Stochastik

Skript zur Vorlesung von
Prof. Dr. Michael Kohler
Sommersemester 2007

1. Grundbegriffe der Maßtheorie

Ziel: Konstruktion von Maßzahlen (wie z. B. Wahrscheinlichkeit / Länge / Fläche / Volumen / ...) von Mengen $A \subseteq \Omega$, wobei $\Omega \neq \emptyset$ fest.

Im Folgenden ist

$$\mathcal{P}(\Omega) := \{A : A \subseteq \Omega\}$$

die sogenannte **Potenzmenge** von Ω , die “abgeschlossen” ist bzgl. beliebigen Mengenoperationen (z. B. Vereinigung, Durchschnitt, Komplement, ...).

Im Allgemeinen ist es nicht möglich, eine sinnvolle Maßzahl für jede Menge aus $\mathcal{P}(\Omega)$ festzulegen. Statt dessen bestimmt man diese nur für Mengen aus einem vorgegebenen Mengensystem $\mathcal{A}$. Für $\mathcal{A}$ wählt man meist sogenannte **σ -Algebren**, die “abgeschlossen” sind bzgl. **abzählbar** vielen der üblichen Mengenoperationen (Vereinigung, Durchschnitt, Komplement).

1.1 Mengensysteme

Schreibweisen:

$\cup$... Vereinigung

$\sum, +$... Vereinigung paarweiser disjunkter Mengen

$\cap$... Schnitt

$\setminus$... Differenz

$A^c = \Omega \setminus A$... Komplement

Zentraler Begriff ist der Begriff der σ -Algebra. Zur Konstruktion und Nachweis von Eigenschaften werden weitere Hilfsbegriffe verwendet.

Definition 1.1

Nichtleere Menge Ω . System $\mathcal{A}$ von Teilmengen von Ω .

a) $\mathcal{A}$ heißt **Ring**, wenn gilt

1.) $\emptyset \in \mathcal{A}$,

2.) $A, B \in \mathcal{A} \Rightarrow A \setminus B \in \mathcal{A}$,

3.) $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$.

b) $\mathcal{A}$ heißt **Algebra**, wenn gilt

1.) $\emptyset \in \mathcal{A}$,

2.) $A \in \mathcal{A} \Rightarrow A^c := \Omega \setminus A \in \mathcal{A}$,

$$3.) A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}.$$

c) $\mathcal{A}$ heißt **σ -Ring**, wenn gilt:

$$1.) \emptyset \in \mathcal{A},$$

$$2.) A, B \in \mathcal{A} \Rightarrow A \setminus B \in \mathcal{A}$$

$$3.) A_n \in \mathcal{A} (n \in \mathbb{N}) \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$$

d) $\mathcal{A}$ heißt **σ -Algebra**, wenn gilt

$$1.) \emptyset \in \mathcal{A}$$

$$2.) A \in \mathcal{A} \Rightarrow A^c := \Omega \setminus A \in \mathcal{A}$$

$$3.) A_n \in \mathcal{A} (n \in \mathbb{N}) \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$$

Lemma 1.2

$$a) \mathcal{A} \text{ Ring, } A_1, \dots, A_m \in \mathcal{A} \Rightarrow \bigcup_{n=1}^m A_n \in \mathcal{A} \text{ und } \bigcap_{n=1}^m A_n \in \mathcal{A}$$

$$b) \mathcal{A} \text{ Algebra} \Leftrightarrow \mathcal{A} \text{ Ring und } \Omega \in \mathcal{A}$$

Beweis

a) Für $m = 2$ gilt $A_1 \cup A_2 \in \mathcal{A}$ nach Definition und $A_1 \cap A_2 = A_1 \setminus (A_1 \setminus A_2) \in \mathcal{A}$ nach zweimaliger Anwendung von 2.)

Allgemeiner Fall folgt mit vollständiger Induktion.

b) “ $\Rightarrow$ ”: Sei $\mathcal{A}$ Algebra. Dann gilt

$$\Omega = \Omega \setminus \emptyset \in \mathcal{A} \quad \text{nach 1.) und 2.),}$$

und für $A, B \in \mathcal{A}$ ist

$$A \setminus B = A \cap B^c = (A^c \cup B)^c \in \mathcal{A}$$

nach 2.) und 3.), wobei die zweite Gleichheit aus den de Morganschen Regeln folgt.

“ $\Leftarrow$ ”: Sei $\mathcal{A}$ Ring mit $\Omega \in \mathcal{A}$. Dann gilt für $A \in \mathcal{A}$

$$A^c = \Omega \setminus A \in \mathcal{A}$$

nach 2.) (da $A, \Omega \in \mathcal{A}$). □

Folgerung 1.3

Eine Algebra ist ein Mengensystem $\mathcal{A}$, das $\emptyset$ und Ω enthält, und bei dem Differenzbildung, endliche Vereinigungsbildung und endliche Durchschnittsbildung nicht aus $\mathcal{A}$ herausführt. Wendet man also endlich viele der üblichen Mengenoperationen auf Mengen aus $\mathcal{A}$ an, so erhält man eine Menge, die wieder in $\mathcal{A}$ liegt.

Bemerkung 1.4

- a) $\mathcal{A}$ σ -Ring $\Rightarrow \mathcal{A}$ Ring
- b) $\mathcal{A}$ σ -Algebra $\Rightarrow \mathcal{A}$ Algebra

Beweis

Folgt aus $A \cup B = A \cup B \cup \emptyset \cup \emptyset \cup \dots$ und $\emptyset \in \mathcal{A}$ □

Umkehrung gilt i. A. nicht!

Lemma 1.5

- a) $\mathcal{A}$ σ -Ring, $A_n \in \mathcal{A}$ ($n \in \mathbb{N}$) $\Rightarrow \bigcap_{n \in \mathbb{N}} A_n \in \mathcal{A}$
- b) $\mathcal{A}$ σ -Algebra $\Leftrightarrow \mathcal{A}$ σ -Ring und $\Omega \in \mathcal{A}$.

Beweis

- a) Wegen $\bigcap_{n \in \mathbb{N}} A_n \subseteq A_1$ gilt nach den de Morganschen Regeln

$$\bigcap_{n \in \mathbb{N}} A_n = A_1 \setminus (A_1 \setminus \bigcap_{n \in \mathbb{N}} A_n) = A_1 \setminus \left(\bigcup_{n \in \mathbb{N}} A_1 \setminus A_n \right) \in \mathcal{A},$$

(denn $A_1 \setminus A_n \in \mathcal{A} \Rightarrow \bigcup_{n \in \mathbb{N}} A_1 \setminus A_n \in \mathcal{A}$ nach 3.), dann noch 2.) anwenden).

- b) Folgt wie in Lemma 1.2 b) □

Folgerung 1.6

Eine σ -Algebra ist ein Mengensystem $\mathcal{A}$, das $\emptyset$ und Ω enthält, und bei dem Differenzbildung, abzählbare Vereinigungsbildung und abzählbare Durchschnittsbildung nicht aus $\mathcal{A}$ herausführen. Bei einer σ -Algebra führen also endlich oder abzählbar unendlich viele der üblichen Mengenoperationen nicht aus der σ -Algebra heraus.

Zusammenhang der Begriffe:

$$\begin{array}{ccc} \sigma\text{-Algebra} & \Rightarrow & \text{Algebra} \\ \Downarrow & & \Downarrow \\ \sigma\text{-Ring} & \Rightarrow & \text{Ring} \end{array}$$

Definition 1.7

Ist Ω eine nichtleere Menge und $\mathcal{A}$ eine σ -Algebra in Ω , so heißt $(\Omega, \mathcal{A})$ **Messraum** und die Elemente $A \in \mathcal{A}$ heißen **messbare Mengen**.

Beispiele

- 1.) $\{\emptyset, \Omega\}$ ist σ -Algebra (und damit auch σ -Ring, Algebra und Ring).
- 2.) Für $\emptyset \subsetneq A \subsetneq \Omega$ ist $\{\emptyset, A, A^c, \Omega\}$ σ -Algebra, aber $\{\emptyset, A, A^c\}$ keine Algebra und kein Ring (da $A \cup A^c \notin \{\emptyset, A, A^c\}$).
- 3.) $\mathcal{P}(\Omega)$ ist σ -Algebra.
- 4.) Das System aller endlichen Teilmengen von Ω ist ein Ring, aber i. A. kein σ -Ring (vgl. Übungen).
- 5.) Sei $\Omega = \mathbb{R}^n$ und für $a = (a_1, \dots, a_n)^T \in \Omega, b = (b_1, \dots, b_n)^T \in \Omega$ mit $a_i \leq b_i$ ($i = 1, \dots, n$) sei

$$(a, b] := \{(x_1, \dots, x_n)^T \in \mathbb{R}^n : a_i < x_i \leq b_i \ (i = 1, \dots, n)\}.$$

Sei $\mathcal{E}_n$ das System aller elementaren Figuren, d. h. aller endlicher Summen von halboffenen Intervallen $(a, b]$ definiert wie oben. Hierbei ist eine Summe eine Vereinigung, bei der paarweise Disjunktheit der zu vereinigenden Mengen vorausgesetzt wird.

Lemma 1.8

Das System $\mathcal{E}_n$ der elementaren Figuren in $\mathbb{R}^n$ (definiert als System aller endlicher Summen von halboffenen Intervallen $(a, b]$) ist ein Ring.

Beweis

- 1.) $\emptyset = (a, a] \in \mathcal{E}_n$
- 2.) Seien $A, B \in \mathcal{E}_n$. Zu zeigen: $A \setminus B \in \mathcal{E}_n$.
Es gelte $A = A_1 + \dots + A_p, B = B_1 + \dots + B_q$, wobei “+” für Vereinigung bei paarweiser Disjunktheit steht. Wir zeigen $A \setminus B \in \mathcal{E}_n$ mit Induktion nach q :
“ $q = 1$ ” : $A \setminus B_1 = (A_1 \setminus B_1) + \dots + (A_p \setminus B_1) \in \mathcal{E}_n$
da $(a_1, b_1] \setminus (a_2, b_2] = (a_1, b_1] \setminus \underbrace{((a_1, b_1] \cap (a_2, b_2])}_{\text{wieder von der Form } (a, b]}}$
sich als endliche Summe halboffener Intervalle darstellen lässt.
“ $q \rightarrow q + 1$ ” : $A \setminus (B_1 + \dots + B_{q+1}) = (A \setminus (B_1 + \dots + B_q)) \setminus B_{q+1} \in \mathcal{E}$ nach Induktionsvoraussetzung sowie dem Fall $q = 1$.

3.) Seien $A, B \in \mathcal{E}_n$. Zu zeigen: $A \cup B \in \mathcal{E}_n$.

Folgt mit $A \cup B = A + B \setminus A, B \setminus A \in \mathcal{E}_n$ nach 2. und Definition von $\mathcal{E}_n$. $\square$

Lemma 1.9

$\mathcal{A}$ σ -Algebra $\Leftrightarrow \mathcal{A}$ Algebra und für paarweise disjunkte $A_n \in \mathcal{A}$ ($n \in \mathbb{N}$)
gilt immer auch $\sum_{n \in \mathbb{N}} A_n \in \mathcal{A}$

Beweis: Siehe Übungen $\square$

Nachweis der Eigenschaften einer σ -Algebra oft möglich über Umweg über Dynkin-Systeme.

Definition 1.10

$\Omega \neq \emptyset$. System $\mathcal{D}$ von Teilmengen von Ω heißt **Dynkin-System**, wenn gilt

1. $\Omega \in \mathcal{D}$
2. $A, B \in \mathcal{D}, A \subseteq B \Rightarrow B \setminus A \in \mathcal{D}$
3. $A_n \in \mathcal{D}$ ($n \in \mathbb{N}$) paarweise disjunkt $\Rightarrow \sum_{n \in \mathbb{N}} A_n \in \mathcal{D}$

Beispiel:

Für $\Omega = \{1, 2, 3, 4\}$ ist

$$\mathcal{A} = \{\emptyset, \Omega, \{1, 2\}, \{3, 4\}, \{1, 3\}, \{2, 4\}\}$$

ein Dynkin-System, aber wegen $\{1, 2\} \cup \{2, 4\} \notin \mathcal{A}$ kein Ring, keine Algebra und keine σ -Algebra.

Lemma 1.11

- a) $\mathcal{R}_\alpha$ Ring in Ω ($\alpha \in I \neq \emptyset$) $\Rightarrow \bigcap_{\alpha \in I} \mathcal{R}_\alpha$ Ring in Ω .
- b) $\mathcal{E}$ Mengensystem in Ω . Dann existiert ein "kleinster" Ring, der $\mathcal{E}$ umfasst, d. h. es existiert ein Ring, der $\mathcal{E}$ umfasst und in allen Ringen enthalten ist, die $\mathcal{E}$ umfassen.

Analoge Aussage gilt für σ -Ringe, Algebren, σ -Algebren und Dynkin-Systeme.

Beweis:

- a) 1. $\emptyset \in \bigcap_{\alpha \in I} \mathcal{R}_\alpha$, da $\emptyset \in \mathcal{R}_\alpha$ für alle $\alpha \in I$ und $I \neq \emptyset$.
2. $A, B \in \bigcap_{\alpha \in I} \mathcal{R}_\alpha \Rightarrow A, B \in \mathcal{R}_\alpha$ für alle $\alpha \in I \Rightarrow A \setminus B \in \mathcal{R}_\alpha$ für alle $\alpha \in I \Rightarrow A \setminus B \in \bigcap_{\alpha \in I} \mathcal{R}_\alpha$
3. analog.

b) $I := \{\mathcal{R} \subseteq \mathcal{P}(\Omega) : \mathcal{R} \text{ Ring mit } \mathcal{E} \subseteq \mathcal{R}\}$

$I \neq \emptyset$, da $\mathcal{P}(\Omega) \in I$

Dann ist $\bigcap_{\mathcal{R} \in I} \mathcal{R}$ ein Ring (siehe a)), der $\mathcal{E}$ umfasst (nach Definition), und der in jedem Ring enthalten ist, der $\mathcal{E}$ umfasst (da jeder solcher Ringe in obigem Schnitt auftaucht).

□

Definition 1.12

Sei $\mathcal{E}$ Mengensystem in Ω . Der kleinste Ring, der $\mathcal{E}$ umfasst, heißt **der von $\mathcal{E}$ erzeugte Ring**. $\mathcal{E}$ heißt **Erzeugersystem** dieses Rings.

Analog für σ -Ring, Algebren, σ -Algebren und Dynkin-Systeme.

Bezeichnungen

$\mathcal{F}(\mathcal{E}) = \mathcal{F}_\Omega(\mathcal{E}) \dots$ die von $\mathcal{E}$ (in Ω) erzeugte σ -Algebra.

$\mathcal{D}(\mathcal{E}) = \mathcal{D}_\Omega(\mathcal{E}) \dots$ das von $\mathcal{E}$ (in Ω) erzeugte Dynkin-System.

Beispiel

1. $\mathcal{E} = \sigma\text{-Algebra} \Rightarrow \mathcal{F}(\mathcal{E}) = \mathcal{E}$

2. $\emptyset \subsetneq A \subsetneq \Omega \Rightarrow \mathcal{F}(\{A\}) = \{\emptyset, A, A^c, \Omega\}$.

Definition 1.13

Sei $\mathcal{O}_n$ das System der offenen Mengen im euklidischen Raum $\mathbb{R}^n$. Dann heißt

$$\mathcal{B}_n := \mathcal{F}(\mathcal{O}_n)$$

das System der **Borelschen Mengen** in $\mathbb{R}^n$. $\mathcal{B}_n$ wird als σ -Algebra der **Borelschen Mengen** in $\mathbb{R}^n$ bezeichnet.

Bezeichnung: $\mathcal{B} = \mathcal{B}_1$

Bemerkung 1.14

a) $\mathcal{B}_n$ enthält alle abgeschlossenen Mengen (da jede abgeschlossene Menge das Komplement einer offenen Menge ist), alle Einpunktmengen (da abgeschlossen) und alle abzählbaren Mengen.

b) Man kann zeigen: $\mathcal{B}_n \subsetneq \mathcal{P}(\mathbb{R}^n)$ (ohne Beweis).

Satz 1.15

Die folgenden Mengensysteme sind Erzeuger von $\mathcal{B}_n$:

a) Das System $\mathcal{J}_n$ der halboffenen Intervalle

$$(a, b] = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n : a_i < x_i \leq b_i \ (i = 1, \dots, n) \right\}$$

b) Das System der abgeschlossenen Mengen in $\mathbb{R}^n$.

c) Das System der abgeschlossenen Intervalle in $\mathbb{R}^n$.

d) Das System der offenen Intervalle in $\mathbb{R}^n$.

e) Das System der kompakten Mengen in $\mathbb{R}^n$.

f) Das System der Intervalle (a, ∞) in $\mathbb{R}^n$.

g) Das System der Intervalle $(-\infty, a)$ in $\mathbb{R}^n$.

Beweis von Satz 1.15

Nur Beweis von a) im Fall $n = 2$, Rest analog.

Zu zeigen: $\mathcal{F}(\mathcal{J}_2) = \mathcal{F}(\mathcal{O}_2)$ ($= \mathcal{B}$)

α) Wir zeigen: $\mathcal{F}(\mathcal{J}_2) \subseteq \mathcal{F}(\mathcal{O}_2)$.

Da $\mathcal{F}(\mathcal{O}_2)$ σ -Algebra ist, genügt es zu zeigen:

$$\mathcal{J}_2 \subseteq \mathcal{F}(\mathcal{O}_2)$$

Sei $(a, b] \in \mathcal{J}_2$ beliebig. Dann gilt

$$(a, b] = (a_1, b_1] \times (a_2, b_2] = \bigcap_{k \in \mathbb{N}} (a_1, b_1 + \frac{1}{k}) \times (a_2, b_2 + \frac{1}{k}) \in \mathcal{F}(\mathcal{O}_2),$$

$$\text{nach Lemma 1.5, da } (a_1, b_1 + \frac{1}{k}) \times (a_2, b_2 + \frac{1}{k}) \in \mathcal{O}_2 \subseteq \mathcal{F}(\mathcal{O}_2)$$

für alle $k \in \mathbb{N}$ gilt und $\mathcal{F}(\mathcal{O}_2)$ σ -Algebra ist.

β) Wir zeigen: $\mathcal{F}(\mathcal{O}_2) \subseteq \mathcal{F}(\mathcal{J}_2)$

Wieder genügt es, dazu zu zeigen: $\mathcal{O}_2 \subseteq \mathcal{F}(\mathcal{J}_2)$

Dazu: Sei $D \in \mathcal{O}_2$ beliebig

Da D offen ist, existiert zu jedem $d \in D$ ein offenes Intervall (a, b) mit $d \in (a, b) \subseteq D$.

Durch Verkleinern des Intervalls kann man ε erreichen, dass die Endpunkte des Intervalls rational sind und $(a, b] \subseteq D$ gilt. Dann ist aber

$$D = \bigcup_{d \in D} (\text{zu } d \text{ gehörendes halboffenes Intervall mit rationalen Endpunkten})$$

als **abzählbare** Vereinigung halboffener Intervalle aus $\mathcal{J}_2$ in $\mathcal{F}(\mathcal{J}_2)$ enthalten.

□

Lemma 1.16

- a) $\mathcal{D}$ Dynkin-System $\Rightarrow \emptyset \in \mathcal{D}$ und aus $A \in \mathcal{D}$ folgt $A^c \in \mathcal{D}$
- b)
- $$\mathcal{A} \text{ } \sigma\text{-Algebra} \Leftrightarrow \mathcal{A} \text{ Dynkin-System und } \mathcal{A} \cup\text{-stabil}$$
- (d. h. aus $A, B \in \mathcal{A}$ folgt $A \cup B \in \mathcal{A}$)
- c)
- $$\mathcal{A} \text{ } \sigma\text{-Algebra} \Leftrightarrow \mathcal{A} \text{ Dynkin-System und } \mathcal{A} \cap\text{-stabil}$$
- (d. h. aus $A, B \in \mathcal{A}$ folgt $A \cap B \in \mathcal{A}$).

Beweis

- a) $\emptyset = \Omega \setminus \Omega \in \mathcal{D}$ nach Def. 1.) + 2.)
- $A \in \mathcal{D} \Rightarrow A^c = \Omega \setminus A \in \mathcal{D}$ nach Def. 2.), beachte $A \subseteq \Omega$ und $\Omega \in \mathcal{D}$.
- b) “ $\Rightarrow$ ” klar
- “ $\Leftarrow$ ” Sei $\mathcal{A}$ Dynkin-System und $\cup$ -stabil. Dann gilt:
- (1) $\emptyset \in \mathcal{A}$ nach a)
 - (2) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$ nach a)
 - (3) $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$ (da $\mathcal{A}$ $\cup$ -stabil)
 - (4) $A_n \in \mathcal{A}$ ($n \in \mathbb{N}$) paarweise disjunkt $\Rightarrow \sum_{n=1}^{\infty} A_n \in \mathcal{A}$ nach Definition Dynkin-System.
- Mit Lemma 1.9 folgt die Behauptung.
- c) “ $\Rightarrow$ ” klar
- “ $\Leftarrow$ ” Sei $\mathcal{A}$ $\cap$ -stabiles Dynkin-System. Wegen $A \cup B = (A^c \cap B^c)^c$ ist $\mathcal{A}$ $\cup$ -stabil, und mit b) folgt Beh.

□

Lemma 1.17

Ist $\mathcal{E}$ ein $\cap$ -stabiles Mengensystem, dann ist auch das von $\mathcal{E}$ erzeugte Dynkin-System $\mathcal{D}(\mathcal{E})$ $\cap$ -stabil.

Vor dem Beweis von Lemma 1.17 zeigen wir zunächst die folgende Folgerung aus Lemma 1.17:

Satz 1.18

Seien $\mathcal{A}, \mathcal{E}$ Mengensysteme in Ω .

- a) $\mathcal{E}$ $\cap$ -stabil $\Rightarrow \mathcal{F}(\mathcal{E}) = \mathcal{D}(\mathcal{E})$
- b) $\mathcal{E} \subseteq \mathcal{A}$, $\mathcal{A}$ Dynkin-System, $\mathcal{E}$ $\cap$ -stabil
 $\Rightarrow \mathcal{F}(\mathcal{E}) \subseteq \mathcal{A}$

Anwendung

Ist $\mathcal{E}$ $\cap$ -stabil, und haben die Mengen aus $\mathcal{E}$ die Eigenschaft E , dann haben auch die Mengen aus $\mathcal{F}(\mathcal{E})$ die Eigenschaft E , sofern das System aller Mengen mit Eigenschaft E ein Dynkin-System ist.

Beweis von Satz 1.18

- a) Sei $\mathcal{E}$ $\cap$ -stabil.
 - “ $\supseteq$ ” $\mathcal{E} \subseteq \mathcal{F}(\mathcal{E})$ und $\mathcal{F}(\mathcal{E})$ Dynkin-System nach Lemma 1.16 b)
 $\Rightarrow \mathcal{D}(\mathcal{E}) \subseteq \mathcal{F}(\mathcal{E})$ nach Definition.
 - “ $\subseteq$ ” $\mathcal{E} \subseteq \mathcal{D}(\mathcal{E})$, $\mathcal{D}(\mathcal{E})$ $\cap$ -stabil nach Lemma 1.17, $\mathcal{D}(\mathcal{E})$
Dynkin-System nach Definition $\Rightarrow \mathcal{D}(\mathcal{E})$ σ -Algebra nach
Lemma 1.16 c) $\Rightarrow \mathcal{F}(\mathcal{E}) \subseteq \mathcal{D}(\mathcal{E})$ nach Definition $\mathcal{F}(\mathcal{E})$.
- b) $\mathcal{E} \subseteq \mathcal{A}$ und $\mathcal{A}$ Dynkin-System $\Rightarrow \mathcal{D}(\mathcal{E}) \subseteq \mathcal{A}$ nach Definition. Wegen $\mathcal{E}$ $\cap$ -stabil gilt
nach a) $\mathcal{F}(\mathcal{E}) = \mathcal{D}(\mathcal{E}) \Rightarrow$ Beh.

□

Beweis von Lemma 1.17

Sei $\mathcal{E}$ $\cap$ -stabil.

Zu zeigen: $E, F \in \mathcal{D}(\mathcal{E}) \Rightarrow E \cap F \in \mathcal{D}(\mathcal{E})$

1. Schritt: Sei $E \in \mathcal{E}$. Gezeigt wird:

$$E \cap F \in \mathcal{D}(\mathcal{E}) \text{ für alle } F \in \mathcal{D}(\mathcal{E}).$$

Setze

$$\mathcal{G}_E := \{\tilde{F} \in \mathcal{P}(\Omega) : E \cap \tilde{F} \in \mathcal{D}(\mathcal{E})\}.$$

Behauptung folgt aus

$$\mathcal{D}(\mathcal{E}) \subseteq \mathcal{G}_E \quad (\text{typischer Beweisschritt!})$$

Nach Definition von $\mathcal{D}(\mathcal{E})$ folgt dies wiederum aus:

(1) $\mathcal{E} \subseteq \mathcal{G}_E$

(2) $\mathcal{G}_E$ ist Dynkin-System.

Nachweis von (1): $\tilde{F} \in \mathcal{E}$. Da $\mathcal{E}$ $\cap$ -stabil und $E \in \mathcal{E}$ nach Annahme oben ist, gilt dann $E \cap \tilde{F} \in \mathcal{E} \subseteq \mathcal{D}(\mathcal{E}) \Rightarrow \tilde{F} \in \mathcal{G}_E$

Nachweis von (2):

(1) $\Omega \in \mathcal{G}_E$, da $E \cap \Omega = E \in \mathcal{E} \subseteq \mathcal{D}(\mathcal{E})$

(2) Seien $A, B \in \mathcal{G}_E$ mit $A \subseteq B$.

$E \cap (B \setminus A) = (E \cap B) \setminus (E \cap A) \in \mathcal{D}(\mathcal{E})$ da $E \cap B \in \mathcal{D}(\mathcal{E})$, $E \cap A \in \mathcal{D}(\mathcal{E})$, $E \cap A \subseteq E \cap B$ und $\mathcal{D}(\mathcal{E})$ Dynkin-System
 $\Rightarrow B \setminus A \in \mathcal{G}_E$.

(3) Seien $A_n \in \mathcal{G}_E$ ($n \in \mathbb{N}$) paarweise disjunkt

$\Rightarrow E \cap A_n \in \mathcal{D}(\mathcal{E})$ (nach Def. $\mathcal{G}_E$) und paarweise disjunkt

$\Rightarrow E \cap \left(\sum_{n=1}^{\infty} A_n \right) = \sum_{n=1}^{\infty} (E \cap A_n) \in \mathcal{D}(\mathcal{E})$ (nach Definition Dynkin-System).

$\Rightarrow \sum_{n=1}^{\infty} A_n \in \mathcal{G}_E$ $\rightsquigarrow$ (2).

Bemerkung: Beweis zeigt, dass $\mathcal{G}_E$ auch Dynkin-System ist für $E \in \mathcal{D}(\mathcal{E})$.

2. Schritt: Sei $E \in \mathcal{D}(\mathcal{E})$

Wir zeigen wieder:

$$E \cap F \in \mathcal{D}(\mathcal{E}) \quad \text{für alle } F \in \mathcal{D}(\mathcal{E})$$

.

Dazu genügt es wieder zu zeigen:

$$\mathcal{D}(\mathcal{E}) \subseteq \mathcal{G}_E := \{\tilde{F} \in \mathcal{P}(\Omega) : E \cap \tilde{F} \in \mathcal{D}(\mathcal{E})\}$$

Dies wiederum folgt aus:

(1) $\mathcal{E} \subseteq \mathcal{G}_E$

(2) $\mathcal{G}_E$ Dynkin-System.

Nachweis von (2) erfolgt wie im 1. Schritt.

Nachweis von (1):

Sei $\tilde{F} \in \mathcal{E}$. Nach Schritt 1 gilt dann

$$\tilde{F} \cap E \in \mathcal{D}(\mathcal{E}), \text{ also } \tilde{F} \in \mathcal{G}_E$$

$\rightsquigarrow$ (1) $\rightarrow$ Beh. $\square$

1.2 Inhalte und Maße

Bezeichnungen: $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$

Rechenregeln: $a \pm \infty = \pm\infty = \pm\infty + a$ für $a \in \mathbb{R}$

$$\infty + \infty = \infty, -\infty - \infty = -\infty$$

Nicht definiert sind: $\infty - \infty, \frac{\infty}{\infty}, \frac{0}{0}$.

Im Folgenden: Auflistung von Eigenschaften, die bei Zuweisungen wie

$$A \mapsto \text{Volumen von } A \quad \text{bzw.} \quad A \mapsto \text{Wahrscheinlichkeit von } A$$

sinnvollerweise vorliegen sollten.

Definition 1.19

$\mathcal{C}$ Mengensystem in Ω . $\varphi : \mathcal{C} \rightarrow \overline{\mathbb{R}}$ ist

a) **positiv** : $\Leftrightarrow \varphi(A) \geq 0$ für alle $A \in \mathcal{C}$

b) **monoton**: $\Leftrightarrow$ Für $A, B \in \mathcal{C}$ mit $A \subseteq B$ gilt $\varphi(A) \leq \varphi(B)$

c) **endlich** : $\Leftrightarrow \varphi(A) \in \mathbb{R}$ für alle $A \in \mathcal{C}$

d) **additiv** : $\Leftrightarrow$ Für paarweise disjunkte $A_1, \dots, A_n \in \mathcal{C}$ mit $\sum_{i=1}^n A_i \in \mathcal{C}$ gilt

$$\varphi\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n \varphi(A_i).$$

e) **σ -additiv**: $\Leftrightarrow$ Für paarweise disjunkte $A_1, A_2, \dots \in \mathcal{C}$ mit $\sum_{n \in \mathbb{N}} A_n \in \mathcal{C}$ gilt

$$\varphi\left(\sum_{n \in \mathbb{N}} A_n\right) = \sum_{n=1}^{\infty} \varphi(A_n).$$

f) **σ -subadditiv**: $\Leftrightarrow$ Für alle $A_1, A_2, \dots \in \mathcal{C}$ mit $\bigcup_{n=1}^{\infty} A_n \in \mathcal{C}$ gilt

$$\varphi\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \varphi(A_n).$$

Bemerkung 1.20

a) Endlich $\neq$ beschränkt

b) In e) ist $\sum_{n=1}^{\infty} \varphi(A_n)$ unabhängig von der Summationsreihenfolge, da linke Seite unabhängig von Summationsreihenfolge ist.

Definition 1.21

$\mathcal{R}$ sei Ring über Ω .

Eine Mengenfunktion $\mu : \mathcal{R} \rightarrow \overline{\mathbb{R}}$ heißt

- a) **Inhalt** : $\Leftrightarrow \mu(\emptyset) = 0, \mu$ positiv und μ additiv.
- b) **Maß** : $\Leftrightarrow \mu(\emptyset) = 0, \mu$ positiv und μ σ -additiv

Bemerkung 1.22

- a) Häufig verwendet man σ -Algebren als Definitionsbereich von Maßen.
- b) Jedes Maß ist ein Inhalt, da

$$\mu(A_1 + \dots + A_n) = \mu(A_1 + \dots + A_n + \emptyset + \emptyset + \dots) = \mu(A_1) + \dots + \mu(A_n) + 0 + 0 + \dots$$

- c) Jeder Inhalt ist monoton, da

$$A \subseteq B \Rightarrow \mu(B) = \mu(B \setminus A) + \mu(A) \geq \mu(A),$$

wobei verwendet wurde, dass mit A, B auch $B \setminus A$ im Ring $\mathcal{R}$ liegt.

Beispiele für Inhalte

$$\text{a) } \Omega = \mathbb{N}, \mathcal{R} = \mathcal{P}(\mathbb{N}), \mu(A) = \begin{cases} 0 & , \quad |A| \text{ endlich} \\ \infty & , \quad \text{sonst} \end{cases} \quad (A \in \mathcal{R}).$$

μ ist Inhalt:

- $\mu(\emptyset) = 0$
- $\mu(A) \geq 0$ für alle $A \in \mathcal{R}$
- $\mu(A_1 + \dots + A_n) = \mu(A_1) + \dots + \mu(A_n)$, denn sind alle A_i endlich, so sind beide Seiten Null, andernfalls sind beide Seiten ∞ .

μ ist aber kein Maß, da

$$\infty = \mu(\mathbb{N}) \neq \sum_{n \in \mathbb{N}} \mu(\{n\}) = 0.$$

- b) $\Omega = \mathbb{R}^n, \mathcal{R} = \mathcal{E}_n =$ Ring der elementaren Figuren.

$m(A) =$ elementargeometrisches Volumen von A ($A \in \mathcal{E}_n$).

Man sieht informal sofort: m ist ein Inhalt.

Wir zeigen später: m ist sogar Maß.

Beispiele für Maße

a) triviale Maße:

$$\mu_1(A) = 0 \quad \text{für alle } A \in \mathcal{R}$$
$$\mu_2(A) = \begin{cases} 0 & \text{für } A = \emptyset \\ \infty & \text{für } A \neq \emptyset \end{cases}$$

b) $\mathcal{R}$ Ring über $\Omega, \omega \in \Omega$ fest.

$$\mu(A) = \begin{cases} 1 & \text{für } \omega \in A \\ 0 & \text{für } \omega \notin A \end{cases}$$

(sog. **Dirac-Maß**).

Anschaulich: Maß misst Masse einer Menge, bei Dirac-Maß ist alle Masse im Punkt ω konzentriert.

c) Ist $(\Omega, \mathcal{A})$ ein Meßraum, $Z = \{z_1, z_2, \dots\}$ eine höchstens abzählbare Teilmenge von Ω , so heißt $\mu : \mathcal{A} \rightarrow \overline{\mathbb{R}}$ mit

$$\mu(A) := \text{Anzahl der } z_i \in A, A \in \mathcal{A},$$

ein **abzählendes Maß** (counting measure).

Definition 1.23

Ω Menge, $\mathcal{A}$ σ -Algebra über $\Omega, \mu : \mathcal{A} \rightarrow \overline{\mathbb{R}}$ Maß. Dann heißt $(\Omega, \mathcal{A})$ **Messraum**, die Mengen $A \in \mathcal{A}$ heißen **messbare Mengen**, und $(\Omega, \mathcal{A}, \mu)$ heißt **Maßraum**.

Im Falle $\mu(\Omega) = 1$ heißt μ **Wahrscheinlichkeitsmaß** und $(\Omega, \mathcal{A}, \mu)$ **Wahrscheinlichkeitsraum**.

Im Folgenden behandeln wir einige Folgerungen aus der σ -Additivität.

Lemma 1.24

Sei $\mathcal{A}$ eine Algebra in Ω, μ ein Inhalt auf $\mathcal{A}$.

a) μ ist monoton, d.h. $[A, B \in \mathcal{A}, A \subset B \implies \mu(A) \leq \mu(B)]$.

b) μ ist subadditiv, d.h. $[A_1, \dots, A_m \in \mathcal{A} \implies \mu(\bigcup_{n=1}^m A_n) \leq \sum_{n=1}^m \mu(A_n)]$.

c) Ist μ Maß, dann ist μ sogar σ -subadditiv, d.h. $[A_n \in \mathcal{A} (n = 1, 2, \dots), \bigcup_{n=1}^{\infty} A_n \in \mathcal{A} \implies \mu(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu(A_n)]$.

Beweis:

a) $B = B \setminus A \cup B \implies \mu(B) = \mu(B \setminus A) + \mu(A) \geq \mu(A)$.

b)

$$\begin{aligned} \mu\left(\bigcup_{n=1}^m A_n\right) &= \mu\left(A_1 \cup A_2 \setminus A_1 \cup A_3 \setminus (A_1 \cup A_2) \cup \dots \cup A_m \setminus \left(\bigcup_{n=1}^{m-1} A_n\right)\right) \\ &= \mu(A_1) + \mu(A_2 \setminus A_1) + \dots + \mu\left(A_m \setminus \left(\bigcup_{n=1}^{m-1} A_n\right)\right) \\ &\stackrel{a)}{\leq} \mu(A_1) + \mu(A_2) + \dots + \mu(A_m) \end{aligned}$$

c) Folgt analog wie b) aus der σ -Additivität des Maßes

□

Bemerkung 1.25

$\mathcal{A}$ Algebra in Ω , μ endlicher Inhalt auf $\mathcal{A}$.

$A, B \in \mathcal{A}$, $A \subset B \implies \mu(B \setminus A) = \mu(B) - \mu(A)$... Subtraktivität.

Definition 1.26

A_n ($n \in \mathbb{N}$) und A seien Mengen

a) $(A_n)_{n \in \mathbb{N}}$ **konvergiert von unten gegen** A (kurz: $A_n \uparrow A$) : $\Leftrightarrow A_1 \subseteq A_2 \subseteq \dots$ und $\bigcup_{n=1}^{\infty} A_n = A$

b) $(A_n)_{n \in \mathbb{N}}$ **konvergiert von oben gegen** A (kurz: $A_n \downarrow A$) : $\Leftrightarrow A_1 \supseteq A_2 \supseteq \dots$ und $\bigcap_{n=1}^{\infty} A_n = A$.

Definition 1.27

$\mathcal{C}$ Mengensystem in Ω , $\varphi : \mathcal{C} \rightarrow \overline{\mathbb{R}}$ Abbildung.

a) φ heißt **stetig von unten**, wenn gilt:

$$A_n \in \mathcal{C}, A_n \uparrow A \in \mathcal{C} \implies \varphi(A_n) \rightarrow \varphi(A) \quad (n \rightarrow \infty).$$

b) φ heißt **stetig von oben**, wenn gilt:

$$A_n \in \mathcal{C}, A_n \downarrow A \in \mathcal{C}, \varphi(A_n) \text{ endlich} \implies \varphi(A_n) \rightarrow \varphi(A) \quad (n \rightarrow \infty).$$

c) φ heißt **$\emptyset$ -stetig**, wenn gilt:

$$A_n \in \mathcal{C}, A_n \downarrow \emptyset, \varphi(A_n) \text{ endlich} \implies \varphi(A_n) \rightarrow 0 \quad (n \rightarrow \infty).$$

Satz 1.28

Sei $\mathcal{R}$ Ring über Ω und $\varphi : \mathcal{R} \rightarrow \overline{\mathbb{R}}_+$ ein Inhalt. Dann gilt:

- a) φ σ -additiv $\Leftrightarrow \varphi$ stetig von unten
- b) φ stetig von unten $\Rightarrow \varphi$ stetig von oben $\Rightarrow \varphi$ $\emptyset$ -stetig
- c) φ $\emptyset$ -stetig und φ endlich $\Rightarrow \varphi$ σ -additiv.

Für einen endlichen Inhalt sind also die Begriffe σ -additiv, stetig von unten, stetig von oben und $\emptyset$ -stetig äquivalent.

Beispiel

$(\Omega, \mathcal{A}, \mu)$ sei W -Raum mit $\Omega = \mathbb{R}$ und $\mathcal{A} = \mathcal{B}$, und sei $F : \mathbb{R} \rightarrow \mathbb{R}$ definiert durch $F(x) = \mu((-\infty, x])$ (sog. **Verteilungsfunktion**). Dann ist F rechtsseitig stetig, denn für $x_1 \geq x_2 \geq \dots$ mit $x_n \rightarrow x$ ($n \rightarrow \infty$) gilt $(-\infty, x_n] \downarrow (-\infty, x]$ und daher gilt

$$F(x_n) = \mu((-\infty, x_n]) \xrightarrow{(n \rightarrow \infty)} \mu((-\infty, x]) = F(x).$$

Beweis von Satz 1.28

- a) " $\Rightarrow$ " Sei φ σ -additiv und $A_n, A \in \mathcal{R}$ mit $A_n \uparrow A$
Zu zeigen: $\varphi(A_n) \rightarrow \varphi(A)$ ($n \rightarrow \infty$)

Dazu:

$$A = \bigcup_{n=1}^{\infty} A_n = A_1 + (A_2 \setminus A_1) + (A_3 \setminus A_2) + \dots$$

nach Definition der Konvergenz von unten.

Die Mengen $A_n \setminus A_{n-1}$ sind in Ring $\mathcal{R}$ enthalten, und mit der σ -Additivität von φ folgt:

$$\begin{aligned} \varphi(A) &= \varphi(A_1) + \varphi(A_2 \setminus A_1) + \varphi(A_3 \setminus A_2) + \dots \\ &= \varphi(A_1) + \lim_{n \rightarrow \infty} \sum_{k=2}^n \varphi(A_k \setminus A_{k-1}) \\ &= \lim_{n \rightarrow \infty} (\varphi(A_1) + \varphi(A_n \setminus A_1)) \\ &= \lim_{n \rightarrow \infty} \varphi(A_n). \end{aligned}$$

" $\Leftarrow$ " Sei φ stetig von unten und $A_n \in \mathcal{R}$ ($n \in \mathbb{N}$) paarweise disjunkt mit $\sum_{n=1}^{\infty} A_n \in \mathcal{R}$.

zu zeigen: $\varphi\left(\sum_{n \in \mathbb{N}} A_n\right) = \sum_{n=1}^{\infty} \varphi(A_n)$.

Setze $B_n = \sum_{k=1}^n A_k$. Da $\mathcal{R}$ Ring ist, gilt $B_n \in \mathcal{R}$, weiter gilt nach Konstruktion

$$B_1 \subseteq B_2 \subseteq \dots \text{ und } \bigcup_{n=1}^{\infty} B_n = \sum_{n \in \mathbb{N}} A_n.$$

Mit φ stetig von unten und φ Inhalt folgt

$$\begin{aligned}\varphi\left(\sum_{n \in \mathbb{N}} A_n\right) &= \varphi\left(\bigcup_{n=1}^{\infty} B_n\right) = \lim_{n \rightarrow \infty} \varphi(B_n) \\ &= \lim_{n \rightarrow \infty} \varphi\left(\sum_{k=1}^n A_k\right) = \lim_{n \rightarrow \infty} \sum_{k=1}^n \varphi(A_k) = \sum_{k=1}^{\infty} \varphi(A_k).\end{aligned}$$

- b) $b_1)$ Sei φ stetig von unten und $A_n, A \in \mathcal{R}$ mit $A_1 \supseteq A_2 \supseteq \dots, \bigcap_{n=1}^{\infty} A_n = A$ und $\varphi(A_n) \in \mathbb{R}$ für alle $n \in \mathbb{N}$.

zu zeigen: $\varphi(A_n) \rightarrow \varphi(A)$ ($n \rightarrow \infty$)

Es gilt $A_1 \setminus A_2 \subseteq A_1 \setminus A_3 \subseteq \dots$ und $\bigcup_{n=1}^{\infty} A_1 \setminus A_n = A_1 \setminus \left(\bigcap_{n=1}^{\infty} A_n\right)$

Da φ stetig von unten ist, folgt daraus

$$\varphi(A_1 \setminus A_n) \rightarrow \varphi\left(A_1 \setminus \left(\bigcap_{n=1}^{\infty} A_n\right)\right) \quad (n \in \mathbb{N}).$$

Wegen $\varphi(A_1) \in \mathbb{R}, \varphi$ positiv und

$$\begin{aligned}\varphi(A_1) &= \varphi(A_1 \setminus A_n) + \varphi(A_n), \\ \varphi(A_1) &= \varphi\left(A_1 \setminus \left(\bigcap_{n=1}^{\infty} A_n\right)\right) + \varphi\left(\bigcap_{n=1}^{\infty} A_n\right)\end{aligned}$$

folgt daraus (unter Beachtung von $\varphi(A_n) < \infty$ nach Voraussetzung)

$$\varphi(A_n) = \varphi(A_1) - \varphi(A_1 \setminus A_n) \xrightarrow{n \rightarrow \infty} \varphi(A_1) - \varphi\left(A_1 \setminus \bigcap_{n=1}^{\infty} A_n\right) = \varphi\left(\bigcap_{n=1}^{\infty} A_n\right).$$

- $b_2)$ Ist φ stetig von oben, so ist (mit $A = \emptyset$ und $\varphi(\emptyset) = 0$) φ auch $\emptyset$ -stetig.

- c) Sei φ $\emptyset$ -stetig und φ endlich.

Nach a) genügt es zu zeigen: φ ist stetig von unten.

Dazu: Seien $A_n, A \in \mathcal{R}$ mit $A_1 \subseteq A_2 \subseteq \dots$ und $A = \bigcup_{n=1}^{\infty} A_n$.

Dann ist $A \setminus A_n$ im Ring $\mathcal{R}$ enthalten und erfüllt

$$A \setminus A_1 \supseteq A \setminus A_2 \supseteq A \setminus A_3 \supseteq \dots \text{ sowie } \bigcap_{n=1}^{\infty} A \setminus A_n = A \setminus \left(\bigcup_{n=1}^{\infty} A_n\right) = \emptyset.$$

Also gilt $A \setminus A_n \downarrow \emptyset$ und unter Beachtung von φ endlich folgt mit der $\emptyset$ -Stetigkeit von φ :

$$\varphi(A \setminus A_n) \rightarrow 0 \quad (n \rightarrow \infty).$$

Da φ endlich ist, gilt aber $\varphi(A \setminus A_n) = \varphi(A) - \varphi(A_n)$ (da

$$\varphi(A) = \varphi(A_n + A \setminus A_n) = \varphi(A_n) + \varphi(A \setminus A_n)$$

und alle auftretenden Werte endlich sind), und es folgt

$$\varphi(A_n) \rightarrow \varphi(A) \quad (n \rightarrow \infty)$$

□

Bemerkung 1.29

- a) Satz 1.28 b) gilt nicht, wenn man in der Definition von φ stetig von oben die Bedingung $\varphi(A_n) \in \mathbb{R}$ weglässt.

Beispiel

Für $\Omega = \mathbb{R}$, $\mathcal{A} = \mathcal{B}$ und $m =$ Lebesgue-Borel-Maß (s. u., dasjenige Maß auf $\mathcal{B}$, das Intervallen ihre elementare Länge zuordnet) gilt:

$$A_n = [n, \infty) \downarrow \emptyset, \text{ aber } m(A_n) = \infty \text{ konvergiert nicht gegen } 0 = m(\emptyset) \text{ für } n \rightarrow \infty.$$

- b) Satz 1.28 c) gilt ohne die Voraussetzung “ φ endlich” nicht (vgl. Übungen).

Lemma 1.30

Sei μ ein Inhalt auf einem Ring $\mathcal{R}$ in Ω . Dann gilt

- a) μ ist monoton.
b) μ ist subadditiv, d. h. für $A_1, \dots, A_n \in \mathcal{R}$ gilt immer

$$\mu\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mu(A_i).$$

- c) Ist μ ein Maß, so ist μ sogar σ -subadditiv, d. h. für alle $A_n \in \mathcal{R}$ ($n \in \mathbb{N}$) mit $\bigcup_{n=1}^{\infty} A_n \in \mathcal{R}$ gilt

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n).$$

Beweis

Siehe Übungen. □

Zur Erinnerung:

Elementare Figur = endlich disjunkte Vereinigung beschränkter halboffener Intervalle

$\mathcal{E}_n =$ Ring (!) der elementaren Figuren (d. h. $\mathcal{E}_n$ enthält $\emptyset$ und ist abgeschlossen gegen Vereinigungsbildung, Durchschnittsbildung und Differenzbildung zweier Mengen)

Satz 1.31

Der elementargeometrisch definierte Inhalt m auf dem Ring $\mathcal{E}_n$ der elementaren Figuren in $\mathbb{R}^n$ ist ein Maß.

Beweis von Satz 1.31

Ohne Beweis: m ist Inhalt ... (klar!)

Seien $A_1, A_2, \dots \in \mathcal{E}_n$, paarweise disjunkt, und sei $A = \sum_{n \in \mathbb{N}} A_n \in \mathcal{E}_n$.

zu zeigen: $m\left(\sum_{n \in \mathbb{N}} A_n\right) = \sum_{n=1}^{\infty} m(A_n)$.

1. Schritt: Wir zeigen die Behauptung im Spezialfall $A, A_1, A_2, \dots \in \mathcal{J}_n =$ System der halboffenen, beschränkten Intervalle

$\alpha)$ Wir zeigen: $\sum_{j=1}^{\infty} m(A_j) \leq m(A)$.

Folgt wegen $\sum_{j=1}^N A_j \subseteq A$, m monoton aus

$$m(A) \geq m\left(\sum_{j=1}^N A_j\right) = \sum_{j=1}^N m(A_j) \quad (N \in \mathbb{N})$$

(da m als Inhalt endlich additiv ist) mit $N \rightarrow \infty$.

$\beta)$ Wir zeigen: $m(A) \leq \sum_{j=1}^{\infty} m(A_j)$.

Sei $\varepsilon > 0$ beliebig.

Wähle $B \in \mathcal{J}_n$ mit $\overline{B} \subseteq A$ und $m(B) \geq m(A) - \varepsilon$, wobei $\overline{B}$ die Abschließung von B ist.

Für $j \in \mathbb{N}$ wähle $B_j \in \mathcal{J}_n$ mit $A_j \subseteq \overset{\circ}{B}_j$ und

$$m(B_j) \leq m(A_j) + \frac{\varepsilon}{2^j},$$

wobei $\overset{\circ}{B}$ das Innere von B ist.

Dann gilt

$$\overline{B} \subseteq A = \bigcup_{j=1}^{\infty} A_j \subseteq \bigcup_{j=1}^{\infty} \overset{\circ}{B}_j,$$

und da $\overline{B}$ kompakt ist (da abgeschlossen und beschränkt) und die $\overset{\circ}{B}_j$ offen sind, existiert nach dem Satz von Heine-Borel ein $N \in \mathbb{N}$ mit

$$\overline{B} \subseteq \bigcup_{j=1}^N \overset{\circ}{B}_j,$$

woraus folgt:

$$B \subseteq \bigcup_{j=1}^N B_j$$

$$\begin{aligned} \Rightarrow m(A) - \varepsilon &\leq m(B) \leq m\left(\bigcup_{j=1}^N B_j\right) \\ &\leq \sum_{j=1}^N m(B_j) \quad (\text{Lemma 1.30b))} \\ &\leq \sum_{j=1}^N \left(m(A_j) + \frac{\varepsilon}{2^j}\right) \\ &\leq \sum_{j=1}^N m(A_j) + \varepsilon \end{aligned}$$

Es gilt also insbesondere:

$$m(A) \leq \sum_{j=1}^{\infty} m(A_j) + 2\varepsilon$$

Mit $\varepsilon \downarrow 0$ folgt β).

2. Schritt: Wir zeigen die Behauptung im allgemeinen Fall $A, A_1, A_2, \dots \in \mathcal{E}_n$.

Sei also $A = \sum_{j=1}^{\infty} A_j \in \mathcal{E}_n$

Dann gilt:

$$\begin{aligned} A &= \sum_{k=1}^K I_k \quad \text{für geeignete } I_1, \dots, I_K \in \mathcal{J}_n, \\ A_j &= \sum_{s=1}^{s_j} I_{j,s} \quad \text{für geeignete } I_{j,1}, \dots, I_{j,s_j} \in \mathcal{J}_n, \end{aligned}$$

und damit ist

$$\begin{aligned} I_k &= I_k \cap A = I_k \cap \sum_{j=1}^{\infty} A_j = \sum_{j=1}^{\infty} I_k \cap A_j \\ &= \sum_{j=1}^{\infty} I_k \cap \sum_{s=1}^{s_j} I_{j,s} = \sum_{j=1}^{\infty} \sum_{s=1}^{s_j} I_k \cap I_{j,s}, \end{aligned}$$

wobei $I_k \cap I_{j,s} \in \mathcal{J}_n$ ist, da Durchschnitt halboffener Intervalle wieder ein halboffenes Intervall ergibt.

