Technische Universität Dresden

Skript:

Maßtheorie und Stochastik, Teil II

Verfasser Franziska Kühn

Daten Prof. Dr. Zoltán Sasvári
Sommersemester 2010
Grundstudium

Inhaltsverzeichnis

II Wahrscheinlichkeitstheorie **3**

1 Grundbegriffe **4**
1.1 Wahrscheinlichkeitsraum 4
1.2 Zufallsvariablen, Verteilungen 10
1.3 Unabhängige Zufallsvariablen 15
1.4 Numerische Charakteristik von Zufallsgrößen 20
1.5 Bedingte Wahrscheinlichkeit 28

2 Gesetze **32**
2.1 0-1-Gesetze 32
2.2 Einige Ungleichungen 38
2.3 Konvergenz von Zufallsgrößen 44
2.4 Starkes und schwaches Gesetz der großen Zahlen 53
2.5 Satz von Gliwenko-Cantelli 59

3 Charakteristische Funktionen **61**
3.1 Definition + einige Eigenschaften 61
3.2 Beispiele 64
3.3 Umkehrfunktionen 65

4 Zentraler Grenzwertsatz **67**
4.1 Satz von Moivre und Laplace 67

Teil II

Wahrscheinlichkeitstheorie

1 Grundbegriffe

1.1 Wahrscheinlichkeitsraum

1.1.1 Mathematisches Modell

- Aufgabe der Wahrscheinlichkeitstheorie: Das Problem des Zufalls mit einem exakten mathematischen Modell erfassen. Was ist Zufall? Was sind zufällige Ereignisse?
- Beispiele:
  (i). Werfen eines Würfels, zufälliges Ereignis ist beispielsweise Auftreten einer geraden Augenzahl
  (ii). Anzahl von Telefongesprächen in einer Telefonzentrale in einer bestimmten Stunde
  (iii). Geburten (Junge/Mädchen)
- Schon im 16. Jahrhundert Beschäftigung mit Aufgaben wahrscheinlichkeitstheoretischen Charakters. Lösungen mit kombinatorischen Methoden.

  Ausarbeitung eines exakten mathematischen Modells erst in der Zeit von 1903-1933 (u.a. von Borel, Wiener, Paley, Zygmund, Lommicki, Steinhaus, Kolmogorov: „Grundbegriffe der Wahrscheinlichkeitsrechnung"). Etwa in dieser Zeit: abstraktes Maß, Integral
- Mathematisches Modell:
  (i). mathematisches Modell für Ereignisse:

  Ereignisalgebra (Ereignisse A,B $\Rightarrow$ nicht A, A oder B, A und B). Modell: Algebra von Teilmengen einer Menge (Stone 1933). Eine Familie $\mathcal{A}$ von Teilmengen einer Menge Ω heißt Algebra , falls

  (1) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$
  (2) $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$

  Meist: Bildung einer σ-Algebra $\sigma(\mathcal{A})$

  (ii). mathematisches Modell für die Wahrscheinlichkeit:

  Tritt ein zufälliges Ereignis A in n Versuchen m-mal ein, so heißt m die absolute Häufigkeit und
  $$h_n(A) := \frac{m}{n}$$
  die relative Häufigkeit von A bei n Versuchen. Eigenschaften:

  (1) $0 \leq h_n(A) \leq 1$
  (2) h_n(sicheres Ereignis) $= 1$, h_n(unmögliches Ereignis)$= 0$
  (3) h_n(A oder B) $= h_n(A) + h_n(B)$, falls A und B nicht gleichzeitig eintreten können

Die relative Häufigkeit zeigt eine Stabilität, wenn n groß ist.

Beispiel: Wurf eines Geldstücks, A: Auftreten einer Zahl

	n	$H_n(A)$	$h_n(A)$
Buffon	4040	2048	0.5080
Pearson	12000	6019	0.5016
Pearson	24000	12012	0.5005

„Grenzwert“ $n \to \infty$ von relativen Häufigkeiten: endlich additives Maß h

1.1.2 Definition

- Es sei $\Omega \neq \emptyset$, $\mathcal{A}$ eine σ-Algebra von Teilmengen von Ω und $\mathbb{P}$ ein Maß auf $(\Omega, \mathcal{A})$ mit $\mathbb{P}(\Omega) = 1$. Das Tripel $(\Omega, \mathcal{A}, \mathbb{P})$ heißt Wahrscheinlichkeitsraum. Die Elemente von Ω heißen Elementarereignisse, die Elemente von $\mathcal{A}$ Ereignisse.

  Beachte: Im Allgemeinen $\{w\} \notin \mathcal{A}$ für $w \in \Omega$ (d.h. Elementarereignisse sind i.A. keine Ereignisse).
- $\emptyset$ heißt unmögliches Ereignis, Ω das sichere Ereignis.
- Notation: $\Omega \backslash A$ für $A \in \mathcal{A}$ (nicht A) wird auch mit A^c oder $\bar{A}$ bezeichnet.
- $\mathbb{P}$ heißt Wahrscheinlichkeitsmaß und $\mathbb{P}(A)$ für $A \in \mathcal{A}$ heißt Wahrscheinlichkeit von A.
- Ereignisse A mit $\mathbb{P}(A) = 1$ heißen fast sichere Ereignisse, Ereignisse A mit $\mathbb{P}(A) = 0$ fast unmögliche Ereignisse. Statt $\mathbb{P}$-fast überall sagt man auch fast sicher oder mit Wahrscheinlichkeit 1.

Sprechweise: Seien $A, B \in \mathcal{A}$.

(i). $A \subset B$: Aus A folgt B, A impliziert B

(ii). $A \cap B = \emptyset$: A und B sind unvereinbar

(iii). $A \backslash B$: Es tritt A aber nicht B ein.

(iv). Sei $(A_j)_{j \in \mathbb{N}} \subset \mathcal{A}$. Dann $\bigcup_{j \in \mathbb{N}} A_j$: mindestens 1 A_n tritt ein, $\bigcap_{j \in \mathbb{N}} A_j$: alle A_n treten ein,

$$\lim_{n \to \infty} \inf A_n := \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m$$

ab einem bestimmten Index treten alle A_n ein,

$$\lim_{n \to \infty} \sup A_n := \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$$

unendlich viele der Ereignisse A_n treten ein.

1.1.3 Beispiele

(i). Werfen eines Würfels mit den Augenzahlen $1, \ldots, 6$

$$\Omega := \{1, 2, 3, 4, 5, 6\} \qquad \mathcal{A} := \mathcal{P}(\Omega)$$

erscheinende Augenzahl ist $i \to \{i\}$, erscheinende Augenzahl ist gerade $\to \{2, 4, 6\}$. Wenn es sich um einen idealen Würfel handelt, ist $\mathbb{P}(\{i\}) = \frac{1}{6}$.

(ii). Schießen auf eine Schießscheibe mit Radius $r = 20$ cm

$$\begin{aligned}\Omega &:= \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 20^2\} \\ \mathcal{A} &:= \{B \in \mathcal{B}(\mathbb{R}^2) : B \subset \Omega\}\end{aligned}$$

Nehmen wir an, dass die Schüsse gleichmäßig verteilt sind, d.h.

$$\mathbb{P}(B) = \frac{\lambda^2(B)}{\lambda^2(\Omega)}$$

für $B \in \mathcal{A}$. Andere Möglichkeiten: $\mathcal{A}$ als Familie aller lebesgue-messbaren Mengen $\subset \Omega$, $\mathcal{A}$ Sammlung von endlich vielen Teilmengen von Ω (z.B. Raster)

1.1.4 Aufgabe

Jede endliche σ-Algebra besitzt 2^n Elemente mit $n \in \mathbb{N}$.

1.1.5 Satz

Für jedes Wahrscheinlichkeitsmaß $\mathbb{P}$ gilt:

(i). $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$

(ii). $\mathbb{P}(B \backslash A) = \mathbb{P}(B) - \mathbb{P}(A)$ wenn $A \subset B$

(iii). $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

(iv). Für $A_1 \subset A_2 \ldots$ gilt

$$\mathbb{P}\left(\bigcup_{j \in \mathbb{N}} A_j\right) = \lim_{n \to \infty} \mathbb{P}(A_n)$$

(v). Für $A_1 \supset A_2 \ldots$ gilt

$$\mathbb{P}\left(\bigcap_{j \in \mathbb{N}} A_j\right) = \lim_{n \to \infty} \mathbb{P}(A_n)$$

Beweis: Bekannt aus Maßtheorie

1.1.6 Aufgabe (Poincaré)

$$\mathbb{P}\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \leq i_1 < \ldots < i_k \leq n} \mathbb{P}(A_{i_1} \cap \ldots \cap A_{i_k})$$

1.1.7 Aufgabe

Zwei Spieler S_1, S_2 spielen mit drei (idealen) Würfeln W_1, W_2, W_3 mit den folgenden Augenzahlen:

$$W_1 : 5, 7, 8, 9, 10, 18 \qquad W_2 : 2, 3, 4, 15, 16, 17 \qquad W_3 : 1, 6, 11, 12, 13, 14$$

Das Spiel: Zuerst wählt S_1 einen Würfel, dann S_2. Beide würfeln (mit dem gewählten Würfel); wer die größere Augenzahl hat, bekommt vom anderen 1 Euro. Sie Sind S_1, welchen Würfel würden Sie wählen?

1.1.8 Definition

Zwei Ereignisse $A, B \in \mathcal{A}$ heißen unabhängig (bzgl. $\mathbb{P}$), wenn $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$.

Allgemeiner: Eine Familie $\{A_i\}_{i \in I}$ von Ereignissen in $\mathcal{A}$ heißt unabhängig (bzgl. $\mathbb{P}$), wenn für jede nichtleere Teilmenge $\{i_1, \ldots, i_n\} \subset I$ gilt:

$$\mathbb{P}\left(\bigcap_{j=1}^{n} A_{i_j}\right) = \prod_{j=1}^{n} \mathbb{P}(A_{i_j})$$

1.1.9 Beispiele

(i). Zweimaliges Werfen eines Würfels und den zugehörigen Wahrscheinlichkeitsraum:

$$\Omega := \{1, \ldots, 6\} \times \{1, \ldots, 6\} \qquad \mathcal{A} := \mathcal{P}(\Omega)$$

Im Idealfall gilt $\mathbb{P}(\{(i,j)\}) = \frac{1}{36}$. Betrachte die beiden Ereignisse

- A: Beim 1. Wurf Augenzahl ≤ 3
- B: Beim 2. Wurf Augenzahl $= 6$

Dann:

$$\mathbb{P}(A) = \frac{18}{36} \qquad \mathbb{P}(B) = \frac{6}{36} \qquad \mathbb{P}(A \cap B) = \frac{3}{36} = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

d.h. A und B sind unabhängig.

(ii). Bezeichnungen wie in Beispiel (i) mit den Ereignissen

- A_1: beim 1. Wurf ungerade
- A_2: beim 2. Wurf ungerade
- A_3: Summe ungerade

Kombinatorik ergibt:

$$\begin{aligned}
\mathbb{P}(A_1 \cap A_2) &= \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \\
\mathbb{P}(A_1 \cap A_3) &= \mathbb{P}(A_1) \cdot \mathbb{P}(A_3) \\
\mathbb{P}(A_2 \cap A_3) &= \mathbb{P}(A_2) \cdot \mathbb{P}(A_3) \\
\mathbb{P}(A_1 \cap A_2 \cap A_3) &= 0 \neq \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot \mathbb{P}(A_3) > 0
\end{aligned}$$

d.h. A_1, A_2, A_3 sind paarweise unabhängig, aber nicht als Familie unabhängig. Aus der paarweisen Unabhängigkeit folgt also i.A. nicht die Unabhängigkeit.

1.1.10 Aufgabe

(i). $\{A, B\}$ unabhängig $\Rightarrow \{A, B^c\}$ unabhängig

(ii). $\{A_i\}_{i \in I}$ unabhängig $\Rightarrow \{A_i^c\}_{i \in I}$ unabhängig

1.1.11 Aufgabe

Für jede Folge von unabhängigen Ereignissen $\{A_n\}_{n \in \mathbb{N}}$ gilt:

$$\mathbb{P}\left(\lim_{n \to \infty} \sup A_n\right) = 1 - \lim_{n \to \infty} \lim_{N \to \infty} \prod_{m=n}^{N} (1 - \mathbb{P}(A_m))$$

1.1.12 Aufgabe

Gilt $\lim_{n\to\infty} \mathbb{P}(A_n) = 1$, so ist

$$\lim_{n\to\infty} \mathbb{P}(A \cap A_n) = \mathbb{P}(A)$$

für $A \in \mathcal{A}$.

