=1}^n X_i$, $n \in \mathbb{N}$, *konvergiert* $\mathbb{P}$-f.s. *genau dann, wenn sie stochastisch konvergiert. Andernfalls ist* $(S_n)_{n\in\mathbb{N}}$ *divergent mit Wahrscheinlichkeit Eins.*

(Dreireihensatz von Kolmogorov) $(S_n)_{n\in\mathbb{N}}$ *konvergiert* $\mathbb{P}$-f.s. *genau dann, wenn die folgenden drei Bedingungen für ein* $\varepsilon \in \mathbb{R}^+_{\setminus 0}$ *gelten:* (a) $\sum_{n\in\mathbb{N}} \mathbb{P}(|X_n| > \varepsilon) < \infty$; (b) $\sum_{n\in\mathbb{N}} \mathbb{E}(X_n \mathbb{1}_{\{|X_n|\leqslant\varepsilon\}})$ *konvergiert; und* (c) $\sum_{n\in\mathbb{N}} \mathbb{V}\mathrm{ar}(X_n \mathbb{1}_{\{|X_n|\leqslant\varepsilon\}}) < \infty$.

Sei $(X_{n,j})_{j\in[\![n]\!],n\in\mathbb{N}}$ *ein standardisiertes Dreiecksschema.*

(i) *Erfüllt* $(X_{n,j})_{j\in[\![n]\!],n\in\mathbb{N}}$ *die Lyapunov-Bedingung, so auch die Lindeberg-Bedingung.*

(ii) *(Zentraler Grenzwertsatz nach Lindeberg (1922)) Erfüllt* $(X_{n,j})_{j\in[\![n]\!],n\in\mathbb{N}}$ *die Lindeberg-Bedingung, so gilt für (die Zeilensumme)* $S_n^* = \sum_{j=1}^n X_{nj} \xrightarrow{D} \mathrm{N}_{(0,1)}$. □

Literaturverzeichnis

H. Bauer. *Maß- und Integrationstheorie*. Berlin etc.: Walter de Gruyter, 2., überarbeitete Auflage, 1992.

J. Elstrodt. *Maß- und Integrationstheorie.* Berlin: Springer, 7., überarbeitete und ergänzte Auflage, 2011.

H.-O. Georgii. *Stochastik. Einführung in die Wahrscheinlichkeitstheorie und Statistik.* Berlin: De Gruyter, 5., überarbeitete und ergänzte Auflage, 2015.

O. Kallenberg. *Foundations of Modern Probability*. Springer, Berlin, Heidelberg, 2002.

A. Klenke. *Wahrscheinlichkeitstheorie*. Springer Spektrum, 4., überarbeitete und ergänzte Auflage, 2020.

U. Krengel. *Einführung in die Wahrscheinlichkeitstheorie und Statistik.* Braunschweig: Vieweg, 8., erweiterte Auflage, 2005.

J. Neveu. *Discrete-parameter martingales.*, volume 10 of *North-Holland mathematical library*. Elsevier, 1975.

H. Witting. Mathematische Statistik I: Parametrische Verfahren bei festem Stichprobenumfang. Stuttgart: B. G. Teubner., 1985.