Mit Schritt 1 folgt:

$$\begin{aligned}
m(A) &= m\left(\sum_{k=1}^K I_k\right) &= \sum_{k=1}^K m(I_k) && \text{(da } m \text{ Inhalt)} \\
& &= \sum_{k=1}^K m\left(\sum_{j=1}^{\infty} \sum_{s=1}^{s_j} (I_k \cap I_{j,s})\right) \\
&\stackrel{\text{Schritt 1}}{=} &\sum_{k=1}^K \sum_{j=1}^{\infty} \sum_{s=1}^{s_j} m(I_k \cap I_{j,s}) \\
& &= \sum_{j=1}^{\infty} \sum_{s=1}^{s_j} m\left(\sum_{k=1}^K I_k \cap I_{j,s}\right) \\
& &&& \text{(Reihenfolge der Summation ist bei} \\
& &&& \text{nichtnegativen Summanden egal, } m \text{ Inhalt)} \\
& &= \sum_{j=1}^{\infty} \sum_{s=1}^{s_j} m(A \cap I_{j,s}) \\
& &= \sum_{j=1}^{\infty} \sum_{s=1}^{s_j} m(I_{j,s}) && \text{(da } I_{j,s} \subseteq A_j \subseteq A) \\
& &= \sum_{j=1}^{\infty} m\left(\sum_{s=1}^{s_j} I_{j,s}\right) \\
& &= \sum_{j=1}^{\infty} m(A_j) && \rightarrow \text{Beh.}
\end{aligned}$$

1.3 Fortsetzung von Maßen

Motivation: Der elementargeometrische Inhalt m ist Maß auf dem Ring $\mathcal{E}_n$ der elementaren Figuren. Um auch allgemeineren Mengen ihre Maßzahl zuordnen zu können, ist es wünschenswert, m auf ein größeres Mengensystem $\mathcal{A}$ fortzusetzen, wobei $\mathcal{E}_n \subsetneq \mathcal{A} \subseteq \mathcal{P}(\mathbb{R})$ gilt. D. h. gesucht ist σ -Algebra $\mathcal{A} \supsetneq \mathcal{E}_n$ und **Maß**

$$\bar{m} : \mathcal{A} \rightarrow \overline{\mathbb{R}}$$

mit $\bar{m}|_{\mathcal{E}_n} = m$, also $\bar{m}(A) = m(A)$ für alle $A \in \mathcal{E}_n$.

Wir zeigen im Folgenden, dass m auf $\mathcal{B}_n = \mathcal{F}(\mathcal{E}_n)$ fortsetzbar ist.

Hilfsmittel dabei ist der Begriff des sogenannten äußeren Maßes:

Definition 1.32

Sei $\mathcal{R}$ ein Ring über Ω und $\mu : \mathcal{R} \rightarrow \overline{\mathbb{R}}$ ein Maß. Dann heißt

$$\begin{aligned}
&\mu^* : \mathcal{P}(\Omega) \rightarrow \overline{\mathbb{R}} \\
\mu^*(A) &= \inf \left\{ \sum_{n=1}^{\infty} \mu(B_n) \mid B_n \in \mathcal{R} \text{ mit } A \subseteq \bigcup_{n=1}^{\infty} B_n \right\}
\end{aligned}$$

das **zu μ gehörende äußere Maß**. (Hierbei $\inf \emptyset = \infty$)

Definition 1.33

Eine Mengenfunktion $\nu : \mathcal{P}(\Omega) \rightarrow \overline{\mathbb{R}}$ mit $\nu(\emptyset) = 0$, ν positiv, ν monoton und ν σ -subadditiv heißt **äußeres Maß**.

Hierbei ist ν σ -subadditiv, falls gilt: Für alle $A_1, A_2, \dots \in \mathcal{P}(\Omega)$ ist

$$\nu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \nu(A_n)$$

.

Klar: μ Maß auf $\mathcal{P}(\Omega) \Rightarrow \mu$ äußeres Maß

Die Umkehrung gilt i. A. nicht, z. B. ist μ mit

$$\mu(A) = \begin{cases} 0 & \text{für } A = \emptyset \\ 1 & \text{für } A \neq \emptyset \end{cases}$$

ein äußeres Maß, aber kein Maß.

Satz 1.34

Sei $\mathcal{R}$ ein Ring über Ω , $\mu : \mathcal{R} \rightarrow \overline{\mathbb{R}}$ Maß, und μ^* das zu μ gehörende äußere Maß. Dann gilt:

- a) μ^* ist ein äußeres Maß im Sinne von Definition 1.33.
- b) μ^* stimmt auf $\mathcal{R}$ mit μ überein, d. h. $\mu^*(A) = \mu(A)$ für alle $A \in \mathcal{R}$.

Beweis

- a) – $\mu^*(\emptyset) = 0$ da $\emptyset \subseteq \emptyset \cup \emptyset \cup \emptyset \dots$ und daher
$$\mu^*(\emptyset) \leq \sum_{n=1}^{\infty} \mu(\emptyset) = \sum_{n=1}^{\infty} 0 = 0, \text{ da } \mu \text{ Maß.}$$
- μ^* positiv, da $\sum_{n=1}^{\infty} \mu(B_n) \geq 0$ für alle $A \subseteq \bigcup_{n=1}^{\infty} B_n$ (wegen μ positiv).
- Ist $A \subseteq B$, so folgt aus $B \subseteq \bigcup_{n=1}^{\infty} B_n$ auch $A \subseteq \bigcup_{n=1}^{\infty} B_n$. Daher gilt

$$\begin{aligned} & \left\{ \sum_{n=1}^{\infty} \mu(B_n) \mid B_n \in \mathcal{R} \text{ mit } A \subseteq \bigcup_{n=1}^{\infty} B_n \right\} \\ & \supseteq \left\{ \sum_{n=1}^{\infty} \mu(B_n) \mid B_n \in \mathcal{R} \text{ mit } B \subseteq \bigcup_{n=1}^{\infty} B_n \right\}, \end{aligned}$$

was $\mu^*(A) \leq \mu^*(B)$ impliziert.

– Sei $A_n \in \mathcal{P}(\Omega)$. Wir zeigen: $\mu^* \left(\bigcup_{n=1}^{\infty} A_n \right) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$.

OBdA $\mu^*(A_n) < \infty$ für alle $n \in \mathbb{N}$.

Sei $\varepsilon > 0$ beliebig. Wähle $B_{n,k} \in \mathcal{R}$ ($k \in \mathbb{N}$) mit $A_n \subseteq \bigcup_{k=1}^{\infty} B_{n,k}$ und $\mu^*(A_n) \geq$

$$\sum_{k=1}^{\infty} \mu(B_{n,k}) - \frac{\varepsilon}{2^n}.$$

Dann gilt:

$$\bigcup_{n \in \mathbb{N}} A_n \subseteq \bigcup_{n \in \mathbb{N}, k \in \mathbb{N}} B_{n,k} \quad \text{mit } B_{n,k} \in \mathcal{R}$$

und die Definition von μ^* impliziert:

$$\mu^* \left(\bigcup_{n \in \mathbb{N}} A_n \right) \leq \sum_{n, k \in \mathbb{N}} \mu(B_{n,k}) = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \mu(B_{n,k})$$

(Summationsreihenfolge spielt bei nichtnegativen Zahlen keine Rolle)

$$\leq \sum_{n=1}^{\infty} \left(\mu^*(A_n) + \frac{\varepsilon}{2^n} \right) = \left(\sum_{n=1}^{\infty} \mu^*(A_n) \right) + \varepsilon.$$

Mit $\varepsilon \downarrow 0$ folgt die Behauptung.

b) Sei $A \in \mathcal{R}$ beliebig.

zu zeigen: $\mu^*(A) = \mu(A)$

b₁) Wir zeigen: $\mu^*(A) \leq \mu(A)$

Folgt mit $B_1 = A, B_2 = \emptyset, B_3 = \emptyset, \dots$ unmittelbar aus Definition.

b₂) Wir zeigen: $\mu(A) \leq \mu^*(A)$

Nach Definition von μ^* genügt es dazu zu zeigen:

Ist $B_n \in \mathcal{R}$ ($n \in \mathbb{N}$) mit $A \subseteq \bigcup_{n \in \mathbb{N}} B_n$, so gilt

$$\mu(A) \leq \sum_{n=1}^{\infty} \mu(B_n)$$

Mit

$$A = A \cap \left(\bigcup_{n \in \mathbb{N}} B_n \right) = \bigcup_{n \in \mathbb{N}} (A \cap B_n),$$

wobei $A \cap B_n \in \mathcal{R}$ da $\mathcal{R}$ Ring und $\bigcup_{n \in \mathbb{N}} (A \cap B_n) = A \in \mathcal{R}(!)$, folgt aus der σ -Subadditivität des Maßes μ :

$$\mu(A) \leq \sum_{n=1}^{\infty} \mu(A \cap B_n) \leq \sum_{n=1}^{\infty} \mu(B_n),$$

da μ monoton und $A \cap B_n \subseteq B_n$

→ Beh. □

Wir zeigen im Folgenden:

Die Einschränkung von μ^* auf $\mathcal{F}(\mathcal{R})$ ist ein Maß (und dieses stimmt dann auf $\mathcal{R}$ mit μ überein, da μ^* das schon tut).

Dazu Umweg:

Einführung einer σ -Algebra $\mathcal{M}_\mu^*$, die $\mathcal{F}(\mathcal{R})$ umfaßt und auf der μ^* ein Maß ist.

Motivation der Definition dieses Mengensystems:

Im Falle $\mu^*(\Omega) < \infty$ wird definiert:

$$\begin{aligned}\mathcal{M}_\mu^* &= \{A \in \mathcal{P}(\Omega) \mid \text{Äußeres Maß von } A = \text{Inneres Maß von } A\} \\ &= \{A \in \mathcal{P}(\Omega) \mid \mu^*(A) = \mu^*(\Omega) - \mu^*(A^c)\} \\ &= \{A \in \mathcal{P}(\Omega) \mid \mu^*(\Omega) = \mu^*(A) + \mu^*(A^c)\},\end{aligned}$$

wobei $\mu^*(\Omega) - \mu^*(A^c)$ als Ersatz für ein "inneres Maß" von A verwendet wird.

Im allgemeinen Fall wird definiert:

$$\begin{aligned}\mathcal{M}_\mu^* &= \{A \in \mathcal{P}(\Omega) \mid \mu^*(Q) = \mu^*(Q \cap A) + \mu^*(Q \cap A^c) \\ &\quad \text{für alle } Q \in \mathcal{P}(\Omega)\} \\ \text{vgl. Übungen} &= \{A \in \mathcal{P}(\Omega) \mid \mu^*(S + T) = \mu^*(S) + \mu^*(T) \\ &\quad \text{für alle } S \subseteq A^c, T \subseteq A\}\end{aligned}$$

Im Falle $\mu^*(\Omega) < \infty$ stimmen beide Definitionen überein (vgl. Übungen).

Definition 1.35

Die Mengen aus $\mathcal{M}_\mu^*$ heißen μ^* -messbare Mengen.

Lemma 1.36

Sei μ ein Maß über einem Ring $\mathcal{R}$ über Ω . Dann sind alle Mengen aus $\mathcal{R}$ μ^* -messbare Mengen.

Beweis

Sei $A \in \mathcal{R}$ und seien $S \subseteq A^c, T \subseteq A$.

Zu zeigen:

$$\mu^*(S + T) = \mu^*(S) + \mu^*(T)$$

$\alpha)$ Da μ^* ein äußeres Maß ist (vgl. Satz 1.34 a)) gilt

$$\begin{aligned}\mu^*(S + T) &= \mu^*(S + T + \emptyset + \emptyset + \dots) \leq \mu^*(S) + \mu^*(T) + \mu^*(\emptyset) + \mu^*(\emptyset) + \dots \\ &= \mu^*(S) + \mu^*(T)\end{aligned}$$

$\beta)$ Wir zeigen $\mu^*(S) + \mu^*(T) \leq \mu^*(S + T)$.

oBdA $\mu^*(S + T) < \infty$.

Sei $B_n \in \mathcal{R}$ mit

$$S + T \subseteq \bigcup_{n \in \mathbb{N}} B_n.$$

Wegen $S \subseteq A^c, T \subseteq A$ gilt dann

$$S \subseteq \bigcup_{n \in \mathbb{N}} B_n \cap A^c \quad \text{und} \quad T \subseteq \bigcup_{n \in \mathbb{N}} B_n \cap A$$

wobei $B_n \cap A^c = B_n \setminus A, B_n \cap A \in \mathcal{R}$ gilt wegen $\mathcal{R}$ Ring.

Nach Definition von μ^* gilt dann aber

$$\begin{aligned} \mu^*(S) + \mu^*(T) &\leq \sum_{n=1}^{\infty} \mu(B_n \cap A^c) + \sum_{n=1}^{\infty} \mu(B_n \cap A) \\ &\stackrel{\mu \text{ Ma\ss}}{=} \sum_{n=1}^{\infty} \mu(B_n \cap A^c + B_n \cap A) \\ &= \sum_{n=1}^{\infty} \mu(B_n), \end{aligned}$$

und da $B_n \in \mathcal{R}$ beliebig war mit $S + T \subseteq \bigcup_{n=1}^{\infty} B_n$ folgt daraus

$$\mu^*(S) + \mu^*(T) \leq \mu^*(S + T).$$

□

Lemma 1.37

Sei μ ein Ma\ss \u00fcber einem Ring $\mathcal{R}$ \u00fcber Ω . Dann ist

$$\mathcal{M}_\mu^* = \{A \subseteq \Omega \mid \mu^*(S + T) = \mu^*(S) + \mu^*(T) \text{ f\u00fcr alle } S \subseteq A^c, T \subseteq A\}$$

eine σ -Algebra \u00fcber Ω .

Beweis

Bez.: $\mathcal{M}^* = \mathcal{M}_\mu^*$

1. $\emptyset \in \mathcal{M}^*$, denn $S \subseteq \emptyset^c, T \subseteq \emptyset$ impliziert $T = \emptyset$ und damit

$$\begin{aligned} \mu^*(S + T) &= \mu^*(S) = \mu^*(S) + 0 = \mu^*(S) + \mu^*(\emptyset) \\ &= \mu^*(S) + \mu^*(T), \end{aligned}$$

da $\mu^*(\emptyset) = 0$.

2. Mit $A \in \mathcal{M}^*$ gilt (nach Symmetrie in der Definition von $\mathcal{M}^*$) auch immer $A^c \in \mathcal{M}^*$.

3. Seien $A, B \in \mathcal{M}^*$

Wir zeigen: $A \cup B \in \mathcal{M}^*$

Dazu sei $S \subseteq (A \cup B)^c, T \subseteq A \cup B$.

Zu zeigen: $\mu^*(S + T) = \mu^*(S) + \mu^*(T)$.

“ $\leq$ ”: klar wegen σ -Subadditivität von μ^* .

“ $\geq$ ”: Setze $T' = T \cap A, T'' = T \setminus A$

Dann gilt:

$$\begin{aligned}
 \mu^*(S + T) &= \mu^*(S + T' + T'') \\
 &= \mu^*(S + T'') + \mu^*(T') \\
 &\quad (\text{da } S + T'' \subseteq A^c, T' \subseteq A \text{ und } A \in \mathcal{M}^*) \\
 &= \mu^*(S) + \mu^*(T'') + \mu^*(T') \\
 &\quad (\text{da } S \subseteq B^c, T'' \subseteq B \text{ und } B \in \mathcal{M}^*) \\
 &\geq \mu^*(S) + \mu^*(T' + T'') \\
 &\quad (\text{da } \mu^* \text{ } \sigma\text{-subadditiv}) \\
 &= \mu^*(S) + \mu^*(T).
 \end{aligned}$$

Also ist gezeigt

$$\mu^*(S + T) \geq \mu^*(S) + \mu^*(T),$$

und damit gilt die Zwischenbehauptung.

4. Seien $A_n \in \mathcal{M}^*$ ($n \in \mathbb{N}$) paarweise disjunkt.

Wir zeigen: $\sum_{n \in \mathbb{N}} A_n \in \mathcal{M}^*$.

Dazu sei $S \subseteq \left(\sum_{n \in \mathbb{N}} A_n \right)^c, T \subseteq \sum_{n \in \mathbb{N}} A_n$.

Dann ist zu zeigen:

$$\mu^*(S + T) = \mu^*(S) + \mu^*(T).$$

“ $\leq$ ”: Klar, da μ^* σ -subadditiv.

“ $\geq$ ”: Setze

$$T_n = T \cap A_n.$$

Dann gilt $T = \sum_{n \in \mathbb{N}} T_n$ und damit

$$\begin{aligned}
 \mu(S + T) &= \mu^*(S + T_1 + T_2 + \dots) \\
 &\geq \mu^*(S + T_1 + \dots + T_N) \\
 &\quad (\text{da } \mu^* \text{ monoton}) \\
 &= \mu^*(S + T_1 + \dots + T_{N-1}) + \mu^*(T_N) \\
 &\quad \text{da } S + T_1 + \dots + T_{N-1} \subseteq A_N^c, T_N \subseteq A_N \text{ und } A_N \in \mathcal{M}^* \\
 &= \dots \\
 &= \mu^*(S) + \mu(T_1) + \dots + \mu^*(T_N).
 \end{aligned}$$

Mit $N \rightarrow \infty$ folgt:

$$\begin{aligned}\mu^*(S + T) &\geq \mu^*(S) + \sum_{n=1}^{\infty} \mu^*(T_n) \\ &\geq \mu^*(S) + \mu^*\left(\sum_{n=1}^{\infty} T_n\right) \\ &= \mu^*(S) + \mu^*(T),\end{aligned}$$

wobei die letzte Ungleichung aus der σ -Subadditivität von μ^* folgt.

→ Beh. $\square$

Lemma 1.38

Sei μ ein Maß über einem Ring $\mathcal{R}$ über Ω . Dann ist μ^* ein Maß auf der σ -Algebra $\mathcal{M}_\mu^*$.

Beweis

Da nach Satz 1.34 μ^* ein äußeres Maß auf $\mathcal{P}(\Omega)$ ist, genügt es zu zeigen:

Sind $A_n \in \mathcal{M}_\mu^*$ ($n \in \mathbb{N}$) paarweise disjunkt, so gilt

$$\mu^*\left(\sum_{n \in \mathbb{N}} A_n\right) \geq \sum_{n=1}^{\infty} \mu^*(A_n).$$

Dies folgt aber aus

$$\begin{aligned}\mu^*\left(\sum_{n \in \mathbb{N}} A_n\right) &\geq \mu^*(A_1 + \dots + A_N) \\ &\quad (\text{da } \mu^* \text{ monoton}) \\ &\geq \mu^*(A_1 + \dots + A_{N-1}) + \mu^*(A_N) \\ &\quad (\text{da } A_1 + \dots + A_{N-1} \subseteq A_N^c, A_N \subseteq A_N, A_N \in \mathcal{M}_\mu^*) \\ &\geq \dots \\ &\geq \mu^*(A_1) + \dots + \mu^*(A_N)\end{aligned}$$

mit $N \rightarrow \infty$. $\square$

Satz 1.39 (Fortsetzungssatz von Carathéodory)

Sei μ ein Maß auf einem Ring $\mathcal{R}$ über Ω . Dann existiert mindestens ein Maß $\bar{\mu}$ auf der von $\mathcal{R}$ erzeugten σ -Algebra $\mathcal{F}(\mathcal{R})$, welches μ fortsetzt, d. h. welches auf $\mathcal{R}$ mit μ übereinstimmt.

Beweis

Sei μ^* das zu μ gehörende äußere Maß auf $\mathcal{P}(\Omega)$, und sei

$$\mathcal{M}_\mu^* = \left\{ A \subseteq \Omega \mid \mu^*(S + T) = \mu^*(S) + \mu^*(T) \text{ für alle } S \subseteq A^c, T \subseteq A \right\}.$$

Nach Lemma 1.36 gilt $\mathcal{R} \subseteq \mathcal{M}_\mu^*$. Da $\mathcal{M}_\mu^*$ nach Lemma 1.37 weiter eine σ -Algebra ist, gilt auch

$$\mathcal{F}(\mathcal{R}) \subseteq \mathcal{M}_\mu^*.$$

Nach Lemma 1.38 ist μ^* ein Maß auf $\mathcal{M}_\mu^*$, und damit ist auch

$$\begin{aligned} \bar{\mu} : \mathcal{F}(\mathcal{R}) &\rightarrow \overline{\mathbb{R}} \\ \bar{\mu}(A) &:= \mu^*(A) \end{aligned}$$

ein Maß.

Nach Satz 1.34 b) stimmt μ^* (und damit auch $\bar{\mu}$) auf $\mathcal{R}$ mit μ überein.

→ Beh. $\square$

Im Folgenden: Herleitung von hinreichenden Kriterien zur **eindeutigen** Fortsetzung eines Maßes.

Hilfreich dabei:

Definition 1.40

Sei $\mathcal{C}$ ein Mengensystem über Ω , d. h. $\mathcal{C} \subseteq \mathcal{P}(\Omega)$. Dann heißt $\varphi : \mathcal{C} \rightarrow \overline{\mathbb{R}}$ **σ -endlich**, wenn Mengen $A_k \in \mathcal{C}$ ($k \in \mathbb{N}$) existieren mit $A_k \uparrow \Omega$ und $\varphi(A_k) < \infty$ für alle $k \in \mathbb{N}$.

Beispiel

Der elementargeometrische Inhalt m auf dem Ring der elementaren Figuren ist σ -endlich, da für

$$\begin{aligned} A_k &:= (-k, k] \times \dots \times (-k, k] \in \mathcal{E}_n \quad \text{gilt:} \\ A_k &\uparrow \mathbb{R}^n \quad \text{und} \quad m(A_k) = (2k)^n < \infty \text{ für alle } k \in \mathbb{N}. \end{aligned}$$

Satz 1.41 (Eindeutigkeitssatz)

Seien μ_1, μ_2 zwei Maße auf der σ -Algebra $\mathcal{A}$ über Ω . Sei $\mathcal{C}$ ein $\cap$ -stabiler Erzeuger von $\mathcal{A}$ (d. h. $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$ und es gilt $\mathcal{A} = \mathcal{F}(\mathcal{C})$). Für die Restriktionen $\mu_1|_{\mathcal{C}}$ und $\mu_2|_{\mathcal{C}}$ der Maße auf $\mathcal{C}$ gelte:

$$\mu_1|_{\mathcal{C}} = \mu_2|_{\mathcal{C}} \quad (\text{d. h. } \mu_1(A) = \mu_2(A) \text{ für alle } A \in \mathcal{C})$$

und

$\mu_1|_{\mathcal{C}}, \mu_2|_{\mathcal{C}}$ seien σ -endlich.

Dann gilt auch

$$\mu_1 = \mu_2,$$

d. h. $\mu_1(A) = \mu_2(A)$ für alle $A \in \mathcal{A}$.

Korollar 1.42

Ist μ ein σ -endliches Maß auf einem Ring $\mathcal{R}$ über Ω , so existiert genau eine Fortsetzung $\bar{\mu}$ von μ auf $\mathcal{F}(\mathcal{R})$, d. h. es existiert genau ein Maß $\bar{\mu} : \mathcal{F}(\mathcal{R}) \rightarrow \bar{\mathbb{R}}$ mit

$$\bar{\mu}(A) = \mu(A) \quad \text{für alle } A \in \mathcal{R}.$$

Hierbei gilt:

$$\bar{\mu}(A) = \mu^*(A) = \inf \left\{ \sum_{n=1}^{\infty} \mu(B_n) : B_n \in \mathcal{R} \ (n \in \mathbb{N}) \text{ und } A \subseteq \bigcup_{n=1}^{\infty} B_n \right\}$$

für $A \in \mathcal{F}(\mathcal{R})$.

Beweis

Folgt unmittelbar aus den Sätzen 1.39 und 1.41 (da $\mathcal{R}$ $\cap$ -stabiler Erzeuger von $\mathcal{F}(\mathcal{R})$ ist).

Anwendung

Es existiert genau ein Maß m , das den elementargeometrischen Inhalt von dem Ring $\mathcal{E}_n$ der elementargeometrischen Figuren auf $\mathcal{B}_n = \mathcal{F}(\mathcal{E}_n)$ fortsetzt.

Dieses Maß wird als **Lebesgue-Borel-Maß** (kurz: **LB-Maß**) in $\mathbb{R}^n$ bezeichnet und im Folgenden wieder mit m bezeichnet.

Beweis von Satz 1.41

1. Schritt: Es gelte $\mu_1(\Omega) = \mu_2(\Omega) < \infty$

Wir zeigen: $A \in \mathcal{F}(\mathcal{C}) \Rightarrow \mu_1(A) = \mu_2(A)$

d. h.

$$\mathcal{F}(\mathcal{C}) \subseteq \mathcal{G} := \{G \in \mathcal{F}(\mathcal{C}) : \mu_1(G) = \mu_2(G)\}$$

Dazu genügt es zu zeigen:

(1) $\mathcal{C} \subseteq \mathcal{G}$

(2) $\mathcal{G}$ ist Dynkin-System

Denn daraus folgt:

$$\mathcal{F}(\mathcal{C}) = \mathcal{D}(\mathcal{C}) \subseteq \mathcal{G}$$

da $\mathcal{C}$ $\cap$ -stabil, vgl. Satz 1.18

(1) gilt, da nach Voraussetzung des Satzes $\mu_1|_{\mathcal{C}} = \mu_2|_{\mathcal{C}}$.

Nachweis von (2):

a) $\Omega \in \mathcal{G}$, da $\Omega \in \mathcal{F}(\mathcal{C})$ und $\mu_1(\Omega) = \mu_2(\Omega)$ nach Voraussetzung.

b) Seien $A, B \in \mathcal{G}$ mit $A \subseteq B$.

Dann ist $B \setminus A \in \mathcal{F}(\mathcal{C})$ und

$$\begin{array}{ccccccc} \mu_1(B \setminus A) & = & \mu_1(B) - \mu_1(A) & = & \mu_2(B) - \mu_2(A) & = & \mu_2(B \setminus A), \\ & & \uparrow & & \uparrow & & \uparrow \\ & & \text{da } \mu_1(\Omega) < \infty & & \text{da } A, B \in \mathcal{G} & & \text{da } \mu_2(\Omega) < \infty \end{array}$$

also gilt auch

$$B \setminus A \in \mathcal{G}.$$

c) Seien A_n ($n \in \mathbb{N}$) paarweise disjunkt.

Dann ist $\sum_{n \in \mathbb{N}} A_n \in \mathcal{F}(\mathcal{C})$ und es gilt

$$\begin{array}{ccccccc} \mu_1 \left(\sum_{n \in \mathbb{N}} A_n \right) & = & \sum_{n=1}^{\infty} \mu_1(A_n) & = & \sum_{n=1}^{\infty} \mu_2(A_n) & = & \mu_2 \left(\sum_{n \in \mathbb{N}} A_n \right), \\ & & \uparrow & & \uparrow & & \uparrow \\ & & \mu_1 \text{ Maß} & & A_n \in \mathcal{G} & & \mu_2 \text{ Maß} \end{array}$$

also ist auch $\sum_{n \in \mathbb{N}} A_n \in \mathcal{G}$.

Aus a) - c) folgt (2).

2. Schritt: Allgemeiner Fall

Da $\mu_1|_{\mathcal{C}}$ und $\mu_2|_{\mathcal{C}}$ σ -endlich sind, existieren $A_{i,n} \in \mathcal{C}$ mit $A_{i,n} \uparrow \Omega$ ($n \rightarrow \infty$) und $\mu_i(A_{i,n}) < \infty$ für alle $n \in \mathbb{N}$ ($i \in \{1, 2\}$).

Setze

$$A_n := A_{1,n} \cap A_{2,n}.$$

Da $\mathcal{C}$ $\cap$ -stabil ist, gilt $A_n \in \mathcal{C}$. Weiter erfüllt A_n

$$A_n \uparrow \Omega,$$

da

$$\bigcup_{n \in \mathbb{N}} A_n \supseteq \left(\bigcup_{n \geq N} A_{1,n} \right) \cap A_{2,N} = \Omega \cap A_{2,N}$$

für alle $N \in \mathbb{N}$ und damit

$$\bigcup_{n \in \mathbb{N}} A_n \supseteq \bigcup_{N \in \mathbb{N}} A_{2,N} = \Omega.$$

Es gilt außerdem

$$\mu_1(A_n) \leq \mu_1(A_{1,n}) < \infty \text{ und } \mu_2(A_n) \leq \mu_2(A_{2,n}) < \infty,$$

da μ_1, μ_2 als Maße monoton sind.

Wir zeigen im Folgenden:

$$(*) \quad \forall n \in \mathbb{N} \forall A \in \mathcal{F}(\mathcal{C}) : \mu_1(A \cap A_n) = \mu_2(A \cap A_n).$$

Wegen $A \cap A_n \uparrow A$ (folgt aus $A_n \uparrow \Omega$) folgt daraus mit der Stetigkeit des W-Maßes von unten:

$$\mu_1(A) = \lim_{n \rightarrow \infty} \mu_1(A \cap A_n) = \lim_{n \rightarrow \infty} \mu_2(A \cap A_n) = \mu_2(A)$$

für alle $A \in \mathcal{F}(\mathcal{C})$, was zu zeigen war.

Nachweis von (*): Sei $n \in \mathbb{N}$ fest.

Setze

$$\mathcal{C}_n = A_n \cap \mathcal{C} = \{A_n \cap A : A \in \mathcal{C}\}.$$

Da $\mathcal{C}$ $\cap$ -stabil ist, gilt $\mathcal{C}_n \subseteq \mathcal{C}$.

$\mathcal{C}_n$ ist ein $\cap$ -stabiler Erzeuger von $\mathcal{F}_{A_n}(\mathcal{C}_n)$, und die Maße μ_1 und μ_2 stimmen auf $\mathcal{C}_n$ überein (da sie sogar auf $\mathcal{C}$ übereinstimmen) und erfüllen

$$\mu_1(A_n) = \mu_2(A_n) \quad (\text{da } A_n \in \mathcal{C}_n)$$

und

$$\mu_1(A_n) < \infty \quad (\text{nach Konstruktion von } A_n).$$

Mit Schritt 1 folgt daraus

$$\mu_1(A) = \mu_2(A) \text{ für alle } A \in \mathcal{F}_{A_n}(\mathcal{C}_n).$$

Wegen

$$\mathcal{F}_{A_n}(\mathcal{C}_n) = \mathcal{F}_{A_n}(A_n \cap \mathcal{C}) = A_n \cap \mathcal{F}_\Omega(\mathcal{C})$$

(vgl. Übungen), folgt daraus aber (*). □

Definition 1.43

Eine Funktion $G : \mathbb{R} \rightarrow \mathbb{R}$ heißt **maßdefinierende Funktion**, falls sie monoton wachsend (d. h. $x_1 \leq x_2 \Rightarrow G(x_1) \leq G(x_2)$) und rechtsseitig stetig ist.

Gilt außerdem

$$\lim_{x \rightarrow -\infty} G(x) = 0$$

und

$$\lim_{x \rightarrow \infty} G(x) = 1,$$

so heißt G (eindimensionale) **Verteilungsfunktion**.

Beispiele

a) $G(x) = \begin{cases} 3x + 7 & \text{für } x \geq 0 \\ -x^2 & \text{für } x < 0 \end{cases}$ ist maßdefinierende Funktion.

b) $G(x) = x$ ($x \in \mathbb{R}$) ist maßdefinierende Funktion.

c) $G(x) = 1_{[u, \infty)}(x)$ ist Verteilungsfunktion.

d) $G(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$ ist Verteilungsfunktion.

Satz 1.44

a) Zu jedem Maß $\mu : \mathcal{B} \rightarrow \overline{\mathbb{R}}$ mit

$$\mu((a, b]) < \infty \text{ für alle } a, b \in \mathbb{R}, a \leq b$$

existiert eine – bis auf eine additive Konstante eindeutige – maßdefinierende Funktion $G : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$\mu((a, b]) = G(b) - G(a) \quad \text{für alle } a, b \in \mathbb{R}, a \leq b,$$

und umgekehrt.

b) Zu jedem Wahrscheinlichkeitsmaß $\mu : \mathcal{B} \rightarrow \mathbb{R}$ existiert genau eine Verteilungsfunktion $F : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$\mu((a, b]) = F(b) - F(a) \text{ für alle } a, b \in \mathbb{R}, a \leq b$$

und umgekehrt.

Hierbei gilt: $F(x) = \mu((-\infty, x])$ ($x \in \mathbb{R}$).

Beweis

a) a_1) Sei $\mu : \mathcal{B} \rightarrow \overline{\mathbb{R}}$ Maß mit $\mu((a, b]) < \infty$ ($a, b \in \mathbb{R}, a \leq b$).

Wir zeigen: Es existiert eine maßdefinierende Funktion G mit

$$\mu((a, b]) = G(b) - G(a) \quad (a, b \in \mathbb{R}, a \leq b).$$

Dazu definiere

$$G(x) = \begin{cases} \mu((0, x]) & \text{falls } x \geq 0, \\ -\mu((x, 0]) & \text{falls } x < 0. \end{cases}$$

Dann gilt:

(i) G ist monoton wachsend, denn:

Sei $x_1, x_2 \in \mathbb{R}$ mit $x_1 \leq x_2$.

Ist $x_1, x_2 \geq 0$, so ist $(0, x_1] \subseteq (0, x_2]$

und daher $G(x_1) = \mu((0, x_1]) \leq \mu((0, x_2]) = G(x_2)$.

Ist $x_1 < 0, x_2 \geq 0$, so ist

$$G(x_1) = -\mu((x_1, 0]) \leq 0 \leq \mu((0, x_2]) = G(x_2).$$

Ist $x_1, x_2 < 0$, so gilt

$$(x_1, 0] \supseteq (x_2, 0]$$

und daher

$$G(x_1) = -\mu((x_1, 0]) \leq -\mu((x_2, 0]) = G(x_2).$$

(ii) G ist rechtsseitig stetig, denn:

Sei

$$x_1 > x_2 > \dots > x_n \rightarrow x \quad (n \rightarrow \infty)$$

Fall 1: $x \geq 0$

Dann gilt $(0, x_n] \downarrow (0, x]$, und mit Stetigkeit des W-Maßes von oben folgt

$$G(x_n) = \mu((0, x_n]) \rightarrow \mu((0, x]) = G(x) \quad (n \rightarrow \infty).$$

Fall 2: $x < 0$

Dann gilt $(x_n, 0] \uparrow (x, 0]$, und mit Stetigkeit des W-Maßes von unten folgt wieder

$$G(x_n) = -\mu((x_n, 0]) \rightarrow -\mu((x, 0]) = G(x) \quad (n \rightarrow \infty)$$

Da G monoton ist, folgt daraus die rechtsseitige Stetigkeit.

(iii) Für G gilt

$$\mu((a, b]) = G(b) - G(a) \quad (a, b \in \mathbb{R}, a \leq b).$$

Fall 1: $a, b \geq 0$:

$$\begin{aligned} G(b) - G(a) &= \mu((0, b]) - \mu((0, a]) \\ &= \mu((0, b] \setminus (0, a]) = \mu((a, b]) \end{aligned}$$

Fall 2: $a < 0, b \geq 0$:

$$G(b) - G(a) = \mu((0, b]) + \mu((a, 0]) = \mu((a, b])$$

.

Fall 3: $a < 0, b < 0$:

$$\begin{aligned} G(b) - G(a) &= -\mu((b, 0]) + \mu((a, 0]) \\ &= \mu((a, 0] \setminus (b, 0]) = \mu((a, b]) \rightsquigarrow a_1). \end{aligned}$$

(iv) G ist bis auf eine additive Konstante eindeutig.

Seien G_1, G_2 maßdefinierende Funktionen mit

$$G_1(b) - G_1(a) = \mu((a, b]) = G_2(b) - G_2(a) \text{ für } a, b \in \mathbb{R}, a \leq b$$

Dann gilt für $x \geq 0$ (mit $b = x, a = 0$):

$$G_1(x) = G_2(x) - G_2(0) + G_1(0)$$

und für $x < 0$ (mit $b = 0, a = x$)

$$G_1(x) = G_2(x) - G_2(0) + G_1(0)$$

$$\rightsquigarrow G_1 = G_2 + \text{const} \quad \text{mit} \quad \text{const} = G_1(0) - G_2(0).$$

a_2) Sei $G : \mathbb{R} \rightarrow \mathbb{R}$ eine maßdefinierende Funktion.

Dann lässt sich die durch

$$\mu((a, b]) := G(b) - G(a) \geq 0 \quad (\text{da } G \text{ monoton wachsend})$$

$(a, b \in \mathbb{R}, a \leq b)$

definierte Funktion eindeutig zu einem Inhalt (!) $\mu : \mathcal{E}_1 \rightarrow \mathbb{R}$ fortsetzen.

Analog zu Satz 1.31 (im Spezialfall $G(x) = x$) folgt aus der rechtsseitigen Stetigkeit und der Monotonie von G , dass μ sogar ein **Maß** auf $\mathcal{E}_1$ ist.

Wegen

$$(-n, n] \uparrow \mathbb{R} \quad \text{und} \quad \mu((-n, n]) = G(n) - G(-n) < \infty$$

für alle $n \in \mathbb{R}$ lässt sich μ auf eindeutige Weise auf $\mathcal{B}$ fortsetzen (vgl. Satz 1.34 und Satz 1.36). Dieses μ erfüllt nach Konstruktion

$$\mu((a, b]) = G(b) - G(a) < \infty \text{ für alle } a, b \in \mathbb{R}, a < b.$$

$\rightsquigarrow a)$

b) b_1) Setze $F(x) = \mu((-\infty, x])$. Analog zu a_1) sieht man, dass F Verteilungsfunktion ist, und darüber hinaus für $a \leq b$ gilt:

$$\begin{aligned} F(b) - F(a) &= \mu((-\infty, b]) - \mu((-\infty, a]) \\ &= \mu((-\infty, b] \setminus (-\infty, a]) \\ &= \mu((a, b]). \end{aligned}$$

F ist eindeutig, da für Verteilungsfunktionen F_1, F_2 aus

$$F_1(b) - F_1(a) = F_2(b) - F_2(a) \text{ für alle } a, b \in \mathbb{R}, a \leq b$$

mit $a \rightarrow -\infty$

$$F_1(b) = F_2(b) \quad (b \in \mathbb{R})$$

folgt.

b₂) Existenz und Eindeutigkeit des Maßes

$$\mu : \mathcal{B} \rightarrow \mathbb{R}$$

mit

$$\mu((a, b]) = F(b) - F(a) \quad (a, b \in \mathbb{R}, a \leq b)$$

folgt aus a).

Wegen $(-n, n] \uparrow \mathbb{R}$, μ stetig von unten gilt

$$\mu(\mathbb{R}) = \lim_{n \rightarrow \infty} \mu((-n, n]) = \lim_{n \rightarrow \infty} (F(n) - F(-n)) = 1 - 0 = 1,$$

also ist μ sogar W-Maß.

Aus

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

folgt mit $a \rightarrow -\infty$:

$$\mu((-\infty, x]) = \lim_{a \rightarrow -\infty} \mu((a, x]) = \lim_{a \rightarrow -\infty} (F(x) - F(a)) = F(x) - 0$$

↪ Beh. $\square$

Beispiele

a) Das zu $G(x) = x$ gehörende Maß ist das LB-Maß.

b) Das zu

$$G(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

gehörende Maß heißt (standardisierte) Normalverteilung.

1.4 Anwendung in der Wahrscheinlichkeitstheorie

Definition 1.45

Sei $\mathcal{A}$ eine σ -Algebra in Ω . Eine Abbildung $P : \mathcal{A} \rightarrow \mathbb{R}$ heißt ein **Wahrscheinlichkeitsmaß** (W-Maß) auf $\mathcal{A}$, wenn

$$P(\emptyset) = 0,$$

$$P \text{ positiv, d.h. } \forall A \in \mathcal{A} : P(A) \geq 0,$$

P σ -additiv, d.h. für alle paarweise disjunkte $A_n \in \mathcal{A}$ ($n = 1, 2, \dots$) gilt

$$P\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n),$$

$$P \text{ normiert, d.h. } P(\Omega) = 1.$$

$(\Omega, \mathcal{A}, P)$ heißt dann ein **Wahrscheinlichkeitsraum** (probability space).

Ein Wahrscheinlichkeitsraum ist ein normierter Maßraum.

Ist $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum, so bezeichnet man die Grundmenge Ω auch als Merkmalraum, Stichprobenraum (sample space), die Mengen $A \in \mathcal{A}$ als Ereignisse, $P(A)$ als Wahrscheinlichkeit von A , die Elemente $\omega \in \Omega$ bzw. die 1-Punkt-Mengen $\{\omega\} \subset \Omega$ (nicht notwendig $\in \mathcal{A}$) als Elementarereignisse (auch Realisierungen).

Ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ mit Ω endlich, $\mathcal{A} = \mathcal{P}(\Omega)$, $P(\{\omega\}) = \frac{1}{|\Omega|}$ ($\omega \in \Omega$) ($|\Omega|$ = Anzahl der Elemente in Ω) heißt **Laplacescher Wahrscheinlichkeitsraum**:

$$\text{Für } A \in \mathcal{A} \text{ gilt } P(A) = \frac{|A|}{|\Omega|} = \frac{\text{Anzahl der "günstigen" Fälle}}{\text{Anzahl der "möglichen" Fälle}}$$

Es sei $(\Omega, \mathcal{A}) = (\mathbb{R}, \mathcal{B})$ und $(p_k)_{k \in \mathbb{N}_0}$ ($\mathbb{N}_0 := \{0, 1, \dots\}$) eine Zahlenfolge mit $p_k \geq 0$ ($k \in \mathbb{N}_0$), $\sum_{k=0}^{\infty} p_k = 1$. Dann ist $P : \mathcal{B} \rightarrow \mathbb{R}$ mit

$$P(A) := \sum_{k \in A \cap \mathbb{N}_0} p_k, \quad A \in \mathcal{A},$$

ein W-Maß. (p_k) heißt **Zähldichte** zu P .

Ist speziell $p_k = e^{-\lambda} \frac{\lambda^k}{k!}$, ($k \in \mathbb{N}_0$) bei festem $\lambda > 0$, so heißt P **Poisson-Verteilung mit Parameter λ** ... Bezeichnung $\pi(\lambda)$.

Satz 1.46

(1. Lemma von Borel und Cantelli).

Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$; $A_n \in \mathcal{A}$ ($n \in \mathbb{N}$). Mit

$$\begin{aligned}\overline{\lim} A_n &:= \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \{\omega \in \Omega \mid \text{existieren unendlich viele } n \in \mathbb{N} \text{ mit } \omega \in A_n\} \\ &= \text{“Ereignis, dass unendlich viele } A_n \text{ eintreten”}\end{aligned}$$

gilt dann:

$$\sum_{n=1}^{\infty} P(A_n) < \infty \implies P(\overline{\lim} A_n) = 0, \quad \text{d.h. P-f.s. gilt: } A_n \text{ tritt nur endlich oft auf}$$

Beispiel

Lernen aus Erfahrung

Student macht bei n -ter Aufgabe Fehler mit Wk. $\frac{1}{n^2} = P(A_n)$.

D. g.: $P(\text{“Student macht unendlich viele Fehler”}) = 0$,

da $\sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$.

Beweis von Satz 1.46

Für jedes $n \in \mathbb{N}$ gilt wegen $\overline{\lim} A_n \subset \bigcup_{k=n}^{\infty} A_k$.

$$P(\overline{\lim} A_n) \leq P\left(\bigcup_{k=n}^{\infty} A_k\right) \leq \sum_{k=n}^{\infty} P(A_k) \rightarrow 0 \quad (n \rightarrow \infty), \quad \text{da } \sum_{k=1}^{\infty} P(A_k) < \infty. \quad \square$$

Definition 1.47

Eine Funktion $F : \mathbb{R} \rightarrow \mathbb{R}$, die monoton wachsend (im Sinne von nicht fallend) und rechtsseitig stetig ist mit

$$\lim_{x \rightarrow -\infty} F(x) = 0, \quad \lim_{x \rightarrow +\infty} F(x) = 1,$$

heißt (eindimensionale) **Verteilungsfunktion** (VF).

Satz 1.48

Zu jedem W-Maß P auf $\mathcal{B}$ gibt es genau eine VF $F : \mathbb{R} \rightarrow \mathbb{R}$, so dass

$$(*) \quad F(b) - F(a) = P((a, b]), \quad a \leq b, \quad a, b \in \mathbb{R}$$

gilt, und umgekehrt ... Bijektion $P \longleftrightarrow F$. Hierbei $F(x) = P((-\infty, x])$, $x \in \mathbb{R}$.

Beweis

Siehe Satz 1.44 b) □

Beispiel: $\Phi : \mathbb{R} \rightarrow [0, 1]$ mit

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt, \quad x \in \mathbb{R},$$

ist eine VF. Das zugehörige W-Maß auf $\mathcal{B}$ ist die sogenannte **standardisierte Normalverteilung** ... Bezeichnung $N(0, 1)$.

Verallgemeinerung: **Normalverteilung (Gauß-Verteilung)** $N(a, \sigma^2)$ ($a, \sigma \in \mathbb{R}$, $\sigma > 0$) mit durch

$$F(x) := \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^x e^{-\frac{(t-a)^2}{2\sigma^2}} dt, \quad x \in \mathbb{R},$$

gebener VF F .

Bemerkung 1.49

Definition 1.47, Satz 1.48 lassen sich auf $\mathcal{B}_n$ und $\mathbb{R}^n$ als Definitionsbereiche von P bzw. F verallgemeinern statt $\mathcal{B}$ bzw. $\mathbb{R}$. Dabei ist die VF F zu einem W-Maß P auf $\mathcal{B}_n$ gegeben durch

$$F(x_1, \dots, x_n) = P(\{(u_1, \dots, u_n) \in \mathbb{R}^n \mid u_1 \leq x_1, \dots, u_n \leq x_n\})$$

für $(x_1, \dots, x_n) \in \mathbb{R}^n$.

Definition 1.50

Es sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum, $A \in \mathcal{A}$, $B \in \mathcal{A}$ mit $P(B) > 0$. Dann heißt

$$P(A \mid B) := \frac{P(A \cap B)}{P(B)}$$

die **bedingte Wahrscheinlichkeit** (conditional probability) von A unter der Bedingung B .

Satz 1.51

Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum.

a) Bei festem $B \in \mathcal{A}$ mit $P(B) > 0$ ist $P(\cdot | B)$ ein W-Maß auf $\mathcal{A}$, das auf B konzentriert ist [d.h. $P(B | B) = 1$].

b) Für $A, B \in \mathcal{A}$ mit $P(A) > 0, P(B) > 0$ gilt

$$P(A | B) = P(B | A) \cdot \frac{P(A)}{P(B)}.$$

c) Für $A_n \in \mathcal{A}$ ($n = 1, \dots, m$) mit $P(\bigcap_{n=1}^{m-1} A_n) > 0$ gilt

$$P(\bigcap_{n=1}^m A_n) = P(A_1) \cdot P(A_2 | A_1) \cdot P(A_3 | A_1 \cap A_2) \cdot \dots \cdot P(A_m | \bigcap_{n=1}^{m-1} A_n).$$

Beweis

Vgl. Elementare Wahrscheinlichkeitstheorie und Statistik □

Satz 1.52

Es sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum, $\{B_n; n \in I\}$ eine höchstens abzählbare Familie paarweise disjunkter Ereignisse $B_n \in \mathcal{A}$ mit $P(B_n) > 0$ ($n \in I$) und $\sum_{n \in I} B_n = \Omega$.

Dann gilt

a)
$$P(A) = \sum_{n \in I} P(B_n)P(A | B_n) \quad (A \in \mathcal{A})$$

(Formel von der totalen Wahrscheinlichkeit),

b)
$$P(B_k | A) = \frac{P(B_k) \cdot P(A | B_k)}{\sum_{n \in I} P(B_n)P(A | B_n)}, \quad k \in I \quad (A \in \mathcal{A} \text{ mit } P(A) > 0)$$

(Formel von Bayes).

Beweis: Vgl. Elementare Wahrscheinlichkeitstheorie und Statistik □

Bezeichnungen [Maßraum $(\Omega, \mathcal{A}, \mu)$, Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$]:

μ -fast überall (μ -f.ü.) in Ω bzw. P -fast sicher (P -f.s.) in $\Omega \dots$ überall bis auf eine Menge

$\in \mathcal{A}$ vom μ -Maß bzw. P -Maß Null [oder eine Teilmenge hiervon].

L -fast überall (L -f.ü.) in $\mathbb{R}^n \dots$ überall bis auf eine Menge $\in \mathcal{B}_n$ vom L-B-Maß Null [oder eine Teilmenge hiervon].

2. Messbare Abbildungen, Zufallsvariablen und Verteilungen

Beispiel: Werfen zweier echter Würfel.

Wie groß ist die Wahrscheinlichkeit, dass die Summe der Augenzahlen größer oder gleich 10 ist?

Stochastische Modellierung:

$$\Omega = \{(\omega_1, \omega_2) : \omega_1, \omega_2 \in \{1, \dots, 6\}\}, \mathcal{A} = \mathcal{P}(\Omega), P(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{36}.$$

Die gesuchte Wahrscheinlichkeit ist

$$P(\{(4, 6), (5, 5), (5, 6), (6, 4), (6, 5), (6, 6)\}) = \frac{6}{36} = \frac{1}{6}.$$

Statt der Werte der beiden Würfel betrachte die Summe der Augenzahlen

$$X((\omega_1, \omega_2)) = \omega_1 + \omega_2.$$

Wahrscheinlichkeitsraum für Zufallsexperiment mit Ergebnis $X((\omega_1, \omega_2))$:

$$\Omega' = \{2, 3, \dots, 12\}, \mathcal{A}' = \mathcal{P}(\Omega') \text{ und}$$

$$P_X : \mathcal{A}' \rightarrow \mathbb{R}$$

definiert durch

$$P_X(A') = P(\{\omega \in \Omega \mid X(\omega) \in A'\}) = P(X^{-1}(A'))$$

Also ist die gesuchte Wahrscheinlichkeit gleich:

$$\begin{aligned} P_X(\{10, 11, 12\}) &= P(X^{-1}(\{10, 11, 12\})) \\ &= P(\{(\omega_1, \omega_2) \in \Omega : \omega_1 + \omega_2 \in \{10, 11, 12\}\}) \\ &= (\{(4, 6), (5, 5), (5, 6), (6, 4), (6, 5), (6, 6)\}) \\ &= \frac{6}{36} = \frac{1}{6}. \end{aligned}$$

Allgemein: Abbildung $X : \Omega \rightarrow \Omega'$.

Damit

$$P_X(A') := P(X^{-1}(A'))$$

wohldefiniert ist für alle $A' \in \mathcal{A}'$, muss gelten:

$$X^{-1}(A') \in \mathcal{A} \quad (A' \in \mathcal{A}').$$

Abbildungen X mit dieser Eigenschaft heißen **Zufallsvariablen**.

2.1 Messbare Abbildungen

Definition 2.1. Zwei nichtleere Mengen Ω, Ω' ; $A' \subset \Omega'$; Abbildung $X : \Omega \rightarrow \Omega'$. Die Menge

$$\{\omega \in \Omega \mid X(\omega) \in A'\} =: X^{-1}(A') =: [X \in A']$$

(in Ω) heißt das **Urbild** von A' bezüglich der Abbildung X ; die somit definierte Abbildung $X^{-1} : \mathcal{P}(\Omega') \rightarrow \mathcal{P}(\Omega)$ heißt **Urbildfunktion** (zu unterscheiden von einer inversen Funktion!).

Bezeichnung. Sei $X : \Omega \rightarrow \Omega'$, $\mathcal{C}'$ Mengensystem in Ω' . $\underbrace{X^{-1}(\mathcal{C}')} := \{X^{-1}(A') \mid A' \in \mathcal{C}'\}$.
ein Mengensystem in Ω .

Satz 2.2. Sei $X : \Omega \rightarrow \Omega'$.

- a) X^{-1} und Mengenoperationen $\cup, \sum, \cap, ^c, \setminus$ sind vertauschbar;
z.B. $X^{-1}(\bigcup_{\alpha \in I} A'_\alpha) = \bigcup_{\alpha \in I} X^{-1}(A'_\alpha)$ ($A'_\alpha \subset \Omega'$, $\alpha \in \text{Indexbereich } I$).
- b) $X^{-1}(\emptyset) = \emptyset$; $X^{-1}(\Omega') = \Omega$; $A' \subset B' \subset \Omega' \implies X^{-1}(A') \subset X^{-1}(B')$.
- c) $\mathcal{A}'$ σ -Algebra in $\Omega' \implies X^{-1}(\mathcal{A}')$ σ -Algebra in Ω .
- d) $\mathcal{C}' \subset \mathcal{P}(\Omega') \implies X^{-1}(\mathcal{F}_{\Omega'}(\mathcal{C}')) = \mathcal{F}_\Omega(X^{-1}(\mathcal{C}'))$.