1.1.13 Aufgabe

Es sei $\Omega := [0,1]$, $\mathcal{A}$ die σ-Algebra der Borelmengen in $[0,1]$ und $\mathbb{P}$ das Lebesgue-Maß auf $[0,1]$. Für jedes $n \in \mathbb{N}$ sei

$$A_n := \left[0, \frac{1}{2^n}\right] \cup \left[\frac{2}{2^n}, \frac{3}{2^n}\right] \cup \ldots \cup \left[\frac{2^n-2}{2^n}, \frac{2^n-1}{2^n}\right]$$

Zeigen Sie: Die Folge $\{A_n\}_{n\in\mathbb{N}}$ dieser Ereignisse ist unabhängig.

1.1.14 Aufgabe

Aus den Zahlen $1, 2, \ldots, n$ wird eine zufällig ausgewählt (jede hat die gleiche Wahrscheinlichkeit). Bestimmen Sie:

(i). $p_n := P(\text{die Zahl ist durch 3 oder 4 teilbar})$

(ii). $\lim_{n\to\infty} p_n$

1.1.15 Definition

Es sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $\mathcal{A}_1, \mathcal{A}_2$ seien Familien von Ereignissen aus $\mathcal{A}$. $\mathcal{A}_1$ und $\mathcal{A}_2$ heißen unabhängig, wenn

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2)$$

für alle $A_1 \in \mathcal{A}_1$, $A_2 \in \mathcal{A}_2$.

Allgemeiner: Eine Familie $\{\mathcal{A}_i\}_{i\in I}$ von Ereignisklassen heißt unabhängig, wenn für jede nichtleere Teilmenge $\{i_1, \ldots, i_n\} \subset I$ gilt:

$$\mathbb{P}\left(\bigcap_{k=1}^{n} A_{i_k}\right) = \prod_{k=1}^{n} \mathbb{P}(A_{i_k})$$

für alle $A_{i_k} \in \mathcal{A}_{i_k}$ $(k = 1, \ldots, n)$.

1.1.16 Lemma: Approximationssatz

Es sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $\mathcal{B}$ die von einer Algebra $\mathcal{B}_0 \subset \mathcal{A}$ erzeugte σ-Algebra. Dann existiert für jedes $A \in \mathcal{B}$ eine Folge $(A_n)_{n\in\mathbb{N}} \subset \mathcal{B}_0$ mit

(i). $\lim_{n\to\infty} \mathbb{P}(A_n \Delta A) = 0$

(ii). $\mathbb{P}(A) = \lim_{n\to\infty} \mathbb{P}(A_n)$

Notation:

$$A\Delta B := (A \backslash B) \cup (B \backslash A)$$

Beweis:

(i). Ist äquivalent zu

$$\lim_{n\to\infty} \mathbb{P}(A \backslash A_n) = \lim_{n\to\infty} \mathbb{P}(A_n \backslash A) = 0 \qquad (1)$$

Definiere $\mathcal{B}^*$ als Gesamtheit aller Ereignisse $A \in \mathcal{B}$ für die eine Folge $(A_n)_{n\in\mathbb{N}} \subset \mathcal{B}_0$ mit (1) bzw. (i) existiert. Dann gilt: $\mathcal{B}_0 \subset \mathcal{B}^* \subset \mathcal{B}$. Wir zeigen, dass $\mathcal{B}^*$ eine σ-Algebra ist, daraus folgt dann $\mathcal{B} = \mathcal{B}^*$ wegen

$$\mathcal{B} = \sigma(\mathcal{B}_0) \subset \underbrace{\sigma(\mathcal{B}^*)}_{\mathcal{B}^*} \subset \mathcal{B}$$

Zunächst: $\mathcal{B}^*$ ist eine Algebra (Übungsaufgabe, Hinweis:

$$(A \cap B)\Delta(C \cap D) \subset (A\Delta C) \cup (B\Delta D)$$

Sei nun $(B_j)_{j\in\mathbb{N}} \subset \mathcal{B}^*$ beliebig. Noch zu zeigen:

$$C := \bigcup_{j\in\mathbb{N}} B_j \in \mathcal{B}^*$$

Definiere $C_n := \bigcup_{j=1}^n B_j \in \mathcal{B}^*$, da B^* eine Algebra ist. Es existieren A_n in $\mathcal{B}_0$ mit

$$\mathbb{P}(A_n \backslash C_n) < \frac{1}{n} \qquad \mathbb{P}(C_n \backslash A_n) < \frac{1}{n} \qquad (2)$$

Behauptung:

$$\lim_{n\to\infty} \mathbb{P}(C\backslash A_n) = \lim_{n\to\infty} \mathbb{P}(A_n \backslash C) = 0$$

d.h. $C \in \mathcal{B}^*$. Die Behauptung folgt aus den Gleichungen

$$\begin{aligned} A_n\backslash C &\subset A_n\backslash C_n \\ C\backslash A_n &\subset (C\backslash C_n) \cup (C_n\backslash A_n) \end{aligned}$$

mit (2) und 1.1.5(v) ergibt sich durch Anwenden von $\mathbb{P}$ die Behauptung.

(ii). Folgt mit (i) aus

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P}(A_n \cap A) + \mathbb{P}(A_n^c \cap A) \\ &= \mathbb{P}(A_n) - \mathbb{P}(A_n \cap A^c) + \mathbb{P}(A_n^c \cap A) \\ &\to \lim_{n\to\infty} \mathbb{P}(A_n) + 0 + 0 \end{aligned}$$

1.1.17 Satz

Die von unabhängigen Ereignisalgebren $\mathcal{A}_0$, $\mathcal{B}_0$ erzeugten σ-Algebren $\mathcal{A} = \sigma(\mathcal{A}_0)$, $\mathcal{B} = \sigma(\mathcal{B}_0)$ sind unabhängig.

Beweis:

- Zu zeigen: Für alle $A \in \mathcal{A}, B \in \mathcal{B}$ gilt

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

  Bekannt für $A \in \mathcal{A}_0$, $B \in \mathcal{B}_0$.

- Wir wählen $(A_j)_{j\in\mathbb{N}} \subset \mathcal{A}_0$,$(B_j)_{j\in\mathbb{N}} \subset \mathcal{B}_0$ mit

$$\lim_{n\to\infty} \mathbb{P}(A\Delta A_n) = \lim_{n\to\infty} \mathbb{P}(B_n\Delta B) = 0$$

  wie in Lemma 1.1.16. Damit folgt wegen

$$\lim_{n\to\infty} \mathbb{P}((A_n \cap B_n)\Delta(A \cap B)) = 0$$

  mit Lemma 1.1.16(ii):

$$\begin{aligned} \lim_{n\to\infty} \mathbb{P}(A_n \cap B_n) &= \mathbb{P}(A \cap B) \\ \Rightarrow \lim_{n\to\infty} (\mathbb{P}(A_n) \cdot \mathbb{P}(B_n)) &= \mathbb{P}(A \cap B) \\ \Rightarrow \mathbb{P}(A) \cdot \mathbb{P}(B) &= \mathbb{P}(A \cap B) \end{aligned}$$

1.2 Zufallsvariablen, Verteilungen

1.2.1 Definition

Sei $(\Omega, \mathcal{A}, \mathcal{P})$ ein Wahrscheinlichkeitsraum und $(S, \mathcal{B})$ ein Maßraum. Eine Zufallsvariable mit Werten in S heißt jede $\mathcal{A}/\mathcal{B}$-messbare Abbildung $X : \Omega \to S$.
In den Fällen $S = \mathbb{R}$, $S = \mathbb{R}^n$, $S = \mathbb{C}$ ist im weiteren immer die σ-Algebra der Borelmengen. Für $S = \mathbb{R}$ heißt X Zufallsgröße, für $S = \mathbb{R}^n$ Zufallsvektor, für $S = \mathbb{C}$ komplexe Zufallsgröße.

1.2.2 Satz

(i). Summe von Zufallsvektoren, Produkt von Zufallsgrößen sind auch Zufallsvektoren- bzw. größen

(ii). Seien X Zufallsgröße, $f : \mathbb{R} \to \mathbb{R}$ borel-messbar. Dann ist $f(X)$ eine Zufallsgröße.

(iii). Sind $X_1, \dots, X_n$ Zufallsgrößen und $f : \mathbb{R}^n \to \mathbb{R}$ borel-messbar, dann ist $f(X_1, \dots, X_n)$ eine Zufallsgröße.

Bemerkung:

(i). Mit $f_1(x, y) = x + y$ und $f_2(x, y) = x \cdot y$ ist (i) Spezialfall von (iii). Auch $Y = (X_1, \dots, X_n)$ ist Zufallsvariable.

Beweis: Bekannt aus Maßtheorie

1.2.3 Definition

Es sei $X : \Omega \to S$ eine Zufallsvariable. Man führt die folgenden Notationen ein:

$$\begin{aligned} [X \in B] &:= \{w \in \Omega : X(w) \in B\} = X^{-1}(B) \\ \mathbb{P}([X \in B]) &=: \mathbb{P}[X \in B] = \mathbb{P}(X^{-1}(B)) \end{aligned}$$

Dann ist $\mu_X(B) := \mathbb{P}[X \in B], B \in \mathcal{B}$ ein Wahrscheinlichkeitsmaß auf $\mathcal{B}$, die Verteilung von X bzgl. $\mathbb{P}$ (Bildmaß). Die Verteilung heißt stetig, wenn $\mu(\{s\}) = 0$ für $s \in S$, $\{s\} \in \mathcal{B}$.
Sei jetzt $S = \mathbb{R}^n$. Die Verteilung μ_X heißt absolut stetig, wenn eine borel-messbare Funktion $p : \mathbb{R}^n \to [0, \infty)$ existiert, sodass

$$\mu_X(A) = \int_A p \, d\lambda$$

für $A \in \mathcal{B}(\mathbb{R}^n)$. Dabei heißt p Dichtefunktion von X oder μ_X.

Bemerkung:

(i). Eine borel-messbare Funktion $p : \mathbb{R}^n \to [0, \infty)$ ist genau dann die Dichte einer Verteilung, wenn sie λ-integrierbar ist und $\int_{\mathbb{R}^n} p \, d\lambda = 1$. Ist p die Dichte von μ_X, so schreibt man $d\mu_X = p \, d\lambda$. Absolut stetige Verteilungen sind stetig, die Umkehrung gilt nicht.

1.2.4 Beispiele

(i). Für jedes $s \in S$ sei ε_s (oder δ_s) das durch die Einheitsmasse in s definierte Wahrscheinlichkeitsmaß auf $\mathcal{B}$ (Einpunkt-Verteilung, Dirac-Maß)

$$\delta_s(B) = \begin{cases} 0 & s \notin B \\ 1 & s \in B \end{cases}$$

Eine Zufallsvariable besitzt eine Einpunkt-Verteilung genau dann, wenn sie fast sicher konstant ist.

(ii). Es sei $\{s_n\}$ eine Folge in S und $\{p_n\}$ eine Folge nichtnegativer Zahlen mit $\sum_{k=1}^{\infty} p_k = 1$. Dann ist

$$\mu := \sum_{n=1}^{\infty} p_n \cdot \varepsilon_{s_n}$$

ein Wahrscheinlichkeitsmaß. Jede solche Verteilung heißt diskret bzw. jede solche Zufallsvariable mit einer solchen Verteilung heißt diskret.

$$\mathbb{P}(X = s_n) = p_n$$

(iii). Es sei $p \in [0, 1]$, $n \in \mathbb{N}$ und

$$B_n^p := \sum_{k=0}^{n} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \cdot \varepsilon_k$$

Wegen

$$1 = (p + 1 - p)^n = \sum_{k=0}^{n} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

ist B_n^p ein diskretes Wahrscheinlichkeitsmaß auf $\mathbb{R}$, Binomialverteilung mit den Parametern n und p.

Beispiel: Wir betrachten eine Folge von Versuchen. In jedem dieser Versuche tritt ein gewisses Ereignis A mit Wahrscheinlichkeit p unabhängig von den Ausgängen der anderen Versuche ein. X_n sei die Anzahl des Eintretens von A in den ersten n Versuchen, mögliche Werte $\{0, \ldots, n\}$, p wie oben. X_n ist binomial-verteilt mit den Parametern n und p.

(iv). Polynomiale Verteilung (Verallgemeinerung von (iii)):

Folge von Versuchen, in jedem dieser Versuche treten gewisse unvereinbare Ereignisse $A_1, \ldots, A_r$ mit den Wahrscheinlichkeiten $p_1, \ldots, p_r$ mit $\sum_{k=1}^{r} p_k = 1$ unabhängig von den Ausgängen der anderen Versuche ein. X_i sei die Anzahl des Eintretens von A_i in den ersten n Versuchen $(i = 1, \ldots, r)$. X_i ist binomialverteilt.

$$X := (X_1, \ldots, X_r)$$

ist ein Zufallsvektor. $(x_1, \ldots, x_r) \in \mathbb{R}^r$ gehört zum Wertebereich von X fast sicher $\Leftrightarrow 0 \leq x_i \leq n$, $x_i \in \mathbb{N}$, $\sum_{i=1}^{r} x_i = n$.

$$\mathbb{P}(X_1 = x_1, \ldots, X_r = x_r) = \frac{n!}{\prod_{i=1}^{r} x_i!} \cdot \prod_{i=1}^{r} p_i^{x_i}$$

Beweis: Permutation mit Wiederholung. (Für $r = 2$: Binomialverteilung) X heißt polynomialverteilt mit den Parametern $n, r, p_1, \ldots, p_r$.

(v). Wegen $e^a = \sum_{k=0}^{\infty} \frac{a^k}{k!}$ ist für jedes $a \geq 0$

$$\Pi_a := \sum_{k=0}^{\infty} e^{-a} \cdot \frac{a^k}{k!} \cdot \varepsilon_k$$

ein diskretes Wahrscheinlichkeitsmaß auf $\mathbb{R}$. Für $a > 0$ heißt Π_a Poisson-Verteilung mit dem Parameter a.

(vi). Die Funktion

$$g_{a,\sigma}(x) := \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{(x-a)^2}{2\sigma^2}}$$

für $x \in \mathbb{R}$ ist eine Dichte.

Beweis:

- Es ist bekannt (Übung):
$$\int_0^\infty e^{-x^2}\,dx = \frac{\sqrt{\pi}}{2}$$
Damit:
$$\begin{aligned}\frac{1}{\sigma\cdot\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-a)^2}{2\sigma^2}}\,dx &= \frac{1}{\sqrt{\pi}}\cdot\int_{-\infty}^{\infty} e^{-u^2}\,du\\ &= \frac{2}{\sqrt{\pi}}\int_0^\infty e^{-u^2}\,du = 1\end{aligned}$$

Das zugehörige Maß $\nu_{a,\sigma}$ heißt Normalverteilung mit den Parametern a und σ.

(vii). Die Funktion
$$p(x) := \frac{1}{\pi\cdot(1+x^2)} \qquad (x\in\mathbb{R})$$
ist eine Dichte:
$$\int_{-\infty}^{\infty}\frac{1}{\pi\cdot(1+x^2)}\,dx = \frac{1}{\pi}\cdot\big[\arctan x\big]_{-\infty}^{\infty} = 1$$
Die zugehörige Verteilung heißt Cauchy-Verteilung.

(viii). Es sei $B\subset\mathbb{R}^n$ eine beliebige Borel-Menge mit $0<\lambda^n(B)<\infty$. Wir definieren
$$p_B(x) := \frac{1}{\lambda^n(B)}\cdot 1_B(x)$$
Dann ist p_B eine Dichte und die zugehörige Verteilung heißt Gleichverteilung.

(ix). Die Exponentialverteilung mit dem Parameter $\lambda>0$:
$$p(x) := \lambda\cdot e^{-\lambda x}\cdot 1_{[0,\infty)}(x) \qquad (x\in\mathbb{R})$$
ist eine Dichte.

(x). Die Gamma-Verteilung mit dem Parameter $a>0$, $\lambda>0$:
$$p(x) := \frac{\lambda^a\cdot x^{a-1}}{\Gamma(a)}\cdot e^{-\lambda x}\cdot 1_{(0,\infty)}(x)$$
ist eine Dichte, wobei
$$\Gamma(a) = \int_0^\infty x^{a-1}\cdot e^{-x}\,dx \qquad (a>0)$$

1.2.5 Aufgabe (Eigenschaften von Γ)

(i). $\Gamma(a)<\infty$ und p aus 1.2.4(x) ist eine Dichte.

(ii). $\Gamma(a+1) = a\cdot\Gamma(a)$

(iii). $\Gamma(n+1) = n!$ für $n\in\mathbb{N}$

(iv). $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

(v). Γ ist stetig.

1.2.6 Aufgabe

Bekannt aus Maßtheorie, Übungsaufgabe 35:
$$\lim_{T\to\infty}\int_0^T \frac{\sin x}{x}\,dx = \frac{\pi}{2}$$

1.2.7 Aufgabe

1.2.8 Aufgabe

1.2.9 Satz

Jedes Wahrscheinlichkeitsmaß μ auf $\mathbb{R}^n$ lässt sich eindeutig zerlegen als

$$\begin{aligned} \mu &= p_1 \cdot \mu_d + (1 - p_1) \cdot \mu_c \\ &= p_1 \cdot \mu_d + p_2 \cdot \mu_{ac} + p_3 \cdot \mu_s \end{aligned}$$

Hierbei sind p_1, p_2, p_3 nichtnegative reelle Zahlen mit $p_1 + p_2 + p_3 = 1$ und $\mu_d, \mu_s, \mu_{ac}, \mu_c$ Wahrscheinlichkeitsmaße mit den folgenden Eigenschaften:

(i). μ_d ist diskret.

(ii). μ_c ist stetig.

(iii). μ_{ac} ist absolut stetig.

(iv). μ_s ist stetig und singulär, d.h. es gibt eine λ-Nullmenge $B \subset \mathbb{R}^n$ mit $\mu_s(B) = 1$.

Beweis: Bekannt aus Maßtheorie (...?)

1.2.10 Satz

Ist μ_X die Verteilung einer Zufallsgröße X, so gilt

$$\int_\Omega g(X(w))\, d\mathbb{P}(w) = \int_\mathbb{R} g(x)\, d\mu_X(x)$$

für $g : \mathbb{R} \to \mathbb{R}$ borel-messbar mit $g \geq 0$ oder μ_X-integrierbar.

Beweis: Bekannt aus Maßtheorie, §14

1.2.11 Definition

Sind $X_1, \ldots, X_n$ Zufallsgrößen, so ist $Y := (X_1, \ldots, X_n)$ ein Zufallsvektor (s.1.2.2(iii)). Die Verteilung μ_Y von Y (auf $\mathbb{R}^n$) wird die gemeinsame Verteilung von $X_1, \ldots, X_n$ genannt.

1.2.12 Satz (Verallgemeinerung von 1.2.10)

Bezeichnungen wie in 1.2.11. Für jede borel-messbare Funktion $g : \mathbb{R}^n \to \mathbb{R}$, die nichtnegativ oder μ_Y-integrierbar ist, gilt

$$\int_\Omega g(Y(w))\, d\mathbb{P}(w) = \int_{\mathbb{R}^n} g(x_1, \ldots, x_n)\, d\mu_Y(x_1, \ldots, x_n)$$

Beweis: Bekannt aus Maßtheorie (§14)

Die Wahrscheinlichkeitsmaße auf $\mathbb{R}$ können mit Hilfe reeller Funktionen beschrieben werden:

1.2.13 Definition

Es sei μ ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Die Funktion

$$F(x) := \mu((-\infty, x)) \qquad (x \in \mathbb{R})$$

heißt die Verteilungsfunktion von μ. Ist μ die Verteilung einer Zufallsgröße X, so nennt man F auch die Verteilungsfunktion von X. Dann gilt für $x \in \mathbb{R}$:

$$\mathbb{P}[X < x] = F(x)$$

Beachte: Auch $F(x) := \mu((-\infty, x])$ als Definition möglich.

1.2.14 Aufgabe

(i). Jede Verteilungsfunktion F besitzt die folgenden Eigenschaften:

(1) $x \leq y \Rightarrow F(x) \leq F(y)$

(2) $\lim_{x\to-\infty} F(x) = 0$ und $\lim_{x\to\infty} F(x) = 1$

(3) F ist linksseitig stetig. (Bemerkung: Mit $(-\infty, x]$ rechtsseitig stetig.)

(ii). Jede reelle Funktion auf $\mathbb{R}$ mit (1)-(3) ist die Verteilungsfunktion einer Zufallsgröße.

1.2.15 Aufgabe

Ist F die Verteilungsfunktion der Zufallsgröße X, so gilt:

(i). $\mathbb{P}(a \leq X < b) = F(b) - F(a)$

(ii). $\mathbb{P}(X = a) = F(a+0) - F(a)$ (Hieraus folgt: X besitzt eine stetige Verteilung $\Leftrightarrow$ F ist stetig)

(iii). $\mathbb{P}(a < X < b) = F(b) - F(a+0)$

(iv). $\mathbb{P}(a \leq X \leq b) = F(b+0) - F(a)$

(v). $\mathbb{P}(a < X \leq b) = F(b+0) - F(a+0)$

(vi). Besitzt X eine Dichte p, so gilt für $x \in \mathbb{R}$:

$$F(x) = \int_{-\infty}^{x} p(t)\,dt$$

1.2.16 Satz

Die Zufallsgröße X besitze eine Dichte p und sei F die Verteilungsfunktion von X. Dann ist F λ-fast überall differenzierbar und

$$F'(x) = p(x)$$

λ-fast überall.

Beweis: Bekannt aus Maßtheorie, Analysis (§10?)

Bemerkung: Umkehrung gilt nicht.

1.2.17 Aufgabe

1.2.18 Aufgabe

Berechnen Sie die Dichtefunktion der Zufallsgröße X^2, $X \in N(0,1)$.

1.2.19 Aufgabe

Notation:

$$\begin{aligned}\varphi(x) &:= \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} \\ \Phi(x) &:= \int_{-\infty}^{x} \varphi(y)\,dy\end{aligned}$$

1.2.20 Aufgabe

Ist $x > 0$, so ist die Differenz

$$1 - \Phi(x) - \varphi(x) \cdot \left(\frac{1}{x} - \frac{1}{x^3} + \frac{1 \cdot 3}{x^5} - \frac{1 \cdot 3 \cdot 5}{x^7} + \ldots + \frac{(-1)^k \cdot \prod_{j=1}^k (2k-1)}{x^{2k+1}} \right)$$

positiv oder negativ, je nachdem, ob k ungerade oder gerade ist. Folgerung: Näherungen für x groß

$$\begin{aligned} \varphi(x) \cdot \left(\frac{1}{x} - \frac{1}{x^3} \right) &< 1 - \Phi(x) < \varphi(x) \cdot \frac{1}{x} \\ 1 - \Phi(x) &\sim \frac{\varphi(x)}{x} \end{aligned}$$

1.2.21 Aufgabe

Die Funktion φ_σ $(\sigma > 0)$ mit

$$\varphi_\sigma(x) := \frac{1}{(\sqrt{2\pi} \cdot \sigma)^d} \cdot \exp\left(-\frac{\|x\|^2}{2\sigma^2} \right) \qquad (x \in \mathbb{R}^d)$$

ist eine Dichtefunktion.

1.3 Unabhängige Zufallsvariablen

1.3.1 Definition

Zwei Zufallsvariablen $X_1 : \Omega \to (S_1, \mathcal{B}_1)$, $X_2 : \Omega \to (S_2, \mathcal{B}_2)$ heißen unabhängig, falls die Ereignisse $[X_1 \in B_1]$, $[X_2 \in B_2]$ für beliebige $B_i \in \mathcal{B}_i$ unabhängig sind, d.h.

$$\mathbb{P}[X_1 \in B_1, X_2 \in B_2] = \mathbb{P}[X_1 \in B_1] \cdot \mathbb{P}[X_2 \in B_2]$$

Allgemeiner: Eine Familie $\{X_i\}_{i \in I}$ von Zufallsvariablen $X_i : \Omega \to (S_i, \mathcal{B}_i)$ heißen unabhängig, falls für jede nichtleere Teilmenge $\{i_1, \ldots, i_n\} \subset I$ und beliebige $B_{i_j} \in \mathcal{B}_{i_j}$ gilt:

$$\mathbb{P}[X_{i_1} \in B_{i_1}, \ldots, X_{i_n} \in B_{i_n}] = \prod_{j=1}^n \mathbb{P}[X_{i_j} \in B_{i_j}]$$

d.h. falls die Ereignisse $[X_{i_j} \in B_{i_j}]$ unabhängig sind $(j = 1, \ldots, n)$.

1.3.2 Aufgabe

(i). Die Ereignisse $A_1, \ldots, A_n$ sind unabhängig $\Leftrightarrow$ die Zufallsgrößen $1_{A_1}, \ldots, 1_{A_n}$ sind unabhängig.

(ii). Seien X,Y Zufallsgrößen mit X fast sicher konstant. Dann X, Y unabhängig.