Beweis:

a) , b) elementar

- c) – $\emptyset \in X^{-1}(\mathcal{A}')$ da $\emptyset = X^{-1}(\emptyset)$ und $\emptyset \in \mathcal{A}'$
– Sei $A = X^{-1}(A') \in X^{-1}(\mathcal{A}')$ mit $A' \in \mathcal{A}'$.

D. g.

$$A^c = \Omega \setminus X^{-1}(A') \stackrel{a)}{=} X^{-1}(\Omega' \setminus A') \in X^{-1}(\mathcal{A}')$$

da $\Omega' \setminus A' \in \mathcal{A}'$

- Sei $A_n = X^{-1}(A'_n) \in X^{-1}(\mathcal{A}')$ mit $A'_n \in \mathcal{A}'$ ($n \in \mathbb{N}$)

D. g.

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} X^{-1}(A'_n) \stackrel{a)}{=} X^{-1}\left(\bigcup_{n=1}^{\infty} A'_n\right) \in X^{-1}(\mathcal{A}')$$

da $\bigcup_{n=1}^{\infty} A'_n \in \mathcal{A}'$.

d) Zu zeigen:

$$X^{-1}(\mathcal{F}_{\Omega'}(\mathcal{C}')) = \mathcal{F}_\Omega(X^{-1}(\mathcal{C}'))$$

“ $\supset$ ” Folgt aus $X^{-1}(\mathcal{C}') \subset X^{-1}(\mathcal{F}_{\Omega'}(\mathcal{C}'))$ und $X^{-1}(\mathcal{F}_{\Omega'}(\mathcal{C}'))$ σ -Algebra nach c).

“ $\subset$ ” Man sieht leicht:

$$\bar{\mathcal{A}} := \{\bar{A} \subset \Omega' : X^{-1}(\bar{A}) \in \mathcal{F}_{\Omega}(X^{-1}(\mathcal{C}'))\}$$

ist σ -Algebra in Ω' mit $\mathcal{C}' \subset \bar{\mathcal{A}}$.

$\Rightarrow \mathcal{F}_{\Omega'}(\mathcal{C}') \subset \bar{\mathcal{A}} \Rightarrow X^{-1}(\mathcal{F}_{\Omega'}(\mathcal{C}')) \subset \mathcal{F}_{\Omega}(X^{-1}(\mathcal{C}'))$.

□

Def. 2.3 $(\Omega, \mathcal{A})$ und $(\Omega', \mathcal{A}')$ seien Messräume, weiter sei $f : \Omega \rightarrow \Omega'$.

f heißt $\mathcal{A} - \mathcal{A}'$ -**messbar** (kurz: messbar), falls gilt:

$$f^{-1}(A') := \{\omega \in \Omega \mid f(\omega) \in A'\} \in \mathcal{A}$$

für alle $A' \in \mathcal{A}'$. d. h.:

$$f^{-1}(\mathcal{A}') = \{f^{-1}(A') : A' \in \mathcal{A}'\} \subseteq \mathcal{A}$$

bzw. die Urbilder messbarer Mengen in Ω' sind messbare Mengen in Ω .

In diesem Falle Schreibweise $f: (\Omega, \mathcal{A}) \rightarrow (\Omega', \mathcal{A}')$.

Bemerkung 2.4. In Satz 2.2c ist $X^{-1}(\mathcal{A}')$ die kleinste der σ -Algebren $\mathcal{A}$ in Ω mit $\mathcal{A} - \mathcal{A}'$ -Messbarkeit von X .

Beispiele:

a) Jede konstante Funktion ist messbar (da Urbilder immer gleich Ω oder gleich $\emptyset$ sind).

b) $(\Omega, \mathcal{A}) = (\Omega', \mathcal{A}') = (\mathbb{R}, \mathcal{B})$, $A \subseteq \mathbb{R}$ und $f = \chi_A$ mit

$$\chi_A(x) = \begin{cases} 1 & , \quad x \in A \\ 0 & , \quad x \notin A \end{cases}$$

(sog. Indikatorfunktion). Für $B \in \mathcal{B}$ gilt

$$\chi_A^{-1}(B) = \begin{cases} A & \text{falls } 0 \notin B, 1 \in B \\ A^c & \text{falls } 0 \in B, 1 \notin B \\ \emptyset & \text{falls } 0 \notin B, 1 \notin B \\ \mathbb{R} & \text{falls } 0 \in B, 1 \in B. \end{cases}$$

daher ist χ_A genau dann messbar, wenn $A \in \mathcal{B}$ gilt.

Lemma 2.5

$(\Omega_i, \mathcal{A}_i)$ ($i = 1, 2, 3$) seien Messräume. Für

$$f_1 : (\Omega_1, \mathcal{A}_1) \rightarrow (\Omega_2, \mathcal{A}_2) \text{ und } f_2 : (\Omega_2, \mathcal{A}_2) \rightarrow (\Omega_3, \mathcal{A}_3)$$

gilt immer

$$f_2 \circ f_1 : (\Omega_1, \mathcal{A}_1) \rightarrow (\Omega_3, \mathcal{A}_3),$$

d.h. Verkettung messbarer Funktionen ergibt messbare Funktion.

Beweis Für $A_3 \in \mathcal{A}_3$ gilt

$$f_2^{-1}(A_3) \in \mathcal{A}_2$$

(da f_2 messbar), woraus folgt

$$(f_2 \circ f_1)^{-1}(A_3) = f_1^{-1}(f_2^{-1}(A_3)) \in \mathcal{A}_1$$

(da f_1 messbar). □

Satz 2.6: $(\Omega, \mathcal{A})$ und $(\Omega', \mathcal{A}')$ seien Messräume, $\mathcal{C}$ sei Erzeugersystem von $\mathcal{A}'$, d.h. $\mathcal{A}' = \mathcal{F}_{\Omega'}(\mathcal{C})$. Dann sind äquivalent:

- (i) $f : \Omega \rightarrow \Omega'$ ist $\mathcal{A} - \mathcal{A}'$ messbar
- (ii) $f^{-1}(\mathcal{C}) \subseteq \mathcal{A}$, d. h. $f^{-1}(A') \in \mathcal{A}$ für alle $A' \in \mathcal{C}$.

Beweis:

“(i) $\Rightarrow$ (ii)”: klar

“(ii) $\Rightarrow$ (i)”: Da $\mathcal{A}$ σ -Algebra ist, folgt aus

$$f^{-1}(\mathcal{C}) \subseteq \mathcal{A}$$

immer auch $\mathcal{F}_{\Omega}(f^{-1}(\mathcal{C})) \subseteq \mathcal{A}$. Behauptung folgt mit

$$\mathcal{F}_{\Omega}(f^{-1}(\mathcal{C})) = f^{-1}(\mathcal{F}_{\Omega'}(\mathcal{C})) = f^{-1}(\mathcal{A}')$$

(vgl. Satz 2.2 d)). □

Beispiel: Jede stetige Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ist $\mathcal{B}_n - \mathcal{B}_m$ -messbar (da Urbild offener Mengen immer offen ist)

Satz 2.6 ermöglicht eine einfache Charakterisierung messbarer numerischer Funktionen, d. h. von Funktionen $f : (\Omega, \mathcal{A}) \rightarrow (\Omega', \mathcal{A}')$ mit $(\Omega', \mathcal{A}') = (\mathbb{R}, \mathcal{B})$ bzw.

$$\Omega' = \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\},$$

$$\mathcal{A}' = \overline{\mathcal{B}} = \{B, B \cup \{+\infty\}, B \cup \{-\infty\}, B \cup \{+\infty, -\infty\} \mid B \in \mathcal{B}\}$$

Korollar 2.7 Sei $(\Omega, \mathcal{A})$ Messraum und f eine Funktion $f : \Omega \rightarrow \mathbb{R}$ oder $f : \Omega \rightarrow \overline{\mathbb{R}}$. Dann sind äquivalent:

- (i) f ist $\mathcal{A} - \mathcal{B}$ - bzw. $\mathcal{A} - \overline{\mathcal{B}}$ -messbar

(ii) $[f < \alpha] := \{\omega \in \Omega : f(\omega) < \alpha\} = f^{-1}((-\infty, \alpha)) \in \mathcal{A}$ für alle $\alpha \in \mathbb{R}$

(iii) $[f \leq \alpha] \in \mathcal{A}$ für alle $\alpha \in \mathbb{R}$

(iv) $[f > \alpha] \in \mathcal{A}$ für alle $\alpha \in \mathbb{R}$

(v) $[f \geq \alpha] \in \mathcal{A}$ für alle $\alpha \in \mathbb{R}$.

Beweis: Für $f : \Omega \rightarrow \mathbb{R}$ folgt die Behauptung unmittelbar aus Satz 2.6 und den in Satz 1.15 angegebenen Erzeugendensystemen von $\mathcal{B}$. Für $f : \Omega \rightarrow \overline{\mathbb{R}}$ wendet man zu Satz 1.15 analoge Aussagen für $\overline{\mathcal{B}}$ an. $\square$

Korollar 2.8. Für zwei Abbildungen $X, Y : (\Omega, \mathcal{A}) \rightarrow (\overline{\mathbb{R}}, \overline{\mathcal{B}})$ gilt

$$\underbrace{[X < Y]}, \quad [X \leq Y], \quad [X = Y], \quad [X \neq Y] \in \mathcal{A}.$$

$$:= \{\omega \in \Omega \mid X(\omega) < Y(\omega)\}.$$

Beweis:

- $[X < Y] = \bigcup_{r \in \mathbb{Q}} [X < r] \cap [r < Y]$
- $[X \leq Y] = [Y < X]^c$
- $[X = Y] = [X \leq Y] \cap [Y \leq X]$
- $[X \neq Y] = [X = Y]^c$

$\square$

Lemma 2.9: Sei $(\Omega, \mathcal{A})$ Messraum und seien $f_n : \Omega \rightarrow \overline{\mathbb{R}}$ $\mathcal{A} - \overline{\mathcal{B}}$ -messbar ($n \in \mathbb{N}$). Dann sind auch

$$\inf_n f_n, \sup_n f_n, \overline{\lim}_n f_n = \limsup_n f_n, \underline{\lim}_n f_n = \liminf_n f_n$$

sowie (falls existent)

$$\lim_n f_n$$

$\mathcal{A} - \overline{\mathcal{B}}$ -messbar.

Beweis: Mit Korollar 2.7 folgt die Behauptung aus

1. $[\inf_n f_n < \alpha] = \bigcup_{n \in \mathbb{N}} [f_n < \alpha] \in \mathcal{A}$
(da $[f_n < \alpha] \in \mathcal{A}$ und $\mathcal{A}$ σ -Algebra)
2. $[\sup_n f_n > \alpha] = \bigcup_{n \in \mathbb{N}} [f_n > \alpha] \in \mathcal{A}$
(analog)

3. $\overline{\lim}_n f_n = \inf_n \sup_{k \geq n} f_k$ messbar nach 1. und 2.

4. $\underline{\lim}_n f_n = \sup_n \inf_{k \geq n} f_k$ messbar nach 1. und 2.

5. $\lim_n f_n = \underline{\lim}_n f_n$ (falls existent) messbar nach 4. □

Lemma 2.10

$(\Omega, \mathcal{A})$ sei Messraum.

a) Sind $f_n : \Omega \rightarrow \mathbb{R}$ $\mathcal{A}$ - $\mathcal{B}$ -messbar ($n = 1, \dots, m$) und ist $g : \mathbb{R}^m \rightarrow \mathbb{R}$ $\mathcal{B}_m$ - $\mathcal{B}$ -messbar, so ist $g \circ (f_1, \dots, f_m)$ $\mathcal{A}$ - $\mathcal{B}$ -messbar.

b) Sind $f_i : \Omega \rightarrow \mathbb{R}$ $\mathcal{A}$ - $\mathcal{B}$ -messbar ($i = 1, 2$) und sind $\alpha, \beta \in \mathbb{R}$, so sind

$$\alpha f_1 + \beta \cdot f_2, f_1 \cdot f_2 \text{ und } \frac{f_1}{f_2} \text{ (sofern existent)}$$

$\mathcal{A}$ - $\mathcal{B}$ -messbar.

Beweis:

a) Nach Satz 2.6 ist $(f_1, \dots, f_m) : \Omega \rightarrow \mathbb{R}^m$ $\mathcal{A}$ - $\mathcal{B}_m$ -messbar, da für alle $\alpha_1, \dots, \alpha_m \in \mathbb{R}$ gilt

$$[(f_1, \dots, f_m) \leq (\alpha_1, \dots, \alpha_m)] = \bigcap_{n=1}^m [f_n \leq \alpha_n] \in \mathcal{A}$$

Mit Lemma 2.5 folgt die Behauptung.

b) Behauptung folgt aus a) unter Beachtung der Messbarkeit der stetigen (!) Abbildungen

$$(x_1, x_2) \mapsto \alpha x_1 + \beta x_2$$

$$(x_1, x_2) \mapsto x_1 \cdot x_2,$$

$$(x_1, x_2) \mapsto \frac{x_1}{x_2}$$

Die letzte Abbildung ist dabei als punktwiser Grenzwert der (stetigen) messbaren Abbildungen

$$g_n(x_1, x_2) = \begin{cases} \frac{x_1}{x_2}, & \text{falls } |x_2| > \frac{1}{n} \\ n \cdot x_1, & \text{sonst} \end{cases}$$

messbar. □

Lemma 2.11

Mit $f : \Omega \rightarrow \overline{\mathbb{R}}$ sind auch

$$f^+ = \max\{f, 0\}$$

$$f^- = \max\{-f, 0\}$$

und

$$|f|$$

$\mathcal{A} - \overline{\mathcal{B}}$ -messbar.

Beweis: Wegen Lemma 2.9 sind

$$f^+ = \sup\{f, 0, 0, \dots\}$$

und

$$f^- = \sup\{-f, 0, 0, \dots\}$$

messbar (wobei Messbarkeit von $-f$ z.B. aus Lemma 2.10 b) folgt). Wegen obigem und Lemma 2.10 b) ist analog auch

$$|f| = f^+ + f^-$$

messbar. □

Lemma 2.12. Für $X : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}^m, \mathcal{B}_m)$ und eine stetige Abbildung $g : \mathbb{R}^m \rightarrow \mathbb{R}^n$ ist $g \circ X : \Omega \rightarrow \mathbb{R}^n$ $\mathcal{A} - \mathcal{B}_n$ -messbar.

Beweis: Folgt unmittelbar aus Lemma 2.5 unter Beachtung der Messbarkeit von stetigen Abbildungen (vgl. Beispiel oben). □

2.2 Zufallsvariablen, Bildmaße und Verteilungen

Definition 2.13. Es sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum, $(\Omega', \mathcal{A}')$ ein Messraum. Die Abbildung $X : (\Omega, \mathcal{A}) \rightarrow (\Omega', \mathcal{A}')$ heißt **Zufallsvariable (ZV)** auf $(\Omega, \mathcal{A}, P)$ [mit Zustandsraum Ω'] (ausführlich: $(\Omega', \mathcal{A}')$ -ZV auf $(\Omega, \mathcal{A}, P)$; random variable).

$X : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B})$ heißt **reelle ZV** (oft auch nur ZV);

$X : (\Omega, \mathcal{A}) \rightarrow (\overline{\mathbb{R}}, \overline{\mathcal{B}})$ heißt **erweitert-reelle ZV**;

$X : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}^n, \mathcal{B}_n)$ heißt **n -dimensionaler Zufallsvektor**.

$X(\omega)$ für ein $\omega \in \Omega$ heißt eine **Realisierung** der ZV X .

Bezeichnung $\overline{\mathcal{B}} := \{B, B \cup \{+\infty\}, B \cup \{-\infty\}, B \cup \{-\infty, +\infty\} \mid B \in \mathcal{B}\}$.

Satz 2.14. Messräume $(\Omega, \mathcal{A}), (\Omega', \mathcal{A}')$; Abbildung $X : (\Omega, \mathcal{A}) \rightarrow (\Omega', \mathcal{A}')$; Maß μ auf $\mathcal{A}$. Durch

$$\begin{aligned} \mu_X(A') &:= \mu(X^{-1}(A')) \\ &= \mu(\{\omega \in \Omega \mid X(\omega) \in A'\}) =: \mu[X \in A']; \quad A' \in \mathcal{A}', \end{aligned}$$

wird ein Maß (das sogenannte **Bildmaß**) μ_X auf $\mathcal{A}'$ definiert.

Ist μ ein W-Maß auf $\mathcal{A}$, dann ist μ_X ein W-Maß auf $\mathcal{A}'$.

Beweis: 0) μ_X ist wohldefiniert, da $X^{-1}(A') \in \mathcal{A}$ für $A' \in \mathcal{A}'$ gilt.

1. $\mu_X(\emptyset) = \mu(X^{-1}(\emptyset)) = \mu(\emptyset) = 0$
2. $\mu_X(A') = \mu(X^{-1}(A')) \geq 0$, da $\mu(A) \geq 0$ ($A \in \mathcal{A}$).
3. Sind $A'_1, A'_2, \dots \in \mathcal{A}'$ paarweise disjunkt, so sind auch $X^{-1}(A'_n), X^{-1}(A'_2), \dots \in \mathcal{A}$ paarweise disjunkt, und da μ Maß ist, gilt:

$$\begin{aligned}
 \mu_X\left(\bigcup_{n=1}^{\infty} A'_n\right) &= \mu\left(X^{-1}\left(\bigcup_{n=1}^{\infty} A'_n\right)\right) && \text{(nach Definition)} \\
 &= \mu\left(\bigcup_{n=1}^{\infty} X^{-1}(A'_n)\right) && \text{(nach Satz 2.2 a)} \\
 &= \sum_{n=1}^{\infty} \mu(X^{-1}(A'_n)) && \text{(da } \mu \text{ } \sigma\text{-additiv)} \\
 &= \sum_{n=1}^{\infty} \mu_X(A'_n) && \text{(nach Definition).}
 \end{aligned}$$

Ist μ Wahrscheinlichkeitsmaß, so gilt weiter:

$$\mu_X(\Omega') = \mu(X^{-1}(\Omega')) = \mu(\Omega) = 1$$

□

Definition 2.15. Sei X eine $(\Omega', \mathcal{A}')$ -ZV auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$. Das W-Maß P_X im Bild-Wahrscheinlichkeitsraum $(\Omega', \mathcal{A}', P_X)$ heißt **Verteilung** der ZV X .

Sprechweise: Die ZV X liegt in $A' \in \mathcal{A}'$ bzw. nimmt Werte in $A' \in \mathcal{A}'$ an mit Wahrscheinlichkeit $P_X(A') = P[X \in A']$.

$P[X \in A'] = 1 \dots P$ -fast sicher (P -f.s.) liegt X in A .

Bemerkung 2.16. Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$, Messräume $(\Omega', \mathcal{A}')$, $(\Omega'', \mathcal{A}'')$; $X : (\Omega, \mathcal{A}) \rightarrow (\Omega', \mathcal{A}')$, $Y : (\Omega', \mathcal{A}') \rightarrow (\Omega'', \mathcal{A}'')$. Dann gilt

$$P_{Y \circ X} = (P_X)_Y.$$

Begründung:

$$\begin{aligned}
 P_{Y \circ X}(A'') &= P((Y \circ X)^{-1}(A'')) && \text{(nach Definition)} \\
 &= P(X^{-1}(Y^{-1}(A''))) \\
 &= P_X(Y^{-1}(A'')) && \text{(nach Definition)} \\
 &= (P_X)_Y(A'') && \text{(nach Definition)}
 \end{aligned}$$

□

Definition 2.17. Messräume $(\Omega_i, \mathcal{A}_i)$ ($i = 1, \dots, n$).

Die **Produkt- σ -Algebra** $\bigotimes_{i=1}^n \mathcal{A}_i$ wird definiert als die von dem Mengensystem

$\left\{ \prod_{i=1}^n A_i \mid A_i \in \mathcal{A}_i \ (i = 1, \dots, n) \right\}$ in $\prod_{i=1}^n \Omega_i$ erzeugte σ -Algebra.

$\bigotimes_{i=1}^n (\Omega_i, \mathcal{A}_i) := \left(\prod_{i=1}^n \Omega_i, \bigotimes_{i=1}^n \mathcal{A}_i \right)$ heißt **Produkt-Messraum**.

Beispiel: Mit Identifizierung von $\mathbb{R}^n \times \mathbb{R}^m$ und $\mathbb{R}^{n+m}$ gilt

$$\mathcal{B}_n \otimes \mathcal{B}_m = \mathcal{B}_{n+m},$$

vgl. Übungen.

Bemerkung 2.18. Abbildung $X_i : (\Omega, \mathcal{A}) \rightarrow (\Omega_i, \mathcal{A}_i)$ ($i = 1, \dots, n$). Die Abbildung

$X : \Omega \rightarrow \prod_{i=1}^n \Omega_i$ mit

$$X(\omega) := (X_1(\omega), \dots, X_n(\omega)), \ \omega \in \Omega$$

ist $\mathcal{A}$ - $\bigotimes_{i=1}^n \mathcal{A}_i$ -messbar.

Begründung: Behauptung folgt aus Satz 2.6 unter Beachtung von

$$X^{-1} \left(\prod_{i=1}^n A_i \right) = \bigcap_{i=1}^n X_i^{-1}(A_i) \in \mathcal{A}$$

für alle $A_i \in \mathcal{A}_i$ (vgl. Lemma 2.10). □

Bemerkung 2.19. Messräume $(\Omega_i, \mathcal{A}_i)$ ($i = 1, \dots, n$). Ein W -Maß auf $\bigotimes_{i=1}^n \mathcal{A}_i$ ist eindeutig festgelegt durch seine Werte auf dem Mengensystem $\left\{ \prod_{i=1}^n A_i \mid A_i \in \mathcal{A}_i \ (i = 1, \dots, n) \right\}$.

Begründung: Folgt aus Korollar 1.42 (betreffend Fortsetzung von Maßen) unter Beachtung der Tatsache, daß das Wahrscheinlichkeitsmaß als endliches Maß auf $\left\{ \prod_{i=1}^n A_i : A_i \in \mathcal{A}_i \right\}$ immer σ -endlich ist. □

Definition 2.20. Seien X_i $(\Omega_i, \mathcal{A}_i)$ -ZVn auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ ($i = 1, \dots, n$). Die Verteilung P_X der ZV $X := (X_1, \dots, X_n)$ auf $(\Omega, \mathcal{A}, P)$ — erklärt durch

$$P_X(A) := P[(X_1, \dots, X_n) \in A]$$

für $A \in \bigotimes_{i=1}^n \mathcal{A}_i$ oder auch nur für $A = \prod_{i=1}^n A_i$ ($A_i \in \mathcal{A}_i$, $i = 1, \dots, n$) — heißt die **gemeinsame Verteilung** der ZVn X_i . Die Verteilungen P_{X_i} — erklärt durch

$$P_{X_i}(A_i) := P[X_i \in A_i] = P[X \in \Omega_1 \times \dots \times \Omega_{i-1} \times A_i \times \Omega_{i+1} \times \dots \times \Omega_n]$$

für $A_i \in \mathcal{A}_i$ — heißen **Randverteilungen** von P_X ($i = 1, \dots, n$).

Wichtiger Spezialfall in Definition 2.20: $(\Omega_i, \mathcal{A}_i) = (\mathbb{R}, \mathcal{B})$ ($i = 1, \dots, n$).

Bemerkung 2.21. Bezeichnungen wie in Definition 2.20.

a) Die Verteilung P_X ist ohne Zusatzvoraussetzung durch ihre Randverteilungen **nicht** eindeutig festgelegt.

b) Ist $\pi_i : \prod_{k=1}^n \Omega_k \rightarrow \Omega_i$ die (messbare!) Projektionsabbildung

$$(\omega_1, \dots, \omega_n) \rightarrow \omega_i,$$

so gilt $P_{X_i} = (P_X)_{\pi_i}$ ($i = 1, \dots, n$).

Begründung:

a) Siehe Übungen

b) $(P_X)_{\pi_i} = P_{\pi_i \circ X} = P_{X_i}$ nach Bemerkung 2.16. □

Bemerkung 2.22.

a) Sei Q ein W-Maß auf $(\mathbb{R}^n, \mathcal{B}_n)$. Dann existieren reelle ZVn $X_1, \dots, X_n$ auf einem geeigneten Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ so, dass ihre gemeinsame Verteilung mit Q übereinstimmt:

$(\Omega, \mathcal{A}, P) := (\mathbb{R}^n, \mathcal{B}_n, Q)$, $X_i : \Omega \rightarrow \mathbb{R}$ wird definiert als die Projektionsabbildung

$$(x_1, \dots, x_n) \rightarrow x_i \quad (i = 1, \dots, n);$$

hierbei ist also $X := (X_1, \dots, X_n)$ die auf $\mathbb{R}^n$ definierte identische Abbildung.

b) Möglichkeit der unmittelbaren Verallgemeinerung von a) auf einen Produkt-Meßraum

$$\left(\prod_{i=1}^n \Omega_i, \bigotimes_{i=1}^n \mathcal{A}_i \right) \text{ statt } (\mathbb{R}^n, \mathcal{B}_n).$$

c) Sonderfall zu b): Ist $(\Omega, \mathcal{A}, Q)$ ein Wahrscheinlichkeitsraum, $X : \Omega \rightarrow \Omega$ die identische Abbildung, so gilt $P_X = Q$. **Jedes W-Maß lässt sich somit als eine Verteilung auffassen (und — definitionsgemäß — umgekehrt).**

Begründung von a):

Sei $P = Q$ und $X_i = \pi_i$ mit π_i wie in Bemerkung 2.22 b). Dann gilt

$$\begin{aligned} X : \Omega &\rightarrow \Omega \\ X((\omega_1, \dots, \omega_n)) &= (\pi_1(\omega_1, \dots, \omega_n), \dots, \pi_n(\omega_1, \dots, \omega_n)) \\ &= (\omega_1, \dots, \omega_n) \end{aligned}$$

und damit

$$P_X(A) = Q(X^{-1}(A)) = Q(A).$$

□

3 Integrale, Erwartungswerte und Dichten

3.1 Das Maßintegral

Definition 3.1 $(\Omega, \mathcal{A})$ sei Messraum.

Eine Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$ heißt **einfach**, wenn

$$f = \sum_{i=1}^N \alpha_i \cdot \chi_{A_i}$$

für ein $N \in \mathbb{N}$, $A_1, \dots, A_N \in \mathcal{A}$ paarweise disjunkt und $\alpha_1, \dots, \alpha_N \in \overline{\mathbb{R}}$ gilt, d. h. wenn f nur endlich viele Werte annimmt und diese jeweils auf einer Menge aus $\mathcal{A}$ angenommen werden.

Satz 3.2: $(\Omega, \mathcal{A})$ sei Messraum.

- a) $f : \Omega \rightarrow \overline{\mathbb{R}}$ ist genau dann $\mathcal{A} - \overline{\mathcal{B}}$ -messbar, wenn f^+ und f^- $\mathcal{A} - \overline{\mathcal{B}}$ -messbar sind.
- b) Eine Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}_+$ ist genau dann messbar, wenn sie Limes einer monoton wachsenden Folge nichtnegativer (beschränkter) einfacher Funktionen ist.

Beweis.

- a) “ $\Rightarrow$ ” Folgt aus Lemma 2.11
“ $\Leftarrow$ ” Mit f^+, f^- ist nach Lemma 2.10 b) auch $f = f^+ - f^-$ messbar.
- b) “ $\Leftarrow$ ” Für $A \in \mathcal{A}$ ist χ_A messbar, und mit Lemma 2.10 folgt, dass auch jede einfache Funktion und damit nach Lemma 2.9 auch der Limes einfacher Funktionen messbar ist.

“ $\Rightarrow$ ” Sei $f : \Omega \rightarrow \overline{\mathbb{R}}_+$ $\mathcal{A} - \overline{\mathcal{B}}$ -messbar

Dann sind

$$f_n = n \cdot \chi_{[f \geq n]} + \sum_{k=0}^{n \cdot 2^n - 1} \frac{k}{2^n} \cdot \chi_{[\frac{k}{2^n} \leq f < \frac{k+1}{2^n}]}$$

einfach. Man sieht leicht, dass $(f_n)_n$ monoton ist. Weiter gilt

$$\lim_{n \rightarrow \infty} f_n = f,$$

denn für $f(\omega) < \infty$ ist für $n > f(\omega)$

$$|f(\omega) - f_n(\omega)| < \frac{1}{2^n},$$

und für $f(\omega) = \infty$ gilt

$$f_n(\omega) = n \rightarrow \infty \quad (n \rightarrow \infty).$$

□

$(\Omega, \mathcal{A}, \mu)$ Maßraum, $f : (\Omega, \mathcal{A}) \rightarrow (\overline{\mathbb{R}}, \overline{\mathcal{B}})$.

Im Folgenden: Definition des Integrals

$$\int_{\Omega} f \, d\mu$$

in mehreren Schritten:

1. Schritt: f nichtnegativ, einfach
2. Schritt: f nichtnegativ, messbar
3. Schritt: f messbar.

1. Schritt: Sei f nichtnegativ und einfach

Dann existiert Darstellung

$$f = \sum_{i=1}^N \alpha_i \cdot \chi_{A_i}$$

mit $N \in \mathbb{N}$, $A_1, \dots, A_N \in \mathcal{A}$ bilden Partition von Ω (d. h. $A_i \cap A_j = \emptyset \, \forall i \neq j$ und $\bigcup_{j=1}^N A_j = \Omega$), $\alpha_1, \dots, \alpha_N \in \overline{\mathbb{R}}_+$.

Wir definieren:

$$\int f \, d\mu = \sum_{i=1}^N \alpha_i \cdot \mu(A_i)$$

(mit $\infty \cdot 0 = 0$).

Lemma 3.3: Obiges Integral ist wohldefiniert, d. h. unabhängig von der Wahl der Darstellung $f = \sum_{i=1}^N \alpha_i \cdot \chi_{A_i}$.

Beweis Sei

$$f = \sum_{i=1}^N \alpha_i \cdot \chi_{A_i} = \sum_{j=1}^M \beta_j \cdot \chi_{B_j}$$

mit A_i, B_j wie oben.

zu zeigen: $\sum_{i=1}^N \alpha_i \cdot \mu(A_i) = \sum_{j=1}^M \beta_j \cdot \mu(B_j)$.

Aus $\Omega = \sum_{i=1}^N A_i = \sum_{j=1}^M B_j$ folgt

$$\Omega = \sum_{\substack{i=1, \dots, N, \\ j=1, \dots, M}} A_i \cap B_j$$

mit $A_i \cap B_j \in \mathcal{A}$ (da $A_i, B_j \in \mathcal{A}$). Daher existiert Darstellung

$$f = \sum_{i,j} \gamma_{i,j} \cdot \chi_{A_i \cap B_j}$$

für $\gamma_{i,j} \in \overline{\mathbb{R}}_+$, und für $A_i \cap B_j \neq \emptyset$ ist

$$\gamma_{i,j} = \alpha_i = \beta_j$$

(denn aus $\omega \in A_i \cap B_j$ folgt:

$$\begin{aligned}\alpha_i &= \sum_{k=1}^N \alpha_k \cdot \chi_{A_k}(\omega) = f(\omega), \\ \beta_j &= \sum_{k=1}^M \beta_k \cdot \chi_{B_k}(\omega) = f(\omega) \\ \text{und } \gamma_{i,j} &= f(\omega),\end{aligned}$$

also ist $\alpha_i = f(\omega) = \beta_j = \gamma_{i,j}$).

Mit

$$A_i = \sum_{j=1}^M A_i \cap B_j \text{ und } B_j = \sum_{i=1}^N A_i \cap B_j$$

folgt:

$$\begin{aligned}\sum_{i=1}^N \alpha_i \cdot \mu(A_i) &= \sum_{i=1}^N \alpha_i \cdot \mu\left(\sum_{j=1}^M A_i \cap B_j\right) = \sum_{i=1}^N \sum_{j=1}^M \alpha_i \cdot \mu(A_i \cap B_j) \\ &= \sum_{i=1}^N \sum_{j=1}^M \gamma_{i,j} \cdot \mu(A_i \cap B_j)\end{aligned}$$

und analog:

$$\sum_{j=1}^M \beta_j \cdot \mu(B_j) = \sum_{j=1}^M \sum_{i=1}^N \gamma_{i,j} \cdot \mu(A_i \cap B_j) \rightsquigarrow \text{Beh. } \square$$

Lemma 3.4 (Eigenschaften des Integrals für nichtnegative einfache Funktionen)

Seien f, f_1 und f_2 nichtnegativ einfach. Dann gilt:

- a) $\int f \, d\mu \geq 0$
- b) $\forall c \in \mathbb{R}_+ : \int (c \cdot f) \, d\mu = c \cdot \int f \, d\mu$
- c) $\int (f_1 + f_2) \, d\mu = \int f_1 \, d\mu + \int f_2 \, d\mu$
- d) $f_1 \leq f_2 \Rightarrow \int f_1 \, d\mu \leq \int f_2 \, d\mu$.

Beweis:

a) Klar nach Definition.

b) $f = \sum_{i=1}^N \alpha_i \cdot \chi_{A_i} \Rightarrow c \cdot f = \sum_{i=1}^N (c \cdot \alpha_i) \cdot \chi_{A_i}$ mit $c \cdot \alpha_i \geq 0$ wegen $c \geq 0, \alpha_i \geq 0$.

Somit:

$$\begin{aligned}\int (c \cdot f) \, d\mu &= \sum_{i=1}^N (c \cdot \alpha_i) \cdot \mu(A_i) = c \cdot \sum_{i=1}^N \alpha_i \cdot \mu(A_i) \\ &= c \cdot \int f \, d\mu.\end{aligned}$$

$$c) f_1 = \sum_{i=1}^N \alpha_i \cdot \chi_{A_i}, \quad f_2 = \sum_{j=1}^M \beta_j \cdot \chi_{B_j}$$

Wie im Beweis von Lemma 3.3 sieht man:

$$f_1 = \sum_{\substack{i=1, \dots, N \\ j=1, \dots, M}} \alpha_i \cdot \chi_{A_i \cap B_j}, \quad f_2 = \sum_{\substack{i=1, \dots, N \\ j=1, \dots, M}} \beta_j \cdot \chi_{A_i \cap B_j}$$

woraus folgt

$$f_1 + f_2 = \sum_{\substack{i=1, \dots, N \\ j=1, \dots, M}} (\alpha_i + \beta_j) \cdot \chi_{A_i \cap B_j}.$$

Damit:

$$\begin{aligned} \int (f_1 + f_2) d\mu &\stackrel{Def.}{=} \sum_{\substack{i=1, \dots, N \\ j=1, \dots, M}} (\alpha_i + \beta_j) \cdot \mu(A_i \cap B_j) \\ &= \sum_{\substack{i=1, \dots, N \\ j=1, \dots, M}} \alpha_i \cdot \mu(A_i \cap B_j) + \sum_{\substack{i=1, \dots, N \\ j=1, \dots, M}} \beta_j \cdot \mu(A_i \cap B_j) \\ &\stackrel{Def.}{=} \int f_1 d\mu + \int f_2 d\mu \end{aligned}$$

d) $f_1 \leq f_2$ mit f_1, f_2 nichtnegativ einfach

$\Rightarrow$ Es existiert nichtnegative einfache Funktion g mit

$$f_2 = f_1 + g.$$

Damit:

$$\begin{aligned} \int f_2 d\mu &= \int (f_1 + g) d\mu \stackrel{c)}{=} \int f_1 d\mu + \int g d\mu \\ &\stackrel{a)}{\geq} \int f_1 d\mu + 0 = \int f_1 d\mu. \end{aligned}$$

□

2. Schritt Sei f nichtnegativ und messbar.

Nach Satz 3.2 existieren nichtnegative einfache Funktionen f_n mit $f_n \uparrow f$. Wir definieren

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

Lemma 3.5 Obiges Integral ist wohldefiniert, d. h. der obige Grenzwert existiert und ist unabhängig von der Wahl der Folge $(f_n)_n$ mit f_n nichtnegativ einfach und $f_n \uparrow f$.

Beweis

a) Wegen $f_n \leq f_{n+1}$ gilt nach Lemma 3.4 d):

$$\int f_n d\mu \leq \int f_{n+1} d\mu,$$

also existiert der Grenzwert der monoton wachsenden Folge $(\int f_n d\mu)_{n \in \mathbb{N}}$.

b) Sei f nichtnegativ messbar, seien f_n, g_n nichtnegativ einfach mit $f_n \uparrow f$
 $g_n \uparrow f$.

Zu zeigen:

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

Aus Symmetriegründen genügt es zu zeigen:

$$\lim_{n \rightarrow \infty} \int f_n d\mu \geq \lim_{n \rightarrow \infty} \int g_n d\mu.$$

Da die Folgen auf der rechten und der linken Seite monoton wachsend sind, ist dies wiederum äquivalent zu

$$(*) \quad \sup_n \int f_n d\mu \geq \int g_k d\mu$$

für alle $k \in \mathbb{N}$.

Wegen $f_n \uparrow f, g_n \uparrow f$ gilt

$$g_k \leq f = \sup_n f_n.$$

Also folgt (*) aus

$$(**) \quad \left\{ \begin{array}{l} \text{Sind } f_n, h \text{ nichtnegative einfache Funktionen} \\ \text{mit } h \leq \sup_n f_n \text{ und } f_n \uparrow, \text{ so gilt} \\ \int h d\mu \leq \sup_n \int f_n d\mu \end{array} \right.$$

Im Folgenden zeigen wir (**).

Fall 1: $0 < m := \min_{\omega \in \Omega} h(\omega) \leq \max_{\omega \in \Omega} h(\omega) =: M < \infty$

Sei $\varepsilon > 0$ beliebig mit $0 < \varepsilon < m$.

Setze

$$\Omega_n := \{\omega \in \Omega \mid f_n(\omega) \geq h(\omega) - \varepsilon\} \in \mathcal{A}$$

Wegen $f_n \uparrow, \sup_n f_n \geq h$ und $\max_{\omega \in \Omega} h(\omega) < \infty$ gilt dann

$$\Omega_n \uparrow \Omega.$$

Nach Definition von Ω_n ist

$$f_n(\omega) \geq (h(\omega) - \varepsilon) \cdot \chi_{\Omega_n}(\omega) \quad (\omega \in \Omega),$$

wobei die Funktionen auf beiden Seiten nichtnegativ und einfach sind.

Mit Lemma 3.4 d) folgt

$$\int f_n d\mu \geq \int (h - \varepsilon) \cdot \chi_{\Omega_n} d\mu.$$

Im Falle $\mu(\Omega) = +\infty$ erhält man daraus

$$\begin{aligned} \int f_n d\mu &\geq \int (h - \varepsilon) \cdot \chi_{\Omega_n} d\mu \\ &\geq \int (m - \varepsilon) \cdot \chi_{\Omega_n} d\mu \text{ (nach Lemma 3.4 d)} \\ &= (m - \varepsilon) \cdot \mu(\Omega_n) \xrightarrow{(n \rightarrow \infty)} (m - \varepsilon) \cdot \mu(\Omega) = \infty \end{aligned}$$

da $m - \varepsilon > 0, \mu(\Omega) = \infty$, und damit hat man gezeigt:

$$\sup_n \int f_n d\mu = \infty \geq \int h d\mu.$$

Im Falle $\mu(\Omega) < \infty$ folgt daraus unter Beachtung von

$$\begin{aligned} \int h d\mu &= \int ((h - \varepsilon) \cdot \chi_{\Omega_n} + h \cdot \chi_{\Omega_n^c} + \varepsilon \cdot \chi_{\Omega_n}) d\mu \\ &\stackrel{\text{Lemma 3.4 c)}}{=} \int (h - \varepsilon) \cdot \chi_{\Omega_n} d\mu + \int h \cdot \chi_{\Omega_n^c} d\mu + \varepsilon \cdot \mu(\Omega_n), \end{aligned}$$

dass gilt:

$$\begin{aligned} \int f_n d\mu &\geq \int (h - \varepsilon) \cdot \chi_{\Omega_n} d\mu \\ &= \int h d\mu - \int h \cdot \chi_{\Omega_n^c} d\mu - \varepsilon \cdot \mu(\Omega_n) \\ &\geq \int h d\mu - M \cdot \mu(\Omega_n^c) - \varepsilon \cdot \mu(\Omega_n) \\ &= \int h d\mu - M \cdot (\mu(\Omega) - \mu(\Omega_n)) - \varepsilon \cdot \mu(\Omega_n) \\ &\longrightarrow \int h d\mu - M \cdot (\mu(\Omega) - \mu(\Omega)) - \varepsilon \cdot \mu(\Omega) \\ &= \int h d\mu - 0 - \varepsilon \cdot \mu(\Omega) \end{aligned}$$

Mit

$$\sup_n \int f_n d\mu = \lim_{k \rightarrow \infty} \sup_{n \geq k} \int f_n d\mu$$

(wegen Monotonie von $(\int f_n d\mu)_{n \in \mathbb{N}}$) folgt daraus

$$\sup_n \int f_n d\mu \geq \int h d\mu - \varepsilon \cdot \mu(\Omega).$$

Mit $\varepsilon \downarrow 0$ folgt die Behauptung.

Fall 2: $0 < m \leq M = \infty$

h lässt sich dann darstellen als

$$h = \sum_{i=1}^N \alpha_i \cdot \chi_{A_i} + \infty \cdot \chi_{[h=\infty]} \text{ mit } \alpha_1, \dots, \alpha_N \in \mathbb{R}, A_1, \dots, A_n \in \mathcal{A}.$$

Setze

$$h^{(c)} = \sum_{i=1}^N \alpha_i \cdot \chi_{A_i} + c \cdot \chi_{[h=\infty]}$$

für $c > 0$.

Dann gilt $h^{[c]} \uparrow h, \int h^{(c)} d\mu \xrightarrow{(c \rightarrow \infty)} \int h d\mu$ und aus

$$h^{(c)} \leq h \leq \sup_n f_n$$

folgt nach Fall 1:

$$\sup_n \int f_n d\mu \geq \int h^{(c)} d\mu \rightarrow \int h d\mu \quad (c \rightarrow \infty).$$

Daraus folgt die Behauptung.

Fall 3: $0 \leq m \leq M \leq \infty$.

oBdA sei $h \not\equiv 0$.

Setze

$$\Omega' = \{\omega \in \Omega \mid h(\omega) > 0\}$$

Da h einfach, gilt

$$\min_{\omega \in \Omega'} h(\omega) > 0.$$

Betrachte Maßraum

$$(\Omega', \Omega' \cap \mathcal{A}, \mu|_{\Omega' \cap \mathcal{A}}).$$

Dann gilt:

$$\begin{aligned} \sup_n \int f_n d\mu &\geq \sup_n \int f_n \cdot \chi_{\Omega'} d\mu \\ &= \sup_n \int f_n d\mu|_{\Omega' \cap \mathcal{A}} \\ &\stackrel{\text{Fälle 1 und 2}}{\geq} \int h d\mu|_{\Omega' \cap \mathcal{A}} \\ &= \int h d\mu, \end{aligned}$$

da $h(\omega) = 0$ für $\omega \notin \Omega'$.

→ Beh. $\square$

Lemma 3.6 (Eigenschaften des Integrals für nichtnegative messbare Funktionen)

Seien f, g nichtnegativ messbar. Dann gilt:

- a) $\int f d\mu \geq 0$
- b) $\forall c \in \mathbb{R}_+ : \int (c \cdot f) d\mu = c \cdot \int f d\mu$
- c) $\int (f + g) d\mu = \int f d\mu + \int g d\mu$
- d) $f \leq g \Rightarrow \int f d\mu \leq \int g d\mu$.

Beweis:

- a) klar.

b) Seien f_n nichtnegativ einfach mit $f_n \uparrow f$.

Wegen $c \geq 0$ sind dann $c \cdot f_n$ nichtnegativ einfach mit $c \cdot f_n \uparrow c \cdot f$.

$$\begin{aligned} \Rightarrow \int c \cdot f d\mu &= \lim_{n \rightarrow \infty} \int c \cdot f_n d\mu \quad (\text{Definition}) \\ &= \lim_{n \rightarrow \infty} c \cdot \int f_n d\mu \quad (\text{nach Lemma 3.4 b))} \\ &= c \cdot \int f d\mu \quad (\text{Definition}) \end{aligned}$$

c) Folgt analog zu b) mit Lemma 3.4 c).

d) $f \leq g \Rightarrow g = f + (g - f)$ mit $g - f \geq 0$

$$\stackrel{c)}{\Rightarrow} \int g d\mu = \int f d\mu + \int (g - f) d\mu \stackrel{a)}{\geq} \int f d\mu. \quad \square$$

Der folgende Satz zeigt, dass der bisher eingeführte Integralbegriff "abgeschlossen" ist bzgl. der Grenzwertbildung im Schritt 2.

Satz 3.7 (Satz von der monotonen Konvergenz bzw. Satz von B. Levi (1875 - 1961))

Seien f_n nichtnegative messbare Funktionen ($n \in \mathbb{N}$) mit $f_n \uparrow f$. Dann ist auch f nichtnegativ messbar, und es gilt

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

Beweis: f ist nichtnegativ und messbar, da f Grenzwert nichtnegativer und messbarer Funktionen ist (vgl. Lemma 2.9).

Setze

$$g_j = f_j - f_{j-1} \quad (j \in \mathbb{N}), \text{ wobei } f_0 := 0.$$

Dann ist g_j nichtnegativ messbar, und es gilt

$$f_n = f_n - f_0 = \sum_{j=1}^n g_j.$$

Nach Satz 3.2 ist g_j Limes einer monoton wachsenden Folge einfacher Funktionen. Die dabei auftretenden Funktionen lassen sich darstellen als Partialsummen von nichtnegativ einfachen Funktionen $h_{j,k}$, d. h. es gilt

$$g_j = \sum_{k=1}^{\infty} h_{j,k} \text{ mit } h_{j,k} \text{ nichtnegativ einfach}$$

Damit

$$f_n = \sum_{j=1}^n \sum_{k=1}^{\infty} h_{j,k}$$

und

$$f = \lim_{n \rightarrow \infty} f_n = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} h_{j,k} = \sum_{j,k} h_{j,k},$$

wobei im letzten Schritt benutzt wurde, dass Summationsreihenfolge bei nichtnegativen reellen Zahlen keine Rolle spielt.

Für endliche Mengen $\mathcal{J}, K \subseteq \mathbb{N}$ gilt nach Lemma 3.4 c)

$$\int \sum_{j \in \mathcal{J}, k \in K} h_{j,k} d\mu = \sum_{j \in \mathcal{J}, k \in K} \int h_{j,k} d\mu,$$

also folgt mit der Definition des Integrals:

$$\int f d\mu = \sum_{j,k} \int h_{j,k} d\mu.$$

Mit

$$\begin{aligned} \int f_n d\mu &\stackrel{\text{Lemma 3.6 c)}}{=} \sum_{j=1}^n \int \sum_{k=1}^{\infty} h_{j,k} d\mu \\ &\stackrel{\text{Def.}}{=} \sum_{j=1}^n \lim_{m \rightarrow \infty} \int \sum_{k=1}^m h_{j,k} d\mu \\ &\stackrel{\text{Lemma 3.4 c)}}{=} \sum_{j=1}^n \sum_{k=1}^{\infty} \int h_{j,k} d\mu \\ &\rightarrow \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \int h_{j,k} d\mu \quad (n \rightarrow \infty) \end{aligned}$$

und

$$\sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \int h_{j,k} d\mu = \sum_{j,k} \int h_{j,k} d\mu$$

(da Integrale nichtnegativ und Summationsreihenfolge deshalb keine Rolle spielt) folgt die Behauptung. $\square$

3. Schritt: Sei f messbar. Nach Satz 3.2 sind dann auch $f^+ = \max\{f, 0\}$ und $f^- = \max\{-f, 0\}$ nichtnegativ und messbar und damit existieren.

$$\int f^+ d\mu \text{ und } \int f^- d\mu \in \overline{\mathbb{R}}_+.$$

Falls nicht beide Integrale unendlich sind, definieren wir:

$$\int f d\mu := \int f^+ d\mu - \int f^- d\mu.$$

Bem.: Ist f nichtnegativ messbar, so ist

$$\int f^- d\mu = \int 0 d\mu = 0 \text{ und } f = f^+.$$

Damit stimmt die obige Definition mit der aus Schritt 2 überein, sofern f nichtnegativ ist.

Lemma 3.8: $(\Omega, \mathcal{A}, \mu)$ sei Maßraum, $u, v : \Omega \rightarrow \overline{\mathbb{R}}_+$ seien messbar. Sind

$$u - v \text{ und } \int u d\mu - \int v d\mu$$

definiert, dann existiert $\int (u - v) d\mu$ und es gilt

$$\int (u - v) d\mu = \int u d\mu - \int v d\mu.$$

Beweis:

$$\begin{aligned} u - v &= (u - v)^+ - (u - v)^- \\ \Rightarrow u + (u - v)^- &= v + (u - v)^+ \\ \stackrel{\text{Lemma 3.6 c)}}{\Rightarrow} \int u d\mu + \int (u - v)^- d\mu &= \int v d\mu + \int (u - v)^+ d\mu \\ \Rightarrow \int u d\mu - \int v d\mu &= \int (u - v)^+ d\mu - \int (u - v)^- d\mu \quad (*) \end{aligned}$$

denn:

Die letzte Umformung ist klar für $\int v d\mu < \infty$ und $\int (u - v)^- d\mu < \infty$.

Im Falle $\int v d\mu = +\infty$ ist aber $\int u d\mu < \infty$ und damit (wegen der Nichtnegativität von u und v) auch

$$\int (u - v)^+ d\mu \leq \int u^+ d\mu = \int u d\mu < \infty$$

Wegen

$$\int u d\mu + \int (u - v)^- d\mu = \int v d\mu + \int (u - v)^+ d\mu$$

folgt daraus

$$\int (u - v)^- d\mu = \infty,$$

womit (*) wieder gilt.

Im Falle $\int (u - v)^- d\mu = \infty$ gilt (wegen der Nichtnegativität von u und v)

$$\int v d\mu \geq \int (u - v)^- d\mu = \infty$$

woraus $\int u d\mu < \infty$ und $\int (u - v)^+ d\mu \leq \int u d\mu < \infty$ folgt, also gilt wieder (*).

Weiter folgt wegen

$$\begin{aligned} \int (u - v)^+ d\mu &\leq \int u d\mu \\ \int (u - v)^- d\mu &\leq \int v d\mu \end{aligned}$$

aus der Existenz von $\int u d\mu - \int v d\mu$ auch die Existenz von

$$\int (u - v)^+ d\mu - \int (u - v)^- d\mu = \int (u - v) d\mu.$$

□

Lemma 3.9

Seien $a, b, c, d \in \overline{\mathbb{R}}_+$. Sind $a - b, c - d$ und $(a - b) + (c - d)$ definiert, so ist auch

$$(a + b) - (b + d)$$

definiert, und es gilt

$$(a + c) - (b + d) = (a - b) + (c - d)$$

Beweis: Folgt mit Fallunterscheidung $< \infty, = \infty$.

Fall 1. $a, b, c, d \in \mathbb{R}_+ \Rightarrow$ Aussage ist trivial.

Fall 2. $a = \infty \Rightarrow b < \infty, d < \infty$

Fall 1.1: $c = \infty$

$$\text{li. S.} = \infty, \text{ re. S.} = \infty + \infty = \infty$$

Fall 1.2: $c < \infty$

$$\text{li. S.} = \infty, \text{ re. S.} = \infty$$

Fall 3. $a < \infty, b = \infty \Rightarrow c < \infty$

Fall 2.1: $d = \infty$

$$\text{li. S.} = -\infty, \text{ re. S.} = -\infty - \infty = -\infty$$

Fall 2.2: $d < \infty$

$$\text{li. S.} = -\infty, \text{ re. S.} = -\infty$$

Rest durch Vertauschen von a, b mit c, d . □

Def. 3.10 Maßraum $(\Omega, \mathcal{A}, \mu)$.

a) Jedes $N \in \mathcal{A}$ mit $\mu(N) = 0$ heißt μ -**Nullmenge**.

b) Eine Eigenschaft gilt μ -**fast überall** (kurz: μ -**f.ü.**), wenn die Menge der ω , für die sie nicht gilt, *Teilmenge* einer μ -Nullmenge ist.