(iii). X ist fast sicher konstant $\Leftrightarrow F_X(\mathbb{R}) = \{0, 1\}$

(iv). X unabhängig von sich selbst $\Leftrightarrow$ X fast sicher konstant.

1.3.3 Satz

Es seien $X_1, \ldots, X_n$ Zufallsgrößen. Die folgenden Aussagen sind äquivalent:

(i). $X_1, \ldots, X_n$ sind unabhängig.

(ii). ihre gemeinsame Verteilung $\mu_Y = \mu_{X_1, \ldots, X_n}$ ist das Produkt ihrer einzelnen Verteilungen μ_{X_i}:

$$\mu_Y = \bigtimes_{i=1}^n \mu_{X_i} \qquad (1)$$

(iii). $\mathbb{P}(X_1 < t_1, \ldots, X_n < t_n) = \mathbb{P}(X_1 < t_1) \cdots \mathbb{P}(X_n < t_n)$ für beliebige Zahlen $t_i \in \mathbb{R}$ $(i = 1, \ldots, n)$

Beweis:

- Es gilt:

$$\begin{aligned} & \mu_Y = \bigtimes_{i=1}^{n} \mu_{X_i} \\ \overset{\text{Def}}{\Leftrightarrow} \quad & \mu_Y(B_1 \times \ldots \times B_n) = \prod_{i=1}^{n} \mu_{X_i}(B_i) \qquad (B_i \in \mathcal{B}(\mathbb{R})) \qquad (2) \\ \overset{\text{Erzeuger}}{\Leftrightarrow} \quad & \mu_Y(B_1 \times \ldots \times B_n) = \prod_{i=1}^{n} \mu_{X_i}(B_i) \qquad (B_i = (-\infty, t_i)) \qquad (3) \end{aligned}$$

Da

$$\mathbb{P}[(X_1, \ldots, X_n) \in B_1 \times \ldots \times B_n] = \mathbb{P}[X_1 \in B_1, \ldots, X_n \in B_n]$$

und

$$\prod_{i=1}^{n} \mu_{X_i}(B_i) = \prod_{i=1}^{n} \mathbb{P}[X_i \in B_i]$$

erhalten wir: (2) $\Leftrightarrow$ (i), (1) $\Leftrightarrow$ (ii), (3) $\Leftrightarrow$ (iii).

1.3.4 Satz

Es seien $X_1, \ldots, X_n$ unabhängige Zufallsgrößen und $g : \mathbb{R}^k \to \mathbb{R}^l$ $(1 \leq k \leq n)$ eine borel-messbare Funktion. Dann sind die Zufallsvektoren $g(X_1, \ldots, X_k)$, $X_{k+1}, \ldots, X_n$ unabhängig.

Beweis:

- Es seien $B \subset \mathbb{R}^l$, $B_{k+1} \subset \mathbb{R}, \ldots, B_n \subset \mathbb{R}$ beliebige borel-messbare Mengen.

$$\begin{aligned} & \mathbb{P}[g(X_1, \ldots, X_k) \in B, X_{k+1} \in B_{k+1}, \ldots, X_n \in B_n] \\ = \quad & \mathbb{P}[(X_1, \ldots, X_k) \in g^{-1}(B); X_{k+1} \in B_{k+1}, \ldots, X_n \in B_n] \\ = \quad & \mathbb{P}[(X_1, \ldots, X_n) \in g^{-1}(B) \times B_{k+1} \times \ldots \times B_n] \\ = \quad & \mu_{(X_1, \ldots, X_n)}(g^{-1}(B) \times B_{k+1} \times \ldots \times B_n) \\ \overset{1.3.3}{=} \quad & \left(\prod_{i=k+1}^{n} \mu_{X_i}(B_i) \right) \cdot \mu_{(X_1, \ldots, X_k)}(g^{-1}(B)) \\ = \quad & \underbrace{\mathbb{P}[(X_1, \ldots, X_k) \in g^{-1}(B)]}_{\mathbb{P}[g(X_1, \ldots, X_k) \in B]} \cdot \prod_{i=k+1}^{n} \mathbb{P}[X_i \in B_i] \end{aligned}$$

1.3.5 Folgerung

$\{X, Y, Z\}$ unabhängig $\Rightarrow$ $\{X + Y, Z\}$, $\{X \cdot Y, Z\}$ unabhängig

Beweis: Klar mit 1.3.4 ($g(x, y) = x + y$ bzw. $g(x, y) = x \cdot y$).

1.3.6 Definition

Bekannt aus Maßtheorie (§14): Es seien μ, ν Wahrscheinlichkeitsmaße auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Durch die Gleichung

$$\mu * \nu(B) := \int_{\mathbb{R}^n} \mu(B - y) \, d\nu(y)$$

wird ein Wahrscheinlichkeitsmaß $\mu * \nu$ auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ definiert. $\mu * \nu$ heißt Faltung von μ und ν.

Es gilt:

(i). $\mu * \nu = \nu * \mu$

(ii). $\mu * (\nu_1 + \nu_2) = \mu * \nu_1 + \mu * \nu_2$

(iii). Wenn μ eine Dichte p besitzt, dann hat $\mu * \nu$ auch eine Dichte h:

$$h(x) = \int_{\mathbb{R}^n} p(x-y)\, d\nu(y)$$

(iv). Wenn auch ν eine Dichte q besitzt, dann

$$h(x) = \int_{\mathbb{R}^n} p(x-y) \cdot q(y)\, \lambda(dy)$$

h heißt Faltung von p und q. Ist p oder q stetig, so auch h.

(v). Für $f \in \mathcal{L}^1(\mu * \nu)$:

$$\int_{-\infty}^{\infty} f(t)\, d(\mu * \nu)(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t+s)\, d\mu(t)\, d\nu(s)$$

(vi). $\varepsilon_x * \varepsilon_y = \varepsilon_{x+y}$

1.3.7 Satz

Sind $X, Y : \Omega \to \mathbb{R}^d$ unabhängige Zufallsvektoren, so gilt:

$$\mu_{X+Y} = \mu_X * \mu_Y$$

Beweis:

- Setze $Z := (X, Y)$, $\varphi(x, y) := x + y$ für $x, y \in \mathbb{R}^d$. Für jede borel-messbare Menge $B \subset \mathbb{R}^n$ gilt:

$$\begin{aligned} \mu_{X+Y}(B) &= \mathbb{P}[X+Y \in B] = \mathbb{P}[\varphi(Z) \in B] \\ &= \mathbb{P}[Z \in \varphi^{-1}(B)] = \mu_Z(\varphi^{-1}(B)) \end{aligned}$$

Nach Definition des Integrals folgt:

$$\begin{aligned} \mu_Z(\varphi^{-1}(B)) &= \int_{\mathbb{R}^d \times \mathbb{R}^d} 1_{\varphi^{-1}(B)}(x, y)\, d\mu_Z(x, y) \\ &= \int_{\mathbb{R}^d \times \mathbb{R}^d} 1_B(x+y)\, d\mu_Z(x, y) \\ &\overset{1.3.3}{=} \int_{\mathbb{R}^d \times \mathbb{R}^d} 1_B(x+y)\, d(\mu_X \times \mu_Y)(x, y) \\ &\overset{\text{Tonelli}}{=} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} 1_B(x+y)\, d\mu_X(x)\, d\mu_Y(y) \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} 1_{B-y}(x)\, d\mu_X(x)\, d\mu_Y(y) \\ &= \int_{\mathbb{R}^d} \mu_X(B-y)\, d\mu_Y(y) = \mu_X * \mu_Y(B) \end{aligned}$$

1.3.8 Bemerkung

Die Umkehrung von 1.3.7 gilt nicht. Beispiel: Seien $X = Y$ Cauchy-verteilt mit der Dichte

$$f(x) = \frac{1}{\pi \cdot (1 + x^2)} \qquad (x \in \mathbb{R})$$

Dann sind X und Y nicht unabhängig (1.3.2(iii)), aber

$$\mu_{X+Y} = \mu_{2X} \overset{\text{Aufgabe}}{=} \mu_X * \mu_X$$

1.3.9 Beispiel

Es seien X, Y unabhängige gamma-verteilte Zufallsgrößen mit der Dichte

$$\begin{aligned} p_X(x) &= \frac{\lambda^a \cdot x^{a-1}}{\Gamma(a)} \cdot e^{-\lambda x} \cdot 1_{(0,\infty)}(x) \\ p_Y(x) &= \frac{\lambda^b \cdot x^{b-1}}{\Gamma(b)} \cdot e^{-\lambda x} \cdot 1_{(0,\infty)}(x) \end{aligned}$$

mit $a, b, \lambda > 0$. Dann ist $X + Y$ auch gamma-verteilt mit den Parametern $a + b, \lambda$:

$$\begin{aligned} p_{X+Y} &\overset{1.3.6(iv)}{=} \int_{-\infty}^{\infty} p_X(x-y) \cdot p_Y(y)\, \lambda(dy) \\ &= \frac{\lambda^{a+b}}{\Gamma(a) \cdot \Gamma(b)} \cdot e^{-\lambda x} \int_{-\infty}^{\infty} (x-y)^{a-1} \cdot y^{b-1} \cdot 1_{(0,\infty)}(x-y) \cdot 1_{(0,\infty)}(y)\, \lambda(dy) \\ &= \frac{\lambda^{a+b}}{\Gamma(a) \cdot \Gamma(b)} \cdot e^{-\lambda x} \int_0^x (x-y)^{a-1} \cdot y^{b-1}\, \lambda(dy) \\ &\overset{y=tx}{=} \frac{\lambda^{a+b}}{\Gamma(a) \cdot \Gamma(b)} \cdot x^{a+b-1} \cdot e^{-\lambda x} \int_0^1 (1-t)^{a-1} \cdot t^{b-1}\, dt \end{aligned}$$

also $p_{X+Y} = c \cdot p_{a+b,\lambda}$, wobei c eine Konstante ist. Da p_{X+Y} und $p_{a+b,\lambda}$ Dichtefunktionen sind, muss $c = 1$ sein. Folgerung:

$$B(a,b) := \int_0^1 (1-t)^{a-1} \cdot t^{b-1}\, dt = \frac{\Gamma(a) \cdot \Gamma(b)}{\Gamma(a+b)} \qquad (a, b > 0)$$

B ist die sogenannte Beta-Funktion.

1.3.10 Die χ^2-Verteilung

Seien $X_1, \ldots, X_n$ unabhängige, $N(0,1)$-verteilte Zufallsgrößen. Die Verteilung der Zufallsgrößen

$$Y_n := X_1^2 + \ldots + X_n^2$$

heißt χ_n^2-Verteilung mit dem Freiheitsgrad n. Die Verteilung von $\sqrt{Y_n}$ heißt χ_n-Verteilung mit dem Freiheitsgrad n. Behauptung: Für die Dichtefunktion h_n von Y_n gilt

$$h_n(x) = \frac{x^{\frac{n}{2}-1} \cdot e^{-\frac{x}{2}}}{2^{\frac{n}{2}} \cdot \Gamma\left(\frac{n}{2}\right)} \cdot 1_{(0,\infty)}(x)$$

Beweis: (per Induktion)

- Induktionsanfang: Laut 1.2.18 (Aufgabe 53) gilt diese Formel für $n = 1$
- Induktionsschritt:

$$\begin{aligned} h_{n+1}(x) &= \int_{-\infty}^{\infty} h_n(x-y) \cdot h_1(y)\, dy \\ &= \frac{1}{2^{\frac{n}{2}} \cdot \Gamma\left(\frac{n}{2}\right) \cdot \sqrt{2\pi}} \int_0^x (x-y)^{\frac{n}{2}-1} \cdot e^{-\frac{x-y}{2}} \cdot e^{-\frac{y}{2}} \cdot y^{-\frac{1}{2}}\, dy \\ &\overset{y=tx}{=} \frac{1}{2^{\frac{n}{2}} \cdot \Gamma\left(\frac{n}{2}\right) \cdot \sqrt{2\pi}} \cdot e^{-\frac{x}{2}} \cdot x^{\frac{n+1}{2}-1} \underbrace{\int_0^x (1-t)^{-\frac{1}{2}} \cdot t^{\frac{n}{2}-1}\, dt}_{B\left(\frac{1}{2}, \frac{n}{2}\right) = \frac{\Gamma(\frac{1}{2}) \cdot \Gamma(\frac{n}{2})}{\Gamma(\frac{n+1}{2})}} \\ &= \frac{x^{\frac{n+1}{2}-1} \cdot e^{-\frac{x}{2}}}{2^{\frac{n+1}{2}} \cdot \Gamma\left(\frac{n+1}{2}\right)} \cdot 1_{(0,\infty)}(x) \end{aligned}$$

Bemerkungen:

(i). Exponentialverteilung:

$$h_2(x) = \frac{1}{2} \cdot e^{-\frac{x}{2}}$$

(ii). Für die Dichte von $\sqrt{Y_n}$ gilt:

$$g_n(x) = \frac{x^{n-1}}{2^{\frac{n}{2}-1} \cdot \Gamma\left(\frac{n}{2}\right)} \cdot e^{-\frac{x^2}{2}} \cdot 1_{(0,\infty)}(x)$$

(Aufgabe). Hinweis: $\mathbb{P}(\sqrt{Y_n} < x) = \mathbb{P}(Y_n < x^2)$ für $x > 0$.

1.3.11 Satz

Für eine beliebige Familie $\{\mu_i\}_{i \in I}$ von Verteilungen auf $\mathbb{R}$ existiert ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ und darauf definierte Zufallsgrößen X_i, die unabhängig sind und die gegebenen Verteilungen μ_i besitzen.

Beweis: Für endliches $I = \{1, \ldots, n\}$: Aufgabe. Hinweis: Produktmaß

1.3.12 Aufgabe

(i). Es seien $X_1, \ldots, X_n$ unabhängige Zufallsgrößen mit $\mathbb{P}[X_i = 1] = p$, $\mathbb{P}[X_i = 0] = 1 - p$ mit $0 \leq p \leq 1$. Zu zeigen: $X_1 + \ldots + X_n$ ist binomialverteilt mit den Parametern n und p.

(ii). $B_n^p * B_m^p = B_{n+m}^p$

1.3.13 Aufgabe

(i). $\varepsilon_x * \varepsilon_y = \varepsilon_{x+y}$

(ii). Summe von unabhängigen poissonverteilten Zufallsgrößen ist poissonverteilt.

Hinweis: Mit Hilfe von (i) zeigen: $\Pi_a * \Pi_b = \Pi_{a+b}$

(iii). X, Y unabhängig, identisch verteilt mit der Dichte $f(x) = 1_{[0,1]}(x)$. Gesucht ist die Dichte von $X + Y$.

1.3.14 Aufgabe

Der Zufallsvektor $Y = (X_1, X_2) = (r \cdot \cos\varphi, r \cdot \sin\varphi)$ sei auf der Einheitskreisscheibe $K := \{(x_1, x_2) : x_1^2 + x_2^2 \leq 1\} \subset \mathbb{R}^2$ gleichverteilt. Man zeige, dass die Polarkoordinaten r und φ unabhängig sind, die kartesischen X_1, X_2 aber nicht.

1.3.15 Aufgabe

1.3.16 Aufgabe

Seien $X_1, \ldots, X_n$ Zufallsgrößen, μ ihre gemeinsame Verteilung auf $\mathbb{R}^n$, $h : \mathbb{R}^n \to \mathbb{R}$ borel-messbar, $x = (x_1, \ldots, x_n)$. Verteilungsfunktion der Zufallsgröße $Z := h(X_1, \ldots, X_n)$:

$$\mathbb{P}[Z < t] = \mu(\{x : h(x) < t\}) = \int_{h(x)<t} 1 \, d\mu(x)$$

(i). Spezialfall: $n = 2$, $Z = X \cdot Y$ mit X, Y unabhängig mit Dichtefunktionen f und g. Aufgabe: $X \cdot Y$ besitzt eine Dichte p mit

$$p(x) = \int_{-\infty}^{\infty} \frac{1}{|y|} \cdot g(y) \cdot f\left(\frac{x}{y}\right) \, dy$$

(ii). Analog: Dichte von $\frac{X}{Y}$ (falls $Y \neq 0$ fast sicher):

$$q(x) = \int_{-\infty}^{\infty} |y| \cdot g(y) \cdot f(x \cdot y)\, dy$$

1.3.17 Aufgabe

1.3.18 Aufgabe

Sei $(X_n)_{n\in\mathbb{N}}$ unabhängig, $\mathbb{P}[X_i = 0] = \mathbb{P}[X_i = 1] = \frac{1}{2}$ für alle $i \in \mathbb{N}$. Dann ist

$$S_r = \sum_{n=1}^{\infty} r^n \cdot X_n \qquad (r \in (0,1))$$

absolut konvergent für alle $w \in \Omega$.

(i). $S_{\frac{1}{2}}$ ist auf dem Intervall $[0,1]$ gleichmäßig verteilt.

(ii). Für $r = 2^{-\frac{1}{k}}$ $(k \in \mathbb{N})$ ist S_r absolut stetig.

(iii). Für $r \in \left(0, \frac{1}{2}\right)$ ist S_r stetig und singulär.

Beweis:

- Wir zeigen, dass der Wertebereich $W(S_r)$ von S_r das Lebesgue-Maß 0 hat. Dazu sei λ^* das äußere Lebesgue-Maß, definiert für beliebige $A \subseteq \mathbb{R}^n$, ist subadditiv

$$\lambda^*(A \cup B) \leq \lambda^*(A) + \lambda^*(B)$$

und verschiebungsinvariant. Damit:

$$\begin{aligned} W(S_r) &= W(0 + r \cdot S_r) \cup W(r + r \cdot S_r) \\ \Rightarrow \lambda^*(W(S_r)) &\leq \lambda^*(W(0 + r \cdot S_r)) + \lambda^*(W(r + r \cdot S_r)) \\ &= \underbrace{2r}_{<1} \cdot \lambda^*(W(S_r)) \\ \Rightarrow \lambda^*(W(S_r)) &= 0 \\ \Rightarrow \lambda(W(S_r)) &= 0 \end{aligned}$$

- Stetigkeit: $q \in W(S_r)$ lässt sich darstellen als

$$q = \sum_{n=1}^{\infty} \varepsilon_n \cdot r^n$$

Darstellung eindeutig für $r < \frac{1}{2}$. Damit $\mathbb{P}(S_r = q) = 0$.

1.4 Numerische Charakteristik von Zufallsgrößen

1.4.1 Definition

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, X eine $\mathbb{P}$-integrierbare oder nichtnegative Zufallsgröße. Dann heißt

$$\mathbb{E}(X) := \int_{\Omega} X \, d\mathbb{P}$$

der Erwartungswert von X. Es gilt:

(i). $\mathbb{E}(a \cdot X + b \cdot Y) = a \cdot \mathbb{E}(X) + b \cdot \mathbb{E}(Y)$

(ii). Ist $X = c$ fast sicher, so gilt $\mathbb{E}(X) = c$.

(iii). Aus $a \leq X \leq b$ fast sicher folgt $a \leq \mathbb{E}(X) \leq b$.

(iv). $|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$

(v). Aus $X \geq 0$ fast sicher und $\mathbb{E}(X) = 0$ folgt $X = 0$ fast sicher.

Analoge Definition für komplexe Zufallsgrößen:

$$\mathbb{E}(X) = \mathbb{E}(Re\,X) + \imath \cdot \mathbb{E}(Im\,X)$$

für Im X, Re $X \in \mathcal{L}^1(\mathbb{P})$. Es gilt:

$$\mathbb{E}(X) = \int_{\mathbb{R}} x\, d\mu_X(x)$$

(Folgt aus 1.2.10 bzw. §14) Ist X diskret mit den Werten $x_1, \dots, x_n$ und den zugehörigen Wahrscheinlichkeiten $p_1, \dots, p_n$ so gilt

$$\mathbb{E}(X) = \sum_k x_k \cdot p_k$$

(falls existent). Wenn X eine Dichte p besitzt, dann

$$\mathbb{E}(X) = \int_{\mathbb{R}} x \cdot p(x)\, \lambda(dx)$$

1.4.2 Satz

Es seien X, Y unabhängige $\mathbb{P}$-integrierbare Zufallsgrößen. Dann ist $X \cdot Y \in \mathcal{L}^1(\mathbb{P})$ und

$$\mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$$

Beweis:

- Bezeichne μ_X (μ_Y) die Verteilung von X (Y), ν sei gemeinsame Verteilung von X, Y auf $\mathbb{R}^2$. Dann gilt:

$$\begin{aligned}
\mathbb{E}(X) \cdot \mathbb{E}(Y) &= \left(\int_{\mathbb{R}} x\, d\mu_X(x)\right) \cdot \left(\int_{\mathbb{R}} y\, d\mu_Y(y)\right) \\
&\overset{\text{Fubini}}{=} \int_{\mathbb{R}^2} x \cdot y\, d(\mu_X \times \mu_Y)(x, y) \\
&\overset{1.3.3}{=} \int_{\mathbb{R}^2} x \cdot y\, d\nu(x, y) \overset{1.2.12}{=} \mathbb{E}(X \cdot Y)
\end{aligned}$$

1.4.3 Bemerkung

(i). Die Umkehrung von 1.4.2 ist falsch. Sei zum Beispiel $X \in \mathcal{L}^3(\mathbb{P})$ symmetrisch zum Punkt 0 verteilt (d.h. $\mu_X(B) = \mu_X(-B)$ bzw.X und $-X$ haben die gleiche Verteilung). Definiere $Y := X^2$. Dann ist auch X^3 symmetrisch bzgl. 0,

$$\begin{aligned}
\mathbb{E}(X) &= 0 \\
\mathbb{E}(X \cdot Y) &= \mathbb{E}(X^3) = 0 = \mathbb{E}(X) \cdot \mathbb{E}(Y)
\end{aligned}$$

aber i.a. X und X^2 nicht unabhängig.

(ii). Satz 1.4.2 gilt auch für komplexwertige Zufallsgrößen. (Aufgabe)

1.4.4 Beispiele

(i). X binomialverteilt mit den Parametern n und p. Dann

$$\mathbb{E}(X) = \sum_{k=0}^{n} \underbrace{k}_{x_k} \cdot \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

Einfacher: Es seien $X_1, \dots, X_n$ unabhängige Zufallsgrößen mit $\mathbb{P}(X_k = 1) = p$, $\mathbb{P}(X_k = 0) = 1-p$. Nach 1.3.12 ist $\tilde{X} := \sum_{i=1}^{n} X_i$ binomialverteilt mit den Parametern n und p. Damit:

$$\mathbb{E}(X) = \mathbb{E}(\tilde{X}) = n \cdot \mathbb{E}(X_1) = n \cdot (1 \cdot p + 0 \cdot (1-p)) = n \cdot p$$

(ii). X habe in $[a,b]$ gleichmäßige Verteilung. Dann:

$$\begin{aligned}
\mathbb{E}(X) &= \int_{\mathbb{R}} x \cdot p(x)\, \lambda(dx) \\
&= \int_a^b x \cdot \frac{1}{b-a}\, dx \\
&= \frac{1}{b-a} \cdot \frac{b^2-a^2}{2} = \frac{a+b}{2}
\end{aligned}$$

(iii). X poisson-verteilt mit Parameter a:

$$\begin{aligned}
\mathbb{P}[X = k] &= \frac{a^k \cdot e^{-a}}{k!} \\
\Rightarrow \mathbb{E}(X) &= \sum_{k \in \mathbb{N}_0} k \cdot \frac{a^k \cdot e^{-a}}{k!} = \sum_{k \in \mathbb{N}} \frac{a^k \cdot e^{-a}}{(k-1)!} \\
&= e^{-a} \cdot a \cdot \sum_{m \in \mathbb{N}_0} \frac{a^m}{m!} \\
&= a
\end{aligned}$$

(iv). X normalverteilt mit der Dichte

$$p_{a,\sigma}(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

Dann:

$$\begin{aligned}
\mathbb{E}(X) &= \int_{\mathbb{R}} x \cdot p_{a,\sigma}(x)\, \lambda(dx) \\
&\overset{z := \frac{x-a}{\sigma}}{=} \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \sigma \int_{\mathbb{R}} (\sigma \cdot z + a) \cdot e^{-\frac{z^2}{2}}\, \lambda(dz) \\
&= \frac{\sigma}{\sqrt{2\pi}} \underbrace{\int_{\mathbb{R}} e^{-\frac{z^2}{2}} \cdot z\, \lambda(dz)}_{0} + \frac{a}{\sqrt{2\pi}} \underbrace{\int_{\mathbb{R}} e^{-\frac{z^2}{2}}\, \lambda(dz)}_{\sqrt{2\pi}} \\
&= a
\end{aligned}$$

(v). Aufgabe: χ_n und χ_n^2-Verteilung.

1.4.5 Aufgabe

(i). Wie lange muss man im Durchschnitt eine Münze werfen, bis man beide Seiten erhalten hat?