Satz 3.11 (Eigenschaften des Integrals)

Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und seien $f, g : \Omega \rightarrow \overline{\mathbb{R}}$ messbare Funktionen, für die $\int f d\mu$ und $\int g d\mu$ existieren. Dann gilt

a) $f \geq 0 \Rightarrow \int f d\mu \geq 0$

b) Für $\alpha \in \mathbb{R}$ existiert $\int (\alpha \cdot f) d\mu$ und es gilt

$$\int (\alpha \cdot f) d\mu = \alpha \cdot \int f d\mu.$$

c) Sind $f + g$ und $\int f d\mu + \int g d\mu$ definiert, dann existiert $\int (f + g) d\mu$ und es gilt

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

- d) $f \geq g \Rightarrow \int f d\mu \geq \int g d\mu$.
- e) $f = g \mu$ -f.ü. $\Rightarrow \int f d\mu = \int g d\mu$
- f) $|\int f d\mu| \leq \int |f| d\mu$.

Beweis:

a) Folgt aus Lemma 3.6 a).

b) Trivial für $\alpha = 0$.

Fall $\alpha > 0$: $(\alpha \cdot f)^+ = \alpha \cdot f^+, (\alpha \cdot f)^- = \alpha \cdot f^-$

$$\begin{aligned} \Rightarrow \int (\alpha \cdot f) d\mu &\stackrel{Def.}{=} \int (\alpha \cdot f)^+ d\mu - \int (\alpha \cdot f)^- d\mu \\ &\stackrel{s.o.}{=} \int \alpha \cdot f^+ d\mu - \int \alpha \cdot f^- d\mu \\ &\stackrel{Lemma 3.6 b)}{=} \alpha \cdot \int f^+ d\mu - \alpha \cdot \int f^- d\mu \\ &= \alpha \cdot (\int f^+ d\mu - \int f^- d\mu) \\ &\stackrel{Def.}{=} \alpha \cdot \int f d\mu. \end{aligned}$$

Fall $\alpha < 0$: $(\alpha \cdot f)^+ = (-\alpha) \cdot f^-, (\alpha \cdot f)^- = (-\alpha) \cdot f^+$

Wegen $(-\alpha) > 0$ folgt mit der Definition des Integrals und Lemma 3.6 b):

$$\begin{aligned} \int (\alpha \cdot f) d\mu &= \int (\alpha \cdot f)^+ d\mu - \int (\alpha \cdot f)^- d\mu \\ &\stackrel{s.o.}{=} \int (-\alpha) \cdot f^- d\mu - \int (-\alpha) \cdot f^+ d\mu \\ &= (-\alpha) \cdot (\int f^- d\mu - \int f^+ d\mu) \\ &= (-\alpha) \cdot (-1) \cdot \int f d\mu \\ &= \alpha \cdot \int f d\mu. \end{aligned}$$

c) Nach Voraussetzung existieren

$$f^+ - f^- = f$$

$$g^+ - g^- = g$$

$$(f^+ - f^-) + (g^+ - g^-) = f + g.$$

Mit Lemma 3.9 folgt, dass auch

$$(*) \quad (f^+ + g^+) - (f^- + g^-) = f + g$$

existiert.

Analog sieht man, dass auch

$$(**) \left\{ \begin{array}{l} \int (f^+ + g^+) d\mu - \int (f^- + g^-) d\mu \\ = (\int f^+ d\mu - \int f^- d\mu) + (\int g^+ d\mu - \int g^- d\mu) \\ = \int f d\mu + \int g d\mu \end{array} \right.$$

existiert.

Mit Lemma 3.8 (setze $u = f^+ + g^+, v = f^- + g^-$) folgt die Existenz von

$$\int (u - v) d\mu \stackrel{(*)}{=} \int (f + g) d\mu$$

sowie

$$\begin{aligned} \int (f + g) d\mu &= \int (u - v) d\mu = \int u d\mu - \int v d\mu \\ &\stackrel{(**)}{=} \int f d\mu + \int g d\mu. \end{aligned}$$

d) $f \geq g \Rightarrow f^+ \geq g^+$ und $f^- \leq g^-$

$$\begin{aligned} \Rightarrow \int f d\mu &= \int f^+ d\mu - \int f^- d\mu \\ &\stackrel{\text{Lemma 3.6 d)}}{\geq} \int g^+ d\mu - \int g^- d\mu \\ &= \int g d\mu. \end{aligned}$$

e) Sei $f = g$ f.ü.

Da f und g messbar sind, gilt $A = [f = g] \in \mathcal{A}$.

Nach Voraussetzung ist $\mu(A^c) = 0$.

Mit

$$\begin{aligned} \int f^+ d\mu &= \int f^+ \chi_A d\mu + \int f^+ \chi_{A^c} d\mu \\ &= \int f^+ \chi_A d\mu \end{aligned}$$

(wegen $0 \leq \int f^+ \chi_{A^c} d\mu \leq \infty \cdot \mu(A^c) = 0$) und den analogen Aussagen für f^-, g^+ und g^- folgt:

$$\begin{aligned} \int f d\mu &= \int f^+ d\mu - \int f^- d\mu \\ &= \int f^+ \chi_A d\mu - \int f^- \chi_A d\mu \\ &= \int g^+ \chi_A d\mu - \int g^- \chi_A d\mu \text{ (nach Definition von } A) \\ &= \int g^+ d\mu - \int g^- d\mu \\ &= \int g d\mu. \end{aligned}$$

f) $\int f d\mu = \int f^+ d\mu - \int f^- d\mu$

$$\begin{aligned} \Rightarrow |\int f d\mu| &\leq \int f^+ d\mu + \int f^- d\mu \\ &\stackrel{c)}{=} \int (f^+ + f^-) d\mu \\ &= \int |f| d\mu \end{aligned}$$

wegen $|f| = f^+ + f^-$. □

Def. 3.12 Maßraum $(\Omega, \mathcal{A}, \mu)$, $f : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \overline{\mathcal{B}})$.

f heißt μ -**integrierbar** (kurz: integrierbar), falls $\int f d\mu$ existiert und endlich ist.

Satz 3.13 Maßraum $(\Omega, \mathcal{A}, \mu)$, $f, g : \Omega \rightarrow \overline{\mathbb{R}}$.

a) Sei f messbar. Dann gilt

$$\begin{aligned} f \text{ integrierbar} &\Leftrightarrow f^+, f^- \text{ integrierbar} \Leftrightarrow |f| \text{ integrierbar} \\ &\Rightarrow f \text{ f.ü. endlich.} \end{aligned}$$

b) f messbar, g integrierbar mit $|f| \leq g$ f.ü.

$\Rightarrow f$ integrierbar

Beweis:

a) f integrierbar

$$\Leftrightarrow \int f^+ d\mu < \infty \text{ und } \int f^- d\mu < \infty$$

$$\Leftrightarrow f^+, f^- \text{ integrierbar}$$

$$\stackrel{(*)}{\Leftrightarrow} |f| \text{ integrierbar}$$

Begr. von (*): " $\Rightarrow$ ":

$$\begin{aligned} \int |f| d\mu &= \int (f^+ + f^-) d\mu \\ &\stackrel{\text{Satz 3.11 c)}}{=} \int f^+ d\mu + \int f^- d\mu \\ &< \infty \text{ nach Vor.} \end{aligned}$$

$$\text{"} \Leftarrow \text{"} \quad f^+ \leq |f|, f^- \leq |f|$$

$$\rightsquigarrow \int f^+ d\mu \leq \int |f| d\mu < \infty \text{ und } \int f^- d\mu \leq \int |f| d\mu < \infty$$

nach Satz 3.11 d)

$$\rightsquigarrow f^+, f^- \text{ integrierbar.}$$

Ist nun $|f|$ integrierbar, so gilt

$$\infty > \int |f| d\mu \geq \int \infty \cdot 1_{\{|f|=\infty\}} d\mu$$

(nach Satz 3.11 d)

$$= \infty \cdot \mu(\{|f| = \infty\})$$

$$\Rightarrow \mu(\{|f| = \infty\}) = 0$$

$$\Rightarrow f \text{ f.ü. endlich}$$

b) Sei $|f| \leq g$ f.ü. mit g integrierbar.

$\Rightarrow f^+ \leq g$ f.ü. und $f^- \leq g$ f.ü.

Wähle Nullmenge N mit $f^+ \cdot 1_{N^c} \leq g \cdot 1_{N^c}$ und $f^- \cdot 1_{N^c} \leq g \cdot 1_{N^c}$

$\Rightarrow \int f^+ d\mu \stackrel{\text{Satz 3.11 e)}}{=} \int f^+ \cdot 1_{N^c} d\mu \leq \int g \cdot 1_{N^c} d\mu = \int g d\mu < \infty$

und analog

$$\int f^- d\mu \leq \int g d\mu < \infty$$

$\stackrel{a)}{\Rightarrow} f$ integrierbar. □

Def. 3.14 $(\Omega, \mathcal{A}, \mu)$ Maßraum, $f : (\Omega, \mathcal{A}) \rightarrow (\overline{\mathbb{R}}, \overline{\mathcal{B}})$, $A \in \mathcal{A}$. Wir setzen

$$\int_A f d\mu := \int f \cdot \chi_A d\mu$$

Def. 3.15 Maßraum $(\mathbb{R}^n, \mathcal{B}_n, m)$ mit $m = LB$ -Maß,

$$f : (\mathbb{R}^n, \mathcal{B}_n) \rightarrow (\overline{\mathbb{R}}, \overline{\mathcal{B}}), A \in \mathcal{A}.$$

Dann heißt

$$\int_A f dm \quad (\text{falls existent})$$

Lebesgue-Integral (kurz: L -Integral)

Für $n = 1$ auch Bezeichnung: $\int_A f(x) dx$ oder $\int_A f dx$.

Satz 3.16:

Sei $f : [a : b] \rightarrow \mathbb{R}$ Riemann-integrierbar und messbar. Dann ist f auch Lebesgue-integrierbar und das Lebesgue-Integral stimmt mit dem Riemann-Integral überein.

Bem.:

Die Umkehrung von Satz 3.16 gilt aber i. A. nicht:

$$f : [0, 1] \rightarrow \mathbb{R}, \quad f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

ist Lebesgue-integrierbar (da $f = 0$ m -f.ü.), aber nicht Riemann-integrierbar.

Beweis von Satz 3.16:

Ist f Riemann-integrierbar, so ist f auch beschränkt (da sonst jede Obersumme $+\infty$ wäre), daher ist f nach Satz 3.13 b) auch Lebesgue-integrierbar.

Also genügt es zu zeigen: Riemann- und Lebesgue-Integral stimmen überein.

Dazu sei oBdA $f \geq 0$ (sonst zerlegen wir f in Positiv- und Negativanteil und zeigen die Aussagen für die beiden Anteile separat).

Setze

$$x_{k,n} = a + k \cdot \frac{b-a}{n} \text{ für } k = 0, 1, \dots, n$$

und

$$I_{k,n} = \begin{cases} [x_{k-1,n}, x_{k,n}) & \text{für } k \leq n-1, \\ [x_{n-1,n}, x_{n,n}] & \text{für } k = n. \end{cases}$$

Dann gilt

$$\begin{aligned} \text{Riemann-Integral von } f &= \lim_{n \rightarrow \infty} \sum_{k=1}^n \left(\inf_{x \in I_{k,n}} f(x) \right) \cdot \frac{1}{n} \\ &= \lim_{n \rightarrow \infty} \sum_{k=1}^n \left(\sup_{x \in I_{k,n}} f(x) \right) \cdot \frac{1}{n}. \end{aligned}$$

Mit

$$\begin{aligned} \sum_{k=1}^n \left(\inf_{x \in I_{k,n}} f(x) \right) \cdot \chi_{I_{k,n}}(t) &\leq f(t) \cdot \chi_{[a,b]}(t) \\ &\leq \sum_{k=1}^n \left(\sup_{x \in I_{k,n}} f(x) \right) \cdot \chi_{I_{k,n}}(t) \end{aligned}$$

folgt aus der Monotonie des (Lebesgue-)Integrals:

$$\begin{aligned} \sum_{k=1}^n \left(\inf_{x \in I_{k,n}} f(x) \right) \cdot \frac{1}{n} &\leq \int_{\mathbb{R}} f(t) \cdot \chi_{[a,b]}(f) dt \\ &\leq \sum_{k=1}^n \left(\sup_{x \in I_{k,n}} f(x) \right) \cdot \frac{1}{n}. \end{aligned}$$

Da rechts und linke Seite beide für $n \rightarrow \infty$ gegen das Riemann-Integral von f konvergieren, folgt daraus die Behauptung. $\square$

3.2 Integralkonvergenzsätze

Wir zeigen zunächst folgende leichte Verallgemeinerung des Satzes von Beppo Levi:

Satz 3.17: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $f_n : (\Omega, \mathcal{A}) \rightarrow (\overline{\mathbb{R}}, \overline{\mathcal{B}})$ ($n \in \mathbb{N}$). Es gelte

$$\int f_n d\mu > -\infty \text{ für mindestens ein } n$$

und

$$f_n \uparrow f.$$

Dann existiert $\int f d\mu$ und es gilt

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

Beweis:

(*Beweisidee:* Bisherige Fassung des Satzes auf $h_n = f_n - f_1 \geq 0$ anwenden).

oBdA $\int f_n d\mu > -\infty$ für alle $n \in \mathbb{N}$

(denn $\int f_n d\mu \leq \int f_{n+1} d\mu$ und Aussage ist asymptotisch)

Also

$$\int f_n^- d\mu < \infty \text{ für alle } n \in \mathbb{N}.$$

Mit Satz 3.13 a) folgt daraus

$$\forall n \in \mathbb{N} : \text{f.ü. } f_n^- < \infty$$

was

$$\forall n \in \mathbb{N} \text{ f.ü. } f_n > -\infty$$

impliziert.

Da abzählbare Vereinigung von Nullmengen eine Nullmenge ergibt, gilt sogar:

$$\text{f.ü. } [\forall n \in \mathbb{N} f_n > -\infty]$$

Ändert man die Integranden auf einer Nullmenge ab, so ändert sich der Wert des Integrals nicht (vgl. Satz 3.11 e)). Daher sogar oBdA:

$$\forall n \in \mathbb{N} : f_n > -\infty$$

Setze

$$h_n(\omega) = \begin{cases} f_n(\omega) - f_1(\omega) & , \text{ falls } f_1(\omega) < \infty \\ 0 & , \text{ falls } f_1(\omega) = \infty \end{cases}$$

und

$$h(\omega) = \begin{cases} f(\omega) - f_1(\omega) & , \text{ falls } f_1(\omega) < \infty \\ 0 & , \text{ falls } f_1(\omega) = \infty. \end{cases}$$

Dann gilt:

$$h_n, h \text{ nichtnegativ messbar} \quad (\text{da } f_n \uparrow)$$

$$h_n \uparrow h \quad (\text{denn für } f_1(\omega) = \infty \text{ gilt trivialerweise } h_n(\omega) \uparrow h(\omega) \text{ (} n \rightarrow \infty))$$

Mit Satz 3.7 folgt:

$$\int h_n d\mu \rightarrow \int h d\mu \quad (n \rightarrow \infty).$$

Wegen

$$f_n = f_1 + h_n$$

$$f = f_1 + h$$

(denn im Falle $f_1(\omega) < \infty$ ist diese Beziehung klar, für $f_1(\omega) = \infty$ gilt $f_n(\omega) = f(\omega) = \infty$ wegen $f_n \uparrow f$ und daher sind dann oben beide Seiten unendlich) folgt daraus mit Satz 3.11 c):

$$\int f_n d\mu = \int f_1 d\mu + \int h_n d\mu \xrightarrow{\text{s.o.}} \int f_1 d\mu + \int h d\mu = \int f d\mu \quad (n \rightarrow \infty).$$

Hierbei existiert $\int f d\mu$, da $\int f_1 d\mu$ existiert nach Voraussetzung, $\int h d\mu$ existiert nach Satz 3.7 und $\int f_1 d\mu + \int h d\mu$ existiert wegen Voraussetzung $\int f_1 d\mu > -\infty$ (vgl. Satz 3.11 c)). $\square$

Bemerkung 3.18: Die Voraussetzung

a) $\int f_n d\mu > -\infty$ für ein $n \in \mathbb{N}$

bzw.

b) $f_n \uparrow$

in Satz 3.17 darf nicht weggelassen werden.

Begründung:

Setze $(\Omega, \mathcal{A}, \mu) = (\mathbb{R}, \mathcal{B}, m)$ mit $m = \text{LB-Ma\ss}$.

Zu a):

$$f_n := -\frac{1}{n} \uparrow 0$$

aber

$$\int f_n d\mu = -\infty \not\rightarrow \int 0 d\mu = 0.$$

Zu b):

$$f_n := \frac{1}{n} \downarrow 0$$

aber

$$\int f_n d\mu = +\infty \not\rightarrow \int 0 d\mu = 0.$$

Satz 3.19: (Lemma von Fatou)

Sei $(\Omega, \mathcal{A}, \mu)$ ein Ma\ssraum.

a) Sind f_n messbar ($n \in \mathbb{N}$) mit $f_n \geq g$ f.ü. für eine integrierbare Funktion g . Dann existiert

$$\int \underline{\lim} f_n d\mu \text{ und es gilt}$$

$$\int \underline{\lim} f_n d\mu \leq \underline{\lim} \int f_n d\mu.$$

b) Sind f_n messbar ($n \in \mathbb{N}$) mit $f_n \leq h$ f.ü. für eine integrierbare Funktion h . Dann existiert $\int \overline{\lim} f_n d\mu$ und es gilt

$$\int \overline{\lim} f_n d\mu \geq \overline{\lim} \int f_n d\mu.$$

Beweis.

a) **Fall 1:** $f_n \geq 0$ für alle $n \in \mathbb{N}$

Setze

$$g_n = \inf_{k \geq n} f_k.$$

Dann ist g_n nichtnegativ messbar mit

$$g_n \uparrow \lim_{n \rightarrow \infty} g_n = \lim_{n \rightarrow \infty} \inf_{k \geq n} f_k = \underline{\lim} f_k.$$

Nach dem Satz von der monotonen Konvergenz existiert dann $\int \underline{\lim} f_k d\mu$ und es gilt

$$\begin{aligned} \int \underline{\lim} f_k d\mu &= \lim_{n \rightarrow \infty} \int g_n d\mu \\ &= \lim_{n \rightarrow \infty} \int \inf_{k \geq n} f_k d\mu \\ &\leq \lim_{n \rightarrow \infty} \int \inf_{k \geq n} f_k d\mu \\ &\quad \text{(Monotonie des Integrals)} \\ &= \underline{\lim} \int f_k d\mu. \end{aligned}$$

Fall 2: Allgemeines f_n .

Da g integrierbar gilt nach Satz 3.13 a) $g(\omega) \in \mathbb{R}$ f. ü.

Da Integral sich bei Abändern des Integranden auf einer Nullmenge nicht ändert (vgl. Satz 3.11 e)) gilt

$$\text{oBdA } g(\omega) \in \mathbb{R} \quad \text{für alle } \omega \in \Omega$$

und

$$\text{oBdA } f_n(\omega) \geq g(\omega) \quad \text{für alle } \omega \in \Omega.$$

Setze

$$F_n(\omega) = f_n(\omega) - g(\omega) \geq 0.$$

Nach dem 1. Fall existiert dann $\int \underline{\lim} F_n d\mu$ und es gilt

$$\int \underline{\lim} F_n d\mu \leq \underline{\lim} \int F_n d\mu.$$

Mit

$$f_n = F_n + g \quad \text{und} \quad \underline{\lim} f_n = \underline{\lim} F_n + g$$

folgt, da g integrierbar und reellwertig ist, dass auch

$$\int \underline{\lim} f_n d\mu = \int (\underline{\lim} f_n + g) d\mu$$

existiert (vgl. Satz 3.11 c)), und es folgt:

$$\begin{aligned}
 \int \underline{\lim} f_n d\mu &= \int (\underline{\lim} F_n + g) d\mu \\
 &= \int \underline{\lim} F_n d\mu + \int g d\mu \\
 &\quad \text{(nach Satz 3.11 c))} \\
 &\leq \underline{\lim} \int F_n d\mu + \int g d\mu \\
 &\quad \text{(nach Fall 1)} \\
 &= \underline{\lim} (\int F_n d\mu + \int g d\mu) \\
 &= \underline{\lim} \int (F_n + g) d\mu \\
 &\quad \text{(nach Satz 3.11 c))} \\
 &= \underline{\lim} \int f_n d\mu.
 \end{aligned}$$

b) Ist

$$f_n \leq h \text{ f.ü.},$$

so ist auch

$$-f_n \geq -h \text{ f.ü.}$$

und nach Teil a) existiert $\int \underline{\lim} (-f_n) d\mu$ und es gilt

$$\int \underline{\lim} (-f_n) d\mu \leq \underline{\lim} \int (-f_n) d\mu.$$

Mit

$$\begin{aligned}
 \underline{\lim} (-a_n) &= \sup_{n \in \mathbb{N}} \inf_{k \geq n} (-a_n) = - \inf_{n \in \mathbb{N}} \sup_{k \geq n} a_n \\
 &= -\overline{\lim} a_n
 \end{aligned}$$

folgt daraus und mit der Linearität des Integrals die Existenz von $\int \overline{\lim} f_n d\mu$ sowie

$$- \int \overline{\lim} f_n d\mu \leq -\overline{\lim} \int f_n d\mu,$$

was die Behauptung impliziert. □

Bemerkung 3.20: Die Voraussetzung

$$f_n \geq g \text{ f.ü. mit } g \text{ integrierbar}$$

in a) ist notwendig.

Denn für $(\Omega, \mathcal{A}, \mu) = (\mathbb{R}, \mathcal{B}, m)$ mit $m = \text{LB-Maß}$ und

$$f_n = -\frac{1}{n}$$

gilt

$$\int \underline{\lim} f_n d\mu = \int 0 d\mu = 0$$

aber

$$\underline{\lim} \int f_n d\mu = \underline{\lim}(-\infty) = -\infty,$$

also ist hier

$$\int \underline{\lim} f_n d\mu > \underline{\lim} \int f_n d\mu.$$

Satz 3.21 (Satz von der majorisierten Konvergenz; Lebesgue)

Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und seien

$$f_n, f \text{ messbare Funktionen } (n \in \mathbb{N}).$$

Sei g integrierbar und es gelte

$$|f_n| \leq g \text{ f. ü. } (n \in \mathbb{N})$$

und

$$f_n \rightarrow f \text{ } (n \rightarrow \infty) \text{ f.ü.}$$

Dann sind f_n, f integrierbar und es gilt

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

Beweis:

oBdA gelten die Voraussetzungen überall statt f.ü.

oBdA sei g reellwertig.

f_n, f sind messbar mit

$$\int |f_n| d\mu \leq \int |g| d\mu < \infty$$

und

$$\int |f| d\mu \leq \int |g| d\mu < \infty,$$

also sind f_n, f nach Satz 3.13 b) integrierbar.

Wegen

$$\begin{aligned} |\int f_n d\mu - \int f d\mu| &= |\int (f_n - f) d\mu| \\ &\leq \int |f_n - f| d\mu \end{aligned}$$

genügt es zu zeigen

$$\int |f_n - f| d\mu \rightarrow 0 \text{ } (n \rightarrow \infty).$$

Setze

$$h_n := |f_n - f| \geq 0.$$

Nach Voraussetzung gilt $h_n \rightarrow 0$ ($n \rightarrow \infty$) und

$$h_n \leq |f_n| + |f| \leq 2 \cdot g$$

mit $2g$ integrierbar.

Aus $h_n > 0$ folgt

$$0 \leq \int h_n d\mu$$

und mit dem Lemma von Fatou folgt weiter

$$\overline{\lim} \int h_n d\mu \leq \int \overline{\lim} h_n d\mu = \int 0 d\mu = 0.$$

Also gilt

$$\lim_{n \rightarrow \infty} \int h_n d\mu = 0, \text{ w.z.z.w.}$$

□

Bemerkung 3.22: Die Voraussetzung $|f_n| \leq g$ f.ü. darf in Satz 3.21 nicht weggelassen werden, denn für $(\Omega, \mathcal{A}, \mu) = (\mathbb{R}, \mathcal{B}, m)$ mit $m = \text{LB-Ma\ss}$ und

$$f_n(x) = \begin{cases} n^2 & , \quad 0 < x < \frac{1}{n} \\ 0 & , \quad \text{sonst} \end{cases}$$

gilt $f_n \rightarrow 0$ f. ü. aber $\int f_n d\mu = n \rightarrow \infty$ ($n \rightarrow \infty$).

3.3 Ma\ss und Integral in Produkträumen

Fragestellungen:

- a) Wie kann man zwei Ma\ssen ein Produkt zuordnen?
- b) Wie integriert man bzgl. einem solchen Produktma\ss?

zu a):

Genauer:

Sind $(\Omega, \mathcal{A}_i, \mu_i)$ ($i = 1, 2$) Ma\ssräume, so ist ein Ma\ss μ definiert auf einer σ -Algebra $\mathcal{A} \subseteq \mathcal{P}(\Omega_1 \times \Omega_2)$ gesucht mit

$$\mu(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2)$$

für alle $A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2$.

Definitionsbereich dieses Maßes?

Def. 3.23 Seien $(\Omega_1, \mathcal{A}_1), (\Omega_2, \mathcal{A}_2)$ zwei Messräume. Dann wird das Produkt der σ -Algebren $\mathcal{A}_1$ und $\mathcal{A}_2$ definiert durch

$$\mathcal{A}_1 \otimes \mathcal{A}_2 := \mathcal{F}_{\Omega_1 \times \Omega_2}(\{A_1 \times A_2 : A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\}),$$

d. h. $\mathcal{A}_1 \otimes \mathcal{A}_2$ ist die kleinste σ -Algebra in $\Omega_1 \times \Omega_2$, die alle Mengen der Form $A_1 \times A_2$ ($A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2$) enthält.

Satz 3.24 Sind $(\Omega_i, \mathcal{A}_i, \mu_i)$ ($i = 1, 2$) zwei σ -endliche Maßräume, so existiert höchstens ein Maß μ auf $\mathcal{A}_1 \otimes \mathcal{A}_2$ mit

$$(*) \quad \mu(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2) \text{ für alle } A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2$$

Beweis:

$$\mathcal{C} := \{A_1 \times A_2 : A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\}$$

ist $\cap$ -stabiler Erzeuger von $\mathcal{A}_1 \otimes \mathcal{A}_2$ (denn

$$(A_1 \times A_2) \cap (\overline{A_1} \times \overline{A_2}) = (A_1 \cap \overline{A_1}) \times (A_2 \cap \overline{A_2}).$$

Die Mengenfunktion

$$A_1 \times A_2 \mapsto \mu_1(A_1) \cdot \mu_2(A_2) \quad (A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2)$$

ist σ -endlich auf $\mathcal{C}$. Denn da $\mathcal{A}_i$ σ -endlich ist, existieren Mengen $A_n^i \uparrow \Omega_i$ mit $\mu_i(A_n^i) < \infty$ für alle $n \in \mathbb{N}$ ($i = 1, 2$). Dann gilt aber auch

$$A_n^1 \times A_n^2 \uparrow \Omega_1 \times \Omega_2$$

(da $A_n^1 \times A_n^2 \subseteq A_{n+1}^1 \times A_{n+2}^2$ und

$$\begin{aligned} \bigcup_{n \in \mathbb{N}} A_n^1 \times A_n^2 &\supseteq \bigcup_{j \in \mathbb{N}} \left(\bigcup_{n \in \mathbb{N}} A_n^1 \times A_j^2 \right) = \bigcup_{j \in \mathbb{N}} (\Omega_1 \times A_j^2) \\ &= \Omega_1 \times \Omega_2 \quad) \end{aligned}$$

und

$$\mu_1(A_n^1) \cdot \mu_2(A_n^2) < \infty \text{ für alle } n \in \mathbb{N}.$$

Nach Satz 1.41 ist daher μ (sofern überhaupt existent) durch $(*)$ bereits eindeutig festgelegt. $\square$

Def. 3.25 Ω_1, Ω_2 Mengen, $A \subseteq \Omega_1 \times \Omega_2$, $\omega_1 \in \Omega_1, \omega_2 \in \Omega_2$.

Dann heißt

$$A_{\omega_1} = \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in A\}$$

bzw.

$$A_{\omega_2} = \{\omega_1 \in \Omega_1 : (\omega_1, \omega_2) \in A\}$$

der ω_1 -Schnitt von A bzw. der ω_2 -Schnitt von A :

Lemma 3.26

Seien $(\Omega_i, \mathcal{A}_i)$ ($i = 1, 2$) Messräume und $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$.

Dann gilt für alle $\omega_1 \in \Omega_1$

$$A_{\omega_1} \in \mathcal{A}_2$$

und für alle $\omega_2 \in \Omega_2$

$$A_{\omega_2} \in \mathcal{A}_1.$$

Beweis:

Wegen Symmetrie genügt der Nachweis der Aussage für ω_1 . Dies wiederum folgt aus

$$\mathcal{A}_1 \otimes \mathcal{A}_2 \subseteq \mathcal{G} := \{A \subseteq \Omega_1 \times \Omega_2 : A_{\omega_1} \in \mathcal{A}_2 \text{ für alle } \omega_1 \in \Omega_1\}$$

Nun gilt:

1.) $\mathcal{C} = \{A_1 \times A_2 : A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\} \subseteq \mathcal{G}$

da

$$(A_1 \times A_2)_{\omega_1} = \begin{cases} A_2 & \text{falls } \omega_1 \in A_1 \\ \emptyset & \text{falls } \omega_1 \notin A_1 \end{cases}$$

2.) $\mathcal{G}$ ist σ -Algebra in $\Omega_1 \times \Omega_2$, denn:

(a) $\emptyset \in \mathcal{G}$, da $\emptyset_{\omega_1} = \emptyset \in \mathcal{A}_2$ für alle $\omega_1 \in \Omega_1$.

(b) Ist $A \in \mathcal{G}$, so gilt auch $A^c = (\Omega_1 \times \Omega_2) \setminus A \in \mathcal{G}$ da

$$\begin{aligned} & ((\Omega_1 \times \Omega_2) \setminus A)_{\omega_1} \\ &= \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in (\Omega_1 \times \Omega_2) \setminus A\} \\ &= \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \notin A\} \\ &= \{\omega_2 \in \Omega_2 : \omega_2 \notin A_{\omega_1}\} \\ &= \Omega_2 \setminus A_{\omega_1} \in \mathcal{A}_2 \end{aligned}$$

(denn $A_{\omega_1} \in \mathcal{A}_2$ und $\mathcal{A}_2$ ist σ -Algebra).

(c) Sind $A_n \in \mathcal{G}$ ($n \in \mathbb{N}$), so ist auch $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{G}$, da

$$\begin{aligned} \left(\bigcup_{n \in \mathbb{N}} A_n \right)_{\omega_1} &= \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in \bigcup_{n \in \mathbb{N}} A_n\} \\ &= \bigcup_{n \in \mathbb{N}} \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in A_n\} \\ &= \bigcup_{n \in \mathbb{N}} (A_n)_{\omega_1} \in \mathcal{A}_2 \end{aligned}$$

da $(A_n)_{\omega_1} \in \mathcal{A}_2$ für alle $n \in \mathbb{N}$ und $\mathcal{A}_2$ σ -Algebra ist.

Damit:

$$\mathcal{G} \stackrel{(1),(2)}{\supseteq} \mathcal{F}(\mathcal{C}) \stackrel{Def.}{=} \mathcal{A}_1 \otimes \mathcal{A}_2$$

□

Lemma 3.27

$(\Omega_i, \mathcal{A}_i, \mu_i)$ ($i = 1, 2$) seien zwei σ -endliche Maßräume. Dann ist für jedes $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ die Funktion

$$\omega_1 \mapsto \mu_2(A_{\omega_1})$$

bzw.

$$\omega_2 \mapsto \mu_1(A_{\omega_2})$$

auf Ω_1 bzw. Ω_2 definiert und $\mathcal{A}_1 - \overline{\mathcal{B}}$ - bzw. $\mathcal{A}_2 - \overline{\mathcal{B}}$ -messbar.

Beweis:

Aus Symmetriegründen genügt es, die Behauptung für die Abbildung

$$\omega_1 \mapsto \mu_2(A_{\omega_1})$$

zu zeigen.

Nach Lemma 3.26 ist $\mu_2(A_{\omega_1})$ definiert.

1. Schritt: Wir zeigen die Behauptung im Spezialfall $\mu_2(\Omega_2) < \infty$.

Bezeichnung: $s_A : \omega_1 \rightarrow \overline{\mathbb{R}}$ sei definiert durch $s_A(\omega_1) = \mu_2(A_{\omega_1})$.

Wir zeigen:

$$\mathcal{A}_1 \otimes \mathcal{A}_2 \subseteq \{A \subseteq \Omega_1 \times \Omega_2 : s_A \text{ } \mathcal{A}_1 - \overline{\mathcal{B}} \text{- messbar}\} =: \mathcal{G}$$

Dazu zeigen wird:

(1) $\mathcal{C} := \{A_1 \times A_2 : A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\} \subseteq \mathcal{G}$

(2) $\mathcal{G}$ ist Dynkin-System.

Daraus folgt die Behauptung, denn:

$$\mathcal{G} \stackrel{(1),(2)}{\supseteq} \mathcal{D}(\mathcal{C}) \stackrel{\mathcal{C} \text{ } \cap\text{-stabil}}{=} \mathcal{F}(\mathcal{C}) \stackrel{Def.}{=} \mathcal{A}_1 \otimes \mathcal{A}_2.$$

Nachweis von (1):

$$\begin{aligned} s_{A_1 \times A_2}(\omega_1) &= \mu_2((A_1 \times A_2)_{\omega_1}) = \begin{cases} \mu_2(A_2) & \text{für } \omega_1 \in A_1 \\ \mu_2(\emptyset) = 0 & \text{für } \omega_1 \notin A_1 \end{cases} \\ &= \mu_2(A_2) \cdot \chi_{A_1}(\omega_1) \end{aligned}$$

Also ist $s_{A_1 \times A_2}$ (als einfache Funktion) messbar für $A_1 \times A_2 \in \mathcal{C}$.

Nachweis von (2):

$\alpha)$ $\Omega_1 \times \Omega_2 \in \mathcal{G}$ da

$$s_{\Omega_1 \times \Omega_2}(\omega_1) = \mu_2(\Omega_2)$$

als konstante Funktion messbar ist.

$\beta)$ Seien $A, B \in \mathcal{G}$ mit $A \subseteq B$.

Dann gilt

$$B_{\omega_1} = (B \setminus A \cup A)_{\omega_1} = (B \setminus A)_{\omega_1} \cup A_{\omega_1},$$

woraus folgt

$$\mu_2(B_{\omega_1}) = \mu_2((B \setminus A)_{\omega_1}) + \mu_2(A_{\omega_1})$$

bzw.

$$\mu_2((B \setminus A)_{\omega_1}) = \mu_2(B_{\omega_1}) - \mu_2(A_{\omega_1})$$

(da $\mu_2(\Omega_2) < \infty!$)

Also ist

$$s_{B \setminus A} \stackrel{s.o.}{=} s_B - s_A$$

als Differenz zweier messbarer Funktionen messbar.

Daraus folgt:

$$B \setminus A \in \mathcal{G}.$$

$\gamma)$ Seien $A_n \in \mathcal{G}$ ($n \in \mathbb{N}$) paarweise disjunkt.

Dann sind auch $(A_n)_{\omega_1} \in \mathcal{A}_2$ paarweise disjunkt, und es gilt

$$\left(\bigcup_n A_n \right)_{\omega_1} = \bigcup_n (A_n)_{\omega_1}.$$

Mit μ_2 Maß folgt daraus

$$\begin{aligned} \mu_2 \left(\left(\bigcup_n A_n \right)_{\omega_1} \right) &= \mu_2 \left(\bigcup_n (A_n)_{\omega_1} \right) \\ &= \sum_{n=1}^{\infty} \mu_2((A_n)_{\omega_1}), \end{aligned}$$

was

$$s_{\bigcup_n A_n} = \sum_{n=1}^{\infty} s_{A_n}$$

impliziert. Da s_{A_n} nach Voraussetzung messbar ist für jedes $n \in \mathbb{N}$, folgt die Messbarkeit von $s_{\bigcup_n A_n}$, was wiederum

$$\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{G}$$

impliziert.

Aus $\alpha), \beta)$, und $\gamma)$ folgt 2), und damit die Behauptung im 1. Schritt.

2. Schritt: Wir zeigen die Behauptung im Fall μ_2 σ -endlich.

Wähle $B_n \in \mathcal{A}_2$ ($n \in \mathbb{N}$) mit $B_n \uparrow \Omega_2$ und $\mu_2(B_n) < \infty$ für alle $n \in \mathbb{N}$.

Dann ist

$$A_2 \mapsto \mu_{2,n}(A_2) = \mu_2(A_2 \cap B_n)$$

ein endliches Maß auf $\mathcal{A}_2$, und damit ist nach Schritt 1

$$\omega_1 \mapsto \mu_2(A_{\omega_1} \cap B_n)$$

für jedes $n \in \mathbb{N}$ eine $\mathcal{A}_1 - \overline{\mathcal{B}}$ -messbare Funktion.

Wegen

$$\mu_2(A_{\omega_1}) = \lim_{n \rightarrow \infty} \mu_2(A_{\omega_1} \cap B_n)$$

(was aus der Stetigkeit von unten von μ_2 und $A_{\omega_1} \cap B_n \uparrow A_{\omega_1}$ folgt), ist dann aber auch

$$\omega_1 \mapsto \mu_2(A_{\omega_1})$$

als punktweiser Grenzwert messbarer Funktionen selbst messbar. □

Satz 3.28 (Vorstufe des Satzes von Fubini):

Seien $(\Omega_i, \mathcal{A}_i, \mu_i)$ ($i = 1, 2$) zwei σ -endliche Maßräume. Dann existiert genau ein Maß μ auf $\mathcal{A}_1 \otimes \mathcal{A}_2$ mit

$$\mu(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2) \quad \text{für alle } A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2.$$

Dieses Maß ist σ -endlich und ist gegeben durch

$$\begin{aligned} \mu(A) &= \int_{\Omega_2} \mu_1(A_{\omega_2}) d\mu_2(\omega_2) \\ &= \int_{\Omega_1} \mu_2(A_{\omega_1}) d\mu_1(\omega_1) \end{aligned}$$

Beweis: Setze

$$\bar{\mu}(A) = \int_{\Omega_2} \mu_1(A_{\omega_2}) d\mu_2(\omega_2).$$

Nach Lemma 3.27 ist $\bar{\mu}$ wohldefiniert.

Wegen

$$\mu_1(A_{\omega_2}) \geq 0 \quad \text{für alle } \omega_2 \in \Omega_2$$

gilt

$$\bar{\mu}(A) \geq 0 \quad \text{für alle } A \in \mathcal{A}_1 \otimes \mathcal{A}_2.$$

Weiter gilt

$$\bar{\mu}(\emptyset) = \int_{\Omega_2} \mu_1(\emptyset) d\mu_2(\omega_2) = \int_{\Omega_2} 0 d\mu_2(\omega_2) = 0$$

und für paarweise disjunkte $A_1, A_2, \dots \in \mathcal{A}_1 \otimes \mathcal{A}_2$ gilt:

$$\begin{aligned}
\bar{\mu} \left(\sum_{n=1}^{\infty} A_n \right) &= \int_{\Omega_2} \mu_1 \left(\underbrace{\left(\sum_{n=1}^{\infty} A_n \right)_{\omega_2}}_{= \sum_{n=1}^{\infty} (A_n)_{\omega_2}} \right) d\mu_2(\omega_2) \\
&= \int_{\Omega_2} \sum_{n=1}^{\infty} \mu_1((A_n)_{\omega_2}) d\mu_2(\omega_2) \\
&\quad (\text{da } \mu_1 \text{ Maß ist}) \\
&= \sum_{n=1}^{\infty} \int_{\Omega_2} \mu_1((A_n)_{\omega_2}) d\mu_2(\omega_2) \\
&\quad (\text{nach dem Satz von der monotonen Konvergenz} \\
&\quad \text{und der Linearität des Integrals)} \\
&= \sum_{n=1}^{\infty} \bar{\mu}(A_n).
\end{aligned}$$

Also ist $\bar{\mu}$ ein Maß. Für dieses gilt weiter:

$$\begin{aligned}
\bar{\mu}(A_1 \times A_2) &= \int_{\Omega_2} \mu_1((A_1 \times A_2)_{\omega_2}) d\mu_2(\omega_2) \\
&= \int_{\Omega_2} \mu_2(A_1) \cdot \chi_{A_2}(\omega_2) d\mu_2(\omega_2) \\
&= \mu_1(A_1) \cdot \mu_2(A_2)
\end{aligned}$$

für alle $A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2$.

Damit ist die Existenz des Maßes μ aus der Behauptung gezeigt.

Die Eindeutigkeit folgt aus Satz 3.24, im Beweis dort wurde auch die σ -Endlichkeit von μ bewiesen.

Analog zu oben folgt, dass auch

$$A \mapsto \hat{\mu}(A) = \int_{\Omega_1} \mu_2(A_{\omega_1}) d\mu_1(\omega_1)$$

ein Maß auf $\mathcal{A}_1 \otimes \mathcal{A}_2$ mit

$$\hat{\mu}(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2) \quad (A_i \in \mathcal{A}_i)$$

ist, woraus mit Satz 3.24 folgt:

$$\hat{\mu} = \bar{\mu}.$$

□

Def. 3.29: Seien $(\Omega_i, \mathcal{A}_i, \mu_i)$ ($i = 1, 2$) zwei σ -endliche Maßräume. Das in Satz 3.28 auftretende Maß

$$\mu =: \mu_1 \otimes \mu_2$$

auf $\mathcal{A}_1 \otimes \mathcal{A}_2$ heißt **Produkt-Maß**.

Der Maßraum $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2, \mu_1 \otimes \mu_2)$ heißt **Produkt-Maßraum**.

Bem.: Es gilt

$$\begin{aligned} \int \chi_A d\mu_1 \otimes \mu_2 &= (\mu_1 \otimes \mu_2)(A) = \int_{\Omega_1} \mu_2(A_{\omega_1}) d\mu(\omega_1) \\ &= \int_{\Omega_1} \int_{\Omega_2} \chi_A(\omega_1, \omega_2) d\mu_2(\omega_2) d\mu_1(\omega_1) \end{aligned}$$

Im Folgenden zeigen wir diese Formel für allgemeines (integrierbares)

$$f : (\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2) \rightarrow (\mathbb{R}, \mathcal{B}).$$

Lemma 3.30:

Messräume $(\Omega_i, \mathcal{A}_i)$ ($i = 1, 2$), $f : (\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2) \rightarrow (\overline{\mathbb{R}}, \overline{\mathcal{B}})$.

Dann ist der sog. ω_1 -Schritt

$$f_{\omega_1} : \Omega_2 \rightarrow \overline{\mathbb{R}}, f_{\omega_1}(\omega_2) = f(\omega_1, \omega_2)$$

von f $\mathcal{A}_2 - \overline{\mathcal{B}}$ -messbar, und der sogenannte ω_2 -Schritt

$$f_{\omega_2} : \Omega_1 \rightarrow \overline{\mathbb{R}}, f_{\omega_2}(\omega_1) = f(\omega_1, \omega_2)$$

von f ist $\mathcal{A}_1 - \overline{\mathcal{B}}$ -messbar (jeweils für alle $\omega_1 \in \Omega_1, \omega_2 \in \Omega_2$).

Beweis:

Für $B \in \overline{\mathcal{B}}$ gilt:

$$\begin{aligned} f_{\omega_1}^{-1}(B) &= \{\omega_2 \in \Omega_2 : f_{\omega_1}(\omega_2) \in B\} \\ &= \{\omega_2 \in \Omega_2 : f(\omega_1, \omega_2) \in B\} \\ &= \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in f^{-1}(B)\} \\ &= (f^{-1}(B))_{\omega_1} \in \mathcal{A}_2 \end{aligned}$$

da $f^{-1}(B) \in \mathcal{A}_1 \otimes \mathcal{A}_2$ und nach Lemma 7.1 für jedes $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ gilt $A_{\omega_1} \in \mathcal{A}_2$.

Analog folgt die Behauptung für f_{ω_2} □

Satz 3.31 (Satz von Tonelli / Satz von Fubini)

Seien $(\Omega_i, \mathcal{A}_i, \mu_i)$ ($i = 1, 2$) zwei σ -endliche Maßräume, und sei

$$f : (\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2) \rightarrow (\overline{\mathbb{R}}_+, \overline{\mathcal{B}}_+).$$

Dann sind die Funktionen

$$\omega_2 \mapsto \int_{\Omega_1} f_{\omega_2} d\mu_1, \quad \omega_1 \mapsto \int_{\Omega_2} f_{\omega_1} d\mu_2$$

$\mathcal{A}_2 - \overline{\mathcal{B}}$ bzw. $\mathcal{A}_1 - \overline{\mathcal{B}}$ -messbar und es gilt:

$$\begin{aligned} \int f d(\mu_1 \otimes \mu_2) &= \int_{\Omega_2} \int_{\Omega_1} (f_{\omega_2}(\omega_1) d\mu_1(\omega_1) d\mu_2(\omega_2) \\ &= \int_{\Omega_1} \int_{\Omega_2} (f_{\omega_1}(\omega_2) d\mu_2(\omega_2) d\mu_1(\omega_1) \end{aligned}$$

Beweis: Gemäß Lemma 3.30 die inneren Integrale wohldefiniert.

Fall 1: $f = \chi_A$

In diesem Fall gilt

$$\int_{\Omega_1} f_{\omega_2} d\mu_1 = \mu_1(A_{\omega_2}), \int_{\Omega_2} f_{\omega_1} d\mu_2 = \mu_2(A_{\omega_2})$$

und

$$\int f d(\mu_1 \otimes \mu_2) = (\mu_1 \otimes \mu_2)(A).$$

Die Behauptung folgt aus Lemma 3.27 und Satz 3.28.

Fall 2: f nichtnegativ einfach.

Unter Beachtung der Messbarkeit von Linearkombinationen messbarer Funktionen und der Linearität des Integrals folgt die Behauptung aus Fall 1.

Fall 3: f nichtnegativ messbar.

Wähle Folge nichtnegativer einfacher Funktionen f_n mit $f_n \uparrow f$.

Dann sind $(f_n)_{\omega_i}$ ebenfalls nichtnegativ einfach und es gilt

$$(f_n)_{\omega_i} \uparrow f_{\omega_i} \quad (i = 1, 2).$$

Insbesondere gilt

$$\int_{\Omega_1} f_{\omega_2} d\mu_1 = \lim_{n \rightarrow \infty} \int_{\Omega_1} (f_n)_{\omega_2} d\mu_1 \quad (*)$$

nach Definition des Integrals und

$$\omega_2 \mapsto \int_{\Omega_1} f_{\omega_2} d\mu_1$$

ist als punktweiser Grenzwert messbarer Funktionen (vgl. Fall 2) selbst messbar.

Weiter folgt aus $((f_n)_{\omega_1})_{n \in \mathbb{N}}$ nichtnegativ und wachsend, dass auch

$$\left(\int (f_n)_{\omega_1} d\mu_2 \right)_{n \in \mathbb{N}}$$

nichtnegativ und wachsend ist, und es gilt:

$$\begin{aligned} \int f d(\mu_1 \otimes \mu_2) &= \lim_{n \rightarrow \infty} \int f_n d(\mu_1 \otimes \mu_2) \\ &\quad \text{(nach Definition des Integrals)} \\ &= \lim_{n \rightarrow \infty} \int_{\Omega_2} \left\{ \int_{\Omega_1} (f_n)_{\omega_2}(\omega_1) d\mu_1(\omega_1) \right\} d\mu_2(\omega_2) \\ &\quad \text{(nach Fall 2)} \\ &= \int_{\Omega_2} \left\{ \lim_{n \rightarrow \infty} \int_{\Omega_1} (f_n)_{\omega_2}(\omega_1) d\mu_1(\omega_1) \right\} d\mu_2(\omega_2) \\ &\quad \text{(nach dem Satz von der monotonen Konvergenz)} \\ &= \int_{\Omega_2} \left\{ \int_{\Omega_1} f_{\omega_2}(\omega_1) d\mu_1(\omega_1) \right\} d\mu_2(\omega_2) \\ &\quad \text{(nach (*))} \end{aligned}$$

Der zweite Teil der Behauptung folgt analog. □

Satz 3.32 (Satz von Fubini)

Seien $(\Omega_i, \mathcal{A}_i, \mu_i)$ ($i = 1, 2$) σ -endliche Maßräume und sei

$$f : (\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2) \rightarrow (\overline{\mathbb{R}}, \overline{\mathcal{B}}).$$

a) Ist f $(\mu_1 \otimes \mu_2)$ -integrierbar, so gilt

$$\begin{aligned} \int f \, d(\mu_1 \otimes \mu_2) &= \int_{\Omega_2} \left(\int_{\Omega_1} f_{\omega_2}(\omega_1) \, d\mu_1(\omega_1) \right) d\mu_2(\omega_2) \\ &= \int_{\Omega_1} \left(\int_{\Omega_2} f_{\omega_1}(\omega_2) \, d\mu_2(\omega_2) \right) d\mu_1(\omega_1), \end{aligned}$$

wobei gilt:

$$f_{\omega_2} : \Omega_1 \rightarrow \overline{\mathbb{R}} \text{ ist für } \mu_2 - f.a. \, \omega_2 \in \Omega_2 \text{ } \mu_1 - \text{integrierbar,}$$

$$f_{\omega_1} : \Omega_2 \rightarrow \overline{\mathbb{R}} \text{ ist für } \mu_1 - f.a. \, \omega_1 \in \Omega_1 \text{ } \mu_2 - \text{integrierbar}$$

und die μ_2 -f.ü. definierte Funktion

$$\omega_2 \mapsto \int_{\Omega_1} f_{\omega_2} \, d\mu_1(\omega_1)$$

ist μ_2 -integrierbar, und die μ_1 -f.ü. definierte Funktion

$$\omega_1 \mapsto \int_{\Omega_2} f_{\omega_1} \, d\mu_2(\omega_2)$$

ist μ_1 -integrierbar.

b) Aus

$$\int_{\Omega_2} \int_{\Omega_1} |f_{\omega_2}(\omega_1)| \, d\mu_1(\omega_1) \, d\mu_2(\omega_2) < \infty$$

bzw.

$$\int_{\Omega_1} \int_{\Omega_2} |f_{\omega_1}(\omega_2)| \, d\mu_2(\omega_2) \, d\mu_1(\omega_1) < \infty$$

folgt f $\mu_1 \otimes \mu_2$ -integrierbar, und in diesem Fall gilt die Behauptung aus a).

Beweis:

a) Sei f $(\mu_1 \otimes \mu_2)$ -integrierbar, und sei $f = f^+ - f^-$ mit $f^+, f^- \geq 0$.

Dann gilt nach Satz 3.31:

$$\int_{\Omega_1} \left[\int_{\Omega_2} (f^+)_{\omega_1}(\omega_2) \, d\mu_2(\omega_2) \right] d\mu_1(\omega_1) = \int f^+ \, d(\mu_1 \otimes \mu_2) < \infty,$$

woraus folgt

- $[\dots] < \infty$ für μ_1 -f.a. ω_1 ,
- $\omega_1 \mapsto [\dots]$ ist μ_1 -integrierbar,
- $(f^+)_{\omega_1}$ ist für μ_1 -f.a. ω_1 μ_2 -integrierbar.