(ii). Analog für einen Würfel bzw. für einen „k-seitigen“ Würfel? ($k \in \mathbb{N}$)

1.4.6 Definition

Als Streuung (Varianz, Dispersion) einer Zufallsgröße $X \in \mathcal{L}^1(\mathbb{P})$ bezeichnet man die Größe

$$\begin{aligned} \sigma^2 &:= \mathbb{V}(X) := \operatorname{var}(X) := D^2(X) \\ &:= \mathbb{E}((X-\mathbb{E}(X))^2) = \int_\Omega (X-\mathbb{E}X)^2 \, d\mathbb{P} \end{aligned}$$

σ heißt Standardabweichung. Es gilt:

$$\mathbb{V}(X) = \int_\mathbb{R} (x-\mathbb{E}(X))^2 \, d\mu_X(x)$$

Ist X diskret, so gilt

$$\mathbb{V}(X) = \sum_{k \in \mathbb{N}} p_k \cdot (x_k - \mathbb{E}(X))^2$$

Besitzt X eine Dichte, dann

$$\mathbb{V}(X) = \int_\mathbb{R} (x-\mathbb{E}(X))^2 \cdot p(x) \, \lambda(dx)$$

1.4.7 Satz

Für $X, Y \in \mathcal{L}^2(\mathbb{P})$ gil:

(i). $\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$

(ii). $\mathbb{V}(a \cdot X + b) = a^2 \cdot \mathbb{V}(X)$

(iii). $\mathbb{V}(X) = 0 \Leftrightarrow X = c$ fast sicher

(iv). $\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2\mathbb{E}[(X-\mathbb{E}(X)) \cdot (Y - \mathbb{E}(Y))]$

(v). Sind X und Y unkorreliert, d.h.

$$\mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$$

dann gilt

$$\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y)$$

(Ist z.B. erfüllt, wenn X und Y unabhängig.)

Beweis: Aufgabe 59

1.4.8 Beispiele

(i). X sei binomialverteilt mit den Parametern n und p. Seien $X_1, \ldots, X_n$ unabhängig mit $\mathbb{P}[X_i = 0] = 1-p$, $\mathbb{P}[X_i = 1] = p$, $\tilde{X} := \sum_{i=1}^n X_i$. Dann $\tilde{X}$ binomialverteilt mit den Parametern n und p,

$$\begin{aligned} \mathbb{E}(X_i^2) &= \mathbb{E}(X_i) = p \\ \mathbb{V}(X) &= \mathbb{E}(X_i^2) - (\mathbb{E}(X_i))^2 = p - p^2 = p \cdot (1-p) \\ \Rightarrow \mathbb{V}(X) &= \mathbb{V}(\tilde{X}) \overset{1.4.7(v)}{=} n \cdot \mathbb{V}(X_1) \\ &= n \cdot p \cdot (1-p) = (1-p) \cdot \mathbb{E}(X) \end{aligned}$$

(ii). X gleichmäßig verteilt in $[a,b]$. Bekannt:

$$\mathbb{E}(X) = \frac{a+b}{2}$$

Somit:

$$\begin{aligned}
\mathbb{E}(X^2) &= \int_a^b x^2 \cdot \frac{1}{b-a}\,dx = \frac{b^3-a^3}{3}\cdot\frac{1}{b-a} \\
&= \frac{a^2+b^2+ab}{3} \\
\Rightarrow \mathbb{V}(X) &= \mathbb{E}(X)^2 - (\mathbb{E}(X))^2 = \frac{a^2+b^2+ab}{3} - \left(\frac{a+b}{2}\right)^2 \\
&= \frac{(b-a)^2}{12}
\end{aligned}$$

(iii). X normalverteilt mit den Parametern a und σ

$$\begin{aligned}
\mathbb{V}(X) &= \frac{1}{\sigma\cdot\sqrt{2\pi}}\int_{\mathbb{R}} (x-a)^2 \cdot e^{-\frac{(x-a)^2}{2\sigma^2}}\,\lambda(dx) \\
&\overset{z=\frac{x-a}{\sigma}}{=} \frac{\sigma^2}{\sqrt{2\pi}}\cdot\int_{\mathbb{R}} z^2\cdot e^{-\frac{z^2}{2}}\,\lambda(dz) \\
&= \frac{\sigma^2}{\sqrt{2\pi}}\cdot\left(\left[-z\cdot e^{-\frac{z^2}{2}}\right]_{-\infty}^{\infty} + \int_{\mathbb{R}} e^{-\frac{z^2}{2}}\,\lambda(dz)\right) \\
&= \frac{\sigma^2}{\sqrt{2\pi}}\cdot(0+\sqrt{2\pi}) = \sigma^2
\end{aligned}$$

(iv). Aufgabe: $\mathbb{V}(\chi_n^2) = 2n$

1.4.9 Definition

Als Kovarianz der Zufallsgrößen $X, Y \in \mathcal{L}^2(\mathbb{P})$ bezeichnet man die Größe

$$\begin{aligned}
\operatorname{cov}(X,Y) &:= \mathbb{E}[(X-\mathbb{E}(X))\cdot(Y-\mathbb{E}(Y))] \\
&= \mathbb{E}(X\cdot Y) - \mathbb{E}(X)\cdot\mathbb{E}(Y)
\end{aligned}$$

X und Y sind unkorreliert $\Leftrightarrow \operatorname{cov}(X,Y)=0$. Man definiert

$$\operatorname{corr}(X,Y) := \frac{\operatorname{cov}(X,Y)}{D(X)\cdot D(Y)}$$

(Korrelationskoeffizienten) für $D(X)\cdot D(Y)\neq 0$. Es gilt:

(i). $|\operatorname{corr}(X,Y)|\leq 1$ wegen Cauchy-Schwarz-Ungleichung

(ii). $\operatorname{corr}(X,X)=1$, $\operatorname{corr}(X,-X)=-1$

(iii). $\operatorname{corr}(X,Y)=0$, falls X,Y unabhängig

(iv). $\mathbb{V}(X_1+\ldots+X_n) = \sum_{i=1}^n \mathbb{V}(X_i) + \sum_{i<j}\operatorname{cov}(X_i,X_j)$

1.4.10 Satz

Seien $X, Y \in \mathcal{L}^2(\mathbb{P})$ Zufallsgrößen mit $D(X)\cdot D(Y)\neq 0$. Dann:

$$|\operatorname{corr}(X,Y)| = 1 \Leftrightarrow \exists a,b\in\mathbb{R}, a\neq 0 : Y = a\cdot X + b\,\text{fast sicher}$$

Beweis:

(i). „⇒“

Ohne Beschränkung der Allgemeinheit sei $\operatorname{corr}(X,Y)=1$. Definiere

$$X' := \frac{X-\mathbb{E}(X)}{D(X)} \qquad Y' := \frac{Y-\mathbb{E}(Y)}{D(Y)}$$

Dann gilt:

$$\begin{aligned}\mathbb{E}(X') &= \mathbb{E}(Y') = 0\\ \mathbb{V}(X') &= \mathbb{E}(X'^2) = \mathbb{E}(Y'^2) = 1\end{aligned}$$

Außerdem

$$1 = \operatorname{corr}(X,Y) = \mathbb{E}(X' \cdot Y')$$

nach Voraussetzung. Damit folgt:

$$\begin{aligned}\mathbb{E}((X'-Y')^2) &= \mathbb{E}(X'^2 - 2X' \cdot Y' + Y'^2)\\ &= 1 - 2 + 1 = 0\end{aligned}$$

Also $X' = Y'$ fast sicher,

$$Y = \mathbb{E}(Y) + D(Y) \cdot \frac{X - \mathbb{E}(X)}{D(X)}$$

(ii). „$\Leftarrow$“: Leicht

1.4.11 Aufgabe

In einer Lotterie werden n Zahlen von den Zahlen $1, 2, \dots, N$ ausgewählt ($1 \le n \le N$). Man berechne die Streuung der Summe S_n der gezogenen Zahlen. Lösung:

- Sei $S_n := \sum_{i=1}^n X_i$, wobei X_i die i-te gezogene Zahl bezeichnet.

$$\mathbb{V}(S_n) = \sum_{i=1}^n \mathbb{V}(X_i) + 2\sum_{i<j} \operatorname{cov}(X_i, X_j)$$

X_i und (X_i, X_j) sind gleichmäßig verteilt auf $\{1,\dots,N\}$ bzw. $\{1,\dots,N\} \times \{1,\dots,N\} \setminus \{(k,k); k \in \{1,\dots,N\}\}$.

- Definiere

$$\sigma^2 := \mathbb{V}(X_i) \qquad \varrho := \operatorname{cov}(X_i, X_j)$$

(Diese Größen hängen nicht von i, j ab.) Damit

$$\mathbb{V}(S_n) = n \cdot \sigma^2 + n \cdot (n-1) \cdot \varrho$$

für $n \in \{1,\dots,N\}$. Speziell für $n = N$:

$$\begin{aligned}\mathbb{V}(S_N) &= 0 = N \cdot \sigma^2 + N \cdot (N-1) \cdot \varrho\\ \Rightarrow \varrho &= -\frac{\sigma^2}{N \cdot (N-1)}\end{aligned}$$

Damit:

$$\mathbb{V}(S_n) = n \cdot \sigma^2 \cdot \left(1 - \frac{n-1}{N-1}\right)$$

und

$$\begin{aligned}\sigma^2 &= \mathbb{E}(X_i^2) - (\mathbb{E}(X_i))^2\\ &= \frac{1}{N} \cdot \sum_{k=1}^N k^2 - \left(\frac{1}{N} \cdot \sum_{k=1}^N k\right)^2\\ &= \frac{1}{6} \cdot (N+1) \cdot (2N+1) - \frac{1}{4} \cdot (N+1)^2\end{aligned}$$

1.4.12 Aufgabe

1.4.13 Aufgabe

1.4.14 Definition

Als Moment k-ter Ordnung einer Zufallsgröße $X \in \mathcal{L}^k(\mathbb{P})$ bezeichnet man den Erwartungswert

$$M_k(X) := \mathbb{E}(X^k) = \int_{\mathbb{R}} x^k \, d\mu_X(x)$$

mit $k \in \mathbb{N}$. Als absolutes Moment bezeichnet man

$$A_k(X) := \mathbb{E}(|X|^k) = \int_{\mathbb{R}} |x|^k \, d\mu_X(x)$$

Bemerkungen:

(i). Aus der Existenz von M_k folgt die Existenz von M_l für $l \leq k$, da

$$|x|^l \leq |x|^k + 1_{[-1,1]}(x)$$

(ii). Sei $n > k \geq 1$ und nehmen wir an, dass A_n existiert. Höldersche Ungleichung:

$$\begin{aligned} A_k &= \int_{\mathbb{R}} |x|^k \cdot 1 \, d\mu_X(x) \\ &\overset{p>1}{\leq} \left(\int_{\mathbb{R}} |x|^{k \cdot p} \, d\mu_X(x) \right)^{\frac{1}{p}} \end{aligned}$$

Für $p = \frac{n}{k} > 1$:

$$A_k^{\frac{1}{k}} \leq A_n^{\frac{1}{n}}$$

1.4.15 Aufgabe

Man berechne die (absoluten) Momente der Normalverteilung mit $a = 0$:

$$\begin{aligned} A_k &= \sqrt{\frac{2}{\pi}} \cdot \sigma^k \int_0^\infty x^k \cdot e^{-\frac{x^2}{2}} \, dx \\ &\overset{z:=\frac{x^2}{2}}{=} \sqrt{\frac{2}{\pi}} \cdot \sigma^k \cdot 2^{\frac{k-1}{2}} \int_0^\infty z^{\frac{k-1}{2}} \cdot e^{-z} \, dz \\ &= \sqrt{\frac{2}{\pi}} \cdot \sigma^k \cdot 2^{\frac{k-1}{2}} \cdot \Gamma\left(\frac{k+1}{2}\right) \\ &= \frac{1}{\sqrt{\pi}} \cdot (\sqrt{2} \cdot \sigma)^k \cdot \Gamma\left(\frac{k+1}{2}\right) \end{aligned}$$

Es gilt: $M_k = A_k$ falls k gerade, $M_k = 0$ falls k ungerade.

1.4.16 Aufgabe

Verteilungen sind i.A. nicht eindeutig durch ihre Momente bestimmt. Beispiel: Die Zufallsgröße X sei logarithmisch normalverteilt, d.h. sie hat die Dichte

$$p(x) = \frac{1}{\sqrt{2\pi}} \cdot x^{-1} \cdot e^{-\frac{(\ln x)^2}{2}} \cdot 1_{(0,\infty)}(x)$$

Weiterhin sei $|a| \leq 1$ und X_a besitze die Dichte

$$p_a(x) = p(x) \cdot (1 + a \cdot \sin(2\pi \ln z))$$

Zeigen Sie: $\mathbb{E}(X^k) = \mathbb{E}(X_a^k) = e^{\frac{k^2}{2}}$ für $k \in \mathbb{N}_0$.