Unter Verwendung der analogen Aussagen für f^- und Beachtung von

$$(f^+)_{\omega_1}(\omega_2) = (f_{\omega_1})^+(\omega_2), (f^-)_{\omega_2}(\omega_1) = (f_{\omega_2})^-(\omega_1)$$

folgt der zweite Teil der Behauptung aus

$$\int f d(\mu_1 \otimes \mu_2) = \int f^+ d(\mu_1 \otimes \mu_2) - \int f^- d(\mu_1 \otimes \mu_2).$$

Vertauschen von ω_1 und ω_2 liefert analog den ersten Teil der Behauptung.

b) Aus

$$\int |f| d(\mu_1 \otimes \mu_2) = \int_{\Omega_2} \int_{\Omega_1} |f_{\omega_2}(\omega_1)| d\mu_1(\omega_1) d\mu_2(\omega_2) < \infty$$

(wobei die erste Gleichheit aus Satz 3.31 folgt) folgt $|f|$ $(\mu_1 \otimes \mu_2)$ -integrierbar, was nach Satz 3.13 f integrierbar impliziert. $\square$

Bemerkung:

Die Definition und Sätze dieses Abschnittes gelten entsprechend auch für endliche Produkte von Maßräumen.

3.4 Erwartungswerte und Dichten

Beispiel: Glücksrad mit 64 Feldern, davon 1 rotes, 8 blaue und 55 weiße.

Als Gewinn erhält man bei

- rotem Feld: EURO 8
- blauem Feld: EURO 4
- weißem Feld: EURO 0

Wie groß ist der Gewinn bei Drehen an diesem Glücksrad "im Mittel"?

Glücksrad wird N -mal gedreht, N_r bzw. N_b bzw. N_w sei die Anzahl des Auftretens eines roten bzw. blauen bzw. weißen Feldes.

Ausgezahlter mittlerer Gewinn:

$$\frac{8 \cdot N_r + 4 \cdot N_b + 0 \cdot N_w}{N} = 8 \cdot \frac{N_r}{N} + 4 \cdot \frac{N_b}{N} + 0 \cdot \frac{N_w}{N}.$$

Für N "groß" strebt das ganze gegen:

$$8 \cdot \frac{1}{64} + 4 \cdot \frac{8}{64} + 0 \cdot \frac{55}{64} = \frac{40}{64} = \frac{5}{8}$$

Verallgemeinerung: X sei reelle ZV, die die Werte $x_1, \dots, x_k$ mit den Wahrscheinlichkeiten $p_1, \dots, p_K \geq 0$ annimmt, weiter $p_1 + \dots + p_K = 1$.

Mittlerer Wert von X ist

$$EX = \sum_{k=1}^N x_k \cdot p_k = \sum_{k=1}^N x_k \cdot P[X = x_k].$$

Hierbei werden die auftretenden Werte mit den Wahrscheinlichkeiten gewichtet gemittelt.

Beispiel: (Kontinuierliche Verallgemeinerung des obigen Beispiels)

X sei normalverteilte Zufallsvariable, d. h. P_X hat die Verteilungsfunktion

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(t-a)^2}{2\sigma^2}} dt$$

↑
sogenannte Dichtefunktion,

Wert entspricht den Wk.

$p_1, \dots, p_K$ von oben.

Dann setzen wird

$$\begin{aligned} EX &= \int_{-\infty}^{\infty} t \cdot \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(t-a)^2}{2\sigma^2}} dt \\ &= \int_{-\infty}^{\infty} (t-a) \cdot \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(t-a)^2}{2\sigma^2}} dt + a \cdot \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(t-a)^2}{2\sigma^2}} dt \\ &= 0 + a \cdot 1 = a, \end{aligned}$$

wobei beim dritten Gleichheitszeichen verwendet wurde, dass der erste Integrand punktsymmetrisch bzgl. $t = a$, ist, und das zweite Integral Eins ergibt (wegen $F(x) \rightarrow 1$ ($x \rightarrow \infty$)).

Allgemeiner

Gegeben seien ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ und eine reelle ZV X auf $(\Omega, \mathcal{A}, P)$ mit Verteilung P_X (auf $\mathcal{B}$) und VF $F : \mathbb{R} \rightarrow \mathbb{R}$.

Bezeichnungen:

$$X^+(\omega) := \begin{cases} X(\omega), & \text{falls } X(\omega) > 0 \\ 0 & \text{sonst} \end{cases} \quad X^-(\omega) := \begin{cases} -X(\omega), & \text{falls } X(\omega) < 0 \\ 0 & \text{sonst.} \end{cases}$$

Der Erwartungswert EX der ZV X gibt den "mittleren Wert" von X bezüglich P an, die Varianz $V(X)$ gibt die "Schwankung" der ZV X als "mittleren Wert" ihrer quadratischen Abweichung von EX an.

Definition 3.33: Sei X reelle ZV oder erweitert reellwertige ZV auf W-Raum $(\Omega, \mathcal{A}, P)$.

Dann heißt

$$EX := \int_{\Omega} X(\omega) dP(\omega) =: \int_{\Omega} X dP$$

Erwartungswert (kurz: EW) von X .

Bemerkung 3.34

Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$. Für $A \in \mathcal{A}$ gilt $P(A) = E\chi_A$, da $\chi_A = 1 \cdot \mathcal{X}_A + 0 \cdot \mathcal{X}_{A^c}$ und damit nach Schritt 0 der Definition des Integrals

$$E\chi_A = 1 \cdot P(A) + 0 \cdot P(A^c) = P(A).$$

Definition 3.35 Existiert für die reelle ZV X auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ ein endlicher Erwartungswert EX , so heißt X integrierbar (bezüglich P).

Lemma 3.36 Für eine reelle ZV $X \geq 0$ mit VF F gilt

$$EX = \int_0^{\infty} (1 - F(x)) dx .$$

Beweis:

Nach dem Satz von Fubini gilt:

$$\begin{aligned} \int_0^{\infty} (1 - F(t)) dt &= \int_0^{\infty} P[X > t] dt = \int_0^{\infty} \int_{\Omega} \mathcal{X}_{(t, \infty)}(X) dP dt \\ &= \int_{\Omega} \int_0^{\infty} \mathcal{X}_{(t, \infty)}(X) dt dP = \int_{\Omega} X dP = EX, \end{aligned}$$

wobei die zweitletzte Gleichheit aus

$$\mathcal{X}_{(t, \infty)}(X) = \begin{cases} 1 & \text{für } t \leq X \\ 0 & \text{für } t > X \end{cases}$$

folgt. □

Satz 3.37 Seien X, Y reelle integrierbare ZVn auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$.

- a) Es existiert $E(X + Y) = EX + EY$.
- b) Es existiert $E(\alpha X) = \alpha EX$ ($\alpha \in \mathbb{R}$).
- c) $X \geq Y \implies EX \geq EY$.
- d) $|EX| \leq E|X|$.
- e) $X = Y$ P-f.s. $\implies EX = EY$.

f) $X \geq 0$, $EX = 0 \implies X = 0$ P-f.s.

Beweis:

a)- e) folgt aus Satz 3.11

f) Sei $X \geq 0$ und $EX = 0$.

Setze $A = [X > 0]$ und $A_n = [X \geq \frac{1}{n}]$.

Dann gilt: $A_n \uparrow A$

und daher

$$\begin{aligned} P[X > 0] &= P(A) = \lim_{n \rightarrow \infty} P(A_n) = \lim_{n \rightarrow \infty} P[X \geq \frac{1}{n}] \\ &\leq \lim_{n \rightarrow \infty} E(n \cdot X) \quad (\text{da } \mathcal{X}_{[X \geq \frac{1}{n}]} \leq n \cdot X) \\ &= \lim_{n \rightarrow \infty} n \cdot EX = \lim_{n \rightarrow \infty} n \cdot 0 = 0. \end{aligned}$$

$\rightsquigarrow X = 0$ f.s.

□ V

Definition 3.38 Sei X ein n -dimensionaler Zufallsvektor mit Verteilung P_X auf $\mathcal{B}_n$ und VF $F : \mathbb{R}^n \rightarrow \mathbb{R}$.

a) Nimmt (eventuell nach Vernachlässigung einer P -Nullmenge in Ω) X höchstens abzählbar viele Werte an, so ist P_X eine sogenannte **diskrete W-Verteilung**.

Ist X eine reelle ZV und P_X auf $\mathbb{N}_0$ konzentriert (d.h. $P_X(\mathbb{N}_0) = 1$), so heißt die Folge (p_k) mit $p_k := P_X(\{k\}) = P[X = k]$, $k \in \mathbb{N}_0$, (wobei $\sum p_k = 1$) **Zähldichte** von X bzw. P_X .

b) Jede Funktion $f : (\mathbb{R}^n, \mathcal{B}_n) \rightarrow (\mathbb{R}, \mathcal{B})$ mit

$$f(x) \geq 0 \quad (x \in \mathbb{R}) \quad \text{und} \quad \int_{\mathbb{R}^n} f(x) dx = 1$$

heißt **Dichte**.

Gilt

$$P_X(B) = \int_B f(x) dx := \int f(x) \cdot \mathcal{X}_B(x) dx \quad (B \in \mathcal{B}_n),$$

so heißt f **Dichte von P_X bzw. X bzw. F** . In diesem Fall heißen P_X und F **totalstetig**.

Bemerkung 3.39

a) F totalstetig $\Rightarrow F$ stetig.

Begründung: oBdA $n = 1$. Sei f Dichte von P_X .

Dann gilt für $(x_n)_n$ mit $x_n \rightarrow x$ ($n \rightarrow \infty$):

$$f(t) \cdot \mathcal{X}_{(-\infty, x_n]}(t) \rightarrow f(t) \cdot \mathcal{X}_{(-\infty, x]}(t)$$

für L-f.a. $t \in \mathbb{R}$ und $f(t) \cdot \mathcal{X}_{(-\infty, x_n]}(t)$ ist betragsmäßig durch die integrierbare Majorante f beschränkt.

Nach dem Satz von der majorisierten Konvergenz gilt dann

$$\begin{aligned} \lim_{n \rightarrow \infty} F(x_n) &= \lim_{n \rightarrow \infty} P_X((-\infty, x_n]) \\ &= \lim_{n \rightarrow \infty} \int_{(-\infty, x_n]} f(t) dt \\ &= \lim_{n \rightarrow \infty} \int f(t) \cdot \mathcal{X}_{(-\infty, x_n]}(t) dt \\ &= \int f(t) \cdot \mathcal{X}_{(-\infty, x]}(t) dt \quad (\text{majorisierte Konvergenz}) \\ &= P_X((-\infty, x]) \\ &= F(x). \end{aligned}$$

□

b) Besitzt die n -dimensionale VF $F : \mathbb{R}^n \rightarrow \mathbb{R}$ eine Dichtefunktion $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$, so existiert L-f.ü. $\frac{\partial^n F}{\partial x_1 \dots \partial x_n}$, und L-f.ü. — insbesondere an den Stetigkeitsstellen von f — gilt $\frac{\partial^n F}{\partial x_1 \dots \partial x_n} = f$.

c) **Begründung für $n = 1$:**

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{n} \int_x^{x+h} f(t) dt \rightarrow f(x) \quad (h \rightarrow 0)$$

für L-f.a. x (nach dem Dichtetheorem von Lebesgue). Trivialerweise gilt die obige Grenzwertbeziehung in allen Stetigkeitspunkten von f . □

d) f Dichte von $P_X \iff \forall x \in \mathbb{R}^n : F(x) = \int_{(-\infty, x]} f(t) dt$.

Begründung: “ $\Rightarrow$ ” folgt durch Wahl von $B = (-\infty, x]$ in Def. 3.38 b)

“ $\Leftarrow$ ” Nach Voraussetzung gilt

$$P_X((-\infty, x]) = \int_{(-\infty, x]} f(t) dt$$

für alle $x \in \mathbb{R}^n$. Da $\{(-\infty, x] : x \in \mathbb{R}^n\}$ ein Erzeugersystem für $\mathcal{B}_n$ ist, folgt die Behauptung mit dem Fortsetzungssatz für Maße. □

Satz 3.39 Sei X eine reelle ZV mit Verteilung P_X .

a) Ist P_X auf $\mathbb{N}_0$ konzentriert mit Zähldichte (p_k) , dann gilt

$$EX = \sum_k k \cdot p_k.$$

b) Besitzt X eine Dichtefunktion $f : \mathbb{R} \rightarrow \mathbb{R}_+$, so existiert das Lebesgue-Integral $\int_{\mathbb{R}} xf(x) dx =: \int_{\mathbb{R}} xf(x) \lambda(dx)$ [λ L-B-Maß auf $\mathcal{B}$] genau dann, wenn EX existiert, und es gilt hierbei

$$EX = \int_{\mathbb{R}} xf(x) dx.$$

Beweis: Folgt aus Satz 3.40 unten □

Beispiel zu Satz 3.39 ZV X auf $(\Omega, \mathcal{A}, P)$.

a) X sei $b(n, p)$ -verteilt, d.h. $P[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$ ($k = 0, 1, \dots, n$) mit $n \in \mathbb{N}$, $p \in [0, 1]$. Dann gilt

$$EX = \sum_{k=0}^n k \cdot \binom{n}{k} \cdot p^k (1-p)^{n-k}.$$

Nach dem binomischen Lehrsatz gilt für die sogenannte erzeugende Funktion g von X :

$$g(s) = \sum_{k=0}^n \binom{n}{k} \cdot p^k (1-p)^{n-k} \cdot s^k = (ps + (1-p))^n.$$

Ableiten aller drei Ausdrücke liefert

$$g'(s) = \sum_{k=0}^n k \cdot \binom{n}{k} p^k (1-p)^{n-k} \cdot s^{k-1} = n \cdot (ps + (1-p))^{n-1} \cdot p,$$

woraus folgt

$$\underline{EX} = g'(1) = n \cdot (p \cdot 1 + (1-p))^{n-1} \cdot p = \underline{n \cdot p}.$$

b) X sei $\pi(\lambda)$ -verteilt, d.h. $P[X = k] = e^{-\lambda} \frac{\lambda^k}{k!}$ ($k \in \mathbb{N}_0$), mit $\lambda > 0$. Dann gilt

$$\begin{aligned} EX &= \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} \cdot e^{-\lambda} \\ &= \lambda \cdot e^{-\lambda} \cdot \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\ &= \lambda \cdot e^{-\lambda} \cdot e^{\lambda} = \lambda. \end{aligned}$$

c) X sei $N(a, \sigma^2)$ -verteilt, d.h. X habe Dichtefunktion $f : \mathbb{R} \rightarrow \mathbb{R}_+$ mit

$$f(x) := \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}} \quad x \in \mathbb{R},$$

mit $a \in \mathbb{R}$, $\sigma > 0$. Dann gilt $EX = a$ (siehe oben).

Satz 3.40 X sei eine reelle ZV mit Verteilung P_X und VF F ; $g : (\mathbb{R}, \mathcal{B}) \rightarrow (\mathbb{R}, \mathcal{B})$. $E(g \circ X)$ existiert genau dann, wenn $\int_{\mathbb{R}} g dP_X$ existiert, und es gilt hierbei nach dem Transformationsatz für Integrale

$$E(g \circ X) = \int_{\mathbb{R}} g dP_X =: \int_{\mathbb{R}} g dF =: \int_{\mathbb{R}} g(x) \underbrace{dF(x)}_{F(dx)},$$

sowie

$$E(g \circ X) = \begin{cases} \sum_k g(k)p_k & \text{unter der Voraussetzung von Satz 3.39 a)} \\ \underbrace{\int_{\mathbb{R}} g(x)f(x) dx}_{\text{Lebesgue-Integral}} & \text{unter der Voraussetzung von Satz 3.39 b)} \end{cases}$$

Insbesondere gilt (mit $g = \chi_A$, wobei $A \in \mathcal{B}$)

$$P[X \in A] = P_X(A) = \begin{cases} \sum_{k \in A \cap \mathbb{N}_0} p_k & \text{unter Vorauss. von Satz 3.39 a)} \\ \underbrace{\int_{\mathbb{R}} \chi_A(x)f(x) dx =: \int_A f(x) dx}_{\text{Lebesgue-Integral}} & \text{unter Vorauss. von Satz 3.39 b)} \end{cases}$$

Bemerkung: Satz 3.39 folgt aus Satz 3.40 mit $g(x) = x$.

Beweis von Satz 3.40

a) **Fall 0:** $g = \chi_A$ für ein $A \in \mathcal{B}_n$.

D. g.

$$\begin{aligned} E\chi_A \circ X &= E\chi_{X^{-1}(A)} \stackrel{\text{Def. Integral}}{=} P(X^{-1}(A)) \\ &\stackrel{\text{Def. } P_X}{=} P_X(A) \stackrel{\text{Def. Integral}}{=} \int \chi_A dP_X. \end{aligned}$$

Fall 1: g nicht-negativ einfach.

Folgt mit Linearität des Erwartungswertes aus Fall 0.

Fall 2: g nichtnegativ.

Wähle nichtnegative einfache g_n mit $g_n \uparrow g$.

Dann sind $g_n \circ X$ nichtnegativ einfach mit $g_n \circ X \uparrow g \circ X$, und mit der Definition des Integrals folgt:

$$Eg \circ X \stackrel{\text{Def.}}{=} \lim_{n \rightarrow \infty} Eg_n \circ X \stackrel{\text{Fall 1}}{=} \lim_{n \rightarrow \infty} \int g_n dP_X \stackrel{\text{Def.}}{=} \int g dP_X.$$

Fall 3: g messbar

Dann gilt $g = g^+ - g^-$ und $(g \circ X)^+ = g^+ \circ X$, $(g \circ X)^- = g^- \circ X$, woraus folgt

$$\begin{aligned} E(g \circ X) &\stackrel{\text{Def.}}{=} E(g \circ X)^+ - E(g \circ X)^- \\ &\stackrel{\text{s.o.}}{=} E(g^+ \circ X) - E(g^- \circ X) \\ &\stackrel{\text{Fall 2}}{=} \int g^+ dP_X - \int g^- dP_X \\ &\stackrel{\text{Def.}}{=} \int g dP_X \end{aligned}$$

b) Nach Satz 3.37 e) bzw. nach dem Satz von der monotonen Konvergenz gilt

$$\begin{aligned}
 \int g^+ dP_X &= \int g^+ \mathcal{X}_{\mathbb{N}_0} dP_X \\
 &= \int \sum_{k=0}^{\infty} g^+(k) \cdot \mathcal{X}_{\{k\}} dP_X \\
 &= \sum_{k=0}^{\infty} \int g^+(k) \cdot \mathcal{X}_{\{k\}} dP_X \\
 &= \sum_{k=0}^{\infty} g^+(k) \cdot \int \mathcal{X}_{\{k\}}(x) dP_X(x) \\
 &= \sum_{k=0}^{\infty} g^+(k) \cdot P[X = k]
 \end{aligned}$$

und analog

$$\int g^- dP_X = \sum_{k=0}^{\infty} g^-(k) \cdot P[X = k]$$

woraus folgt

$$\begin{aligned}
 \int g dP_X &= \int g^+ dP_X - \int g^- dP_X = \sum_{k=0}^{\infty} (g^+(k) - g^-(k)) \cdot P[X = k] \\
 &= \sum_{k=0}^{\infty} g(k) \cdot P[X = k].
 \end{aligned}$$

c) Für $g = \mathcal{X}_A$ gilt nach Definition der Dichte:

$$\begin{aligned}
 \int g dP_X &= \int \mathcal{X}_A dP_X = P_X(A) \\
 &= \int_A f(x) dx = \int \mathcal{X}_A(x) \cdot f(x) dx \\
 &= \int g(x) \cdot f(x) dx.
 \end{aligned}$$

Analog zu a) folgt daraus die Behauptung im allgemeinen Fall (wobei in Fall 2 der Satz von der monotonen Konvergenz angewendet wird). $\square$

Beispiel Ein Punkt P wird rein zufällig aus der Einheitskreisscheibe ausgewählt. Wie groß ist der zufällige Abstand vom Ursprung im Mittel?

Zur stochastischen Modellierung betrachten wir einen 2-dimensionalen Zufallsvektor (X_1, X_2) auf W-Raum $(\Omega, \mathcal{A}, P)$, wobei $P_{(X_1, X_2)}$ konzentriert ist auf

$$K = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}$$

und

$$P_{(X_1, X_2)}|_K = \frac{1}{\pi} \cdot m|_K \text{ mit } m = LB - \text{Ma\ss auf } \mathcal{B}_2.$$

Gesucht ist dann

$$E \left(\sqrt{X_1^2 + X_2^2} \right).$$

1. Lösungsweg: Verteilung von $Y = \sqrt{X_1^2 + X_2^2}$ ermitteln.

F sei die Verteilungsfunktion von Y .

Dann gilt für $0 \leq y \leq 1$:

$$\begin{aligned} F(y) &= P(Y \leq y) \\ &= P(X_1^2 + X_2^2 \leq y^2) \\ &= P_{(X_1, X_2)}(\underbrace{\{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq y^2\}}_{\text{Kreis um } 0 \text{ mit Radius } y}) \\ &= \frac{1}{\pi} \cdot \pi y^2 = y^2, \end{aligned}$$

woraus folgt

$$F(y) = \begin{cases} 0 & , y < 0 \\ y^2 & , 0 \leq y \leq a \\ 1 & , y > 1 \end{cases}$$

Zu F existiert Dichte f mit

$$f(y) = F'(y) = \begin{cases} 0 & , y < 0 \text{ oder } y > 1 \\ 2y & , 0 \leq y \leq 1 \end{cases}$$

(denn man sieht sofort, dass die angegebene Funktion

$$\int_{-\infty}^y f(t) dt = F(y) \quad (y \in \mathbb{R})$$

erfüllt). Damit folgt

$$\begin{aligned} E\sqrt{X_1^2 + X_2^2} &= EY \stackrel{\text{Satz 2.4b)}}{=} \int_{-\infty}^{\infty} y \cdot f(y) dy \\ &= \int_0^1 2y^2 dy = \frac{2}{3}. \end{aligned}$$

2. Lösungsweg: Anwendung von Satz 3.40

(X_1, X_2) hat Dichte

$$f(x) = \begin{cases} \frac{1}{\pi} & , x \in K \\ 0 & , \text{sonst} \end{cases}$$

Also gilt nach Satz 3.40:

$$\begin{aligned} E\sqrt{X_1^2 + X_2^2} &= \int_K \sqrt{x_1^2 + x_2^2} \cdot \frac{1}{\pi} d(x_1, x_2) \\ &\stackrel{\text{Polarkoordinaten}}{=} \int_0^{2\pi} \int_0^1 r \cdot \frac{1}{\pi} \cdot r \, dr d\varphi \\ &= \int_0^1 2r^2 dr = \frac{2}{3}. \end{aligned}$$

Bemerkung 3.41 $\int_A f(x) dx$ in Satz 3.39 ist gleich dem L-B-Maß der Ordinatenmenge $\{(x, y) \in \mathbb{R}^2 \mid x \in A, 0 \leq y \leq f(x)\}$ von $f|_A$ in $\mathbb{R}^2$.

Lemma 3.42 X sei eine reelle ZV; $0 \leq \alpha < \beta < \infty$. Dann gilt

$$E|X|^\beta < \infty \implies E|X|^\alpha < \infty.$$

Beweis:

$$|X|^\alpha \leq 1 + |X|^\beta \Rightarrow E(|X|^\alpha) \leq 1 + E(|X|^\beta) \quad \square$$

Definition 3.43 X sei eine reelle ZV; $k \in \mathbb{N}$. Im Falle der Existenz heißt EX^k das **k -te Moment** von X und $E(X - EX)^k$ das **k -te zentrale Moment** von X . Ist X integrierbar, dann heißt $V(X) := E(X - EX)^2$ die **Varianz** von X und $\sigma(X) := \sqrt{V(X)}$ die **Streuung** von X .

Satz 3.44 X sei eine reelle integrierbare ZV.

- a) $V(X) = EX^2 - (EX)^2$.
- b) $V(aX + b) = a^2V(X)$ ($a, b \in \mathbb{R}$).

Beweis:

a)

$$\begin{aligned} V(X) &= E(X^2 - 2X \cdot EX + (EX)^2) \\ &= E(X^2) - 2 \cdot EX \cdot EX + (EX)^2 \\ &= E(X^2) - (EX)^2. \end{aligned}$$

b)

$$\begin{aligned} V(a \cdot X + b) &= E[(a \cdot X + b - E(a \cdot X + b))^2] \\ &= E[(a \cdot X + b - a \cdot EX - b)^2] \\ &= E[a^2 \cdot (X - EX)^2] \\ &= a^2 \cdot V(X) \end{aligned}$$

□

Beispiel zu Satz 3.44

- a) X sei $b(n, p)$ -verteilt mit $n \in \mathbb{N}$, $p \in [0, 1]$. Dann gilt $V(x) = np(1 - p)$.

Begründung: Es gelte

$$P[X = k] = \binom{n}{k} \cdot p^k (1 - p)^{n-k} \quad (k \in \{0, \dots, n\}).$$

Setze

$$g(s) = \sum_{k=0}^n \binom{n}{k} \cdot p^k (1 - p)^{n-k} \cdot s^k = (ps + 1 - p)^n.$$

Zweimaliges Ableiten aller drei Ausdrücke liefert

$$\begin{aligned} g''(s) &= \sum_{k=0}^n k(k-1) \cdot \binom{n}{k} \cdot p^k (1-p)^{n-k} \cdot s^{k-2} \\ &= n \cdot (n-1) \cdot (ps + 1 - p)^{n-2} \cdot p^2, \end{aligned}$$

also

$$\begin{aligned} g''(1) &= \sum_{k=0}^n k \cdot (k-1) \binom{n}{k} \cdot p^k (1-p)^{n-k} \\ &= n \cdot (n-1) \cdot (p + 1 - p)^{n-2} \cdot p^2 \\ &= n \cdot (n-1) \cdot p^2. \end{aligned}$$

Damit

$$\begin{aligned} V(X) &= E(X^2) - (EX)^2 = E(X \cdot (X-1)) + EX - (EX)^2 \\ &= n \cdot (n-1) \cdot p^2 + n \cdot p - (n \cdot p)^2 \\ &= n^2 \cdot p^2 - n \cdot p^2 + n \cdot p - n^2 \cdot p^2 \\ &= n \cdot (p - p^2) = n \cdot p \cdot (1-p). \end{aligned}$$

b) X sei $\pi(\lambda)$ -verteilt mit $\lambda > 0$. Dann gilt $V(X) = \lambda$.

Begründung:

Sei X $\pi(\lambda)$ -verteilt, also

$$P[X = k] = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \quad (k \in \mathbb{N}_0).$$

Dann gilt

$$\begin{aligned} E(X \cdot (X-1)) &= \sum_{k=0}^{\infty} k \cdot (k-1) \cdot \frac{\lambda^k}{k!} \cdot e^{-\lambda} \\ &= \sum_{k=2}^{\infty} \frac{\lambda^k}{(k-2)!} \cdot e^{-\lambda} = \lambda^2 \cdot e^{-\lambda} \cdot \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} \\ &= \lambda^2 \cdot e^{-\lambda} \cdot e^{\lambda} = \lambda^2, \end{aligned}$$

woraus folgt

$$\begin{aligned} V(X) &= E(X \cdot (X-1)) + EX - (EX)^2 \\ &= \lambda^2 + \lambda - (\lambda)^2 = \lambda. \end{aligned}$$

c) X sei $N(a, \sigma^2)$ -verteilt mit $a \in \mathbb{R}$, $\sigma > 0$.

Dann gilt $E(X-a)^{2k-1} = 0$, $E(X-a)^{2k} = \sigma^{2k} \prod_{j=1}^k (2j-1)$ ($k \in \mathbb{N}$).

Satz 3.45 X sei eine reelle ZV auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$. Dann gilt für jedes $\varepsilon > 0$, $r > 0$:

$$P[|X| \geq \varepsilon] \leq \varepsilon^{-r} E|X|^r$$

(Markoffsche Ungleichung).

Für $r = 2$ erhält man die sog. **Tschebyschevsche Ungleichung**, die bei integrierbarem X auch in der Variante

$$P[|X - EX| \geq \varepsilon] \leq \varepsilon^{-2}V(X)$$

angegeben wird.

Beweis.

a)

$$\begin{aligned} P(|X| \geq \varepsilon) &= \int_{\{\omega : |X(\omega)| \geq \varepsilon\}} 1 \, dP \\ &\leq \int_{\{\omega : |X(\omega)| \geq \varepsilon\}} \frac{|X(\omega)|^r}{\varepsilon^r} \, dP \\ &\quad \text{(da Integrand auf dem Integrationsbereich } \geq 1 \text{ ist)} \\ &\leq \int_{\omega} \frac{|X(\omega)|^r}{\varepsilon^r} \, dP \\ &\quad \text{(da Integrand nichtnegativ ist)} \\ &= \varepsilon^{-r} \cdot E(|X|^r). \end{aligned}$$

b) Mit $Y = |X - EX|$ und $r = 2$ folgt aus a):

$$\begin{aligned} P(|X - EX| \geq \varepsilon) &= P(|Y| \geq \varepsilon) \\ &\leq \varepsilon^{-2} \cdot E(|Y|^2) \\ &= \frac{V(X)}{\varepsilon^2} \end{aligned}$$

□

Satz 3.46 (Transformationssatz für Dichten). [Verallgemeinerung s. Hinderer S. 148]
 Es sei Q ein W-Maß auf $\mathcal{B}_2$ und $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ eine injektive stetig-differenzierbare Abbildung. Es sei $R := Q_T$ das Bild-W-Maß von Q bzgl. T [d.h. $R(B) = Q(T^{-1}(B))$, $B \in \mathcal{B}_2$] und Δ der Betrag der Funktionaldeterminante von T . Hat R die Dichte g , so hat Q die Dichte $(g \circ T)\Delta$.

4. Unabhängigkeit

W-Raum $(\Omega, \mathcal{A}, P)$, $A, B \in \mathcal{A}$ mit $P(A) > 0$ und $P(B) > 0$.

Frage: Wann beeinflussen sich A und B gegenseitig nicht?

Naheliegende Forderung dafür:

Wenn gilt

$$P(A|B) = P(A) \text{ und } P(B|A) = P(B).$$

Äquivalent, und auch im Falle $P(A) = 0$ oder $P(B) = 0$ formulierbar:

$$P(A \cap B) = P(A) \cdot P(B).$$

Beispiel: Werfen zweier echter Würfel.

A = "Erste Augenzahl ist gerade"

B = "Zweite Augenzahl ist gerade"

C = "Summe der Augenzahl ist gerade".

Dann ist

$$P(A) = P(B) = P(C) = \frac{1}{2}$$

und

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{1}{4},$$

also gilt hier

$$P(A \cap B) = P(A) \cdot P(B), P(A \cap C) = P(A) \cdot P(C), P(B \cap C) = P(B)P(C).$$

Definition 4.1. Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$. Eine Familie $\{A_i \mid i \in I\}$ von Ereignissen $A_i \in \mathcal{A}$ heißt **unabhängig** (ausführlich: stochastisch unabhängig bzgl. P), falls für jede nichtleere endliche Menge $K \subset I$ gilt

$$P\left(\bigcap_{k \in K} A_k\right) = \prod_{k \in K} P(A_k).$$

Definition 4.2. Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$. Eine Familie $\{X_i \mid i \in I\}$ von $(\Omega_i, \mathcal{A}_i)$ - ZVn auf $(\Omega, \mathcal{A}, P)$ heißt **unabhängig**, wenn gilt: Für jede nichtleere endliche Indexmenge $\{i_1, \dots, i_n\} \subset I$ und jede Wahl von Mengen $A_{i_\nu} \in \mathcal{A}_{i_\nu}$ ($\nu = 1, \dots, n$) ist

$$P[X_{i_1} \in A_{i_1}, \dots, X_{i_n} \in A_{i_n}] = \prod_{\nu=1}^n P[X_{i_\nu} \in A_{i_\nu}].$$

Sprechweise: Unabhängigkeit der ZVn statt Unabhängigkeit der Familie der ZVn.

Lemma 4.3. Eine Familie von Ereignissen ist genau dann unabhängig, wenn jede endliche Teilfamilie unabhängig ist. Entsprechendes gilt für Zufallsvariablen.

Beweis. Klar! □

Bemerkung 4.4. zu Definition 4.1 bzw. Definition 4.2.

Paarweise Unabhängigkeit $\not\Rightarrow$ Unabhängigkeit

Begründung:

Im Beispiel oben sind $\{A, B\}$, $\{A, C\}$ und $\{B, C\}$ jeweils unabhängig, aber $\{A, B, C\}$ ist wegen

$$P(A \cap B \cap C) = P(\emptyset) = 0 \neq P(A) \cdot P(B) \cdot P(C)$$

nicht unabhängig. □

Bemerkung 4.5. Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$; Ereignisse $A_i \in \mathcal{A}$, $i \in I$, mit (reellen) ZVn χ_{A_i} (Indikatorfunktionen) auf $(\Omega, \mathcal{A}, P)$.

$$\{A_i \mid i \in I\} \text{ unabhängig} \iff \{\chi_{A_i} \mid i \in I\} \text{ unabhängig}$$

Begründung:

“ $\Rightarrow$ ” Wegen

$$[\chi_{A_i} \in B] \in \{\emptyset, A_i, A_i^c, \Omega\}$$

genügt es zu zeigen:

Sind $\{A_i \mid i \in I\}$ unabhängig und $\bar{A}_i \in \{\emptyset, A_i, A_i^c, \Omega\}$, so ist auch $\{\bar{A}_i \mid i \in I\}$ unabhängig. Da bei der Definition der Unabhängigkeit nur **endliche** Indexteilmengen in I betrachtet werden, können wir *oBdA* I endlich ansetzen. Dann genügt es zu zeigen: Aus $\{A_1, A_2, \dots, A_n\}$ unabhängig und $\bar{A}_1 \in \{\emptyset, A_1, A_1^c, \Omega\}$ folgt $\{\bar{A}_1, A_2, \dots, A_n\}$ unabhängig. Dies ist trivial für $\bar{A}_1 \in \{\emptyset, A_1, \Omega\}$.

Ist nun $\bar{A}_1 = A_1^c$ so gilt wegen $\{A_1, \dots, A_n\}$ unabhängig:

$$\begin{aligned} & P(A_1^c \cap A_{i_2} \cap \dots \cap A_{i_k}) \\ &= P(A_{i_2} \cap \dots \cap A_{i_k}) - P(A_1 \cap A_{i_2} \cap \dots \cap A_{i_k}) \\ &= P(A_{i_2}) \cdot \dots \cdot P(A_{i_k}) - P(A_1) \cdot P(A_{i_2}) \cdot \dots \cdot P(A_{i_k}) \\ &= (1 - P(A_1)) \cdot P(A_{i_2}) \cdot \dots \cdot P(A_{i_k}) \\ &= P(A_1^c) \cdot P(A_{i_2}) \cdot \dots \cdot P(A_{i_k}) \quad \rightsquigarrow \text{Beh.} \end{aligned}$$

“ $\Leftarrow$ ” Folgt aus

$$\begin{aligned} & P(A_{i_1} \cap \dots \cap A_{i_k}) = P(\chi_{A_{i_1}} \in \{1\}, \dots, \chi_{A_{i_k}} \in \{1\}) \\ &= P(\chi_{A_{i_1}} \in \{1\}) \cdot \dots \cdot P(\chi_{A_{i_k}} \in \{1\}) \\ &\quad (\text{da } \{\chi_{A_i} \mid i \in I\} \text{ unabhängig}) \\ &= P(A_{i_1}) \cdot \dots \cdot P(A_{i_k}). \end{aligned}$$

□

Satz 4.6. Gegeben seien $(\Omega_i, \mathcal{A}_i)$ - ZVn X_i auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ und Abb. $g_i : (\Omega_i, \mathcal{A}_i) \rightarrow (\Omega'_i, \mathcal{A}'_i)$, $i \in I$. Dann gilt:

$$\{X_i \mid i \in I\} \text{ unabhängig} \implies \{g_i \circ X_i \mid i \in I\} \text{ unabhängig}$$

Beweis. Wegen Lemma 4.3 oBdA $I = \{1, \dots, n\}$. Weiter kann oBdA $\{i_1, \dots, i_k\} = I$ angenommen werden.

$$\begin{aligned} & P[g_1 \circ X_1 \in A'_1, \dots, g_n \circ X_n \in A'_n] \\ &= P[X_1 \in g_1^{-1}(A'_1), \dots, X_n \in g_n^{-1}(A'_n)] \\ &= \prod_{i=1}^n P[X_i \in g_i^{-1}(A'_i)] \quad (\text{da } \{X_i : i \in I\} \text{ unabhängig}) \\ &= \prod_{i=1}^n P[g_i \circ X_i \in A'_i] \end{aligned}$$

□

Bsp. Zwei Leute füllen unbeeinflusst voneinander je einen Lottoschein aus. Dann sind das Quadrat des ersten Gewinns und die Wurzel des zweiten Gewinns unabhängig.

Satz 4.7. Gegeben seien eine unabhängige Familie $\{X_i \mid i \in I\}$ reeller ZVn auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ und eine Zerlegung $I = \sum_{j \in J} I_j$ mit $|I_j| < \infty$. Das $|I_j|$ -tupel $\{X_k \mid k \in I_j\}$ sei mit Y_j bezeichnet, $j \in J$.

a) Die Familie $\{Y_j \mid j \in J\}$ ist unabhängig.

b) Sind Abb. $g_j : (\mathbb{R}^{|I_j|}, \mathcal{B}_{|I_j|}) \rightarrow (\mathbb{R}, \mathcal{B})$, $j \in J$, gegeben, so ist (nach Satz 4.1) die Familie $\{g_j \circ Y_j \mid j \in J\}$ unabhängig.

Bemerkung 4.8. Satz 4.7 lässt sich auf $(\Omega_i, \mathcal{A}_i)$ - ZVn und Abb. $g_j : \bigotimes_{i \in I_j} (\Omega_i, \mathcal{A}_i) \rightarrow (\Omega'_j, \mathcal{A}'_j)$ verallgemeinern.

Beispiel:

Sind X, Y, Z unabhängige reelle Zven, dann sind auch

$$(X, Y), Z$$

und

$$X * Y, Z^2$$

jeweils unabhängig.

Beweis von Satz 4.7 und Bemerkung 4.8

b) folgt aus a) und Satz 4.6

In a) kann nach Lemma 4.3 $oBdA$ $|J| < \infty$ angenommen werden. In diesem Fall ist auch $|I| < \infty$.

Zwecks Vereinfachung der Schreibbarkeit wird Beweis nur im Spezialfall $I = \{1, 2, 3, 4\}$, $J = \{1, 2\}$, $I_1 = \{1, 2\}$ und $I_2 = \{3, 4\}$ durchgeführt, allgemeiner Fall folgt analog.

Also

$$Y_1 = (X_1, X_2), Y_2 = (X_3, X_4)$$

und zu zeigen ist

$$P[Y_1 \in A_1, Y_2 \in B] = P[Y_1 \in A] \cdot P[Y_2 \in B]$$

für alle $A \in \mathcal{A}_1 \otimes \mathcal{A}_2, B \in \mathcal{A}_3 \otimes \mathcal{A}_4$.

Fall 1: $A = A_1 \times A_2, B = A_3 \times A_4$ mit $A_i \in \mathcal{A}_i$ ($i \in I$).

Dann gilt

$$\begin{aligned} & P[Y_1 \in A, Y_2 \in B] \\ &= P[(X_1, X_2) \in A_1 \times A_2, (X_3, X_4) \in A_3 \times A_4] \\ &= P[X_1 \in A_1, X_2 \in A_2, X_3 \in A_3, X_4 \in A_4] \\ &= P[X_1 \in A_1] \cdot P[X_2 \in A_2] \cdot P[X_3 \in A_3] \cdot P[X_4 \in A_4] \\ &\quad (\text{da } X_1, \dots, X_4 \text{ unabhängig sind, vgl. Lemma 4.1}) \\ &= P[X_1 \in A_1, X_2 \in A_2] \cdot P[X_3 \in A_3, X_4 \in A_4] \\ &\quad (\text{da } X_1, X_2 \text{ und } X_3, X_4 \text{ jeweils unabhängig sind, vgl. Lemma 4.3}) \\ &= P[Y_1 \in A] \cdot P[Y_2 \in B] \quad \rightsquigarrow \text{Beh.} \end{aligned}$$

Fall 2: $A \in \mathcal{A}_1 \otimes \mathcal{A}_2, B = A_3 \times A_4$ mit $A_3 \in \mathcal{A}_3$ und $A_4 \in \mathcal{A}_4$

zu zeigen:

$$P[Y_1 \in A, Y_2 \in B] = P[Y_1 \in A] \cdot P[Y_2 \in B].$$

Als Funktion von A sind beide Seiten endliche W-Maße, die auf dem Erzeuger

$$\{A_1 \times A_2 : A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\}$$

von $\mathcal{A}_1 \otimes \mathcal{A}_2$ übereinstimmen (nach Fall 1). Mit dem Fortsetzungs- und Eindeutigkeitsatz für Maße folgt die Behauptung.

Fall 3: $A \in \mathcal{A}_1 \otimes \mathcal{A}_2, B \in \mathcal{A}_3 \otimes \mathcal{A}_4$

Analog zu Fall 2, indem man die nach Fall 2 auf $\{A_3 \times A_4 : A_3 \in \mathcal{A}_3, A_4 \in \mathcal{A}_4\}$ übereinstimmenden Maße eindeutig auf $\mathcal{A}_3 \otimes \mathcal{A}_4$ fortsetzt. $\square$

Satz 4.9. Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und seien $X : (\Omega, \mathcal{A}_i) \rightarrow (\Omega_i, \mathcal{A}_i)$ Zufallsvariablen ($i = 1, \dots, n$). Dann gilt:

$$X_1, \dots, X_n \text{ unabhängig} \Leftrightarrow P_{(X_1, \dots, X_n)} = P_{X_1} \otimes \dots \otimes P_{X_n},$$

d. h. die Zufallsvariablen sind genau dann unabhängig, wenn ihre gemeinsame Verteilung mit dem Produkt der Randverteilungen übereinstimmt.

Beweis. $X_1, \dots, X_n$ sind genau dann unabhängig, wenn für alle $A_i \in \mathcal{A}_i$ gilt:

$$P[X_1 \in A_1, \dots, X_n \in A_n] = P[X_1 \in A_1] \cdot \dots \cdot P[X_n \in A_n].$$

Beachtet man, dass die linke Seite gleich

$$P_{(X_1, \dots, X_n)}(A_1 \times \dots \times A_n)$$

ist, und dass die rechte Seite nach Definition des Produktmaßes mit

$$P_{X_1}(A_1) \cdot \dots \cdot P_{X_n}(A_n) = (P_{X_1} \otimes \dots \otimes P_{X_n})(A_1 \times \dots \times A_n)$$

übereinstimmt, so folgt “ $\Leftarrow$ ” unmittelbar, während sich “ $\Rightarrow$ ” aus der eindeutigen Bestimmtheit des Produktmaßes durch seine Werte auf dem $\cap$ -stabilen Erzeugersystem

$$\{A_1 \times \dots \times A_n : A_1 \in \mathcal{A}_1, \dots, A_n \in \mathcal{A}_n\}$$

der Produkt- σ -Algebra ergibt. □

Satz 4.10. Seien X_i ($i = 1, \dots, n$) reelle ZVn auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ mit Verteilungsfunktionen F_i . Der Zufallsvektor $X := (X_1, \dots, X_n)$ habe die n -dimensionale VF F .

a) $\{X_1, \dots, X_n\}$ unabhängig $\Leftrightarrow (*) \quad \forall_{(x_1, \dots, x_n) \in \mathbb{R}^n} F(x_1, \dots, x_n) = \prod_{i=1}^n F_i(x_i).$

b) Existieren zu F_i ($i = 1, \dots, n$), F Dichten f_i bzw. f , dann gilt

$$(*) \Leftrightarrow \begin{cases} f(x_1, \dots, x_n) = \prod_{i=1}^n f_i(x_i) \\ \text{für Lebesgue - fast alle } (x_1, \dots, x_n) \in \mathbb{R}^n. \\ \lambda\text{-fast alle mit } \lambda = \text{L-B-Maß auf } \mathcal{B}_n \end{cases}$$

Beweis

a) “ $\Rightarrow$ ” klar, wähle $A_i = (-\infty, x_i]$ in Definition 4.2

“ $\Leftarrow$ ” Aus $(*)$ folgt:

$$P[(X_1, \dots, X_n) \in I_1 \times \dots \times I_n] = P[X_1 \in I_1] \cdot \dots \cdot P[X_n \in I_n]$$

für alle Intervalle $I_j \in \{(-\infty, x] : x \in \mathbb{R}\}$.

Sukzessives Anwenden des Eindeutigkeitsatzes für endliche Maße liefert zuerst

$$P[(X_1, \dots, X_n) \in B_1 \times I_2 \times \dots \times I_n] = P[X_1 \in B_1] \cdot P[X_2 \in I_2] \cdot \dots \cdot P[X_n \in I_n]$$

für alle $B_1 \in \mathcal{B}, I_2, \dots, I_n \in \{(-\infty, x] : x \in \mathbb{R}\}$, und nach n -maligem Anwenden

$$P[(X_1, \dots, X_n) \in B_1 \times B_2 \times \dots \times B_n] = P[X_1 \in B_1] \cdot \dots \cdot P[X_n \in B_n]$$

für alle $B_1, \dots, B_n \in \mathcal{B}$, was zu zeigen war.

b) “ $\Rightarrow$ ” Wendet man auf (*) die Differentiation

$$\frac{\partial^n}{\partial x_1 \dots \partial x_n}$$

L -f.ü. an, so folgt die Behauptung aus Bemerkung 3.39 b).

“ $\Leftarrow$ ” Es gelte

$$f(x_1, \dots, x_n) = \prod_{i=1}^n f_i(x_i) \quad \text{für L-f.a. } (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Dann gilt auch

$$\begin{aligned} F(x_1, \dots, x_n) &= \int_{(-\infty, x_1] \times \dots \times (-\infty, x_n]} f(u_1, \dots, u_n) d(u_1, \dots, u_n) \\ &\stackrel{\text{Vor.}}{=} \int_{(-\infty, x_1] \times \dots \times (-\infty, x_n]} f_1(u_1) \cdot \dots \cdot f_n(u_n) d(u_1, \dots, u_n) \\ &\stackrel{\text{Fubini}}{=} \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f_1(u_1) \cdot \dots \cdot f_n(u_n) du_n \dots du_1 \\ &= \int_{-\infty}^{x_1} f_1(u_1) du_1 \cdot \dots \cdot \int_{-\infty}^{x_n} f_n(u_n) du_n \\ &= F_1(x_1) \cdot \dots \cdot F_n(x_n). \end{aligned}$$

□

Satz 4.11. Seien $X_1, \dots, X_n$ unabhängige reelle ZVn auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ mit endlichen Erwartungswerten. Dann existiert $E(X_1 \cdot \dots \cdot X_n)$ und es gilt:

$$E(X_1 \cdot \dots \cdot X_n) = EX_1 \cdot \dots \cdot EX_n.$$

— Entsprechendes gilt für komplexe ZVn $X_k : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}^2, \mathcal{B}_2)$, wobei

$$EX_k := E \operatorname{Re} X_k + iE \operatorname{Im} X_k \quad (k = 1, \dots, n)$$

Beweis: oBdA $n = 2$ und X_i reellwertig.

1. Weg

Im Falle $X_1 = \mathcal{X}_{A_1}, X_2 = \mathcal{X}_{A_2}$ mit A_1, A_2 unabhängig gilt

$$\begin{aligned} E(X_1 \cdot X_2) &= E(\mathcal{X}_{A_1 \cap A_2}) = P(A_1 \cap A_2) \\ &= P(A_1) \cdot P(A_2) \quad (\text{da } A_1, A_2 \text{ unabhängig}) \\ &= E\mathcal{X}_{A_1} \cdot E\mathcal{X}_{A_2} = EX_1 \cdot EX_2. \end{aligned}$$

Dann gemäß der schrittweisen Definition des Maßintegrals den allgemeinen Fall darauf zurückführen.

2. Weg: Eleganter, mit Maßtheorie.

ZVen $X_i : (\Omega, \mathcal{A}, P) \rightarrow (\Omega_i, \mathcal{A}_i) \quad (i \in \{1, 2\})$.

Produkt Messraum $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$, wobei $(\mathcal{A}_1 \otimes \mathcal{A}_2)$ die kleinste σ -Algebra in $\Omega_1 \times \Omega_2$ ist, die alle Mengen der Form $A_1 \times A_2$ ($A_i \in \mathcal{A}_i$ ($i \in \{1, 2\}$)) enthält.

Man kann zeigen:

Es existiert ein eindeutiges Maß Q auf $\mathcal{A}_1 \otimes \mathcal{A}_2$ mit

$$Q(A_1 \times A_2) = P_{X_1}(A_1) \cdot P_{X_2}(A_2)$$

für alle $A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2$.

Bezeichnung: Q ist das Produkt-Maß zu P_{X_1} und $P_{X_2}, Q = P_{X_1} \otimes P_{X_2}$.

Nun gilt:

(X_1, X_2) unabhängig

$$\begin{aligned} &\stackrel{Def.}{\Leftrightarrow} \forall A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2 : P[X_1 \in A_1, X_2 \in A_2] = P[X_1 \in A_1] \cdot P[X_2 \in A_2] \\ &\Leftrightarrow \forall A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2 : P_{(X_1, X_2)}(A_1 \times A_2) = (P_{X_1} \otimes P_{X_2})(A_1 \times A_2) \\ &\Leftrightarrow P_{(X_1, X_2)} = P_{X_1} \otimes P_{X_2} \end{aligned}$$

wobei die letzte Äquivalenz aus der Eindeutigkeit des Produkt-Maßes folgt.

In anderen Worten:

Im Falle der Unabhängigkeit von X_1 und X_2 ist die gemeinsame Verteilung $P_{(X_1, X_2)}$ also gerade gleich dem Produkt der Randverteilungen P_{X_1} und P_{X_2} .

Weiter gilt:

Satz von Fubini (für Produkt W-Maße, analog für Produkte von σ -endlichen Maßen).

- a) $f : (\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2) \rightarrow (\mathbb{R}_+, \mathcal{B}_+)$, Q_1 bzw. Q_2 seien W-Maße auf $\mathcal{A}_1$ bzw. $\mathcal{A}_2$.
 $Q = Q_1 \otimes Q_2$ sei das Produkt W-Maß von Q_1 und Q_2 .

Dann gilt:

$$\begin{aligned}
\int_{\Omega_1 \times \Omega_2} f \, dQ &= \int_{\Omega_1} \underbrace{\left[\int_{\Omega_2} f(\omega_1, \omega_2) dQ_2(\omega_2) \right]}_{\text{als Funktion von } \omega_1 \text{ messbar (!)}} dQ_1(\omega_1) \\
&= \int_{\Omega_2} \underbrace{\left[\int_{\Omega_1} f(\omega_1, \omega_2) dQ_1(\omega_1) \right]}_{\text{als Funktion von } \omega_2 \text{ messbar (!)}} dQ_2(\omega_2)
\end{aligned}$$

b) Für

$$f : (\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2) \rightarrow (\mathbb{R}, B)$$

mit

$$\int_{\Omega_1 \times \dots \times \Omega_2} |f| \, dQ \stackrel{\text{gemäß a)}}{=} \int_{\Omega_1} \int_{\Omega_2} |f(\omega_1, \omega_2)| dQ_2(\omega_2) dQ_1(\omega_1) < \infty$$

gilt

$$\begin{aligned}
\int_{\Omega_1 \times \Omega_2} f \, dQ &= \int_{\Omega_1} \left[\int_{\Omega_2} f(\omega_1, \omega_2) dQ_2(\omega_2) \right] dQ_1(\omega_1) \\
&= \int_{\Omega_2} \left[\int_{\Omega_1} f(\omega_1, \omega_2) dQ_1(\omega_1) \right] dQ_2(\omega_2),
\end{aligned}$$

wobei jeweils [...] als Funktion von ω_1 bzw. ω_2 integrierbar ist.

Damit Nachweis von $E(X_1 \cdot X_2) = EX_1 \cdot EX_2$:

Gemäß a) gilt für die messbare Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x_1, x_2) = x_1 \cdot x_2 :$$

$$\begin{aligned}
\int |f| d(P_{X_1} \otimes P_{X_2}) &= \int_{\mathbb{R}} \int_{\mathbb{R}} |x_1| \cdot |x_2| dP_{X_1} dP_{X_2} \\
&= \int_{\mathbb{R}} |x_2| \cdot \underbrace{\int_{\mathbb{R}} |x_1| \cdot dP_{X_1}}_{=E|X_1| < \infty} dP_{X_2} \\
&= E|X_1| \cdot \int |x_2| dP_{X_2} \\
&= E|X_1| \cdot E|X_2| < \infty
\end{aligned}$$

nach Voraussetzung.

Damit erhalten wir

$$\begin{aligned}
 E(X_1 \cdot X_2) &= \int_{\mathbb{R} \times \mathbb{R}} x_1 \cdot x_2 dP_{(X_1, X_2)}(x_1, x_2) \\
 &= \int_{\mathbb{R} \times \mathbb{R}} x_1 \cdot x_2 d(P_{X_1} \otimes P_{X_2})(x_1, x_2) \\
 &\quad \text{(da } X_1, X_2 \text{ unabhängig)} \\
 &= \int_{\mathbb{R}} \left[\int_{\mathbb{R}} x_1 \cdot x_2 dP_{X_1}(x_1) \right] dP_{X_2}(x_2) \\
 &\quad \text{(nach Teil b) des Satzes von Fubini,} \\
 &\quad \text{der wegen } \int |x_1 \cdot x_2| d(P_{X_1} \otimes P_{X_2}) < \infty \\
 &\quad \text{anwendbar ist)} \\
 &= \int_{\mathbb{R}} x_2 \cdot \underbrace{\int_{\mathbb{R}} x_1 dP_{X_1}(x_1)}_{=EX_1 \in \mathbb{R}} dP_{X_2}(x_2) \\
 &= EX_1 \cdot \int_{\mathbb{R}} x_2 dP_{X_2}(x_2) \\
 &= EX_1 \cdot EX_2
 \end{aligned}$$

↔ Satz 4.11. $\square$

Satz 4.12. Seien $X_1, \dots, X_n$ unabhängige reelle ZVn mit endlichen Erwartungswerten. Dann gilt

$$V\left(\sum_{j=1}^n X_j\right) = \sum_{j=1}^n V(X_j).$$

Beweis:

$$\begin{aligned}
 V(X_1 + \dots + X_n) &= E\left[\left(\sum_{i=1}^n (X_i - EX_i)\right)^2\right] \\
 &= \sum_{i=1}^n E[(X_i - EX_i)^2] + 2 \sum_{1 \leq i < j \leq n} E[(X_i - EX_i) \cdot (X_j - EX_j)]
 \end{aligned}$$

Für $i \neq j$ sind $X_i - EX_i$ und $X_j - EX_j$ unabhängig, also gilt nach Satz 4.11:

$$\begin{aligned}
 &E[(X_i - EX_i) \cdot (X_j - EX_j)] \\
 &= E[X_i - EX_i] \cdot E[X_j - EX_j] \\
 &= (EX_i - EX_i) \cdot (EX_j - EX_j) \\
 &= 0 \cdot 0 = 0.
 \end{aligned}$$

Damit

$$V(X_1 + \dots + X_n) = \sum_{i=1}^n E[(X_i - EX_i)^2] + 0 = \sum_{i=1}^n V(X_i).$$

$\square$

Bemerkung 4.13. In Satz 4.12 genügt statt der Unabhängigkeit von $\{X_1, \dots, X_n\}$ die sog. paarweise Unkorreliertheit, d.h.

$$\forall_{j \neq k} E(X_j - EX_j)(X_k - EX_k) = 0.$$

Vorbemerkung zu Definition 4.14:

Sind X, Y unabhängige reelle ZVen, so ist

$$P_{(X,Y)} \stackrel{s.o.}{=} P_X \otimes P_Y$$

bereits durch P_X und P_Y eindeutig festgelegt, und daher ist auch

$$P_{X+Y}$$

bereits eindeutig durch P_X und P_Y festgelegt, da

$$P_{X+Y}(A) = P_{(X,Y)}(\{(x, y) : x + y \in A\})$$

für $A \in \mathcal{B}$.

Definition 4.14. Sind X, Y zwei unabhängige reelle ZVn auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ mit Verteilungen P_X, P_Y , so wird die Verteilung $P_{X+Y} =: P_X * P_Y$ der ZV $X + Y$ als **Faltung** von P_X und P_Y bezeichnet.

Bemerkung 4.15. Die Faltungsoperation ist kommutativ und assoziativ. (Klar, da Addition in $\mathbb{R}$ kommutativ und assoziativ ist).

Satz 4.16. Seien X, Y zwei unabhängige reelle ZVn.

- a) Besitzen X und Y Dichten f bzw. g , so besitzt $X + Y$ eine [als Faltung von f und g bezeichnete] Dichte h mit

$$h(t) = \underbrace{\int_{\mathbb{R}} f(t-y)g(y) dy}_{\text{Lebesgue-Integrale}} = \underbrace{\int_{\mathbb{R}} g(t-x)f(x) dx}_{\text{Lebesgue-Integrale}}, \quad t \in \mathbb{R}.$$

- b) Besitzen X und Y Zähldichten $(p_k)_{k \in \mathbb{N}_0}$ bzw. $(q_k)_{k \in \mathbb{N}_0}$, so besitzt $X + Y$ eine [als Faltung von (p_k) und (q_k) bezeichnete] Zähldichte $(r_k)_{k \in \mathbb{N}_0}$ mit

$$r_k = \sum_{j=0}^k p_{k-j}q_j = \sum_{i=0}^k q_{k-i}p_i, \quad k \in \mathbb{N}_0.$$

Beweis (von a), Teil b) geht analog):

Wir zeigen: $P[X + Y \leq z] = \int_{-\infty}^z h(t) dt \quad (z \in \mathbb{R})$

(vgl. Bemerkung 3.39 d))

Dazu:

$$\begin{aligned}
 P[X + Y \leq z] &= P_{(X,Y)}(\underbrace{\{(x, y) \in \mathbb{R}^2 : x + y \leq z\}}_{=:A}) \\
 &= \int \mathcal{X}_A(x, y) dP_{(X,Y)}(x, y) \\
 &= \int \mathcal{X}_A(x, y) (P_X \otimes P_Y)(dx, dy) \\
 &\quad (\text{da } X, Y \text{ unabhängig}) \\
 &\stackrel{\text{Fubini}}{=} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathcal{X}_{(-\infty, z-y]}(x) P_X(dx) P_Y(dy) \\
 &= \int_{\mathbb{R}} \int_{\mathbb{R}} \mathcal{X}_{(-\infty, z-y]}(x) \cdot f(x) \cdot g(y) dx dy \\
 &\quad (\text{nach Satz 3.40}) \\
 &= \int_{\mathbb{R}} \int_{\mathbb{R}} \mathcal{X}_{(-\infty, z-y]}(t-y) \cdot f(t-y) \cdot g(y) dt dy \\
 &\quad (\text{mit Substitution } x = t-y) \\
 &\stackrel{\text{Fubini}}{=} \int_{-\infty}^z [\int_{\mathbb{R}} f(t-y) \cdot g(y) dy] dt
 \end{aligned}$$

□

Bemerkung 4.17. Gegeben seien Messräume $(\Omega_i, \mathcal{A}_i)$ und W-Maße Q_i auf $\mathcal{A}_i$, $i \in I$. Dann existiert ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ und $(\Omega_i, \mathcal{A}_i)$ -ZVn X_i auf $(\Omega, \mathcal{A}, P)$, $i \in I$, mit Unabhängigkeit von $\{X_i \mid i \in I\}$ und $P_{X_i} = Q_i$.

Begründung:

Wähle

$$\begin{aligned}
 \Omega &:= \prod_{i \in I} \Omega_i = \{(\omega_i)_{i \in I} : \omega_i \in \Omega_i \ (i \in I)\} \\
 \mathcal{A} &:= \bigotimes_{i \in I} \mathcal{A}_i = \text{kleinste } \sigma\text{-Algebra, die alle Mengen der Form} \\
 &\quad \prod_{i \in I} A_i, \text{ mit } A_i \in \mathcal{A}_i \ (i \in I) \text{ und} \\
 &\quad A_i \neq \Omega_i \text{ nur für endlich viele } i \in I, \text{ enthält.}
 \end{aligned}$$

$$P := \bigotimes_{i \in I} Q_i = \text{Produkt-W-Maß der } Q_i.$$

Hierbei ist $\bigotimes_{i \in I} Q_i$ dasjenige W-Maß $P : \mathcal{A} \rightarrow \mathbb{R}$ mit

$$P\left(\prod_{i \in I} A_i\right) = \prod_{\substack{i \in I \\ A_i \neq \Omega_i}} P(A_i)$$

für alle $\prod_{i \in I} A_i$ aus der Definition von $\bigotimes_{i \in I} \mathcal{A}_i$.

Ein solches W-Maß P existiert eindeutig nach dem Satz von **Andersen und Jessen**.
Wir definieren nun

$$X_i : \Omega \rightarrow \Omega_i$$

als Projektion von Ω auf Ω_i .

Dann gilt:

- $P_{X_i} = Q_i \quad (i \in I)$
- $\{X_i : i \in I\}$ ist unabhängig

Anwendung

In Beweisen kann man annehmen, dass unabhängige ZVen auf dem gleichen W-Raum definiert sind, und dann Aussagen über ihre Verteilung herleiten.

Satz 4.18. (2. Lemma von Borel und Cantelli) Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$. Die Familie $\{A_n; n \in \mathbb{N}\}$ von Ereignissen ($\in \mathcal{A}$) sei unabhängig. Dann gilt:

$$\sum_{n=1}^{\infty} P(A_n) = \infty \quad (\Leftrightarrow) \quad \underbrace{P(\overline{\lim} A_n) = 1}_{\text{P-f.s. tritt } A_n \text{ unendlich oft auf}}$$

Bemerkung:

Ohne Unabhängigkeitsvoraussetzung gilt nach dem 1. Lemma von Borel und Cantelli:

$$\sum_{n=1}^{\infty} P(A_n) < \infty \Rightarrow P(\overline{\lim} A_n) = 0.$$

Also folgt “ $\Leftarrow$ ” in Satz 4.18 aus dem 1. Lemma von Borel und Cantelli.

Beispiel:

Unabhängige Folge von Versuchen mit Misserfolgswahrscheinlichkeit $\frac{1}{n}$ im n -ten Versuch.
Wegen

$$\sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

gilt dann nach Satz 4.18: P-f.s. treten unendlich viele Misserfolge auf.

Beweis von Satz 4.18

“ $\Leftarrow$ ”: Folgt aus dem 1. Lemma von Borel und Cantelli.

“ $\Rightarrow$ ” Wir zeigen: $0 = P((\overline{\lim} A_n)^c) = P\left(\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^c\right)$.

Dazu:

$$\begin{aligned}
P\left(\bigcap_{k=n}^{\infty} A_k^c\right) &= \lim_{N \rightarrow \infty} P\left(\bigcap_{k=n}^N A_k^c\right) \\
&\quad (\text{da W-Ma\ss stetig von oben ist}) \\
&= \lim_{N \rightarrow \infty} \prod_{k=n}^N P(A_k^c) \\
&\quad (\text{da } \{A_n : n\} \text{ unabh\u00e4ngig}) \\
&= \lim_{N \rightarrow \infty} \prod_{k=n}^N (1 - P(A_k)) \\
&\leq \lim_{N \rightarrow \infty} \prod_{k=n}^N e^{-P(A_k)} \\
&\quad (\text{wegen } 1 + x \leq e^x \text{ f\u00fcr } x \in \mathbb{R}) \\
&= \lim_{N \rightarrow \infty} e^{-\sum_{k=n}^N P(A_k)} \\
&= 0 \\
&\quad (\text{da } \sum_{k=n}^{\infty} P(A_k) = \infty)
\end{aligned}$$

Daraus folgt:

$$\begin{aligned}
P\left(\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^c\right) &\leq \sum_{n=1}^{\infty} P\left(\bigcap_{k=n}^{\infty} A_k^c\right) \\
&\quad (\text{da W-Ma\ss e } \sigma \text{ - subadditiv sind}) \\
&= \sum_{n=1}^{\infty} 0 = 0.
\end{aligned}$$

□

Bemerkung 4.19:

a) In Satz 4.18 darf die Unabh\u00e4ngigkeitsvoraussetzung nicht weggelassen werden.

Beispiel dazu:

W\u00e4hlt man $A_n = A$ ($n \in \mathbb{N}$) f\u00fcr ein A mit $P(A) = \frac{1}{2}$, so gilt

$$\sum_{n=1}^{\infty} P(A_n) = \infty \text{ und } P(\overline{\lim} A_n) = P(A) = \frac{1}{2} \neq 1.$$

b) Sind $(A_n)_n$ unabh\u00e4ngig, so gilt nach dem 1. und 2. Lemma von Borell und Cantelli

$$P(\overline{\lim} A_n) \in \{0, 1\}.$$

Im Folgenden verallgemeinern wir diese Aussage.

Definition 4.20. Messraum $(\Omega, \mathcal{A})$. σ -Algebren $\mathcal{A}_n \subset \mathcal{A}$ ($n \in \mathbb{N}$).

$\mathcal{T}_n := \mathcal{F}\left(\bigcup_{k=n}^{\infty} \mathcal{A}_k\right)$... die von $\mathcal{A}_n, \mathcal{A}_{n+1}, \dots$ erzeugte σ -Algebra.

$\mathcal{T}_{\infty} := \bigcap_{n=1}^{\infty} \mathcal{T}_n$ hei\ss t die σ -Algebra der **terminalen Ereignisse** (tail events) der Folge $(\mathcal{A}_n)$.

Beispiele:

- a) $X_n : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B})$
 $\mathcal{A}_n = \mathcal{F}(X_n) = X_n^{-1}(\mathcal{B})$

Hierbei ist $X_n^{-1}(\mathcal{B})$ die kleinste σ -Algebra $\hat{\mathcal{A}}$, für die $X_n \hat{\mathcal{A}} - \mathcal{B}$ -messbar ist.

1. $\left[\sum_n X_n \text{ konvergent} \right] \in \mathcal{T}_\infty$
 (da

$$\left[\sum_n X_n \text{ konvergent} \right] = \left[\sum_{n=k}^\infty X_n \text{ konvergent} \right] \in \mathcal{T}_k \text{ für alle } k \in \mathbb{N}.$$

2. $[(X_n)_n \text{ konvergent}] \in \mathcal{T}_\infty$
 (da

$$[(X_n)_n \text{ konvergent}] = [(X_n)_{n \geq k} \text{ konvergent}] \in \mathcal{T}_k$$

für alle $k \in \mathbb{N}$).

- b) Messraum $(\Omega, \mathcal{A})$, $A_n \in \mathcal{A}$ ($n \in \mathbb{N}$).

Die σ -Algebra $\mathcal{A}_n$ sei definiert durch

$$\mathcal{A}_n = \{\emptyset, \Omega, A_n, A_n^c\}.$$

1. $\overline{\lim} A_n \in \mathcal{T}_\infty$,

denn:

$$\overline{\lim} A_n = \bigcap_{n=1}^\infty \bigcup_{k=n}^\infty A_k = \bigcap_{n=N}^\infty \bigcup_{k=n}^\infty A_k \in \mathcal{T}_N$$

für jedes $N \in \mathbb{N}$, da

$$\bigcup_{k=n}^\infty A_k \in \mathcal{T}_N \text{ für } n \geq N$$

und $\mathcal{T}_n$ σ -Algebra ist.

- 2.

$$\begin{aligned} \underline{\lim} A_n &= \bigcup_{n=1}^\infty \bigcap_{k=n}^\infty A_k = \left(\bigcap_{n=1}^\infty \bigcup_{k=n}^\infty A_k^c \right)^c \\ &= (\overline{\lim} A_n^c)^c \in \mathcal{T}_\infty, \end{aligned}$$

wobei der letzte Schritt analog zu 1. folgt.

Erweiterung von Definition 4.1 / 4.2:

W-Raum $(\Omega, \mathcal{A}, P)$, Mengensysteme $\mathcal{C}_i \subset \mathcal{A}$ ($i \in I \neq \emptyset$).

$\{\mathcal{C}_i : i \in I\}$ heißt **unabhängig**, falls für jede Wahl von Ereignissen $C_i \in \mathcal{C}_i$ ($i \in I$) die Familie $\{C_i : i \in I\}$ unabhängig ist.

Beispiel:

a) Sind $(A_n)_{n \in \mathbb{N}}$ unabhängig, so ist auch

$$(\{\emptyset, \Omega, A_n, A_n^c\} : n \in \mathbb{N})$$

unabhängig (vgl. Begründung von Bemerkung 4.5).

b) Sind die reellen ZVen $(X_n)_{n \in \mathbb{N}}$ unabhängig, so ist auch

$$(X_n^{-1}(\mathcal{B}) : n \in \mathbb{N})$$

unabhängig

(da

$$\begin{aligned} & P[X_{n_1}^{-1}(B_1) \cap \dots \cap X_{n_k}^{-1}(B_k)] \\ &= P[X_{n_1} \in B_1, \dots, X_{n_k} \in B_k] \\ & \text{Unabhängigkeit der } x_n \underline{=} P[X_{n_1} \in B_1] \cdot \dots \cdot P[X_{n_k} \in B_k] \\ &= P(X_{n_1}^{-1}(B_1)) \cdot \dots \cdot P(X_{n_k}^{-1}(B_k)) \end{aligned}$$

ist).

Satz 4.21 (Null–Eins–Gesetz von Kolmogorov) Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$.

Sei $(\mathcal{A}_n)$ eine **unabhängige** Folge von σ -Algebren $\mathcal{A}_n \subset \mathcal{A}$.

Dann gilt für jedes terminale Ereignis A von $(\mathcal{A}_n)$ entweder $P(A) = 0$ oder $P(A) = 1$.

Beispiel: Unabhängige reelle ZVen $(X_n)_n$ auf W-Raum $(\Omega, \mathcal{A}, P)$. Dann gilt

$$\begin{aligned} P[\sum_n X_n \text{ konvergent}] &\in \{0, 1\} \\ P[(X_n)_n \text{ konvergent}] &\in \{0, 1\} \end{aligned}$$

Vorbemerkung zum Beweis von Satz 4.21

Sei $\Omega \neq \emptyset$ und $\mathcal{C} \subseteq \mathcal{P}(\Omega)$.

$\mathcal{C}$ heißt **Dynkin-System**, falls gilt:

1. $\Omega \in \mathcal{C}$
2. $A, B \in \mathcal{C}$ mit $A \subseteq B \Rightarrow B \setminus A \in \mathcal{C}$
3. $A_n \in \mathcal{C}$ ($n \in \mathbb{N}$) paarweise disjunkt $\Rightarrow \sum_{n=1}^{\infty} A_n \in \mathcal{C}$.

$\mathcal{D}(\mathcal{C})$ sei das kleinste Dynkin-System, das $\mathcal{C}$ enthält.

Lemma 4.22.

W-Raum $(\Omega, \mathcal{A}, P), \mathcal{C} \subseteq \mathcal{P}(\Omega), \mathcal{C}_1, \mathcal{C}_2 \subseteq \mathcal{A}$. Dann gilt:

a) Aus $\mathcal{C}$ $\cap$ -stabil (d.h. $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$) folgt

$$\mathcal{D}(\mathcal{C}) = \mathcal{F}(\mathcal{C}).$$

b) Sind $\{\mathcal{C}_1, \mathcal{C}_2\}$ unabhängig, so sind auch

$\{\mathcal{D}(\mathcal{C}_1), \mathcal{C}_2\}$ sowie $\{\mathcal{D}(\mathcal{C}_1), \mathcal{D}(\mathcal{C}_2)\}$ unabhängig .

Beweis:

a) Siehe Satz 1.18 a).

b) Seien $\mathcal{C}_1$ und $\mathcal{C}_2$ unabhängig.

Wir zeigen zuerst: $\{\mathcal{D}(\mathcal{C}_1), \mathcal{C}_2\}$ ist unabhängig.

Äquivalent dazu ist: $\{\{A\}, \mathcal{C}_2\}$ ist unabhängig für alle $A \in \mathcal{D}(\mathcal{C}_1)$.

Um dies nachzuweisen, betrachten wir

$$\mathcal{C}^* = \{A \in \mathcal{A} : \{\{A\}, \mathcal{C}_2\} \text{ unabhängig}\}$$

$\mathcal{C}^*$ ist Dynkin-System, denn es gilt:

- $\Omega \in \mathcal{C}^*$ (klar)
- Sind $A, B \in \mathcal{C}^*$ mit $A \subset B$, so ist auch $B \setminus A \in \mathcal{C}^*$.

Denn für beliebiges $E \in \mathcal{C}_2$ gilt:

$$\begin{aligned} P((B \setminus A) \cap E) &= P((B \cap E) \setminus (A \cap E)) \\ &= P(B \cap E) - P(A \cap E) \\ &\quad (\text{da } A \cap E \subset B \cap E) \\ &= P(B) \cdot P(E) - P(A) \cdot P(E) \\ &\quad (\text{da } A, B \in \mathcal{C}^*) \\ &= (P(B) - P(A)) \cdot P(E) \\ &= P(B \setminus A) \cdot P(E) \\ &\quad (\text{da } A \subset B) \end{aligned}$$

- Sind $A_n \in \mathcal{C}^*$ ($n \in \mathbb{N}$) paarweise disjunkt, so ist auch $\sum_{n=1}^{\infty} A_n \in \mathcal{C}^*$.

Denn für beliebiges $E \in \mathcal{C}_2$ gilt:

$$\begin{aligned} &P\left(\left(\sum_{n \in \mathbb{N}} A_n\right) \cap E\right) \\ &= P\left(\sum_{n \in \mathbb{N}} A_n \cap E\right) \\ &= \sum_{n=1}^{\infty} P(A_n \cap E) \\ &= \sum_{n=1}^{\infty} P(A_n) \cdot P(E) \\ &\quad (\text{da } A_n \in \mathcal{C}^*) \\ &= \left(\sum_{n=1}^{\infty} P(A_n)\right) \cdot P(E) \\ &= P\left(\sum_{n \in \mathbb{N}} A_n\right) \cdot P(E). \end{aligned}$$

Nach Voraussetzung gilt auch $\mathcal{C}_1 \subset \mathcal{C}^*$, und daraus folgt (nach Definition des erzeugten Dynkin-Systems):

$$\mathcal{D}(\mathcal{C}_1) \subset \mathcal{D}(\mathcal{C}^*) = \mathcal{C}^*.$$

Also ist

$$\{\mathcal{D}(\mathcal{C}_1), \mathcal{C}_2\}$$

unabhängig, und nochmaliges Anwenden des gleichen Beweisschrittes ergibt

$$\{\mathcal{D}(\mathcal{C}_1), \mathcal{D}(\mathcal{C}_2)\}$$

unabhängig. □

Beweis von Satz 4.21

Die Behauptung folgt aus

(*) $\{\mathcal{T}_\infty, \mathcal{T}_\infty\}$ unabhängig,

denn dann gilt für jedes $A \in \mathcal{T}_\infty$: $\{A, A\}$ unabhängig, also

$$P(A) = P(A \cap A) \quad = \quad P(A) \cdot P(A),$$

Unabhängigkeit

woraus folgt

$$0 = P(A) - P(A) \cdot P(A) = P(A) \cdot (1 - P(A)),$$

was

$$P(A) = 0 \text{ oder } P(A) = 1$$

impliziert.

Zum Nachweisen von (*) setzen wir

$$\begin{aligned} \mathcal{C}_n &= \mathcal{F}(\mathcal{A}_1, \dots, \mathcal{A}_n) = \mathcal{F}\left(\bigcup_{k=1}^n \mathcal{A}_k\right) \\ \mathcal{T}_{n+1} &= \mathcal{F}(\mathcal{A}_{n+1}, \mathcal{A}_{n+2}, \dots) = \mathcal{F}\left(\bigcup_{k=n+1}^{\infty} \mathcal{A}_k\right) \end{aligned}$$

und beachten

$$\mathcal{C}_n = \mathcal{F}(\mathcal{C})$$

mit

$$\mathcal{C} = \left\{ \bigcap_{i=1}^n A_i : A_i \in \mathcal{A}_i \right\},$$

sowie

$$\mathcal{T}_{n+1} = \mathcal{F}(\mathcal{T})$$

mit

$$\begin{aligned} \mathcal{T} &:= \left\{ \bigcap_{i \in I} A_i : A_i \in \mathcal{A}_i \text{ (} i \in I \text{) und } I \text{ endliche} \right. \\ &\quad \left. \text{nichtleere Teilmenge von } \{n+1, n+2, \dots\} \right\} \end{aligned}$$

Da $(\mathcal{A}_n)_{n \in \mathbb{N}}$ unabhängig sind, sind auch $\{\mathcal{C}, \mathcal{T}\}$ unabhängig.

Mit Lemma 4.22 folgt

$$\{\mathcal{D}(\mathcal{C}), \mathcal{D}(\mathcal{T})\}$$

unabhängig, und da $\mathcal{C}$ und $\mathcal{T}$ beide $\cap$ -stabil sind auch

$$\{\mathcal{F}(\mathcal{C}), \mathcal{F}(\mathcal{T})\}$$

unabhängig, also

$$\{\mathcal{C}_n, \mathcal{T}_{n+1}\}$$

unabhängig.

Mit $\mathcal{T}_\infty = \bigcap_{k=1}^{\infty} \mathcal{T}_k \subseteq \mathcal{T}_{n+1}$ folgt daraus aber

$$\{\mathcal{C}_n, \mathcal{T}_\infty\}$$

unabhängig für alle $n \in \mathbb{N}$.

Dies impliziert

$$\left\{ \bigcup_{n \in \mathbb{N}} \mathcal{C}_n, \mathcal{T}_\infty \right\}$$

unabhängig. Erneute Anwendung von Lemma 4.22 liefert

$$\left\{ \mathcal{D} \left(\bigcup_{n \in \mathbb{N}} \mathcal{C}_n \right), \mathcal{T}_\infty \right\}$$

unabhängig.

Wegen $\mathcal{C}_1 \subseteq \mathcal{C}_2 \subseteq \mathcal{C}_3 \subseteq \dots$ und $\mathcal{C}_n$ σ -Algebra ist $\bigcup_{n \in \mathbb{N}} \mathcal{C}_n$ $\cap$ -stabil, also folgt

$$\begin{aligned} \mathcal{D} \left(\bigcup_{n \in \mathbb{N}} \mathcal{C}_n \right) &= \mathcal{F} \left(\bigcup_{n \in \mathbb{N}} \mathcal{C}_n \right) = \mathcal{F} \left(\bigcup_{k=1}^{\infty} \mathcal{A}_k \right) \\ &= \mathcal{T}_1 \supseteq \mathcal{T}_\infty, \end{aligned}$$

was

$$\{\mathcal{T}_\infty, \mathcal{T}_\infty\}$$

unabhängig impliziert. □

5. Erzeugende und charakteristische Funktionen; Faltungen

Die erzeugenden und charakteristischen Funktionen sind ein methodisches Hilfsmittel zur Untersuchung von W-Maßen.

5.1 Erzeugende Funktionen

Definition 5.1. Es sei $\mu : \mathcal{B} \rightarrow \mathbb{R}$ ein auf $\mathbb{N}_0$ konzentriertes W-Maß mit Zähldichte $(b_k)_{k \in \mathbb{N}_0}$ bzw. X eine reelle ZV auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$, deren Verteilung P_X auf $\mathbb{N}_0$ konzentriert ist, mit Zähldichte $(b_k)_{k \in \mathbb{N}_0}$. Dann heißt die auf $[0, 1]$ (oder auch $\{s \in \mathbb{C} \mid |s| \leq 1\}$) definierte Funktion g mit

$$g(s) := \sum_{k=0}^{\infty} b_k s^k = \begin{cases} \sum \mu(\{k\}) s^k \\ \text{bzw.} \\ \sum P[X = k] s^k = \sum P_X(\{k\}) s^k = E s^X \end{cases}$$

die **erzeugende Funktion** von μ bzw. X .

Bemerkung 5.2. In Definition 5.1 gilt $|g(s)| \leq \sum_{k=0}^{\infty} b_k |s|^k \leq \sum_{k=0}^{\infty} b_k = 1, |s| \leq 1$.

Bemerkung 5.3.

- a) Die Binomialverteilung $b(n, p)$ mit $n \in \mathbb{N}, p \in [0, 1]$ hat die erzeugende Funktion g mit

$$g(s) = \sum_{k=0}^n \binom{n}{k} \cdot p^k (1-p)^{n-k} \cdot s^k = (1-p + p \cdot s)^n.$$

- b) Die Poissonverteilung $\pi(\lambda)$ mit $\lambda \in (0, \infty)$ hat die erzeugende Funktion g mit

$$g(s) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot s^k = e^{-\lambda + \lambda \cdot s} \quad \text{für } s \in \mathbb{R}$$

- c) Ist X geometrisch verteilt mit Parameter $p \in [0, 1]$ (d. h. gilt $P[X = k] = p \cdot (1-p)^k$ ($k \in \mathbb{N}_0$)), so gilt

$$g(s) = \sum_{k=0}^{\infty} p \cdot (1-p)^k \cdot s^k = p \cdot \sum_{k=0}^{\infty} ((1-p) \cdot s)^k = \frac{p}{1 - (1-p) \cdot s}$$

für $s \in [0, 1]$.

d) Als **negative Binomialverteilung** oder **Pascal-Verteilung** $Nb(r, p)$ mit Parametern $r \in \mathbb{N}$, $p \in (0, 1)$ wird ein W-Maß auf $\mathcal{B}$ (oder auch $\overline{\mathcal{B}}$) bezeichnet, das auf $\mathbb{N}_0$ konzentriert ist und — mit $q := 1 - p$ — die Zähldichte

$$k \rightarrow Nb(r, p; k) := \binom{r+k-1}{k} p^r q^k, \quad k \in \mathbb{N}_0$$

besitzt. Speziell ist $Nb(1, p)$ — mit Zähldichte

$$k \rightarrow p(1-p)^k, \quad k \in \mathbb{N}_0$$

— die **geometrische Verteilung** mit Parameter $p \in (0, 1)$. Ist $(X_n)_{n \in \mathbb{N}}$ eine unabhängige Folge von $b(1, p)$ -verteilten ZVn, so ist die erweitert-reelle ZV $X := \inf \{n \in \mathbb{N} \mid \sum_{k=1}^n X_k = r\} - r$ mit $\inf \emptyset := \infty$ $Nb(r, p)$ -verteilt (Beweis siehe Beispiel unten). X gibt die Anzahl der Misserfolge bis zum r -ten Erfolg bei der zu (X_n) gehörigen Folge von Bernoulli-Versuchen an. $EX = \frac{rq}{p}$, $V(X) = \frac{rq}{p^2}$ (siehe unten).

$Nb(r, p)$ hat die erzeugende Funktion g mit $g(s) = p^r(1 - qs)^{-r}$ (siehe unten).

Satz 5.4 (Eindeutigkeitsatz für erzeugende Funktionen).

Besitzen zwei auf $\mathcal{B}$ definierte, aber auf $\mathbb{N}_0$ konzentrierte W-Maße bzw. Verteilungen dieselbe erzeugende Funktion, so stimmen sie überein.

Beweis:

Folgt aus Identitätssatz für Potenzreihen:

$$\sum_{k=0}^{\infty} a_k \cdot x^k = \sum_{k=0}^{\infty} b_k \cdot x^k (x \in [0, 1]) \Rightarrow a_k = b_k \quad (k \in \mathbb{N}_0).$$

□

Satz 5.5. Sei X eine reelle ZV, deren Verteilung auf $\mathbb{N}_0$ konzentriert ist, mit erzeugender Funktion g .

a) g ist in $\{s \in \mathbb{C} \mid |s| < 1\}$ unendlich oft differenzierbar; für $j \in \mathbb{N}$ gilt

$$g^{(j)}(1-) := \lim_{0 \leq s \rightarrow 1-0} g^{(j)}(s) = E[X(X-1)\dots(X-j+1)] \quad (\leq \infty),$$

insbesondere $g'(1-) = EX$.

b) $EX < \infty \implies V(X) = g''(1-) + g'(1-) - (g'(1-))^2$.

Beweis:

Die auf $\{s \in \mathbb{C} : |s| \leq 1\}$ definierte Potenzreihe

$$g(s) = \sum_{k=0}^{\infty} P[X = k] \cdot s^k$$

ist im Inneren $\{s \in \mathbb{C} : |s| < 1\}$ des Definitionsbereichs unendlich oft differenzierbar, und es gilt:

$$g^{(j)}(s) = \sum_{k=j}^{\infty} P[X = k] \cdot k \cdot (k-1) \cdot \dots \cdot (k-j+1) \cdot s^{k-j}.$$

Grenzübergang $s \rightarrow 1-$ ergibt

$$\begin{aligned} \lim_{s \rightarrow 1-} g^{(j)} &= \lim_{s \rightarrow 1-} \sum_{k=j}^{\infty} P[X = k] \cdot k \cdot (k-1) \cdot \dots \cdot (k-j+1) \cdot s^{k-j} \\ &= \sum_{k=j}^{\infty} P[X = k] \cdot k \cdot (k-1) \cdot \dots \cdot (k-j+1) \cdot 1^{k-j} \end{aligned}$$

(Limes und Reihe sind nach dem Satz von der monotonen Konvergenz vertauschbar. Hierbei fasst man die Reihe als Integral bzgl. dem abzählenden Maß auf)

$$\begin{aligned} &= E[X \cdot (X-1) \cdot \dots \cdot (X-j+1)] \\ &\text{(nach Satz 2.5).} \end{aligned}$$

b)

$$\begin{aligned} \text{rechte Seite} &\stackrel{a)}{=} E[X \cdot (X-1)] + EX - (EX)^2 \\ &= E[X^2] - (EX)^2 \\ &= V(X). \end{aligned}$$

□

Satz 5.6. Seien $X_1, \dots, X_n$ unabhängige reelle ZVn, deren Verteilungen jeweils auf $\mathbb{N}_0$ konzentriert sind, mit erzeugenden Funktionen $g_1, \dots, g_n$. Für die erzeugende Funktion g der Summe $X_1 + \dots + X_n$ gilt dann $g = \prod_{j=1}^n g_j$.

Beweis:

$$\begin{aligned} g(s) &= E s^{X_1 + \dots + X_n} \\ &= E \left(\prod_{j=1}^n s^{X_j} \right) \\ &= \prod_{j=1}^n E(s^{X_j}) \\ &\quad \text{(da mit } X_1, \dots, X_n \text{ auch } s^{X_1}, \dots, s^{X_n} \text{ unabhängig sind, und Satz 4.11 gilt)} \\ &= \prod_{j=1}^n g_j(s) \end{aligned}$$

□

Bemerkung:

Zufallsvariablen $X_1, X_2, \dots$ heißen identisch verteilt, falls gilt:

$$P_{X_1} = P_{X_2} = \dots$$

(also falls die Verteilungen der ZVen übereinstimmen).

Beispiel: Die ZVen $X_1, X_2, \dots$ seien unabhängig und identisch verteilt mit

$$P[X_1 = 1] = p, \quad P[X_1 = 0] = 1 - p$$

für ein $p \in (0, 1)$ fest.

Hierbei ist $[X_k = 1]$ das Ereignis, dass im k -ten Bernoulli-Versuch ein Erfolg eintritt.

Die ZV Z_1 (mit Werten in $\mathbb{R}$) gebe die Anzahl der Misserfolge bis zum 1. Erfolg an.

Dann gilt

$$\begin{aligned} P[Z_1 = k] &= P[X_1 = 0, X_2 = 0, \dots, X_k = 0, X_{k+1} = 1] \\ &\stackrel{\text{Unabh.}}{=} P[X_1 = 0] \cdot P[X_2 = 0] \cdot \dots \cdot P[X_k = 0] \cdot P[X_{k+1} = 1] \\ &= (1 - p)^k \cdot p. \end{aligned}$$

Wegen

$$P[Z_1 \in \mathbb{N}_0] = \sum_{k=0}^{\infty} (1 - p)^k \cdot p = 1$$

gilt dann $Z_1 < \infty$ f.s.

Nach dem 1. Erfolg tritt die stochastische Situation von neuem auf. Die ZV Z_2 gebe die Anzahl der Misserfolge nach dem 1. Erfolg bis zum 2. Erfolg an. Entsprechend seien $Z_3, Z_4, \dots$ definiert.

Beh. Die ZVen $Z_1, Z_2, \dots$ sind unabhängig und identisch verteilt.

Begründung:

$$\begin{aligned} P[Z_2 = k] &= \sum_{j=0}^{\infty} P[Z_1 = j, Z_2 = k] \\ &= \sum_{j=0}^{\infty} (1 - p)^j \cdot p (1 - p)^k \cdot p \\ &\quad \text{(folgt analog zur Berechnung von } P[Z_1 = k] \text{ oben)} \\ &= (1 - p)^k \cdot p \cdot \sum_{j=0}^{\infty} (1 - p)^j \cdot p \\ &= (1 - p)^k \cdot p \cdot 1 \end{aligned}$$

Also sind Z_1, Z_2 identisch verteilt.

Weiter gilt

$$P[Z_1 = j, Z_2 = k] = (1 - p)^j \cdot p \cdot (1 - p)^k \cdot p = P[Z_1 = j] \cdot P[Z_2 = k]$$

für alle $j, k \in \mathbb{N}_0$, woraus die Unabhängigkeit von Z_1 und Z_2 folgt.

(denn

$$\begin{aligned}
 P[Z_1 \in A, Z_2 \in B] &= \sum_{\substack{j \in A \cap \mathbb{N}_0 \\ k \in B \cap \mathbb{N}_0}} P[Z_1 = j, Z_2 = k] \\
 &\stackrel{s.o.}{=} \sum_{\substack{j \in A \cap \mathbb{N}_0 \\ k \in B \cap \mathbb{N}_0}} P[Z_1 = j] \cdot P[Z_2 = k] \\
 &= \sum_{j \in A \cap \mathbb{N}_0} P[Z_1 = j] \cdot \sum_{k \in B \cap \mathbb{N}_0} P[Z_2 = k] \\
 &= P[Z_1 \in A] \cdot P[Z_2 \in B]
 \end{aligned}$$

Analog sieht man, dass $Z_1, \dots, Z_n$ unabhängig und identisch verteilt sind. □

Z_1 hat die erzeugende Funktion

$$h(s) = \sum_{k=0}^{\infty} (1-p)^k \cdot p \cdot s^k = \frac{p}{1 - (1-p) \cdot s}$$

Es gilt

$$h'(s) = \frac{p(1-p)}{(1 - (1-p) \cdot s)^2}, \quad h''(s) = \frac{2p(1-p)^2}{(1 - (1-p) \cdot s)^3},$$

woraus mit Satz 5.5 folgt

$$EZ_1 = h'(1) = \frac{p(1-p)}{(1 - (1-p))^2} = \frac{1-p}{p}$$

und

$$\begin{aligned}
 V(Z_1) &= h''(1) + h'(1) - (h'(1))^2 \\
 &= \frac{2p(1-p)^2}{p^3} + \frac{1-p}{p} - \frac{(1-p)^2}{p^2} \\
 &= \frac{(1-p)^2}{p^2} + \frac{1-p}{p} = \frac{1-p}{p} \left(\frac{1-p}{p} + 1 \right) \\
 &= \frac{1-p}{p} \cdot \frac{1}{p} = \frac{1-p}{p^2}.
 \end{aligned}$$

P_{Z_1} ist die sogenannte **geometrische Verteilung** mit Parameter $p \in (0, 1)$.

Für $r \in \mathbb{N}$ betrachten wir die Anzahl der Misserfolge bis zum r -ten Versuch, gegeben durch

$$X = Z_1 + \dots + Z_r.$$

Da $Z_1, \dots, Z_r$ unabhängig und identisch verteilt sind, hat nach Satz 5.6 X die erzeugende Funktion

$$g(s) = (h(s))^r = \frac{p^r}{(1 - (1-p) \cdot s)^r}$$

Die Funktion

$$f(x) = \frac{1}{(1-x)^r} = (1-x)^{-r}$$

hat wegen

$$\begin{aligned}
 f^{(k)}(x) &= (-r) \cdot (-r-1) \cdot \dots \cdot (-r-k+1) \cdot (1-x)^{-r-k} \cdot (-1)^k \\
 &= (r+k-1) \cdot (r+k-2) \cdot \dots \cdot r \cdot (1-x)^{-r-k}
 \end{aligned}$$

die Potenzreihendarstellung

$$\frac{1}{(1-x)^r} = \sum_{k=0}^{\infty} \binom{r+k-1}{k} \cdot x^k.$$

Mit dieser erhält man für g :

$$g(s) = p^r \sum_{k=0}^{\infty} \binom{r+k-1}{k} (1-p)^k \cdot s^k,$$

also gilt (nach dem Eindeutigkeitssatz für erzeugende Funktionen)

$$\underline{\underline{P[X = k] = \binom{r+k-1}{k} \cdot p^r \cdot (1-p)^k \quad (k \in \mathbb{N}_0).}}$$

P_X ist eine sogenannte **negative Binomialverteilung** (oder auch Pascalverteilung) mit Bezeichnung $Nb(r, p)$.

Die Zähldichte von $Nb(r, p)$ ist

$$k \mapsto \binom{r+k-1}{k} \cdot p^r (1-p)^k$$

und wegen $X = Z_1 + \dots + Z_r$ mit $Z_1, \dots, Z_r$ unabhängig identisch verteilt (nach obiger Herleitung) gilt

$$\begin{aligned} EX &= r \cdot EZ_1 = r \cdot \frac{1-p}{p} \\ V(X) &= r \cdot V(Z_1) = r \cdot \frac{1-p}{p^2}. \end{aligned}$$

5.2 Charakteristische Funktionen

Definition 5.7. Es sei μ ein W-Maß auf $\mathcal{B}$ bzw. X eine reelle ZV mit VF F . Dann heißt die auf $-\infty < u < \infty$ definierte (i.a. komplexwertige) Funktion φ mit

$$\varphi(u) := \int_{\mathbb{R}} e^{iux} \mu(dx)$$

bzw.

$$\varphi(u) := Ee^{iuX} = \int_{\mathbb{R}} e^{iux} P_X(dx)$$

die **charakteristische** Funktion von μ bzw. X .

Bemerkung 5.8. In Definition 5.7 gilt

- a) $\varphi(0) = 1$,
- b) $|\varphi(u)| \leq 1, u \in \mathbb{R}$,
- c) φ gleichmäßig stetig in $\mathbb{R}$,
- d) $\varphi(-u) = \overline{\varphi(u)}, u \in \mathbb{R}$.

Begründung: a), b), d) ... trivial.

c)

$$\begin{aligned} |\varphi(u+h) - \varphi(u)| &= \left| \int_{\mathbb{R}} e^{i(u+h)x} - e^{iux} P_X(dx) \right| \\ &= \left| \int_{\mathbb{R}} e^{iux} \cdot (e^{ihx} - 1) P_X(dx) \right| \\ &\leq \int_{\mathbb{R}} |e^{ihx} - 1| P_X(dx) =: f(h) \end{aligned}$$

da

$$|e^{iux}| = 1.$$

Da $f(h)$ nicht von u abhängt, folgt die Behauptung aus

$$f(h) \rightarrow 0 \quad \text{für } h \rightarrow 0.$$

Dies wiederum gilt nach dem Satz von der majorisierten Konvergenz, da

- $|e^{ihx} - 1| \rightarrow |e^0 - 1| = 0$ ($h \rightarrow 0$) für alle $x \in \mathbb{R}$
- $|e^{ihx} - 1| \leq |e^{ihx}| + 1 = 2$ und 2 ist integrierbar.

Also gilt

$$f(h) \rightarrow \int_{\mathbb{R}} |e^{i \cdot 0 \cdot x} - 1| P_X(dx) = \int_{\mathbb{R}} 0 P_X(dx) = 0$$

für $h \rightarrow 0$, w.z.z.w. □

Bemerkung 5.9. Die Normalverteilung $N(a, \sigma^2)$ mit $a \in \mathbb{R}$, $\sigma^2 \in (0, \infty)$ hat die charakteristische Funktion φ mit

$$\varphi(u) = e^{iau} e^{-\frac{\sigma^2 u^2}{2}}.$$

Insbesondere hat $N(0, 1)$ die charakteristische Funktion φ mit $\varphi(u) = e^{-\frac{u^2}{2}}$.

Beweis: Ist $X \mathcal{N}(0, 1)$ -verteilt, so ist nach Übung $Y = \sigma \cdot X + a$ $\mathcal{N}(a, \sigma^2)$ -verteilt.

Es gilt:

$$\begin{aligned} \varphi_Y(u) &= E e^{iu(\sigma \cdot X + a)} = e^{iua} \cdot E e^{i(\sigma \cdot u) \cdot X} \\ &= e^{i \cdot u \cdot a} \cdot \varphi_X(\sigma \cdot u), \end{aligned}$$

also genügt es, die charakteristische Funktion zu $\mathcal{N}(0, 1)$ zu bestimmen.

Hierzu:

$$\begin{aligned} \varphi_X(u) &= \int_{\mathbb{R}} e^{iux} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \cdot \int_{\mathbb{R}} e^{-\frac{(x-iu)^2}{2}} dx \cdot e^{-u^2/2} \end{aligned}$$

Zu zeigen:

$$\int_{\mathbb{R}} e^{-\frac{(x-iu)^2}{2}} dx = \sqrt{2\pi},$$

Dazu betrachten wir für $u \in \mathbb{R}$ und $t > 0$:

$$\int_{-t}^t e^{-\frac{(x-i \cdot u)^2}{2}} dx = \int_{\mathcal{C}_1} e^{-z^2/2} dz$$

wobei das komplexe Kurvenintegral rechts über den Streckenzug $\mathcal{C}_1$ von $-t - i \cdot u$ bis $t - i \cdot u$ berechnet wird.

Nach Funktionentheorie gilt

$$\sum_{i=1}^4 \int_{\mathcal{C}_i} e^{-z^2/2} dz = 0,$$

wobei $\mathcal{C}_2, \mathcal{C}_3$ bzw. $\mathcal{C}_4$ die Streckenzüge von $t - i \cdot u$ bis t , bzw. t bis $-t$, bzw. $-t$ bis $-t - i \cdot u$ sind.

Auf $\mathcal{C}_2$ bzw $\mathcal{C}_4$ ist der Integrand betragsmäßig beschränkt durch

$$\begin{aligned} & \sup_{0 \leq |v| \leq |u|} \left| e^{-\frac{(\pm t - i \cdot v)^2}{2}} \right| \\ &= \sup_{0 \leq |v| \leq |u|} e^{-t^2/2} \cdot \underbrace{|e^{\pm i \cdot t \cdot v}|}_{=1} \cdot e^{v^2/2} \\ &\leq e^{-t^2/2} \cdot e^{u^2/2}. \end{aligned}$$

Damit gilt für $i \in \{2, 4\}$:

$$\left| \int_{\mathcal{C}_i} e^{-z^2/2} dz \right| \leq e^{-t^2/2} \cdot e^{u^2/2} \cdot |u| \rightarrow 0 \text{ für } t \rightarrow \infty.$$

Also gilt:

$$\lim_{t \rightarrow \infty} \int_{\mathcal{C}_1} e^{-z^2/2} dz = \lim_{t \rightarrow \infty} - \int_{\mathcal{C}_3} e^{-z^2/2} dz = \int_{-\infty}^{\infty} e^{-\frac{s^2}{2}} ds = \sqrt{2\pi}$$

□

Bemerkung 5.10 Sei $X \exp(\lambda)$ -verteilt, d.h. X habe die Dichte

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda x} & , \quad x \geq 0 \\ 0 & , \quad x < 0 \end{cases}$$

mit $\lambda > 0$.

X hat die charakteristische Funktion

$$\varphi(u) = \int_0^{\infty} e^{iux} \cdot \lambda \cdot e^{-\lambda x} dx = \frac{\lambda}{iu - \lambda} \cdot e^{(iu-\lambda) \cdot x} \Big|_{x=0}^{\infty} = \frac{\lambda}{\lambda - i \cdot u}.$$

Die Exponentialverteilung ist **gedächtnislos** in dem Sinne, dass gilt:

$$(*) \quad P[X > t + s | X > s] = P[X > t] \quad \text{für alle } s, t > 0,$$

denn

$$\begin{aligned} \text{linke Seite} &= \frac{P[X > t + s]}{P[X > s]} = \frac{\int_{t+s}^{\infty} \lambda \cdot e^{-\lambda x} dx}{\int_s^{\infty} \lambda \cdot e^{-\lambda x} dx} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda \cdot s}} \\ &= e^{-\lambda t} = \text{rechte Seite} . \end{aligned}$$

Ist umgekehrt X eine $\mathbb{R}$ -wertige ZV mit $P[X > 0] > 0$, die die Eigenschaft $(*)$ hat, so ist $X \exp(\lambda)$ -verteilt.

Begründung:

Setze

$$S(x) = 1 - F(x) = 1 - P[X \leq x] = P[X > x].$$

Aus $(*)$ folgt dann

$$S(t + s) = S(t) \cdot S(s) \quad \text{für alle } s, t > 0,$$

woraus wiederum folgt

- $S(k \cdot t) = (S(t))^k$ für alle $k \in \mathbb{N}_0, t \geq 0$
- $S\left(\frac{1}{m}\right) = S(1)^{1/m}$ für alle $m \in \mathbb{N}$ (folgt aus oben mit $t = 1/m, k = m$)
- $S\left(\frac{k}{m}\right) = (S(1))^{\frac{k}{m}}$ für alle $k, m \in \mathbb{N}_0$ (folgt aus den obigen zwei Beziehungen)

Der Fall $S(1) = 1$ kann nicht eintreten, da dies $S(k) = 1$ für alle $k \in \mathbb{N}$ implizieren würde, im Widerspruch zu

$$S(k) = P[X > k] \rightarrow 0 \quad (k \rightarrow \infty).$$

Ebenso kann der Fall $S(1) = 0$ nicht eintreten, da daraus $S\left(\frac{1}{m}\right) = 0$ für alle $m \in \mathbb{N}$ folgen würde, im Widerspruch zu

$$0 < P[X > 0] = \lim_{m \rightarrow \infty} P[X > \frac{1}{m}] = \lim_{m \rightarrow \infty} S\left(\frac{1}{m}\right).$$

Also gilt $S(1) \in (0, 1)$.

Mit $\lambda = -\ln(S(1)) \in (0, \infty)$ folgt

$$P[X > \frac{k}{m}] = S\left(\frac{k}{m}\right) = (S(1))^{\frac{k}{m}} = e^{-\lambda \cdot \frac{k}{m}}$$

für alle $k \in \mathbb{N}_0, m \in \mathbb{N}$.

Damit gilt für $x \in \mathbb{Q}_+$

$$P[X > x] = e^{-\lambda x}$$

bzw.

$$F(x) = P[X \leq x] = 1 - e^{-\lambda x}.$$

Da F als Verteilungsfunktion rechtsseitig stetig ist, folgt daraus

$$F(x) = 1 - e^{-\lambda x} \text{ für alle } x \in \mathbb{R}_+ \quad \rightsquigarrow \text{ Beh.}$$

□

Aufgrund der in (*) formulierten **Gedächtnislosigkeit** wird die Exponentialverteilung häufig bei der Modellierung von Wartesystemen eingesetzt.