1.4.17 Aufgabe

1.4.18 Aufgabe

1.4.19 Satz

Sei X eine Zufallsgröße mit Verteilungsfunktion F. Existiert $\mathbb{E}X$, so gilt

$$\begin{aligned} \lim_{x\to\infty} x\cdot(1-F(x)) &= 0 \\ \lim_{x\to-\infty} x\cdot F(x) &= 0 \\ \mathbb{E}(X) &= \int_{(0,\infty)} (1-F(y))\,\lambda(dy) - \int_{(-\infty,0)} F(y)\,\lambda(dy) \end{aligned}$$

Beweis: (nicht prüfungsrelevant)

- Da $\mathbb{E}X$ existiert, gilt
$$\int_{\mathbb{R}} |x|\,dF(x) := \int_{\mathbb{R}} |x|\,d\mu_X(x) < \infty$$
wobei μ_X die Verteilung von X ist. (Stieltjes-Integral) Es folgt:
$$\begin{aligned} 0 &\leq \lim_{x\to\infty} x\cdot(1-F(x)) \\ &\leq \lim_{x\to\infty} \int_{(x,\infty)} y\,dF(y) = 0 \end{aligned}$$
Zweite Gleichung analog.

- Zur „Erinnerung“, partielle Integration für Stieltjes-Integrale:
$$\int_{(0,x)} f(y)\,dg(y) = f(y)\cdot g(y)\Big|_0^x - \int_{(0,x)} g(y-0)\,df(y)$$
wobei f und g reelle Funktionen mit endlicher Variation (sind Differenz von monoton wachsenden beschränkten Funktionen). Damit:
$$\begin{aligned} \int_{(-x,0)} y\,dF(y) &= x\cdot F(-x) - \int_{(-x,0)} F(y)\,dy \\ \int_{(0,x)} y\,dF(y) &= -x\cdot(1-F(x)) + \int_{(0,x)} (1-F(y))\,dy \end{aligned}$$
Addition der beiden Gleichungen und $x\to\infty$.

1.4.20 Definition

- Sei $X=(X_1,\dots,X_d)$ ein d-dimensionaler Zufallsvektor (reell oder komplex) so, dass $X_j \in \mathcal{L}^1(\mathbb{P})$ für $j=1,\dots,d$. Dann wird der Erwartungsvektor von X definiert durch
$$\mathbb{E}X := (\mathbb{E}X_1,\dots,\mathbb{E}X_d)$$

- Für $X_j \in \mathcal{L}^1(\mathbb{P})$ und $m_j := \mathbb{E}X_j$ definiert man die Matrix
$$\operatorname{cov}(X) := (c_{jk})_{j,k=1}^d$$
wobei
$$\begin{aligned} c_{jk} &:= \mathbb{E}[(X_j-m_j)\cdot\overline{(X_k-m_k)}] \\ &= \mathbb{E}(X_j\cdot\bar{X}_k) - m_j\cdot\bar{m}_k \end{aligned}$$
die Kovarianzmatrix von X. Es gilt:

(i). Falls X_i, X_j paarweise unkorreliert ist $\mathrm{cov}(X)$ eine Diagonalmatrix.

(ii). $\mathrm{cov}(X)$ ist positiv semidefinit:

$$\begin{aligned}\sum_{j=1}^{d}\sum_{k=1}^{d} c_{jk}\cdot z_j\cdot \bar{z_k} &= \mathbb{E}\left(\sum_{j=1}^{d} z_j\cdot (X_j-m_j)\cdot \sum_{k=1}^{d}\overline{z_k\cdot (X_k-m_k)}\right)\\ &= \mathbb{E}\left(\left|\sum_{j=1}^{d} z_j\cdot (X_j-m_j)^2\right|\right)\geq 0\end{aligned}$$

mit $z_j, z_k \in \mathbb{C}$ beliebig.

- Seien Y_{jk} integrierbare Zufallsgrößen $(j,k=1,\ldots,n)$. Wir bilden die Matrix $Y := (Y_{jk})_{j,k=1}^{n}$ (Zufallsmatrix). Definiere
$$\mathbb{E}Y := (\mathbb{E}Y_{jk})_{j,k=1}^{n}$$
Erwartungsmatrix von Y. Mit dieser Bezeichnung gilt:
$$\mathrm{cov}(X) = \mathbb{E}((X-\mathbb{E}X)\cdot \overline{X-\mathbb{E}X}^{T})$$
mit X als Spaltenvektor wie oben.

- Aufgabe:
$$\mathrm{cov}(A\cdot X) = A\cdot \mathrm{cov}(X)\cdot (\bar{A})^{T}$$
wobei $A\in\mathbb{K}^{d\times d}$, $\mathbb{K}\in\{\mathbb{R},\mathbb{C}\}$.

- Seien $\alpha=(\alpha_1,\ldots,\alpha_d)\in\mathbb{N}_0^d$, $x=(x_1,\ldots,x_d)\in\mathbb{R}^d$, μ die Verteilung eines d-dimensionalen Zufallsvektors X. Notation:
$$x^{\alpha} := x_1^{\alpha_1}\cdots x_d^{\alpha_d}$$
Dann heißt
$$M_\alpha = M_\alpha(\mu) = M_\alpha(X) := \int_{\mathbb{R}^d} x^{\alpha}\, d\mu(x)$$
Moment der Ordnung α von μ bzw. X, falls $x\mapsto x^{\alpha}$ bzgl. μ integrierbar ist. Analog: absolute Momente
$$A_\alpha = A_\alpha(\mu) = A_\alpha(X) := \int_{\mathbb{R}^d} |x^{\alpha}|\, d\mu(x)$$

1.5 Bedingte Wahrscheinlichkeit

1.5.1 Motivation

Hat bei N Versuchen das Ereignis B genau n-mal stattgefunden und ist bei diesen n Versuchen k-mal zusammen mit B auch das Ereignis A eingetreten, so wird

$$h_{A|B} := \frac{k}{n}$$

die bedingte relative Häufigkeit von A unter der Bedingung B genannt.

Notation: $h_A, h_{A\cap B}$ - relative Häufigkeit von B bzw. $A\cap B$ in der gesamten Versuchsreihe. Dann:

$$h_{A|B} = \frac{k}{n} = \frac{\frac{k}{N}}{\frac{n}{N}} = \frac{h_{A\cap B}}{h_B} \text{„}\to\text{“} \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}$$

falls $h_B \neq 0$.

1.5.2 Definition

Es sei $B \in \mathcal{A}$ ein Ereignis mit $\mathbb{P}(B) > 0$. Für $A \in \mathcal{A}$ nennt man

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

die bedingte Wahrscheinlichkeit von A unter der Bedingung B. Es gilt:

(i). A, B unabhängig $\Leftrightarrow \mathbb{P}(A|B) = \mathbb{P}(A)$

(ii). Ist $\mathbb{P}(A) > 0$, so gilt

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

1.5.3 Satz

Es seien $(B_n)_{n\in\mathbb{N}}$ unvereinbare Ereignisse (paarweise disjunkt) mit $\mathbb{P}(B_n) > 0$ und $\bigcup_{n\in\mathbb{N}} B_n = \Omega$. Dann gilt:

(i). $\mathbb{P}(A) = \sum_{n\in\mathbb{N}} \mathbb{P}(A \cap B_n)$

(ii). $\mathbb{P}(A) = \sum_{n\in\mathbb{N}} \mathbb{P}(A|B_n) \cdot \mathbb{P}(B_n)$

(iii). Bayes'sche Formel:

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(A|B_k) \cdot \mathbb{P}(B_k)}{\sum_{n\in\mathbb{N}} \mathbb{P}(A|B_n) \cdot \mathbb{P}(B_n)}$$

Beweis: Aufgabe

1.5.4 Definition

Es sei X eine Zufallsgröße und $B \in \mathcal{A}$ ein Ereignis mit $\mathbb{P}(B) > 0$. Dann heißt

$$F(t|B) := \mathbb{P}(X < t|B)$$

für $t \in \mathbb{R}$ die Verteilungsfunktion von X bzgl. B.

1.5.5 Aufgabe

Die Wahrscheinlichkeit p_k, dass eine Familie k Kinder hat, sei gegeben durch

$$\begin{aligned} p_0 &= p_1 = a \\ p_k &= \frac{(1-2a)}{2^{k-1}} \qquad (k \geq 2) \end{aligned}$$

mit $a \in (0, \frac{1}{2})$. Von einer Familie ist bekannt, dass sie genau zwei Jungen hat. Wie groß ist die Wahrscheinlichkeit, dass die Familie

(i). nur zwei Kinder hat?

(ii). genau zwei Mädchen hat?

(Annahme: $\mathbb{P}(M) = \mathbb{P}(J) = \frac{1}{2}$)

Lösung:

- Sei $J_2 :=$ 2 Jungen, $M_2 :=$ 2 Mädchen, $K_n := n$ Kinder. Gesucht: $\mathbb{P}(K_2|J_2), \mathbb{P}(M_2|J_2)$

- Nach 1.5.2 gilt:

$$\begin{aligned}
\mathbb{P}(K_2|J_2) &= \frac{\mathbb{P}(K_2)\cdot\mathbb{P}(J_2|K_2)}{\mathbb{P}(J_2)} \\
\mathbb{P}(M_2|J_2) &= \frac{\mathbb{P}(J_2\cap M_2)}{\mathbb{P}(J_2)}
\end{aligned}$$

mit

$$\mathbb{P}(K_n) = p_n = (1-2a)\cdot 2^{-(n-1)}$$

für $n \geq 2$. Wegen

$$\mathbb{P}(J_2|K_n) = \binom{n}{2}\cdot\frac{1}{2^n}$$

folgt

$$\begin{aligned}
\mathbb{P}(J_2) &= \sum_{n=2}^{\infty}\mathbb{P}(K_n)\cdot\mathbb{P}(J_2|K_n) \\
&= \sum_{n=2}^{\infty}\binom{n}{2}\cdot\frac{1}{2^{2n-1}}\cdot(1-2a) \\
&= \frac{(1-2a)}{4^2}\sum_{n=2}^{\infty} n\cdot(n-1)\cdot\frac{1}{4^{n-2}} \\
&\overset{*}{=} \frac{(1-2a)}{4^2}\,\frac{d^2}{dx^2}\left(\frac{1}{1-x}\right)\bigg|_{x=\frac{1}{4}} = \frac{8}{27}\cdot(1-2a)
\end{aligned}$$

(Beachte: Nutzung der geometrischen Reihe bei $*$ für $x = \frac{1}{4}$). Außerdem:

$$\begin{aligned}
\mathbb{P}(J_2\cap M_2) &= \mathbb{P}(K_4)\cdot\mathbb{P}(J_2\cap M_2|K_4) \\
&= (1-2a)\cdot\frac{1}{2^3}\cdot\binom{4}{2}\cdot\frac{1}{2^4} = \frac{3}{64}\cdot(1-2a)
\end{aligned}$$

Man erhält als Ergebnisse:

(i). $\frac{27}{64}$

(ii). $\frac{81}{512}$

1.5.6 Aufgabe: Problem der besten Wahl

- Situation: n Tankstellen zur Auswahl, Benzin reicht gerade bis zur n-ten Tankstelle. Ziel: an der billigsten Tankstelle tanken.
- Strategie: An $s-1$ ($s \geq 2$) Tankstellen vorbeifahren, den niedrigsten Preis notieren; wenn danach eine billigere Tankstelle kommt dort tanken; wenn nicht bei der letzten. Wie s wählen?

$$\begin{aligned}
p(s) &:= \mathbb{P}[\text{mit dieser Strategie die billigste gewählt}] \\
&= \sum_{k=s}^{n}\mathbb{P}[\text{k-te Tankstelle ist die billigste und wird gewählt}] \\
&= \sum_{k=s}^{n}\mathbb{P}[\text{k-te Tankstelle die billigste}]\cdot\mathbb{P}[\text{k-te TS gewählt}|\text{k-te TS billigste}] \\
&= \sum_{k=s}^{n}\frac{1}{n}\cdot\frac{s-1}{k-1} \to \max
\end{aligned}$$

da

$$\begin{aligned}
\mathbb{P}[\text{k-te TS gewählt}|\text{k-te TS billigste}] &= \mathbb{P}[\text{billigste von ersten k-1 ist unter ersten s-1}] \\
&= \frac{s-1}{k-1}
\end{aligned}$$

Damit

$$
\begin{aligned}
p(s) &= \frac{s-1}{n} \sum_{k=s-1}^{n-1} \frac{1}{k} = \frac{s-1}{n} \cdot \left(\sum_{k=1}^{n-1} \frac{1}{k} - \sum_{k=1}^{s-2} \frac{1}{k} \right) \\
&\overset{s,n \text{ groß}}{\approx} \frac{s-1}{n} \cdot (\ln(n-1) - \ln(s-2)) \\
&= \frac{s-1}{n} \cdot \ln\left(\frac{n-1}{s-2}\right) \approx \frac{s}{n} \cdot \ln \frac{n}{s}
\end{aligned}
$$

Maximum, wenn $s = \frac{n}{e}$ (Ableiten nach s). Dann $p(s) \approx \frac{1}{e} = 0,367$.