Es gilt:

$$\begin{aligned} EX &= \int_0^{\infty} x \cdot \lambda \cdot e^{-\lambda x} dx = \dots = \frac{1}{\lambda} \\ V(X) &= E(X^2) - (EX)^2 = \dots = \frac{1}{\lambda^2}. \end{aligned}$$

Bemerkung 5.11. Als **Gamma-Verteilung** $\Gamma_{\lambda, \nu}$ mit Parametern $\lambda, \nu \in (0, \infty)$ wird ein W-Maß auf $\mathcal{B}$ oder $\overline{\mathcal{B}}$ bezeichnet, das eine Dichtefunktion f mit

$$f(x) = \begin{cases} \frac{\lambda^\nu}{\Gamma(\nu)} x^{\nu-1} e^{-\lambda x} & , x > 0 \\ 0 & , x \leq 0 \end{cases}$$

besitzt. Hierbei $\Gamma_{\lambda, 1} = \exp(\lambda)$. $\Gamma_{\frac{1}{2}, \frac{n}{2}}$ wird als **Chi-Quadrat-Verteilung** χ_n^2 mit $n \in \mathbb{N}$ Freiheitsgraden bezeichnet.

Für eine ZV X mit $P_X = \Gamma_{\lambda, \nu}$ gilt $EX = \frac{\nu}{\lambda}$, $V(X) = \frac{\nu}{\lambda^2}$ (Beweis siehe unten).

$\Gamma_{\lambda, \nu}$ hat die charakteristische Funktion φ mit $\varphi(u) = \left(\frac{\lambda}{\lambda - iu}\right)^\nu$ (Beweis siehe unten).

Satz 5.12. (Eindeutigkeitssatz für charakteristische Funktionen)

Besitzen zwei auf $\mathcal{B}$ definierte W-Maße bzw. Verteilungen dieselbe charakteristische Funktion, so stimmen sie überein. — Ist μ ein W-Maß auf $\mathcal{B}$ bzw. P_X die Verteilung einer reellen ZV X mit zugehöriger VF F und zugehöriger charakteristischer Funktion φ und sind a, b mit $a < b$ Stetigkeitsstellen von F , so gilt die Umkehrformel

$$\begin{aligned} \underbrace{F(b) - F(a)}_{= \mu((a, b]) \text{ bzw. } P_X((a, b])} &= \lim_{U \rightarrow \infty} \frac{1}{2\pi} \int_{-U}^U \frac{e^{-iua} - e^{-iub}}{iu} \varphi(u) du. \end{aligned}$$

Beweis: Da F überall bis auf höchstens abzählbar viele Stellen stetig ist, und als rechtsseitig stetige Funktion durch seine Werte auf einer dichten Teilmenge von $\mathbb{R}$ eindeutig bestimmt ist, genügt es, die Umkehrformel zu zeigen.

Dazu beachten wir für Stetigkeitsstellen a und b von F :

$$\begin{aligned}
& \frac{1}{2\pi} \int_{-U}^U \frac{e^{-iua} - e^{-iub}}{iu} \cdot \varphi(u) du \\
&= \frac{1}{2\pi} \int_{-U}^U \frac{e^{-iua} - e^{-iub}}{i \cdot u} \int_{\mathbb{R}} e^{iux} P_X(dx) du \\
&= \int_{\mathbb{R}} \frac{1}{2\pi} \int_{-U}^U \frac{e^{iu(x-a)} - e^{iu(x-b)}}{iu} du P_X(dx) \\
&\quad (\text{nach dem Satz von Fubini}) \\
&= \int_{\mathbb{R}} \frac{1}{2\pi} \int_{-U}^U \left(\underbrace{\frac{\cos(u(x-a)) - \cos(u \cdot (x-b))}{iu}}_{\substack{\text{ungerade als} \\ \text{Funktion von } u}} + \underbrace{\frac{\sin(u \cdot (x-a)) - \sin(u \cdot (x-b))}{u}}_{\substack{\text{gerade als} \\ \text{Funktion von } u}} \right) du P_X \\
&= \int_{\mathbb{R}} \frac{1}{\pi} \left[\underbrace{\int_0^U \frac{\sin(u \cdot (x-a))}{u} du}_{\substack{t=\underline{u \cdot (x-a)} \\ \int_0^{U \cdot (x-a)} \frac{\sin t}{t} dt}} - \underbrace{\int_0^U \frac{\sin(u \cdot (x-b))}{u} du}_{\substack{t=\underline{u \cdot (x-b)} \\ \int_0^{U \cdot (x-b)} \frac{\sin t}{t} dt}} \right] P_X(dx) \\
&= \int_{\mathbb{R}} \frac{1}{\pi} \int_{U \cdot (x-b)}^{U \cdot (x-a)} \frac{\sin t}{t} dt P_X(dx).
\end{aligned}$$

Nach der Dirichletschen Formel gilt

$$\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin v}{v} dv = 2 \cdot \frac{1}{\pi} \int_0^{\infty} \frac{\sin v}{v} dv = 1.$$

Daraus folgt

$$\lim_{U \rightarrow \infty} \frac{1}{\pi} \int_{U \cdot (x-b)}^{U \cdot (x-a)} \frac{\sin t}{t} dt = \begin{cases} 1 & , \text{ falls } a < x < b \\ 0 & , \text{ falls } x < a \text{ oder } x > b \\ \frac{1}{2} & , \text{ falls } x = a \text{ oder } x = b, \end{cases}$$

wobei die Integrale links wegen der Oszillation des Sinus und

$$\lim_{t \rightarrow 0} \frac{\sin t}{t} = \lim_{t \rightarrow 0} \frac{\cos t}{1} = 1$$

beschränkt sind.

Mit dem Satz von der majorisierten Konvergenz und $P_X(\{a\}) = P_X(\{b\}) = 0$ (da a und b Stetigkeitsstellen von F sind) folgt:

$$\begin{aligned}
& \lim_{U \rightarrow \infty} \frac{1}{2\pi} \int_{-U}^U \frac{e^{-iua} - e^{-iub}}{iu} \cdot \varphi(u) du = \int_{(a,b)} 1 P_X(dx) \\
&= P_X((a, b)) = P_X((a, b]) = F(b) - F(a),
\end{aligned}$$

wobei bei der 3. Gleichheit erneut $P_X(\{b\}) = 0$ verwendet wurde. □

Satz 5.13. Gegeben sei eine VF F mit charakteristischer Funktion φ ; es sei $\int_{-\infty}^{\infty} |\varphi(u)| du < \infty$. Dann besitzt F eine Dichte f , die stetig und beschränkt ist, und es gilt die Umkehrformel

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iux} \varphi(u) du, \quad x \in \mathbb{R}$$

[vgl. die Formel $\varphi(u) = \int_{\mathbb{R}} e^{iux} f(x) dx$, $u \in \mathbb{R}$, bei existierender Dichte f].

Beweis: Siehe Übungen □

Satz 5.14. Sei X eine reelle ZV mit Verteilung P_X und charakteristischer Funktion φ . Für $j \in \mathbb{N}$ gilt:

$$E |X|^j < \infty$$

$$\implies \varphi \text{ hat eine stetige } j\text{-te Ableitung } \varphi^{(j)} \text{ mit } \varphi^{(j)}(u) = i^j \int_{\mathbb{R}} x^j e^{iux} P_X(dx), \quad u \in \mathbb{R},$$

$$\text{insbesondere } \varphi^{(j)}(0) = i^j EX^j.$$

Beweis:

Fall $j=1$

$$\frac{\varphi(u+h) - \varphi(u)}{h} = \int_{\mathbb{R}} \underbrace{e^{iux} \cdot \left(\frac{e^{ixh} - 1}{h} \right)}_{(h \rightarrow 0) e^{iux} \cdot i \cdot x} P_X(dx).$$

Dabei ist der Integrand betragsmäßig beschränkt durch

$$\begin{aligned} \left| \frac{e^{ixh} - 1}{h} \right| &= \left| \frac{\cos(hx) - 1}{h} + i \cdot \frac{\sin(hx) - 0}{h} \right| \\ &\leq \left| \frac{\cos(hx) - 1}{h} \right| + \left| \frac{\sin(hx) - 0}{h} \right| \leq |x| \cdot |-\sin \xi| + |x| \cdot |\cos \bar{\xi}| \\ &\leq 2 \cdot |x|, \end{aligned}$$

wobei die zweite Ungleichung aus dem Mittelwertsatz folgt. Die Funktion $x \mapsto 2 \cdot |x|$ ist wegen $E|X| < \infty$ bzgl P_X integrierbar. Mit dem Satz von der majorisierten Konvergenz folgt daraus:

$$\varphi'(u) = \lim_{h \rightarrow 0} \frac{\varphi(u+h) - \varphi(u)}{h} = \int_{\mathbb{R}} e^{iux} \cdot ix P_X(dx).$$

φ ist gleichmäßig stetig, da

$$\begin{aligned} |\varphi'(u+h) - \varphi'(u)| &= \left| \int_{\mathbb{R}} ix \cdot (e^{i(u+h)x} - e^{iux}) P_X(dx) \right| \\ &\leq \int_{\mathbb{R}} |x| \cdot |e^{ixh} - 1| P_X(dx) \\ &\rightarrow \int_{\mathbb{R}} 0 P_X(dx) = 0, \end{aligned}$$

wobei das Integral

$$\int_{\mathbb{R}} |x| \cdot |e^{ihx} - 1| P_X(dx)$$

unabhängig von u ist und für $h \rightarrow 0$ wegen

$$|x| \cdot |e^{ihx} - 1| \rightarrow 0 \quad (h \rightarrow 0)$$

$$|x| \cdot |e^{ihx} - 1| \leq 2|x|$$

$$\text{und} \quad \int 2|x| P_X(dx) = 2E|X| < \infty,$$

nach dem Satz von der majorisierten Konvergenz gegen Null konvergiert.

Der allgemeine Fall folgt daraus mit Induktion. □

Satz 5.15. Seien $X_1, \dots, X_n$ unabhängige reelle ZVn mit charakteristischen Funktionen $\varphi_1, \dots, \varphi_n$. Für die charakteristische Funktion der Summe $X_1 + \dots + X_n$ gilt dann

$$\varphi = \prod_{j=1}^n \varphi_j.$$

Beweis.

$$\begin{aligned} \varphi(u) &= E e^{iu(X_1 + \dots + X_n)} = E (e^{iuX_1} \cdot e^{iuX_2} \cdot \dots \cdot e^{iuX_n}) \\ &= E e^{iuX_1} \cdot E e^{iuX_2} \cdot \dots \cdot E e^{iuX_n} \\ &\quad (\text{da mit } X_1, \dots, X_n \text{ auch } e^{iuX_1}, \dots, e^{iuX_n} \\ &\quad \text{unabhängig sind und da Satz 4.11 gilt).} \\ &= \varphi_{X_1}(u) \cdot \varphi_{X_2}(u) \cdot \dots \cdot \varphi_{X_n}(u). \end{aligned}$$

□

Beispiel:

$$f(x) = \begin{cases} \frac{\lambda^\nu}{\Gamma(\nu)} \cdot x^{\nu-1} \cdot e^{-\lambda x} & , x > 0 \\ 0 & x \leq 0 \end{cases}$$

heißt **Gamma-verteilt** mit Parameter $\lambda, \nu > 0$ (kurz: $\Gamma_{\lambda, \nu}$ -verteilt).

Hierbei ist

$$\Gamma(\nu) = \int_0^\infty x^{\nu-1} \cdot e^{-x} dx$$

die sogenannte Gamma-Funktion, für die gilt:

$$\Gamma(\nu + 1) = \nu \cdot \Gamma(\nu),$$

$$\Gamma(n) = (n - 1)! \quad (n \in \mathbb{N})$$

$$\text{sowie} \quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

Spezialfälle sind:

$$\Gamma_{\lambda,1} = \exp(\lambda)$$

$\Gamma_{\frac{1}{2}, \frac{n}{2}} \dots$ sogenannte Chi-Quadrat-Verteilung χ_n^2 mit n Freiheitsgraden.

Für die charakteristische Funktion φ von $\Gamma_{\lambda,\nu}$ gilt:

$$\varphi(u) = \left(\frac{\lambda}{\lambda - i \cdot u} \right)^\nu \quad (u \in \mathbb{R}).$$

Begründung:

$$\begin{aligned} \varphi(u) &= \int_0^\infty e^{iux} \frac{\lambda^\nu}{\Gamma(\nu)} \cdot x^{\nu-1} \cdot e^{-\lambda x} dx \\ &= \frac{\lambda^\nu}{(\lambda - i \cdot u)^\nu} \frac{1}{\Gamma(\nu)} \cdot \int_0^\infty ((\lambda - i \cdot u) \cdot x)^{\nu-1} \cdot e^{-(\lambda - i \cdot u) \cdot x} \cdot (\lambda - i \cdot u) dx \\ &= \frac{\lambda^\nu}{(\lambda - i \cdot u)^\nu} \frac{1}{\Gamma(\nu)} \cdot \lim_{t \rightarrow \infty} \int_{\mathcal{C}_1} z^{\nu-1} \cdot e^{-z} dz \end{aligned}$$

wobei $\mathcal{C}_1$ der (komplexe) Streckenzug von 0 bis $(\lambda - iu) \cdot t$ ist. Ist $\mathcal{C}_2$ der (komplexe) Streckenzug von $(\lambda - iu) \cdot t$ bis λt und $\mathcal{C}_3$ der von λt bis 0, so gilt:

$$\sum_{i=1}^3 \int_{\mathcal{C}_i} z^{\nu-1} e^{-z} dz = 0,$$

und mit

$$\lim_{t \rightarrow \infty} \int_{\mathcal{C}_2} z^{\nu-1} \cdot e^{-z} dz = 0 \quad (!)$$

folgt

$$\begin{aligned} \varphi(u) &= \frac{\lambda^\nu}{(\lambda - iu)^\nu} \cdot \frac{1}{\Gamma(\nu)} \cdot \lim_{t \rightarrow \infty} \left(- \int_{\mathcal{C}_3} z^{\nu-1} e^{-z} dz \right) \\ &= \frac{\lambda^\nu}{(\lambda - iu)^\nu} \cdot \frac{1}{\Gamma(\nu)} \int_0^\infty x^{\nu-1} \cdot e^{-x} dx \\ &= \left(\frac{\lambda}{\lambda - iu} \right)^\nu. \end{aligned}$$

□

Daraus folgt

- $\varphi'(u) = \lambda^\nu (-\nu) (\lambda - iu)^{-(\nu+1)} \cdot (-i)$,
und damit gilt nach Satz 5.14

$$EX = \frac{\varphi'(0)}{i} = \frac{\lambda^\nu (-\nu) \cdot \lambda^{-(\nu+1)} \cdot (-i)}{i} = \frac{\nu}{\lambda}.$$

•

$$\begin{aligned} \varphi''(u) &= \lambda^\nu \cdot \nu i \cdot (-\nu - 1) \cdot (\lambda - iu)^{-(\nu+2)} \cdot (-i) \\ &= -\nu(\nu + 1) \lambda^\nu (\lambda - iu)^{-(\nu+2)} \end{aligned}$$

und mit Satz 5.14 folgt

$$E(X^2) = \frac{\varphi''(0)}{i^2} = \frac{\nu(\nu+1)}{\lambda^2},$$

also gilt

$$V(X) = E(X^2) - (EX)^2 = \frac{\nu^2 + \nu}{\lambda^2} - \frac{\nu^2}{\lambda^2} = \frac{\nu}{\lambda^2}.$$

Darüber hinaus gilt:

Die Summe von n unabhängigen $\exp(\lambda)$ -verteilten Zufallsvariablen ist $\Gamma_{\lambda,n}$ verteilt.

Begründung:

Sind $X_1, \dots, X_n$ unabhängig und $\exp(\lambda)$ -verteilt, so hat nach Satz 5.15 $X_1 + \dots + X_n$ die charakteristische Funktion

$$\varphi(u) = \frac{\lambda}{\lambda - iu} \cdot \frac{\lambda}{\lambda - iu} \cdot \dots \cdot \frac{\lambda}{\lambda - iu} = \frac{\lambda^n}{(\lambda - iu)^n},$$

und ist daher nach Satz 5.12 $\Gamma_{\lambda,n}$ -verteilt. □

Definition 5.16. Zu zwei W-Maßen P, Q auf $\mathcal{B}$ wird das W-Maß $P * Q$ auf $\mathcal{B}$, die sog. **Faltung** von P und Q , folgendermaßen definiert:

- a) sei $T : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ mit $T(x, y) = x + y$;

$$P * Q := (P \otimes Q)_T$$

($P \otimes Q$ W-Maß auf $\mathcal{B}_2$ mit $(P \otimes Q)(B_1 \times B_2) = P(B_1)P(B_2)$, $B_{1,2} \in \mathcal{B}$, sog. Produkt-W-Maß von P und Q)

oder — äquivalent —

- b) man wähle ein unabhängiges Paar reeller ZVn X, Y auf $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{P})$ mit $\tilde{P}_X = P$, $\tilde{P}_Y = Q$ und setze

$$P * Q := \tilde{P}_{X+Y}.$$

Beweis der Äquivalenz von a) und b):

Für X, Y wie in b) und $Z = (X, Y)$ gilt:

$$\begin{aligned} \tilde{P}_{X+Y} &= \tilde{P}_{T \circ Z} = \left(\tilde{P}_Z \right)_T = \left(\tilde{P}_{(X,Y)} \right)_T \\ &\stackrel{x,y \text{ unabhängig}}{=} \left(\tilde{P}_X \otimes \tilde{P}_Y \right)_T \\ &= (P \otimes Q)_T \end{aligned}$$

da $\tilde{P}_X = P, \tilde{P}_Y = Q$. □

Satz 5.17.

- a) Die Faltungsoperation ist kommutativ und assoziativ.
- b) Es sei (X, Y) ein unabhängiges Paar reeller ZVn. Für $X, Y, X + Y$ seien die Verteilungsfunktionen mit F, G, H , die Dichten — falls existent — mit f, g, h und die Zähldichten — falls P_X, P_Y auf $\mathbb{N}_0$ konzentriert sind — mit $(p_k), (q_k), (r_k)$ bezeichnet. Dann gilt

$$H(t) = \int_{\mathbb{R}} F(t - y) G(dy) = \int_{\mathbb{R}} G(t - x) F(dx), \quad t \in \mathbb{R}.$$

Falls f existiert, so existiert h mit

$$h(t) = \int_{\mathbb{R}} f(t - y) G(dy) \quad \text{für } (L\text{-f.a.}) \quad t \in \mathbb{R};$$

falls zusätzlich g existiert, so gilt

$$h(t) = \int_{\mathbb{R}} f(t - y)g(y) dy = \int_{\mathbb{R}} g(t - x)f(x) dx \quad \text{für } (L\text{-f.a.}) \quad t \in \mathbb{R}.$$

Falls $(p_k)_{k \in \mathbb{N}_0}, (q_k)_{k \in \mathbb{N}_0}$ existieren, so existiert $(r_k)_{k \in \mathbb{N}_0}$ mit

$$r_k = \sum_{i=0}^k p_{k-i}q_i = \sum_{i=0}^k q_{k-i}p_i \quad (k \in \mathbb{N}_0).$$

Beweis: a) klar nach Definition 5.16 b).

b)

$$H(t) = P[X + Y \leq t] = P_{(X,Y)}(B)$$

mit

$$B = \{(x, y) \in \mathbb{R}^2 : x + y \leq t\}.$$

Wegen Unabhängigkeit von (X, Y) folgt weiter

$$\begin{aligned} H(t) &= (P_X \otimes P_Y)(B) \\ &= \int_{\mathbb{R}^2} \chi_B(x, y) d(P_X \otimes P_Y)(x, y) \\ &\stackrel{\text{Fubini}}{=} \int_{\mathbb{R}} \int_{\mathbb{R}} \chi_B(x, y) dP_X(x) dP_Y(y) \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} \chi_{(-\infty, t-y]}(x) P_X(dx) P_Y(dy) \\ &\quad (\text{da } (x, y) \in B \Leftrightarrow x + y \leq t \Leftrightarrow x \leq t - y \Leftrightarrow x \in (-\infty, t - y]) \\ &= \int_{\mathbb{R}} F(t - y) P_Y(dy) \end{aligned}$$

Mit $P_{X+Y} = P_{Y+X}$ folgt auch

$$H(t) = \int_{\mathbb{R}} G(t - x) P_X(dx).$$

Existiert Dichte f zu F , so folgt

$$\begin{aligned}
 H(t) &= \int_{\mathbb{R}} \int_{-\infty}^{t-y} f(\tau) d\tau P_Y(dy) \\
 &= \int_{\mathbb{R}} \int_{-\infty}^t f(s-y) ds P_Y(dy) \\
 &\quad \text{(mit Substitution } s = \tau + y) \\
 &\stackrel{\text{Fubini}}{=} \int_{-\infty}^t \underbrace{\int_{\mathbb{R}} f(s-y) P_Y(dy)}_{\text{ist als Funktion von } s \text{ Dichte von } H} ds
 \end{aligned}$$

Falls zusätzlich Dichte g zu G bzw. P_Y existiert, so gilt für die Dichte h von H :

$$h(t) = \int_{\mathbb{R}} f(s-y) P_Y(dy) = \int_{\mathbb{R}} f(t-y) \cdot g(y) dy$$

Für Zähldichten geht der Beweis analog, vgl. Übungen □

Bemerkung 5.18.

- a) Für $n_{1,2} \in \mathbb{N}$, $p \in [0, 1]$ gilt $b(n_1, p) * b(n_2, p) = b(n_1 + n_2, p)$.
Die Summe von n unabhängigen $b(1, p)$ verteilten ZVn ist also $b(n, p)$ -verteilt.
- b) Für $\lambda_{1,2} > 0$ gilt $\pi(\lambda_1) * \pi(\lambda_2) = \pi(\lambda_1 + \lambda_2)$.
- c) Für $r_{1,2} \in \mathbb{N}$, $p \in (0, 1)$ gilt $Nb(r_1, p) * Nb(r_2, p) = Nb(r_1 + r_2, p)$.
Die Summe von r unabhängigen $Nb(1, p)$ -verteilten ZVn ist also $Nb(r, p)$ -verteilt.
- d) Für $a_{1,2} \in \mathbb{R}$, $\sigma_{1,2}^2 \in (0, \infty)$ gilt $N(a_1, \sigma_1^2) * N(a_2, \sigma_2^2) = N(a_1 + a_2, \sigma_1^2 + \sigma_2^2)$.
- e) Für $\lambda \in (0, \infty)$, $\nu_{1,2} \in (0, \infty)$ gilt $\Gamma_{\lambda, \nu_1} * \Gamma_{\lambda, \nu_2} = \Gamma_{\lambda, \nu_1 + \nu_2}$.
Die Summe von n unabhängigen $\exp(\lambda)$ -verteilten ZVn ist $\Gamma_{\lambda, n}$ -verteilt.
Die Summe der Quadrate von n unabhängigen jeweils $N(0, 1)$ -verteilten ZVn ist χ_n^2 -verteilt.

Beweis: Siehe Übungen. □

6. Spezielle Verteilungen

6.1 Diskrete Verteilungen

Definition 6.1. Als **Binomialverteilung** $b(n, p)$ mit Parametern $n \in \mathbb{N}$, $p \in (0, 1)$ oder auch $[0, 1]$ wird ein W-Maß auf $\mathcal{B}$ bezeichnet, das auf $\{0, 1, \dots, n\}$ konzentriert ist und die Zähldichte

$$k \mapsto b(n, p; k) := \binom{n}{k} p^k (1-p)^{n-k}, \quad k \in \mathbb{N}_0$$

besitzt.

Satz 6.2. Sei $n \in \mathbb{N}$, $p \in [0, 1]$; $q := 1 - p$.

- Ist $(X_1, \dots, X_n)$ ein n -tupel unabhängiger reeller ZVn mit $P[X_i = 1] = p$, $P[X_i = 0] = q$ — d.h. mit jeweiliger $b(1, p)$ -Verteilung — ($i = 1, \dots, n$), so ist $\sum_{i=1}^n X_i$ $b(n, p)$ -verteilt.
- Für eine ZV X mit $P_X = b(n, p)$ gilt $EX = np$, $V(X) = npq$.
- $b(n, p)$ hat die erzeugende Funktion g mit $g(s) = (q + ps)^n$.
- Für $n_{1,2} \in \mathbb{N}$ gilt

$$b(n_1, p) * b(n_2, p) = b(n_1 + n_2, p),$$

d.h. für zwei unabhängige ZVn X_1, X_2 mit $P_{X_1} = b(n_1, p)$, $P_{X_2} = b(n_2, p)$ gilt $P_{X_1+X_2} = b(n_1 + n_2, p)$.

Bemerkung 6.3. Das n -tupel $(X_1, \dots, X_n)$ von ZVn aus Satz 6.2 beschreibt ein n -tupel von Bernoulli-Versuchen, d.h. Zufallsexperimenten ohne gegenseitige Beeinflussung mit Ausgängen 1 (“Erfolg”) oder 0 (“Misserfolg”) und jeweiliger Erfolgswahrscheinlichkeit p ; $\sum_{k=1}^n X_i$ gibt die Anzahl der Erfolge in n Bernoulli-Versuchen an.

Definition 6.4. Als **Poisson-Verteilung** $\pi(\lambda)$ mit Parameter $\lambda > 0$ wird ein W-Maß auf $\mathcal{B}$ bezeichnet, das auf $\mathbb{N}_0$ konzentriert ist und die Zähldichte

$$k \mapsto \pi(\lambda; k) := e^{-\lambda} \frac{\lambda^k}{k!}, \quad k \in \mathbb{N}_0$$

besitzt.

Satz 6.5. Sei $\lambda > 0$.

- Ist $(b(n, p_n))$ eine Folge von Binomialverteilungen mit $np_n \rightarrow \lambda$ ($n \rightarrow \infty$), dann konvergiert

$$b(n, p_n; k) \rightarrow \pi(\lambda; k) \quad (n \rightarrow \infty) \text{ für alle } k \in \mathbb{N}_0.$$

b) Für eine ZV X mit $P_X = \pi(\lambda)$ gilt $EX = V(X) = \lambda$.

c) $\pi(\lambda)$ hat die erzeugende Funktion g mit $g(s) = e^{-\lambda+\lambda s}$.

d) Für $\lambda_{1,2} > 0$ gilt

$$\pi(\lambda_1) * \pi(\lambda_2) = \pi(\lambda_1 + \lambda_2).$$

Beweis von Satz 6.5 a):

Es gelte $n \cdot p_n \rightarrow \lambda$ ($n \rightarrow \infty$), woraus $p_n \rightarrow 0$ ($n \rightarrow \infty$) folgt. Mit

$$\begin{aligned} b(n, p_n; k) &= \binom{n}{k} \cdot p_n^k \cdot (1 - p_n)^{n-k} \\ &= \frac{1}{k!} \cdot \underbrace{(n \cdot p_n)}_{\rightarrow \lambda} \cdot \underbrace{(n-1) \cdot p_n}_{\rightarrow \lambda} \cdot \dots \cdot \underbrace{(n-k+1) \cdot p_n}_{\rightarrow \lambda} \cdot \left((1 - p_n)^{\frac{1}{p_n}} \right)^{(n-k) \cdot p_n} \end{aligned}$$

und

$$(n - k) \cdot p_n \rightarrow \lambda \quad (n \rightarrow \infty)$$

sowie

$$(1 - p_n)^{\frac{1}{p_n}} \rightarrow e^{-1} \quad (n \rightarrow \infty)$$

(wegen $p_n \rightarrow 0$ ($n \rightarrow \infty$), woraus

$$\ln \left((1 - p_n)^{\frac{1}{p_n}} \right) = \frac{\ln(1 - p_n)}{p_n} \rightarrow -1 \quad (n \rightarrow \infty)$$

folgt), folgt die Behauptung. □

Bemerkung 6.6. Bei einer rein zufälligen Aufteilung von n unterscheidbaren Teilchen auf m gleichgroße mehrfach besetzbare Zellen in einem Euklidischen Raum wird die Anzahl der Teilchen in einer vorgegebenen Zelle durch eine $b(n, \frac{1}{m})$ -verteilte ZV angegeben. Der Grenzübergang $n \rightarrow \infty$, $m \rightarrow \infty$ mit *Belegungsintensität* $\frac{n}{m} \rightarrow \lambda > 0$ führt gemäß Satz 6.5a zu einer $\pi(\lambda)$ -verteilten ZV.

Definition 6.7. Als **negative Binomialverteilung** oder **Pascal-Verteilung** $Nb(r, p)$ mit Parametern $r \in \mathbb{N}$, $p \in (0, 1)$ wird ein W-Maß auf $\mathcal{B}$ (oder auch $\overline{\mathcal{B}}$) bezeichnet, das auf $\mathbb{N}_0$ konzentriert ist und — mit $q := 1 - p$ — die Zähldichte

$$k \mapsto Nb(r, p; k) := \binom{r+k-1}{k} p^r q^k, \quad k \in \mathbb{N}$$

besitzt. Speziell wird $Nb(1, p)$ — mit Zähldichte $k \mapsto p(1-p)^k$, $k \in \mathbb{N}_0$ — als **geometrische Verteilung** mit Parameter $p \in (0, 1)$ bezeichnet.

Satz 6.8. Sei $r \in \mathbb{N}$, $p \in (0, 1)$, $q := 1 - p$.

- a) Sei $(X_n)_{n \in \mathbb{N}}$ eine unabhängige Folge von $b(1, p)$ -verteilten ZVn. Die erweitert-reelle ZV $X := \inf\{n \in \mathbb{N} \mid \sum_{k=1}^n X_k = r\} - r$ mit $\inf \emptyset := \infty$ ist $Nb(r, p)$ -verteilt.
- b) Für eine ZV X mit $P_X = Nb(r, p)$ gilt $EX = \frac{rq}{p}$, $V(X) = \frac{rq}{p^2}$.
- c) $Nb(r, p)$ hat die erzeugende Funktion g mit $g(s) = p^r(1 - qs)^{-r}$.
- d) Für $r_{1,2} \in \mathbb{N}$ gilt

$$Nb(r_1, p) * Nb(r_2, p) = Nb(r_1 + r_2, p).$$

Die Summe von r unabhängigen $Nb(1, p)$ -verteilten ZVn ist somit $Nb(r, p)$ -verteilt.

Bemerkung 6.9. Für die Folge (X_n) in Satz 6.8a wird durch die erweitert-reelle ZV $T := \inf\{n \in \mathbb{N} \mid \sum_{k=1}^n X_k = r\}$ mit $\inf \emptyset := \infty$ die *Wartezeit* bis zum r -ten Auftreten der Eins in (X_n) und durch die erweitert-reelle ZV $X = T - r$ die Anzahl der Misserfolge bis zum r -ten Erfolg bei der zu (X_n) gehörigen Folge von Bernoulli-Versuchen angegeben.

Definition 6.10. Als **hypergeometrische Verteilung** mit Parametern $n, r, s \in \mathbb{N}$, wobei $n \leq r + s$, wird ein W-Maß auf $\mathcal{B}$ bezeichnet, das auf $\{0, 1, \dots, n\}$ — sogar auf $\{\max(0, n - s), \dots, \min(n, r)\}$ — konzentriert ist und die Zähldichte

$$k \mapsto \begin{cases} \frac{\binom{r}{k} \binom{s}{n-k}}{\binom{r+s}{n}}, & k = 0, 1, \dots, n \\ 0 & , k = n + 1, n + 2, \dots \end{cases}$$

besitzt.

Bemerkung 6.11. Werden aus einer Urne mit r roten und s schwarzen Kugeln rein zufällig n Kugeln ohne Zurücklegen gezogen (n, r, s wie in Definition 6.10), so wird die Anzahl der gezogenen roten Kugeln durch eine ZV angegeben, die eine hypergeometrische Verteilung mit Parametern n, r, s besitzt. Anwendung der hypergeometrischen Verteilung in der Qualitätskontrolle.

6.2 Totalstetige Verteilungen

Definition 6.12. Als **Gleichverteilung auf (a, b)** mit $-\infty < a < b < \infty$ wird ein W-Maß auf $\mathcal{B}$ bezeichnet, das eine Dichte $f = \frac{1}{b-a} \chi_{(a,b)}$ besitzt.

Satz 6.13. Für eine ZV X mit Gleichverteilung auf (a, b) gilt $EX = \frac{a+b}{2}$, $V(X) = \frac{(b-a)^2}{12}$.

Definition 6.14. Als (eindimensionale) **Normalverteilung** oder Gauß-Verteilung $N(a, \sigma^2)$ mit Parametern $a \in \mathbb{R}$, $\sigma > 0$ wird ein W-Maß auf $\mathcal{B}$ bezeichnet, das eine Dichte f mit

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad x \in \mathbb{R}$$

besitzt. Speziell heißt $N(0, 1)$ **standardisierte Normalverteilung**.

Bemerkung 6.15.

- a) Sei f wie in Definition 6.14. Der Graph von f heißt Gaußsche Glockenkurve. Im Fall $a = 0$ ist $f(0) = \frac{1}{\sqrt{2\pi}\sigma}$ und hat f zwei Wendepunkte $(\pm\sigma; \frac{1}{\sqrt{2\pi e}\sigma})$.
- b) Die Dichtefunktion und die VF ϕ von $N(0, 1)$ sind tabelliert.
- c) Ist die ZV X $N(a, \sigma^2)$ -verteilt, so ist $\frac{X-a}{\sigma}$ $N(0, 1)$ -verteilt. Es gilt hierbei

$$P[a - \sigma \leq X \leq a + \sigma] = 2\phi(1) - 1 \approx 0,683$$

$$P[a - 2\sigma \leq X \leq a + 2\sigma] = 2\phi(2) - 1 \approx 0,954$$

$$P[a - 3\sigma \leq X \leq a + 3\sigma] = 2\phi(3) - 1 \approx 0,997.$$

Zu Bemerkung 6.15 c):

$$\begin{aligned} P[a - k \cdot \sigma \leq X \leq a + k \cdot \sigma] &= P[-k \leq \frac{X-a}{\sigma} \leq k] \\ &= \Phi(k) - \Phi(-k) \\ &= 2\Phi(k) - 1 \end{aligned}$$

da aus Symmetriegründen gilt:

$$\Phi(-k) = 1 - \Phi(k).$$

□

Satz 6.16. Sei $a \in \mathbb{R}$, $\sigma > 0$.

- a) Für eine ZV X mit Verteilung $N(a, \sigma^2)$ gilt:

$$E(X - a)^{2k-1} = 0, \quad E(X - a)^{2k} = \sigma^{2k} \prod_{j=1}^k (2j - 1), \quad k \in \mathbb{N};$$

insbesondere $EX = a$, $V(X) = \sigma^2$; die ZV $cX + b$ mit $0 \neq c \in \mathbb{R}$, $b \in \mathbb{R}$ hat die Verteilung $N(ca + b, c^2\sigma^2)$.

- b) Die charakteristische Funktion von $N(a, \sigma^2)$ ist φ mit

$$\varphi(u) = e^{iau} e^{-\frac{\sigma^2 u^2}{2}}.$$

c) Für $a_{1,2} \in \mathbb{R}$, $\sigma_{1,2} > 0$ gilt

$$N(a_1, \sigma_1^2) * N(a_2, \sigma_2^2) = N(a_1 + a_2, \sigma_1^2 + \sigma_2^2).$$

Definition 6.17. Als **Exponentialverteilung** $\exp(\lambda)$ mit Parameter $\lambda > 0$ wird ein W-Maß auf $\mathcal{B}$ oder $\overline{\mathcal{B}}$ bezeichnet, das eine Dichtefunktion f mit

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

besitzt.

Satz 6.18. Sei $\lambda > 0$.

a) Sei X eine erweitert-reelle ZV, deren Verteilung auf $\overline{\mathbb{R}}_+$ konzentriert sei, mit $P[0 < X < \infty] > 0$. Dann gilt:

$$\forall_{s,t \in (0, \infty)} P[X > t + s \mid X > s] = P[X > t] \iff P_X \text{ ist eine Exponentialverteilung.}$$

“Gedächtnislosigkeit”

b) Für eine ZV X mit $P_X = \exp(\lambda)$ gilt $EX = \frac{1}{\lambda}$, $V(X) = \frac{1}{\lambda^2}$.

c) $\exp(\lambda)$ hat die charakteristische Funktion φ mit

$$\varphi(u) = \frac{\lambda}{\lambda - iu}.$$

Definition 6.19. Als **Gamma-Verteilung** $\Gamma_{\lambda, \nu}$ mit Parametern $\lambda, \nu > 0$ wird ein W-Maß auf $\mathcal{B}$ oder $\overline{\mathcal{B}}$ bezeichnet, das eine Dichtefunktion f mit

$$f(x) = \begin{cases} \frac{\lambda^\nu}{\Gamma(\nu)} x^{\nu-1} e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

besitzt. Hierbei $\Gamma_{\lambda, 1} = \exp(\lambda)$. $\Gamma_{\frac{1}{2}, \frac{n}{2}}$ wird als Chi-Quadrat-Verteilung χ_n^2 mit $n \in \mathbb{N}$ Freiheitsgraden bezeichnet.

Satz 6.20. Seien $\lambda, \nu > 0$, $n \in \mathbb{N}$.

a) Für eine ZV X mit $P_X = \Gamma_{\lambda, \nu}$ gilt $EX = \frac{\nu}{\lambda}$, $V(X) = \frac{\nu}{\lambda^2}$.

b) $\Gamma_{\lambda, \nu}$ hat die charakteristische Funktion φ mit

$$\varphi(u) = \left(\frac{\lambda}{\lambda - iu} \right)^\nu.$$

c) Für $\nu_{1,2} > 0$ gilt

$$\Gamma_{\lambda,\nu_1} * \Gamma_{\lambda,\nu_2} = \Gamma_{\lambda,\nu_1+\nu_2}.$$

Insbesondere ist $\Gamma_{\lambda,n}$ die n -fache Faltung von $\exp(\lambda)$.

d) Die Summe der Quadrate von n unabhängigen jeweils $N(0, 1)$ -verteilten reellen ZVn ist χ_n^2 -verteilt.

Zu Satz 6.20 d)

Zu zeigen: $X_1, \dots, X_n$ unabhängig $\mathcal{N}(0, 1)$ -verteilt $\Rightarrow X_1^2 + \dots + X_n^2$ χ_n^2 -verteilt, d.h. $\Gamma_{\frac{1}{2}, \frac{n}{2}}$ -verteilt.

Wegen Faltungseigenschaft von $\Gamma_{\lambda,\nu}$ genügt es zu zeigen:

X_1^2 ist $\Gamma_{1/2, 1/2}$ -verteilt.

Dazu: Für $x \geq 0$ gilt

$$\begin{aligned} F(x) &= P[X_1^2 \leq x] = P[-\sqrt{x} \leq X_1 \leq \sqrt{x}] \\ &= \int_{-\sqrt{x}}^{\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \\ &= \frac{2}{\sqrt{2\pi}} \int_0^{\sqrt{x}} e^{-t^2/2} dt. \end{aligned}$$

Ableiten liefert:

$$\begin{aligned} F'(x) &= \frac{2}{\sqrt{2\pi}} e^{-x/2} \cdot \frac{1}{2\sqrt{x}} = \frac{1}{\sqrt{2\pi}} \cdot x^{-1/2} \cdot e^{-x/2} \\ &= \frac{(1/2)^{1/2}}{\Gamma(1/2)} \cdot x^{1/2-1} \cdot e^{-1/2 \cdot x} \dots \text{Dichte zu } \Gamma_{1/2, 1/2} \\ &\text{da } \Gamma(1/2) = \sqrt{\pi} \end{aligned}$$

$\Rightarrow X_1^2$ ist $\Gamma_{1/2, 1/2}$ -verteilt.

(denn aus der obigen Bezeichnung folgt $\int_0^x F'(t) dt = F(x) - F(0+) = F(x)$ für $x > 0$)

Bemerkung 6.21. Sind X, Y zwei unabhängige reelle ZVn mit $P_X = N(0, 1)$ und $P_Y = \chi_n^2$, so wird die Verteilung der ZV $\frac{X}{\sqrt{Y/n}}$ als **t-Verteilung** oder **Student-Verteilung** t_n bzw. St_n mit n Freiheitsgraden bezeichnet ($n \in \mathbb{N}$). t_1 bzw. St_1 ist eine sogenannte **Cauchy-Verteilung**. — Anwendung von χ_n^2 und St_n in der Statistik.

Beispiel aus der Erneuerungstheorie:

Bauelemente werden nach Ausfall sofort ersetzt, wobei:

- die Bauelemente hinsichtlich der Lebensdauer das gleiche stochastische Verhalten haben.

- die Bauelemente sich hinsichtlich der Lebensdauer unabhängig voneinander verhalten.

Die Bauelemente seien in der Reihenfolge des Einsatzes durchnummeriert.

Zur stochastischen Modellierung fassen wir die Lebensdauern als Realisierungen unabhängig, identisch verteilter, nichtnegativer reeller Zufallsvariablen $T_1, T_2, \dots$ auf, definiert auf W-Raum $(\Omega, \mathcal{A}, P)$.

Es gelte: $P[T_1 = 0] < 1$ (sonst Trivialfall).

Dann beschreibt

$$N_t = \sup\{n \in \mathbb{N} : T_1 + \dots + T_n \leq t\}$$

(mit $\sup \emptyset = 0$) die Anzahl der Erneuerungen im Zeitintervall $[0, t]$.

Hierbei ist

$$(T_1 + \dots + T_n)_{n \in \mathbb{N}} \dots \text{ sog. Erneuerungsproze\ss}$$

$\{N_t : t \in \mathbb{R}_+\}$... sog. Zählproze\ss zum Erneuerungsproze\ss.

N_t ist eine $\bar{\mathbb{N}}_0$ -wertige ZV, da:

$$[N_t = \infty] = \bigcap_{n \in \mathbb{N}} \underbrace{[T_1 + \dots + T_n \leq t]}_{\in \mathcal{A}} \in \mathcal{A}$$

$$[N_t = k] = \underbrace{[T_1 + \dots + T_k \leq t]}_{\in \mathcal{A}} \setminus \underbrace{[T_1 + \dots + T_{k+1} \leq t]}_{\in \mathcal{A}} \in \mathcal{A}.$$

Satz 6.22. Ist — mit den obigen Bezeichnungen — T_n $\exp(\lambda)$ -verteilt ($\lambda > 0$), so ist N_t $\pi(\lambda t)$ -verteilt, $t \in \mathbb{R}_+$.

Bemerkung 6.23. Die Familie $\{N_t \mid t \in \mathbb{R}_+\}$ aus Satz 6.22 ist ein sogenannter **Poisson-Prozess**.

Beweis von Satz 6.22:

$S_k = T_1 + \dots + T_k$ ist $\Gamma_{\lambda, k}$ -verteilt nach Satz 6.20 c).

Daher gilt für $k \in \mathbb{N}$

$$P[N_t = k] = P[S_k \leq t] - P[S_{k+1} \leq t]$$

$$= \int_0^t \underbrace{\frac{\lambda^k}{(k-1)!} \cdot x^{k-1}}_{u'} \cdot \underbrace{e^{-\lambda x}}_v dx - \int_0^t \underbrace{\frac{\lambda^{k+1}}{k!} \cdot x^k \cdot e^{-\lambda x}}_{=-v' \cdot u} dx$$

Mit partieller Integration des ersten Integrals folgt daraus

$$P[N_t = k] = \left[\lambda^k \cdot \frac{x^k}{k!} \cdot e^{-\lambda x} \right]_{x=0}^t = \frac{(\lambda t)^k}{k!} \cdot e^{-\lambda t},$$

d. h. $P[N_t = k]$ stimmt für $k \in \mathbb{N}$ mit dem Wert der Zähldichte von $\pi(\lambda t)$ an der Stelle k überein.

Für $k = 0$ gilt

$$P[N_t = 0] = P[S_1 > t] = P[T_1 > t] = \int_t^{\infty} \lambda \cdot e^{-\lambda x} dx = e^{-\lambda t},$$

also gilt die obige Aussage auch für $k = 0$.

□

7 Konvergenz von Folgen von Zufallsvariablen

Definition 7.1.

a) W-Raum $(\Omega, \mathcal{A}, P)$, $X_n, X : (\Omega, \mathcal{A}, P) \rightarrow (\mathbb{R}, \mathcal{B})$ ($n \in \mathbb{N}$), $p \geq 1$. Man sagt: X_n **konvergiert gegen** X

$a_1)$ **fast sicher** (kurz: $X_n \rightarrow X$ f.s.)
 $:\Leftrightarrow P(\{\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}) = 1.$

$a_2)$ **im p-ten Mittel** (kurz: $X_n \rightarrow X$ im p-ten Mittel)
 $:\Leftrightarrow E\{|X_n - X|^p\} \rightarrow 0$ ($n \rightarrow \infty$)

$a_3)$ **nach Wahrscheinlichkeit** (kurz: $X_n \xrightarrow{P} X$)
 $:\Leftrightarrow \forall \varepsilon > 0 : P[|X_n - X| > \varepsilon] \rightarrow 0$ ($n \rightarrow \infty$)

$a_4)$ **nach Verteilung** (kurz: $X_n \xrightarrow{\mathcal{D}} X$)
 $:\Leftrightarrow$ Für jede beschränkte und stetige Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$ gilt:

$$Eg(X_n) \rightarrow Eg(X) \quad (n \rightarrow \infty)$$

b) W-Maße Q_n, Q ($n \in \mathbb{N}$) auf $\mathcal{B}$. Man sagt: Q_n **konvergiert gegen** Q **schwach** (kurz: $Q_n \rightarrow Q$ schwach)

$:\Leftrightarrow$ Für jede beschränkte und stetige Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$ gilt:

$$\int_{\mathbb{R}} g dQ_n \rightarrow \int_{\mathbb{R}} g dQ \quad (n \rightarrow \infty).$$

Bemerkung 7.2:

a) Zwischen den Konvergenzbegriffen bestehen die folgenden Zusammenhänge:

$a_1)$ $X_n \rightarrow X$ f.s. $\Rightarrow X_n \xrightarrow{P} X$

$a_2)$ $X_n \rightarrow X$ im p-ten Mittel $\Rightarrow X_n \xrightarrow{P} X$

$a_3)$ $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{\mathcal{D}} X$

$a_4)$ $X_n \xrightarrow{P} X \Rightarrow$ Es existiert Teilfolge $(X_{n_k})_k$ mit $X_{n_k} \rightarrow X$ f.s.

Begründung:

$a_1)$ $P[|X_n - X| > \varepsilon] = \int_{\Omega} \chi_{[|X_n - X| > \varepsilon]} dP$

Aus $X_n \rightarrow X$ f.s. folgt nun

$$\chi_{[|X_n - X| > \varepsilon]} \rightarrow 0 \quad \text{f.s.}$$

und da die Funktion links betragsmäßig durch die bzgl. P integrierbare Konstante 1 beschränkt ist, folgt mit dem Satz von der majorisierten Konvergenz die Behauptung.

a₂) Folgt aus der Ungleichung von Markov:

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

a₃) Siehe Satz 7.4.

a₄) Siehe Satz 7.3.

□

b) Für die Konvergenz nach Verteilung (äquivalent: $P_{X_n} \rightarrow P_X$ schwach) müssen die Zufallsvariablen nicht auf dem gleichen W-Raum definiert sein.

c) $X_n \rightarrow^{\mathcal{D}} X$ ist nach Satz 7.5 äquivalent zu:

Für jeden Stetigkeitspunkt $x \in \mathbb{R}$ der Verteilungsfunktion F von X gilt

$$F(x_n) \rightarrow F(x) \quad (n \rightarrow \infty),$$

wobei F_n die Verteilungsfunktion von X_n ist.

Die allgemeinere Definition oben lässt sich aber auf Zufallsvariablen mit Werten in (geeigneten) metrischen Räumen verallgemeinern.

d) Die Grenzverteilung bei Konvergenz nach Verteilung ist eindeutig (und damit auch bei allen anderen Konvergenzarten in Definition 7.1):

$$\left. \begin{array}{l} X_n \rightarrow^{\mathcal{D}} X^* \\ X_n \rightarrow^{\mathcal{D}} X^{**} \end{array} \right\} \Rightarrow^{c)} \left\{ \begin{array}{l} F_n(x) \rightarrow F^*(x) \text{ in allen Stetigkeitsstellen von } F^* \\ F_n(x) \rightarrow F^{**}(x) \text{ in allen Stetigkeitsstellen von } F^{**} \end{array} \right.$$

wobei F^* bzw. F^{**} die Verteilungsfunktion von X^* bzw. X^{**} ist.

Da $\{x \in \mathbb{R} \mid x \text{ ist Unstetigkeitsstelle von } F^* \text{ bzw. } F^{**}\}$ abzählbar ist, folgt mit der rechtsseitigen Stetigkeit der Verteilungsfunktion: $F^*(x) = F^{**}(x)$ für alle $x \in \mathbb{R}$, woraus $P_{X^*} = P_{X^{**}}$ folgt.

e) In der Definition der Konvergenz nach Verteilung lassen sich X_n und X ersetzen durch X'_n und X' , sofern $P_{X_n} = P_{X'_n}$ und $P_{X'} = P_X$.

Satz 7.3:

W-Raum $(\Omega, \mathcal{A}, P)$, $X_n, X : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B})$ ($n \in \mathbb{N}$).

a) Aus $X_n \rightarrow^P X$ folgt, dass eine Teilfolge $(X_{n_k})_k$ von $(X_n)_n$ existiert mit

$$X_{n_k} \rightarrow X \quad \text{f.s.}$$

b) Die beiden folgenden Aussagen sind äquivalent:

b₁) $X_n \rightarrow^P X$

b₂) Für jede Teilfolge $(n_k)_k$ von $(n)_n$ existiert Teilteilstfolge $(n_{k_l})_l$ mit

$$X_{n_{k_l}} \rightarrow X \quad \text{f.s.}$$

Beweis

a) Es gelte $X_n \rightarrow^P X$, d. h.

$$\forall \varepsilon > 0 : P[|X_n - X| > \varepsilon] \rightarrow 0 \quad (n \rightarrow \infty).$$

Daraus folgt:

$$\exists n_1 : P[|X_{n_1} - X| \geq 1] < 1$$

$$\exists n_2 > n_1 : P[|X_{n_2} - X| \geq \frac{1}{2}] < \left(\frac{1}{2}\right)^2$$

$$\exists n_3 > n_2 : P[|X_{n_3} - X| \geq \frac{1}{3}] < \left(\frac{1}{3}\right)^2$$

usw.

Man erhält Teilfolge $(X_{n_k})_k$ von $(X_n)_n$ mit

$$P[|X_{n_k} - X| \geq \frac{1}{k}] \leq \left(\frac{1}{k}\right)^2.$$

Für diese gilt nach dem Satz von der monotonen Konvergenz

$$\begin{aligned} \int_{\Omega} \sum_{k=1}^{\infty} \chi_{[|X_{n_k} - X| \geq \frac{1}{k}]} dP &= \sum_{k=1}^{\infty} \int_{\Omega} \chi_{[|X_{n_k} - X| \geq \frac{1}{k}]} dP \\ &= \sum_{k=1}^{\infty} P[|X_{n_k} - X| \geq \frac{1}{k}] \leq \sum_{k=1}^{\infty} \left(\frac{1}{k}\right)^2 < \infty. \end{aligned}$$

Mit Satz 5.3 folgt

$$\sum_{k=1}^{\infty} \chi_{[|X_{n_k} - X| \geq \frac{1}{k}]} < \infty \quad P - \text{f.s.}$$

Somit gilt für P -fast alle $\omega \in \Omega$:

$$\chi_{[|X_{n_k} - X| \geq \frac{1}{k}]}(\omega) = 1 \quad \text{nur für endlich viele Indices } k,$$

woraus folgt:

$$\exists k_0 = k_0(\omega) \quad \forall k \geq k_0 : |X_{n_k}(\omega) - X(\omega)| < \frac{1}{k}.$$

Dies impliziert für P -f.a. $\omega \in \Omega$:

$$\overline{\lim}_{k \rightarrow \infty} |X_{n_k}(\omega) - X(\omega)| \leq \overline{\lim}_{k \rightarrow \infty} \frac{1}{k} = 0,$$

woraus $X_{n_k} \rightarrow X$ P -f.s. folgt.

b) “ $b_1 \Rightarrow b_2$ ” folgt aus a), da aus $X_n \xrightarrow{P} X$ auch $X_{n_k} \xrightarrow{P} X$ für jede Teilfolge von $(X_n)_n$ folgt.

“ $b_2 \Rightarrow b_1$ ”: Gilt, da es für

$$P[|X_n - X| > \varepsilon] \rightarrow 0 \quad (n \rightarrow \infty)$$

zu zeigen genügt: Zu jeder Teilfolge $(n_k)_k$ von $(n)_n$ existiert Teilteilstolge $(n_{k_l})_l$ von $(n_k)_k$ mit

$$P[|X_{n_{k_l}} - X| > \varepsilon] \rightarrow 0 \quad (l \rightarrow \infty).$$

□

Satz 7.4. Reelle ZVn X_n ($n \in \mathbb{N}$), X auf $(\Omega, \mathcal{A}, P)$.

Aus $X_n \rightarrow X$ nach Wahrscheinlichkeit (Schreibweise $X_n \xrightarrow{P} X$) folgt $X_n \xrightarrow{D} X$.

Beweis: Beliebiges stetiges und beschränktes $g : \mathbb{R} \rightarrow \mathbb{R}$. Zu zeigen:

$$Eg(X_n) \rightarrow Eg(X) \quad (n \rightarrow \infty).$$

Äquivalent (bei Konvergenz reeller Folgen):

Für jede Teilfolge $(n_k)_k$ von $(n)_n$ existiert Teilteilstfolge $(n_{k_l})_l$ mit

$$Eg(X_{n_{k_l}}) \rightarrow Eg(X) \quad (l \rightarrow \infty).$$

Dazu: Sei $(n_k)_k$ beliebige Teilfolge von $(n)_n$.

Nach Satz 7.3 b) existiert dann Teilteilstfolge $(n_{k_l})_l$ mit

$$X_{n_{k_l}} \rightarrow X \quad f.s.$$

Da g stetig ist, folgt daraus

$$g(X_{n_{k_l}}) \rightarrow g(X) \quad f.s.,$$

und der Satz von der majorisierten Konvergenz impliziert

$$Eg(X_{n_{k_l}}) \rightarrow Eg(X) \quad (l \rightarrow \infty)$$

□

Satz 7.5. W-Maße Q_n ($n \in \mathbb{N}$), Q auf $\mathcal{B}$ bzw. reelle ZVn X_n ($n \in \mathbb{N}$), X mit VFn F_n, F . $Q_n \rightarrow Q$ schwach bzw. $X_n \xrightarrow{D} X$ ($n \rightarrow \infty$) gilt genau dann, wenn $F_n(x) \rightarrow F(x)$ ($n \rightarrow \infty$) für alle Stetigkeitspunkte x von F .

Beweis: Siehe Übungen.

Beweisidee:

Approximiere in

$$F_n(x) = \int \chi_{(-\infty, x]}(t) dP_{X_n}(t)$$

die Indikatorfunktion durch stetige und beschränkte Funktionen bzw. approximiere stetige und beschränkte Funktionen durch Linearkombinationen von Indikatorfunktionen. $\square$

Satz 7.6. Reelle ZVn X_n ($n \in \mathbb{N}$), X auf $(\Omega, \mathcal{A}, P)$. Ist X P-f.s. konstant, so gilt:

$$X_n \xrightarrow{\mathcal{D}} X \iff X_n \xrightarrow{P} X.$$

Beweis:

“ $\Leftarrow$ ” Klar nach Satz 7.4.

“ $\Rightarrow$ ” oBdA $X_n \rightarrow^{\mathcal{D}} c$, da Verteilung von X mit der von c übereinstimmt.

Sei F die Verteilungsfunktion von c , also

$$F(x) = \begin{cases} 0 & , x < c \\ 1 & , x \geq c \end{cases}$$

Nach Satz 7.5 gilt für die Verteilungsfunktion F_n von X_n :

$$F_n(x) \rightarrow F(x) \quad (n \rightarrow \infty) \text{ für alle } x \neq c.$$

Daraus folgt für $\varepsilon > 0$:

$$\begin{aligned} & P[|X_n - c| > \varepsilon] \\ &= P[X_n > c + \varepsilon] + P[X_n < c - \varepsilon] \\ &\leq 1 - F_n(c + \varepsilon) + F_n(c - \varepsilon) \\ &\xrightarrow{(n \rightarrow \infty)} 1 - F(c + \varepsilon) - F(c - \varepsilon) = 1 - 1 - 0 = 0. \end{aligned}$$

$\square$

Satz 7.7. Reelle ZVn X_n, Y_n ($n \in \mathbb{N}$), X auf $(\Omega, \mathcal{A}, P)$

$$\left. \begin{array}{l} X_n \xrightarrow{\mathcal{D}} X \\ |X_n - Y_n| \xrightarrow{P} 0 \end{array} \right\} \implies Y_n \xrightarrow{\mathcal{D}} X.$$

Beweis: Beliebiges beschränktes und stetiges $g : \mathbb{R} \rightarrow \mathbb{R}$.

zu zeigen: $Eg(Y_n) - Eg(X) \rightarrow 0$ ($n \rightarrow \infty$)

oBdA $X_n - Y_n \rightarrow 0$ f.s. (sonst mit Teiltonfolgen arbeiten).

Sei $|g(x)| \leq B$ für $x \in \mathbb{R}$. Wähle $M > 0$ mit $-M$ und M sind Stetigkeitspunkte der Verteilungsfunktion F von X .

Dann gilt:

$$\begin{aligned} & |Eg(X_n) - Eg(Y_n)| \\ &\leq E|g(X_n) - g(Y_n)| \cdot \chi_{[-M < X_n \leq M, |X_n - Y_n| \leq 1]} \\ &\quad + 2\beta \cdot P[X_n > M] + 2\beta \cdot P[X_n \leq -M] + 2\beta \cdot P[|X_n - Y_n| > 1]. \end{aligned}$$

Da g auf $[-M - 1, M + 1]$ gleichmäßig stetig ist, folgt aus $X_n - Y_n \rightarrow 0$ f.s.

$$|g(X_n) - g(Y_n)| \cdot \chi_{[-M < X_n \leq M, |X_n - Y_n| \leq 1]} \rightarrow 0 \text{ f.s.}$$

Mit der Beschränktheit von g folgt daraus mit dem Satz von der majorisierten Konvergenz

$$E \left[|g(X_n) - g(Y_n)| \cdot \chi_{[-M < X_n \leq M, |X_n - Y_n| \leq 1]} \right] \rightarrow 0 \quad (n \rightarrow \infty).$$

Da $X_n \xrightarrow{\mathcal{D}} X$ gilt und $-M$ und M Stetigkeitspunkte der Verteilungsfunktion von X sind, gilt

$$P[X_n > M] = 1 - P[X_n \leq M] \xrightarrow{(n \rightarrow \infty)} 1 - P[X \leq M] = P[X > M]$$

und

$$P[X_n \leq -M] \xrightarrow{(n \rightarrow \infty)} P[X \leq -M].$$

Aus $X_n - Y_n \xrightarrow{P} 0$ folgt weiter

$$P[|X_n - Y_n| > 1] \rightarrow 0 \quad (n \rightarrow \infty).$$

Insgesamt erhält man

$$\limsup_{n \rightarrow \infty} |Eg(X_n) - Eg(Y_n)| \leq 0 + 2\beta P[X > M] + 2\beta \cdot P[X \leq -M] + 0.$$

Mit $M \rightarrow \infty$ folgt die Behauptung. □

Satz 7.8 (Spezialfall des **Darstellungssatzes von Skorokhod**).

Reelle ZVn X_n ($n \in \mathbb{N}$), X mit $X_n \xrightarrow{\mathcal{D}} X$. Dann existiert ein Wahrscheinlichkeitsraum $(\Omega^*, \mathcal{A}^*, P^*)$ — wobei man $\Omega^* = [0, 1]$, $\mathcal{A}^* = [0, 1] \cap \mathcal{B}$, $P^* = \text{L-B-Ma\ss auf } \mathcal{A}^*$ wählen kann — und reelle ZVn X_n^* ($n \in \mathbb{N}$), X^* auf $(\Omega^*, \mathcal{A}^*, P^*)$ derart, dass

$$\forall_n P_{X_n} = P_{X_n^*}, \quad P_X = P_{X^*}, \quad X_n^* \rightarrow X^* \quad (n \rightarrow \infty) \text{ } P^*\text{-f.s.}$$

Beweis:

Sei $\Omega^* = (0, 1)$, $\mathcal{A}^* = (0, 1) \cap \mathcal{B}$ und $P^* = \text{LB-Ma\ss } |_{\mathcal{A}^*}$.

Sei F_n bzw. F die Verteilungsfunktion zu X_n bzw. X .

Setze

$$X_n^*(\omega) = \min\{t \in \mathbb{R} : F_n(t) \geq \omega\}$$

$$X^*(\omega) = \min\{t \in \mathbb{R} : F(t) \geq \omega\}$$

für $\omega \in (0, 1)$.

X_n^* bzw. X^* sind wohldefiniert, messbar (da monoton) und es gilt

$$P_{X_n^*}^* = P_{X_n} \text{ und } P_{X^*}^* = P_X$$

(vgl. Übungen)

X^* hat höchstens abzählbar viele Unstetigkeitsstellen (da X^* monotone Abbildung ist, ist die Anzahl der Spünge mit Sprunghöhe $> \varepsilon$ auf jedem kompakten Intervall endlich), diese bilden eine Nullmenge bzgl. des LB-Ma\sses.

Also genügt es zu zeigen:

$$X_n^*(\omega) \rightarrow X^*(\omega) = (n \rightarrow \infty) \text{ für jede Stetigkeitsstelle } \omega \text{ von } X^*.$$

Dazu: Sei ω Stetigkeitsstelle von X^*

Sei $\varepsilon > 0$ beliebig.

Dann existieren Stetigkeitspunkte x_1, x_2 von F mit

$$(*) \quad x_1 < X^*(\omega) < x_2 \quad \text{und} \quad x_2 - x_1 < \varepsilon.$$

Aus $x_1 < X^*(\omega) = \min\{t \in \mathbb{R} : F(t) \geq \omega\}$ folgt

$$F(x_1) < \omega.$$

Da ω Stetigkeitsstelle von X^* ist, ist F rechts von der Stelle $X^*(\omega)$ nicht konstant gleich ω . Daher folgt aus $X^*(\omega) < x_2$: $\omega < F(x_2)$

Insgesamt gilt also:

$$F(x_1) < \omega < F(x_2).$$

Aus $X_n \xrightarrow{\mathcal{D}} X$ folgt

$$F_n(x_1) \rightarrow F(x_1) \quad (n \rightarrow \infty),$$

$$F_n(x_2) \rightarrow F(x_2) \quad (n \rightarrow \infty).$$

Also existiert n_0 so, dass für alle $n \geq n_0$ gilt:

$$F_n(x_1) < \omega < F_n(x_2).$$

Nach Definition von $X_n^*(\omega) = \min\{t \in \mathbb{R} : F_n(t) \geq \omega\}$ folgt daraus aber:

$$(**) \quad x_1 \leq X_n^*(\omega) \leq x_2.$$

Mit (*) und (**) ergibt sich für $n \geq n_0$:

$$|X_n^*(\omega) - X^*(\omega)| \leq x_2 - x_1 < \varepsilon,$$

woraus die Behauptung folgt. □

Satz 7.9 (Satz von der stetigen Abbildung).

Reelle ZVN X_n ($n \in \mathbb{N}$), X mit $X_n \xrightarrow{\mathcal{D}} X$; Abbildung $h : (\mathbb{R}, \mathcal{B}) \rightarrow (\mathbb{R}, \mathcal{B})$ sei P_X -f.ü. stetig. Dann gilt $h(X_n) \xrightarrow{\mathcal{D}} h(X)$.

Bemerkung: Für h stetig ist der Satz trivial.

Beweis von Satz 7.9:

Für

$$D = \{x \in \mathbb{R} : h \text{ unstetig in } x\} \in \mathcal{B}(!)$$

gilt

$$P_X(D) = 0.$$

Nach dem Satz 7.8 existiert W-Raum $(\Omega^*, \mathcal{A}^*, P^*)$ und Zven X_n^* und X^* auf $(\Omega^*, \mathcal{A}^*, P^*)$ mit

$$P_{X_n^*}^* = P_{X_n}, P_{X^*}^* = P_X \text{ und } X_n^*(\omega) \rightarrow X^*(\omega) \text{ f\u00fcr alle } \omega \in \Omega^*.$$

Daraus folgt

$$h(X_n^*(\omega)) \rightarrow h(X^*(\omega)) \quad (n \rightarrow \infty)$$

f\u00fcr alle $\omega \in \Omega^* \setminus \{(\bar{\omega}) : X^*(\bar{\omega}) \in D\}$

Wegen

$$P^*(\{\bar{\omega} \in \Omega^* : X^*(\bar{\omega}) \in D\}) = P_{X^*}^*(D) = P_X(D) = 0$$

gilt also

$$h(X_n^*) \rightarrow h(X^*) \quad P^* - f.s.,$$

was

$$h(X_n^*) \xrightarrow{\mathcal{D}} h(X^*)$$

bzw.

$$P_{h(X_n^*)}^* \rightarrow P_{h(X^*)}^* \text{ schwach}$$

impliziert.

Wegen

$$P_{h(X_n^*)}^* = (P_{X_n^*}^*)_h \stackrel{s.o.}{=} (P_{X_n})_h = P_{h(X_n)}$$

und

$$P_{h(X^*)}^* = (P_{X^*}^*)_h \stackrel{s.o.}{=} (P_X)_h = P_{h(X)}$$

(vergleiche Bemerkung 2.16) ist dies wiederum \u00e4quivalent zu

$$h(X_n) \xrightarrow{\mathcal{D}} h(X) \quad w.z.z.w.$$

□

Satz 7.10 (Satz von Slutsky).

Reelle ZVn X_n, Y_n, X auf $(\Omega, \mathcal{A}, P)$; $c \in \mathbb{R}$.

$$\left. \begin{array}{l} X_n \xrightarrow{\mathcal{D}} X \\ Y_n \xrightarrow{P} c \end{array} \right\} \implies \left\{ \begin{array}{l} X_n + Y_n \xrightarrow{\mathcal{D}} X + c \\ Y_n X_n \xrightarrow{\mathcal{D}} cX. \end{array} \right.$$

Beweis des ersten Teiles:

Aus $Y_n \xrightarrow{P} c$ folgt $Y_n - c \xrightarrow{P} 0$ und damit

$$(X_n + Y_n - c) - X_n \xrightarrow{P} 0.$$

Mit Satz 7.7 folgt daraus

$$X_n + Y_n - c \xrightarrow{\mathcal{D}} X,$$

woraus mit Satz 7.9 (angewendet mit $(h(z) = z + c)$)

$$X_n + Y_n \xrightarrow{\mathcal{D}} X + c$$

folgt. □

Definition 7.11. Eine Menge $\mathcal{Q}$ von W-Maßen auf $\mathcal{B}$ heißt

- a) **relativ-(folgen)kompakt**, wenn jede Folge von Elementen aus $\mathcal{Q}$ eine Teilfolge besitzt, die schwach gegen ein W-Maß auf $\mathcal{B}$ konvergiert,
- b) **gleichmäßig straff** [tight], wenn

$$\forall \varepsilon > 0 \quad \exists \text{ komp. } K \subset \mathbb{R} \quad \forall Q \in \mathcal{Q} \quad Q(K^c) \leq \varepsilon.$$

Satz 7.12 (Spezialfall des **Satzes von Prokhorov**).

Eine Menge $\mathcal{Q}$ von W-Maßen auf $\mathcal{B}$ ist genau dann relativ-kompakt, wenn sie gleichmäßig straff ist.

Im Beweis benötigen wir:

Satz 7.13 (Auswahlsatz von Helly).

Ist (Q_n) eine Folge von W-Maßen auf $\mathcal{B}$ mit zugehörigen VFn $F_n : \mathbb{R} \rightarrow \mathbb{R}$ ($n \in \mathbb{N}$), dann existiert eine Indexfolge (n_i) und ein endliches Maß Q auf $\mathcal{B}$ (nicht notw. W-Maß!) mit zugehöriger maßdefinierender Funktion $F : \mathbb{R} \rightarrow \mathbb{R}$ ($F(x) := Q((-\infty, x]) + \text{const.}$ (geeignet), $x \in \mathbb{R}$) derart, dass $F_{n_i}(x) \rightarrow F(x)$ ($i \rightarrow \infty$) für alle Stetigkeitspunkte x von F .

Bemerkung:

Hierbei ist $F(x) = 0$ ($x \in \mathbb{R}$) möglich, z. B. gilt für

$$F_n(x) = \begin{cases} 0, & x < n \\ 1, & x \geq n \end{cases}$$

(d.h. Q_n ist auf den Punkt $\{n\}$ konzentriert):

$$F_n(x) \rightarrow 0 \quad (n \rightarrow \infty) \quad \text{für alle } x \in \mathbb{R}$$

(“Wegrutschen von Masse ins Unendliche”).

Beweisskizze: Sei $Q = \{x_1, x_2, \dots\}$.

Da jede beschränkte reelle Folge eine konvergente Teilfolge enthält, kann man rekursiv Teilfolgen $(n_{k+1,l})_l$ von $(n_{k,l})_l$ (mit $n_{0,l} = l$ ($l \in \mathbb{N}$)) konstruieren, für die

$$(F_{n_{k+1,l}}(x_i))_{l \in \mathbb{N}}$$

konvergent ist für $i = 1, 2, \dots, k + 1$.

Setzt man dann

$$n_k = n_{k,k},$$

so sieht man leicht, dass

$$(F_{n_k}(x_i))_{k \in \mathbb{N}} \text{ für alle } i \in \mathbb{N} \text{ konvergiert.}$$

Nun definiert man

$$F^* : \mathbb{Q} \rightarrow \mathbb{R}$$

durch

$$F^*(x) = \lim_{n \rightarrow \infty} F_{n_k}(x).$$

Man macht sich leicht klar, dass dann gilt:

1. $F^*(x) \in [0, 1]$ ($x \in \mathbb{Q}$),
2. F^* ist monoton wachsend.

Die Funktion $F : \mathbb{R} \rightarrow \mathbb{R}$ definiert durch

$$F(x) = \inf_{z > x, z \in \mathbb{Q}} F^*(z)$$

erfüllt dann die Behauptung, denn:

- i) $F(x) \in [0, 1]$ ($x \in \mathbb{R}$)
- II) F ist monoton wachsend
- III) F ist rechtsseitig stetig (!)
- IV) Für jeden Stetigkeitspunkt $x \in \mathbb{R}$ von F gilt:

$$F_{n_k}(x) \rightarrow F(x) \quad (k \rightarrow \infty).$$

Nachweis von IV): Für beliebiges $\varepsilon > 0$ wähle $x_1, x_2 \in \mathbb{Q}$ und $\delta > 0$ mit

$$x_1 \leq x \leq x_2 \text{ und } |F(x_2 + \delta) - F(x_1 - \delta)| < \varepsilon.$$

Dann gilt:

$$\begin{aligned} F_{n_k}(x) - F(x) &\stackrel{\text{Monotonie von } F_{n_k}}{\leq} F_{n_k}(x_2) - F(x) \stackrel{(k \rightarrow \infty)}{\rightarrow} F^*(x_2) - F(x) \\ &\leq F(x_2 + \delta) - F(x) \\ &\stackrel{\text{Monotonie von } F}{\leq} F(x_2 + \delta) - F(x_1 - \delta) < \varepsilon \end{aligned}$$

und analog

$$F_{n_k}(x) - F(x) \geq F_{n_k}(x_1) - F(x) \xrightarrow{(k \rightarrow \infty)} F^*(x_1) - F(x) \geq F^*(x_1 - \delta) - F(x_2 + \delta) > -\varepsilon$$

Daraus folgt:

$$\limsup_{n \rightarrow \infty} |F_{n_k}(x) - F(x)| < \varepsilon$$

□

Beweis von Satz 7.12

a) Wir zeigen: $\mathcal{Q}$ relativ kompakt $\Rightarrow \mathcal{Q}$ gleichmäßig straff.

Sei $\mathcal{Q}$ relativ kompakt. Angenommen, $\mathcal{Q}$ ist nicht gleichmäßig straff. Dann gilt:

$$\exists \varepsilon > 0 \quad \forall \text{ kompakten } K \subset \mathbb{R} \exists Q \in \mathcal{Q} : Q(K) < 1 - \varepsilon.$$

Ist $\varepsilon > 0$ entsprechend gewählt, so gilt also:

$$\forall n \in \mathbb{N} \exists Q_n \in \mathcal{Q} : Q_n([-n, n]) < 1 - \varepsilon.$$

Da $\mathcal{Q}$ relativ kompakt ist, existiert Indexteilfolge $(n_i)_i$ und W-Maß Q^* mit

$$Q_{n_i} \rightarrow Q^* \text{ schwach.}$$

Sei F^* die Verteilungsfunktion zu Q^* . Dann existieren Stetigkeitspunkte $a < b$ von F^* mit

$$F^*(b) - F^*(a) \geq 1 - \frac{\varepsilon}{2}.$$

Daraus folgt für die Verteilungsfunktionen F_{n_i} von Q_{n_i}

$$Q_{n_i}((a, b]) = F_{n_i}(b) - F_{n_i}(a) \xrightarrow{(i \rightarrow \infty)} F^*(b) - F^*(a) \geq 1 - \frac{\varepsilon}{2},$$

was ein Widerspruch zu

$$\begin{aligned} Q_{n_i}((a, b]) &\leq Q_{n_i}([-n_i, n_i]) \\ &\text{(für } i \text{ genügend groß)} \\ &\stackrel{s.o.}{<} 1 - \varepsilon \end{aligned}$$

ist.

b) Wir zeigen: $\mathcal{Q}$ gleichmäßig straff $\Rightarrow \mathcal{Q}$ relativ kompakt.

Sei $(Q_n)_n$ beliebige Folge in $\mathcal{Q}$. Sei F_n die Verteilungsfunktion zu Q_n .

Nach dem Auswahlssatz von Helly existiert Indexteilfolge $(n_i)_i$ und maßdefinierende Funktion F mit $F_{n_i}(x) \rightarrow F(x)$ ($i \rightarrow \infty$) in allen Stetigkeitspunkten x von F .

Sei Q das zu F gehörende Maß, d. h.

$$Q((a, b]) = F(b) - F(a) \text{ für alle } a, b \in \mathbb{R}, a < b.$$

Wegen $F_{n_i}(x) \in [0, 1]$ gilt auch $F(x) \in [0, 1]$ für alle $x \in \mathbb{R}$, und mit der Stetigkeit des Maßes Q von unten folgt

$$Q((-\infty, x]) = F(x) - \lim_{a \rightarrow -\infty} F(a)$$

bzw.

$$F(x) = Q((-\infty, x]) + c \quad (x \in \mathbb{R})$$

mit

$$c = \lim_{a \rightarrow -\infty} F(a) \in [0, 1].$$

Wir zeigen im Folgenden: $Q(\mathbb{R}) \geq 1$ (*)

Daraus folgt die Behauptung, denn:

- Wegen $F(x) \in [0, 1]$ ($x \in \mathbb{R}$) und der Stetigkeit von Q von unten gilt

$$Q(\mathbb{R}) = \lim_{n \rightarrow \infty} Q((n, n]) = \lim_{n \rightarrow \infty} (F(n) - F(-n)) \leq 1,$$

also impliziert (*) dass Q ein W-Maß ist.

- Gilt $Q(\mathbb{R}) = 1$, so ist

$$1 \geq \lim_{x \rightarrow \infty} F(x) \stackrel{s.o.}{=} \lim_{x \rightarrow \infty} (Q((-\infty, x]) + c) = Q(\mathbb{R}) + c = 1 + c,$$

woraus mit $c \in [0, 1]$ $c = 0$ folgt, was wiederum

$$F(x) = Q((-\infty, x]) \quad (x \in \mathbb{R})$$

und damit auch

$$Q_n \rightarrow Q \text{ schwach}$$

impliziert.

Nachweis von (*):

Sei $\varepsilon > 0$ beliebig.

Da Q gleichmäßig straff ist, existiert kompaktes $K \subseteq \mathbb{R}$ mit

$$Q_{n_i}(K) \geq 1 - \varepsilon \quad \text{für alle } i \in \mathbb{N}.$$

Wähle Stetigkeitspunkte a, b von F mit $K \subset (a, b]$.

Dann gilt

$$\begin{aligned} Q(\mathbb{R}) &\geq Q((a, b]) \\ &= F(b) - F(a) \\ &= \lim_{i \rightarrow \infty} (F_{n_i}(b) - F_{n_i}(a)) \\ &\quad \text{(nach Wahl der } F_{n_i}) \\ &= \lim_{i \rightarrow \infty} Q_{n_i}((a, b]) \\ &\geq \lim_{i \rightarrow \infty} Q_{n_i}(K) \\ &\quad \text{(da } K \subset (a, b]) \\ &\geq 1 - \varepsilon. \end{aligned}$$

↪ Beh. $\square$.

Der folgende Satz gestattet, Probleme der Verteilungskonvergenz in $\mathbb{R}$ zurückzuführen auf Probleme der punktweisen Konvergenz einer Folge von charakteristischen Funktionen.

Satz 7.14 (Stetigkeitssatz von Lévy-Cramér)

- a) W-Maße Q_n ($n \in \mathbb{N}$) bzw. Q auf $\mathcal{B}$ mit charakteristischen Funktionen φ_n bzw. φ ,
d. h.

$$\varphi_n(u) = \int_{\mathbb{R}} e^{iux} dQ_n(x) \quad (u \in \mathbb{R}).$$

Dann gilt:

$$Q_n \rightarrow Q \text{ schwach} \Leftrightarrow \forall u \in \mathbb{R} : \varphi_n(u) \rightarrow \varphi(u) \quad (n \rightarrow \infty).$$

- b) Genauer gilt:

$b_1)$ $Q_n \rightarrow Q$ schwach $\Rightarrow \forall u \in \mathbb{R} : \varphi_n(u) \rightarrow \varphi(u) \quad (n \rightarrow \infty)$.

- $b_2)$ Wenn für jedes $u \in \mathbb{R}$

$$\varphi(u) = \lim_{n \rightarrow \infty} \varphi_n(u)$$

existiert und $\varphi : \mathbb{R} \rightarrow \mathbb{C}$ im Punkt 0 stetig ist, dann ist φ charakteristische Funktion zu einem W-Maß Q und es gilt:

$$Q_n \rightarrow Q \text{ schwach.}$$

Beweis:

Da die charakteristische Funktion stetig ist, folgt a) aus b).

Nachweis von b_1):

Da $x \mapsto \cos(ux)$ und $x \mapsto \sin(ux)$ stetig und beschränkt sind, folgt aus $Q_n \rightarrow Q$ schwach

$$\begin{aligned}\varphi_n(u) &= \int_{\mathbb{R}} \cos(ux) dQ_n(x) + i \cdot \int_{\mathbb{R}} \sin(ux) dQ_n(x) \\ &\xrightarrow{(n \rightarrow \infty)} \int_{\mathbb{R}} \cos(ux) dQ(x) + i \cdot \int_{\mathbb{R}} \sin(ux) dQ(x) \\ &= \int_{\mathbb{R}} e^{iux} dQ(x) = \varphi(u).\end{aligned}$$

Nachweis von b_2):

Es genügt zu zeigen: $(Q_n)_n$ ist gleichmäßig straff.

Denn nach dem Satz von Prokhorov ist $(Q_n)_n$ dann auch relativ kompakt, also existiert zu $(Q_n)_n$ Teilfolge $(Q_{n_i})_i$ und W-Maß Q mit

$$Q_{n_i} \rightarrow Q \text{ schwach.}$$

Nach b_1) gilt dann für die charakteristischen Funktionen φ_n von Q_n :

$$\varphi_{n_i}(x) \rightarrow \varphi^*(x) \quad (i \rightarrow \infty) \quad \text{für alle } x \in \mathbb{R},$$

wobei φ^* die charakteristische Funktion zu Q ist.

Nach Voraussetzung impliziert dies aber

$$\varphi^* = \varphi,$$

also existiert W-Maß Q mit charakteristischer Funktion φ .

Ist nun $g : \mathbb{R} \rightarrow \mathbb{R}$ stetig und beschränkt, so existiert zu jeder Teilfolge $(n_k)_k$ von $(n)_n$ Teilteilstolge $(n_{k_l})_l$ mit $Q_{n_{k_l}} \rightarrow$ W-Maß schwach, da $(Q_n)_n$ gleichmäßig straff ist. Wie oben sieht man, dass dieses W-Maß gerade Q ist, was

$$\int g(x) dQ_{n_{k_l}}(x) \rightarrow \int g(x) dQ(x) \quad (l \rightarrow \infty) \text{ impliziert.}$$

Nachweis von: $\{Q_n : n \in \mathbb{N}\}$ ist gleichmäßig straff

Aus

$$\varphi(0) = \lim_{n \rightarrow \infty} \varphi_n(0) = \lim_{n \rightarrow \infty} 1 = 1$$

folgt mit der Stetigkeit von φ im Nullpunkt:

$$\frac{1}{v} \int_{-v}^v |1 - \varphi(u)| du \rightarrow 0 \quad \text{für } v \rightarrow 0.$$

Sei $\varepsilon > 0$ beliebig. Dann existiert $v > 0$ mit

$$\frac{1}{v} \int_{-v}^v |1 - \varphi(u)| du < \varepsilon$$

und mit $\varphi_n(u) \rightarrow \varphi(u)$ ($n \rightarrow \infty$) und dem Satz von der majorisierten Konvergenz folgt

$$\frac{1}{v} \int_{-v}^v |1 - \varphi_n(u)| du \xrightarrow{(n \rightarrow \infty)} \frac{1}{v} \int_{-v}^v |1 - \varphi(u)| du < \infty.$$

Also gilt: $\exists n_0 \forall n \geq n_0 : \frac{1}{v} \int_{-v}^v |1 - \varphi_n(u)| du < 2\varepsilon$.

Wir zeigen nun:

$$Q_n \left(\left\{ x \in \mathbb{R} : |x| \geq \frac{2}{v} \right\} \right) \leq \frac{1}{v} \int_{-v}^v |1 - \varphi_n(u)| du.$$

Dazu:

$$\begin{aligned} \frac{1}{v} \int_{-v}^v (1 - \varphi_n(u)) du &= \frac{1}{v} \int_{-v}^v \left(1 - \int_{\mathbb{R}} e^{iux} dQ_n(x) \right) du \\ &= \frac{1}{v} \int_{\mathbb{R}} \left\{ \int_{-v}^v (1 - e^{iux}) du \right\} dQ_n(x) \\ &\quad \text{(nach Fubini und da } Q_n \text{ W-Ma\ss)} \\ &= \frac{1}{v} \int_{\mathbb{R}} \left(2v - 2 \cdot \frac{\sin vx}{x} \right) dQ_n(x) = \int_{\mathbb{R}} \underbrace{\left(2 - 2 \cdot \frac{\sin vx}{vx} \right)}_{\substack{\geq 0 \\ \text{(da } |\sin(u)| \leq u \text{ f\"ur } u \in \mathbb{R})}} dQ_n(x) \end{aligned}$$

impliziert

$$\begin{aligned} \frac{1}{v} \int_{-v}^v |1 - \varphi_n(u)| du &\geq \left| \frac{1}{v} \int_{-v}^v (1 - \varphi_n(u)) du \right| \\ &= \int_{\mathbb{R}} 2 \cdot \left(1 - \frac{\sin vx}{vx} \right) dQ_n(x) \\ &\geq \int_{\{x \in \mathbb{R} : |x| \geq \frac{2}{v}\}} 2 \cdot \left(1 - \underbrace{\frac{\sin vx}{vx}}_{\substack{\leq \frac{1}{v \cdot \frac{2}{v}} = \frac{1}{2} \\ \text{im Integrationsbereich}}} \right) dQ_n(x) \\ &\geq \int_{\{x \in \mathbb{R} : |x| \geq \frac{2}{v}\}} 2 \cdot \left(1 - \frac{1}{2} \right) dQ_n(x) \\ &= Q_n \left(\left\{ x \in \mathbb{R} : |x| \geq \frac{2}{v} \right\} \right), \text{ w.z.z.w.} \end{aligned}$$

Abschluss des Beweises:

Bisher gezeigt: $\forall \varepsilon > 0 \exists v > 0 \exists n_0 \forall n \geq n_0$:

$$Q_n \left(\left\{ x \in \mathbb{R} : |x| \geq \frac{2}{v} \right\} \right) < 2\varepsilon.$$

Sei nun $\varepsilon > 0$ beliebig.

Wir w\u00e4hlen dann $K = [-a, a] \subseteq \mathbb{R}$ so, dass gilt:

1. $K^c \subset \left(-\frac{2}{v}, \frac{2}{v}\right)^c$

2. $Q_n(K^c) \leq 2\varepsilon$ für $n = 1, 2, \dots, n_0$.

Hierbei gilt 1. für a groß genug immer, und 2. ist möglich, da für $a \rightarrow \infty$

$$Q_n([-a, a]^c) \rightarrow Q_n(\emptyset) = 0$$

konvergiert (nach der Stetigkeit des W-Maßes Q_n von oben).

Dann gilt aber für $n \leq n_0 : Q_n(K^c) \leq 2\varepsilon$, und für $n > n_0$:

$$Q_n(K^c) \leq Q_n(\{x \in \mathbb{R} : |x| \geq \frac{2}{v}\}) \stackrel{s.o.}{\leq} 2\varepsilon,$$

womit “ $\{Q_n : n \in \mathbb{N}\}$ gleichmäßig straff” gezeigt ist.

↪ Beh. $\square$

Bemerkung 7.15 Die obigen Definitionen und Sätze lassen sich auf W-Maße auf $\mathcal{B}_k$ bzw. k -dim. Zufallsvektoren übertragen.

8 Gesetze der großen Zahlen

Definition 8.1. Eine Folge $(X_n)_{n \in \mathbb{N}}$ von integrierbaren reellen ZVn auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ genügt dem **schwachen** bzw. **starken Gesetz der großen Zahlen**, wenn gilt:

$$\frac{1}{n} \sum_{k=1}^n (X_k - EX_k) \rightarrow 0 \quad (n \rightarrow \infty) \text{ nach Wahrscheinlichkeit bzw. P-f.s.}$$

Satz 8.2 (Kolmogorovsches starkes Gesetz der großen Zahlen).

Für eine unabhängige Folge $(X_n)_{n \in \mathbb{N}}$ identisch verteilter integrierbarer reeller ZVn gilt

$$\frac{1}{n} \sum_{k=1}^n X_k \rightarrow EX_1 \quad (n \rightarrow \infty) \text{ f.s.}$$

Beweis:

1. Schritt: oBdA $X_1 \geq 0$ f.s.

Denn ist die Aussage für nichtnegative ZV bereits gezeigt, so folgt daraus der allgemeine Fall wie folgt:

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n X_i &= \frac{1}{n} \sum_{i=1}^n X_i^+ - \frac{1}{n} \sum_{i=1}^n X_i^- \\ &\rightarrow E(X_1^+) - E(X_1^-) = EX_1 \quad \text{f.s.,} \end{aligned}$$

da mit $X_1, X_2, \dots$ auch $X_1^+, X_2^+, \dots$ und $X_1^-, X_2^-, \dots$ unabhängig, identisch verteilt und integrierbar sind.

2. Schritt: Sei im Folgenden $X_1 \geq 0$ f.s. und

$$X'_i = X_i \cdot \chi_{[X_i \leq i]}, S'_n = \sum_{i=1}^n X'_i, S_n = \sum_{i=1}^n X_i$$

und $k_n = \lfloor \vartheta^n \rfloor$ für ein beliebiges $\vartheta > 1$.

Dann gilt für $\varepsilon > 0$:

$$\begin{aligned}
& \sum_{n=1}^{\infty} P\left(\frac{|S'_{k_n} - ES'_{k_n}|}{k_n} > \varepsilon\right) \\
&= \sum_{n=1}^{\infty} P(|S'_{k_n} - ES'_{k_n}| > k_n \cdot \varepsilon) \\
&\leq \sum_{n=1}^{\infty} \frac{V(S'_{k_n})}{\varepsilon^2 \cdot k_n^2} \\
&\quad (\text{nach der Ungleichung von Tschebyschev}) \\
&= \sum_{n=1}^{\infty} \frac{1}{\varepsilon^2 \cdot k_n^2} \cdot \sum_{k=1}^{k_n} V(X'_k) \\
&\quad (\text{Unabhängigkeit der } X'_1, \dots, X'_{k_n}) \\
&\leq \sum_{n=1}^{\infty} \frac{1}{\varepsilon^2 \cdot k_n^2} \sum_{k=1}^{k_n} E(X_k^2 \cdot \chi_{[X_k \leq k]}) \\
&\leq \sum_{n=1}^{\infty} \frac{1}{\varepsilon^2 \cdot k_n^2} \cdot \sum_{k=1}^{k_n} E(X_k^2 \cdot \chi_{[X_k \leq k_n]}) \\
&= \sum_{n=1}^{\infty} \frac{1}{\varepsilon^2} \cdot \frac{1}{k_n} \cdot E(X_1^2 \cdot \chi_{[X_1 \leq k_n]})
\end{aligned}$$

3. Schritt: Für $x > 0$ gilt mit $n_0 = \min\{n \in \mathbb{N} : x \leq k_n\} = \min\{n \in \mathbb{N} : x \leq \lfloor \vartheta^n \rfloor\} : k_{n_0}$ erfüllt $\lfloor \vartheta^{n_0} \rfloor \geq x$ und $\vartheta^{n_0} \geq \lfloor \vartheta^{n_0} \rfloor \geq x$:

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{1}{k_n} \cdot \chi_{[x \leq k_n]} &= \sum_{n=n_0}^{\infty} \frac{1}{\lfloor \vartheta^n \rfloor} && \leq && \sum_{n=n_0}^{\infty} \frac{2}{\vartheta^n} = \frac{2}{\vartheta^{n_0}} \cdot \frac{1}{1-\frac{1}{\vartheta}} \\
&&& \uparrow && \\
&&& \lfloor \vartheta^n \rfloor \geq \frac{\vartheta^n}{2} \text{ wegen } \vartheta^n > 1 && \\
&&& (\text{für } z > 1 \text{ gilt } \lfloor z \rfloor > z/2) && \\
&&& \leq \frac{2}{x} \cdot \frac{\vartheta}{\vartheta-1}, &&
\end{aligned}$$

also gilt

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{1}{\varepsilon^2} \cdot \frac{1}{k_n} \cdot E(X_1^2 \cdot \chi_{[X_1 \leq k_n]}) &= \frac{1}{\varepsilon^2} \cdot E\left(X_1^2 \cdot \sum_{n=1}^{\infty} \frac{1}{k_n} \cdot \chi_{[X_1 \leq k_n]}\right) \\
&\stackrel{\text{s.o.}}{\leq} \frac{1}{\varepsilon^2} \cdot E\left(X_1^2 \cdot \frac{2}{X_1} \cdot \frac{\vartheta}{\vartheta-1}\right) \\
&= \frac{1}{\varepsilon^2} \cdot \frac{2\vartheta}{\vartheta-1} \cdot EX_1 < \infty
\end{aligned}$$

wegen $EX_1 < \infty$.

4. Schritt:

Also gilt für jedes $\varepsilon > 0$

$$\sum_{n=1}^{\infty} P\left(\frac{|S'_{k_n} - ES'_{k_n}|}{k_n} > \varepsilon\right) < \infty,$$

woraus mit dem Lemma von Borel-Cantelli folgt:

$$\frac{S'_{k_n} - ES'_{k_n}}{k_n} \rightarrow 0 \quad \text{f.s.}$$

5. Schritt:

Es gilt

$$\begin{aligned} \frac{1}{k_n} E S'_{k_n} &= \frac{1}{k_n} \sum_{k=1}^{k_n} E[X_k \cdot \chi_{[X_k \leq k]}] \\ &= \frac{1}{k_n} \sum_{k=1}^{k_n} E[X_1 \cdot \chi_{[X_1 \leq k]}] \rightarrow E X_1 \end{aligned}$$

da nach dem Satz von der monotonen Konvergenz gilt

$$\lim_{k \rightarrow \infty} E[X_1 \cdot \chi_{[X_1 \leq k]}] = E X_1.$$

6. Schritt:

Weiter gilt:

$$\begin{aligned} \sum_{n=1}^{\infty} P[X_n \neq X'_n] &= \sum_{n=1}^{\infty} P[X_1 > n] \\ &\leq \sum_{n=1}^{\infty} \int_{n-1}^n P[X_1 > t] dt \stackrel{\text{mon. Konvergenz}}{=} \int_0^{\infty} P[X_1 > t] dt \\ &= E X_1 < \infty. \end{aligned}$$

Nach dem Lemma von Borel-Cantelli gilt damit mit Wk. 1, dass X_n mit X'_n für alle bis auf endlich viele Indices übereinstimmt, woraus folgt

$$\frac{S_{k_n} - S'_{k_n}}{k_n} \rightarrow 0 \quad \text{f.s.}$$

Insgesamt erhält man

$$\begin{aligned} \frac{S_{k_n}}{k_n} &= \underbrace{\frac{S_{k_n} - S'_{k_n}}{k_n}}_{\substack{\rightarrow 0 \text{ f.s.} \\ \text{s.o.}}} + \underbrace{\frac{S'_{k_n} - E S'_{k_n}}{k_n}}_{\substack{\rightarrow 0 \text{ f.s.} \\ \text{siehe 4.}}} + \underbrace{\frac{E S'_{k_n}}{k_n}}_{\rightarrow E X_1} \\ &\rightarrow E X_1 \quad \text{f.s.} \end{aligned}$$

7. Schritt:

Wegen $X_i \geq 0$ f. s. gilt für $k_n \leq i \leq k_{n+1}$ mit Wk. 1:

$$\frac{k_n}{k_{n+1}} \cdot \frac{S_{k_n}}{k_n} \leq \frac{k_n \cdot S_i}{(k_{n+1}) \cdot k_n} \leq \frac{S_i}{i} \leq \frac{S_{k_{n+1}}}{i} \leq \frac{S_{k_{n+1}}}{k_{n+1}} \cdot \frac{k_{n+1}}{k_n}.$$

Mit

$$\frac{k_{n+1}}{k_n} = \frac{[\vartheta^{n+1}]}{[\vartheta^n]} \leq \frac{\vartheta^{n+1}}{\vartheta^n - 1} \rightarrow \vartheta \quad (n \rightarrow \infty) \text{ wegen } \vartheta^n \rightarrow \infty \quad (n \rightarrow \infty)$$

und

$$\frac{k_n}{k_{n+1}} = \frac{[\vartheta^n]}{[\vartheta^{n+1}]} \geq \frac{\vartheta^n - 1}{\vartheta^{n+1}} \rightarrow \frac{1}{\vartheta} \quad (n \rightarrow \infty)$$

folgt daraus und aus 6.: $\frac{1}{\vartheta} EX_1 \leq \liminf_{n \rightarrow \infty} \frac{S_n}{n} \leq \limsup_{n \rightarrow \infty} \frac{S_n}{n} \leq EX_1 \cdot \vartheta$ f.s.

Mit $\vartheta \downarrow 1$ folgt die Behauptung. □

Bemerkung 8.3.

- a) In Satz 8.2 darf die Integrierbarkeitsvoraussetzung nicht weggelassen werden.
- b) Die Voraussetzung der Unabhängigkeit in Satz 8.7 kann zur Voraussetzung der paarweisen Unabhängigkeit abgeschwächt werden (N. Etemadi, Z. Wahrscheinlichkeitstheorie verw. Gebiete 55 (1981), 119-122).

Satz 8.4 (Kriterium von Kolmogorov für das starke Gesetz der großen Zahlen).

Eine unabhängige Folge $(X_n)_{n \in \mathbb{N}}$ quadratisch integrierbarer reeller ZVn mit

$$\sum_{n=1}^{\infty} n^{-2} V(X_n) < \infty$$

genügt dem starken Gesetz der großen Zahlen.

Ohne Beweis.

Satz 8.5 (Tschebyschev). Eine Folge $(X_n)_{n \in \mathbb{N}}$ quadratisch integrierbarer paarweise unkorrelierter [d.h. es gilt $\forall_{j \neq k} E[(X_j - EX_j)(X_k - EX_k)] = 0$] reeller ZVn mit

$$n^{-2} \sum_{k=1}^n V(X_k) \rightarrow 0 \quad (n \rightarrow \infty)$$

genügt dem schwachen Gesetz der großen Zahlen.

Beweis: Siehe Übungen. □

Bemerkung 8.6 (Folgerung aus obigen Sätzen). Für eine unabhängige Folge (X_n) identisch verteilter reeller ZVn mit $P[X_1 = 1] = p$, $P[X_1 = 0] = 1 - p$ (festes $p \in [0, 1]$) gilt

$$\frac{1}{n} \sum_{k=1}^n X_k \rightarrow p \quad (n \rightarrow \infty) \text{ P-f.s. und nach Wahrscheinlichkeit}$$

(Borelsches starkes Gesetz der großen Zahlen bzw. Bernoullisches schwaches Gesetz der großen Zahlen).

9 Zentrale Grenzwertsätze

Aussagen über die Approximation von Verteilungen, insbesondere der Verteilungen von Summen von Zufallsvariablen, durch die Normalverteilung im Sinne der schwachen Konvergenz werden als zentrale Grenzwertsätze bezeichnet.

Satz 9.1 (Zentraler Grenzwertsatz von Lindeberg-Lévy im 1-dim. Fall).

Sei $(X_n)_{n \in \mathbb{N}}$ eine unabhängige Folge identisch verteilter quadratisch integrierbarer reeller ZVn mit $EX_1 =: a$, $V(X_1) =: \sigma^2$ mit $\sigma > 0$. Dann gilt:

$$\frac{1}{\sqrt{n}\sigma} \sum_{k=1}^n (X_k - a) \xrightarrow{\mathcal{D}} N(0, 1)\text{-verteilte reelle ZV.}$$

Sprechweise:

$\sum_{k=1}^n X_k$ ist bei großem n annähernd $\mathcal{N}(n \cdot a, n \cdot \sigma^2)$ -verteilt.

Beweis:

oBdA $a = 0$ und $\sigma^2 = 1$, ansonsten ersetzt man X_k durch $\frac{X_k - a}{\sigma}$.

Sei $\varphi_{\frac{1}{\sqrt{n}} \sum_{k=1}^n X_k}$ die charakteristische Funktion von $\frac{1}{\sqrt{n}} \sum_{k=1}^n X_k$. Nach dem Stetigkeitssatz von Lévy-Cramér genügt es dann zu zeigen:

$$\varphi_{\frac{1}{\sqrt{n}} \sum_{k=1}^n X_k}(u) \rightarrow e^{-u^2/2} \quad (u \in \mathbb{R}).$$

Hierbei ist $u \mapsto e^{-u^2/2}$ die charakteristische Funktion einer $\mathcal{N}(0, 1)$ -verteilten Zufallsvariablen.

Da $X_1, X_2, \dots$ unabhängig und identisch verteilt sind, gilt

$$\begin{aligned} \varphi_{\frac{1}{\sqrt{n}} \sum_{k=1}^n X_k}(u) &= E e^{iu \left(\frac{1}{\sqrt{n}} \sum_{k=1}^n X_k \right)} \\ &= E \prod_{k=1}^n e^{i \frac{u}{\sqrt{n}} X_k} \\ &= \prod_{k=1}^n E e^{i \frac{u}{\sqrt{n}} X_k} \\ &= \left(\varphi_{X_1} \left(\frac{u}{\sqrt{n}} \right) \right)^n, \end{aligned}$$

also ist zu zeigen:

$$\left(\varphi_{X_1} \left(\frac{u}{\sqrt{n}} \right) \right)^n \rightarrow e^{-u^2/2} \quad (n \rightarrow \infty).$$

Wegen $EX_1^2 < \infty$ ist φ_{X_1} zweimal stetig differenzierbar, und damit gilt nach Taylor:

$$\begin{aligned} \varphi_{X_1}(u) &= \varphi_{X_1}(0) + \varphi'_{X_1}(0) \cdot u + \frac{1}{2} \cdot \varphi''_{X_1}(0) \cdot u^2 + \rho(u) \\ &= 1 + i \cdot EX_1 \cdot u + \frac{1}{2} \cdot (-1) \cdot EX_1^2 \cdot u^2 + \rho(u) \\ &= 1 - \frac{1}{2} \cdot u^2 + \rho(u) \end{aligned}$$

(wegen $EX_1 = 0$ und $EX_1^2 = 1$ nach Voraussetzung), wobei für das Restglied $\rho(u)$ gilt:

$$\frac{\rho(u)}{u^2} \rightarrow 0 \quad (u \rightarrow 0)$$

Die letzte Beziehung sieht man so ein:

$$\frac{\rho(u)}{\frac{1}{2}u^2} = (\operatorname{Re}\varphi''_{X_1}(\xi_u) - \operatorname{Re}\varphi''_{X_1}(0)) + i \cdot (\operatorname{Im}\varphi''_{X_1}(\mu_u) - \operatorname{Im}\varphi''_{X_1}(0))$$

mit $|\xi_u| \leq u$ und $|\mu_u| \leq u$, woraus mit der Stetigkeit von φ''_{X_1} die Behauptung folgt.

Damit gilt für $u \in \mathbb{R}$ fest

$$\begin{aligned} \left(\varphi_{X_1}\left(\frac{u}{\sqrt{n}}\right)\right) &= \left(1 - \frac{1}{2} \cdot \frac{u^2}{n} + \rho\left(\frac{u}{\sqrt{n}}\right)\right)^n \\ &= \left(1 + \frac{-\frac{1}{2}u^2 + n \cdot \rho\left(\frac{u}{\sqrt{n}}\right)}{n}\right)^n. \end{aligned}$$

Mit

$$n \cdot \rho\left(\frac{u}{\sqrt{n}}\right) = u^2 \cdot \frac{\rho(u/\sqrt{n})}{(u/\sqrt{n})^2} \rightarrow u^2 \cdot 0 = 0 \quad (n \rightarrow \infty)$$

und

$$\left(1 + \frac{x_n}{n}\right)^n \rightarrow e^x \quad (n \rightarrow \infty)$$

für jede Folge $(x_n)_n$ mit $x_n \rightarrow x$ ($n \rightarrow \infty$)

(der letzte Grenzwert folgt z. B. aus

$$\left(1 + \frac{x_n}{n}\right)^n = e^{n \cdot \ln\left(1 + \frac{x_n}{n}\right)} = e^{x_n \frac{\ln\left(1 + \frac{x_n}{n}\right)}{\frac{x_n}{n}}}$$

und

$$\lim_{z \rightarrow 0} \frac{\ln(1+z)}{z} \stackrel{\text{de l'Hospital}}{=} \lim_{z \rightarrow 0} \frac{\frac{1}{1+z}}{1} = 1 \quad)$$

folgt die Behauptung. □

Korollar 9.2 (Zentraler Grenzwertsatz von de Moivre und Laplace).

Für eine unabhängige Folge $(X_n)_{n \in \mathbb{N}}$ identisch verteilter reeller ZVn auf $(\Omega, \mathcal{A}, P)$ mit $P[X_1 = 1] = p$, $P[X_1 = 0] = 1 - p =: q$ ($0 < p < 1$) gilt

$$\forall_{\substack{\alpha < \beta \\ \in \mathbb{R}}} P \left[\alpha \stackrel{(\text{=})}{<} \frac{\sum_{k=1}^n X_k - np}{\sqrt{npq}} \stackrel{(\text{=})}{<} \beta \right] \rightarrow \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-t^2/2} dt \quad (n \rightarrow \infty).$$

Satz 9.3 (Zentraler Grenzwertsatz von Lindeberg).

Die Folge $(X_n)_{n \in \mathbb{N}}$ quadratisch integrierbarer reeller ZVn mit $EX_1^2 > 0$, $\forall_n EX_n = 0$

sei unabhängig und erfülle — mit $s_n^2 := \sum_{i=1}^n EX_i^2$ ($n \in \mathbb{N}$) — die klassische Lindeberg-Bedingung

$$\forall_{\varepsilon > 0} \frac{1}{s_n^2} \sum_{i=1}^n E(X_i^2 \chi_{[|X_i| > \varepsilon s_n]}) \rightarrow 0.$$

Dann gilt:

$$\frac{1}{s_n} \sum_{i=1}^n X_i \xrightarrow{\mathcal{D}} N(0, 1)\text{-verteilte ZV.}$$