2 Gesetze

2.1 0-1-Gesetze

2.1.1 Lemma

Für jede Zahl $p \in [0, 1)$ gilt:

$$\ln(1-p) \leq -p$$

Beweis:

- Mittelwertsatz der Differentialrechnung:

$$\begin{aligned} -\ln(1-p) &= \ln 1 - \ln(1-p) = p \cdot \frac{1}{\eta} \\ &\geq p \end{aligned}$$

  mit $1 - p < \eta < 1$.

2.1.2 Lemma

Für jede Folge $\{p_n\}$ reeller Zahlen mit $0 \leq p_n \leq 1$ gilt:

$$\sum_{k=1}^{\infty} p_k = \infty \Rightarrow \lim_{n \to \infty} \prod_{k=1}^{n} (1 - p_k) = 0$$

Beweis:

- Ohne Beschränkung der Allgemeinheit sei $0 \leq p_n < 1$ für alle $n \in \mathbb{N}$. Nach 2.1.1 ist

$$\ln(1 - p_n) \leq -p_n$$

  und es folgt

$$\begin{aligned} \ln\left(\prod_{k=1}^{n}(1-p_k)\right) &= \sum_{k=1}^{n} \ln(1-p_k) \\ &\leq -\sum_{k=1}^{n} p_k \\ \Rightarrow \prod_{k=1}^{n}(1-p_k) &\leq e^{-\sum_{k=1}^{n} p_k} \to 0 \qquad (n \to \infty) \end{aligned}$$

2.1.3 Satz: Lemma von Borel-Cantelli

Sei $\{A_n\} \subset \mathcal{A}$ eine Folge von Ereignissen und

$$\begin{aligned} A &:= \lim_{n \to \infty} \sup A_n \\ &= \bigcap_{n \in \mathbb{N}} \bigcup_{m \geq n} A_m \end{aligned}$$

Dann gilt:

(i). $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty \Rightarrow \mathbb{P}(A) = 0$

(ii). Ist die Folge $\{A_n\}$ unabhängig, so gilt:

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty \quad \Rightarrow \quad \mathbb{P}(A) = 0$$
$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty \quad \Rightarrow \quad \mathbb{P}(A) = 1$$

Beweis:

(i). $A \subset \bigcup_{m \geq n} A_m$ nach Definition von A (für alle $n \in \mathbb{N}$), also

$$\mathbb{P}(A) \leq \mathbb{P}\left(\bigcup_{m \geq n} A_m\right) \leq \sum_{m=n}^{\infty} \mathbb{P}(A_m) \to 0 \qquad (n \to \infty)$$

(ii). Sei $\{A_n\}$ unabhängig und $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$. Nach 1.1.11 ist

$$\mathbb{P}(A) = 1 - \lim_{n \to \infty} \lim_{N \to \infty} \prod_{m=n}^{N} (1 - \mathbb{P}(A_m))$$

Nach 2.1.2 gilt für alle $n \in \mathbb{N}$:

$$\lim_{N \to \infty} \prod_{m=n}^{N} (1 - \mathbb{P}(A_m)) = 0$$

also $\mathbb{P}(A) = 1$.

2.1.4 Folgerung: 0-1-Gesetz von Borel

Existiert in einer Folge $\{A_n\}$ von Ereignissen eine unabhängige Teilfolge $\{A_{n_k}\}_k$ mit $\sum_{n=1}^{\infty} \mathbb{P}(A_{n_k}) = \infty$, so ist $\mathbb{P}(\limsup A_n) = 1$.

Beweis: Folgt aus $\limsup A_{n_k} \subseteq \limsup A_n$ und aus 2.1.3.

2.1.5 Beispiele

(i). Eine Münze wurde unendlich oft hintereinander geworfen. Man gebe die Wahrscheinlichkeit dafür an, dass unendlich oft zweimal hintereinander Kopf geworfen wird.

Lösung: Sei $A_n :=$ sowohl beim n-ten als auch beim $(n+1)$-ten Wurf Kopf. $\{A_n\}$ ist nicht unabhängig, aber $\{A_{2n}\}$ ist unabhängig. Wegen $\mathbb{P}(A_n) = \frac{1}{4}$ folgt

$$\sum_{n \in \mathbb{N}} \mathbb{P}(A_{2n}) = \infty$$

und aus 2.1.4 somit $\mathbb{P}(\limsup A_{2n}) = 1$.

Bemerkung: Was ist hier $(\Omega, \mathcal{A}, \mathbb{P})$? $\Omega := \{0,1\}^{\infty}$, $\mathcal{A}$: Produkt$-\sigma$-Algebra, $\mathbb{P}$: Produktmaß

(ii). Eine Folge $\{t_n\}$ reeller Zahlen nennen wir eine Unterfolge (Oberfolge) für die Folge der Zufallsgrößen (X_n), wenn die Ereignisse $[X_n > t_n]$ (bzw. $[X_n \leq t_n]$) mit Wahrscheinlichkeit 1 unendlich oft eintreten.

Sind die X_k unabhängig, so ist jede Folge $\{t_n\}$ Ober- oder Unterfolge für (X_n). Das folgt aus 2.1.3 und aus der Tatsache, dass die Reihen

$$\sum_{n=1}^{\infty} \mathbb{P}[X_n > t_n] \qquad\qquad \sum_{n=1}^{\infty} \mathbb{P}[X_n \leq t_n] = \sum_{n=1}^{\infty} (1 - \mathbb{P}[X_n > t_n])$$

nicht gleichzeitig konvergieren können.

Bemerkung: Es ist möglich, dass eine Folge Ober- und Unterfolge ist. Beispiel:

$$X_n = 0 \qquad\qquad t_n = (-1)^n$$

2.1.6 Aufgabe

(i). Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen mit $a_1 \geq a_2 \geq \ldots \geq 0$. Dann gilt:

$$\sum_{n=1}^{\infty} a_n < \infty \Leftrightarrow \sum_{k=1}^{\infty} 2^k \cdot a_{2^k} < \infty$$

(ii). Die Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$$

ist konvergent für $\alpha > 1$.

(iii). Für $a \in (0,1)$, $b > 1$ gilt

$$\sum_{n=1}^{\infty} \frac{1}{(n \cdot \log_b n)^a} = \infty$$

2.1.7 Lemma

Es seien $a, p \in (0,1)$, $\log := \log_{\frac{1}{p}}$ und $k_n := \lfloor n \cdot \log n \rfloor$. Dann ist $(k_n)_2^\infty$ eine streng monoton wachsende Folge von positiven Zahlen mit

(i). $k_n + \lfloor a \cdot \log k_n \rfloor + 1 < k_{n+1}$ für n hinreichend groß

(ii). $\sum_{n=2}^{\infty} \frac{1}{k_n^a} = \infty$

Beweis:

(i). Nach Definition gilt $k_n \leq n \cdot \log n$. Damit folgt:

$$\begin{aligned}
\lfloor a \cdot \log k_n \rfloor &\leq a \cdot \log(n \cdot \log n) = a \cdot (\log n + \log\log n) \\
\Rightarrow k_n + \lfloor a \cdot \log k_n \rfloor &\leq n \cdot \log n + a \cdot (\log n + \log\log n) \\
&= (n+1) \cdot \log n - (1-a) \cdot \log n + a \cdot \log\log n \\
&\overset{*}{\leq} k_{n+1} + 1 \underbrace{-(1-a) \cdot \log n + a \cdot \log\log n}_{\to -\infty (n\to\infty)} \\
&< k_{n+1} - 1
\end{aligned}$$

Bei $(*)$:

$$k_{n+1} + 1 \geq (n+1) \cdot \log(n+1) \geq (n+1) \cdot \log n$$

(ii). Folgt aus 2.1.6(iii)

2.1.8 Aufgabe

Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge unabhängiger Zufallsgrößen mit $\mathbb{P}[X_n = 1] = p$, $\mathbb{P}[X_n = 0] = 1 - p$ mit $p \in (0,1)$. Für $n \in \mathbb{N}$ definieren wir die Zufallsgrößen N_n durch

$$N_n(w) := \begin{cases} 0 & X_n(w) = 0 \\ j & X_n(w) = \ldots = X_{n+j-1}(w) = 1, X_{n+j} = 0 \end{cases}$$

(i). Zu zeigen:

$$\mathbb{P}[N_n = j] = p^j \cdot (1-p) \qquad \mathbb{P}[N_n \geq j] = p^j \qquad \mathbb{E}(N_n) = \frac{p}{1-p}$$

(ii). Ist $n + j \leq m$, so sind die Ereignisse $[N_n \geq j]$ und $[N_m \geq k]$ unabhängig für $k \geq 0$.

2.1.9 Satz

Es gilt:

$$\mathbb{P}\left[\lim_{n\to\infty} \sup \frac{N_n}{\log n} = 1\right] = 1$$

Beweis:

- Für $a > 1$ gilt:

$$\begin{aligned} \mathbb{P}[N_n > a \cdot \log n] &\leq \mathbb{P}[N_n \geq \lfloor a \cdot \log n \rfloor] \\ &\overset{2.1.8(i)}{=} p^{\lfloor a \cdot \log n \rfloor} \leq p^{a \cdot \log n - 1} = \frac{1}{p \cdot n^a} \end{aligned}$$

  Damit

$$\mathbb{P}[N_n > a \cdot \log n, \text{unendlich oft}] = 0$$

  wegen Borel-Cantelli, 2.1.6.(ii). Damit folgt für alle $a > 1$:

$$\begin{aligned} \mathbb{P}\left[\lim_{n\to\infty} \sup \frac{N_n}{\log n} \leq a\right] &= 1 \\ \Rightarrow \mathbb{P}\left[\lim_{n\to\infty} \sup \frac{N_n}{\log n} \leq 1\right] &= 1 \end{aligned}$$

- Die Ereignisse $[N_n > a \cdot \log n]$ sind nicht unabhängig. Sei jetzt $a \in (0,1)$ und $k_n = \lfloor n \cdot \log n \rfloor$ aus 2.1.7. Dann sind die Ereignisse

$$A_n := [N_{k_n} \geq \lfloor a \cdot \log k_n \rfloor + 1]$$

  für $n \geq n_0$ unabhängig, wenn n_0 hinreichend groß ist (2.1.8(ii),2.1.7(i)). Aus

$$\mathbb{P}(A_n) = p^{\lfloor a \cdot \log k_n \rfloor + 1} \geq p^{a \cdot \log k_n + 1} = \frac{p}{k_n^a}$$

  folgt mit 2.1.7(ii):

$$\sum_{n=n_0}^{\infty} \mathbb{P}(A_n) = \infty$$

  und mit Borel-Cantelli:

$$\begin{aligned} \forall a \in (0,1) : \mathbb{P}[N_n \geq a \cdot \log n, \text{unendlich oft}] &\geq \mathbb{P}[N_{k_n} \geq a \cdot \log k_n, \text{unendlich oft}] \\ &\geq \mathbb{P}[N_{k_n} \geq \lfloor a \cdot \log k_n \rfloor + 1, \text{unendlich oft}] \\ &= 1 \\ \Rightarrow \mathbb{P}\left[\lim_{n\to\infty} \sup \frac{N_n}{\log n} \geq 1\right] &= 1 \end{aligned}$$

2.1.10 Satz

Es sei (X_n) eine Folge unabhängiger Zufallsgrößen. Dann gilt:

$$\lim_{n\to\infty} X_n = 0 \,\text{fast sicher} \Leftrightarrow \forall k \in \mathbb{N} : \sum_{n=1}^{\infty} \mathbb{P}\left[|X_n| \geq \frac{1}{k}\right] < \infty$$

Beweis:

- Definiere $A := \{w : X_n(w) \to 0\}$. Aus

$$\begin{aligned} A &= \{w : \forall k \exists n_0 \forall n \geq n_0 : |X_n(w)| < \frac{1}{k}\} \\ &= \bigcap_{k=1}^{\infty} \bigcup_{n_0=1}^{\infty} \bigcap_{n=n_0}^{\infty} \{w : |X_n(w)| < \frac{1}{k}\} \end{aligned}$$

  folgt, dass A ein Ereignis ist, also $A \in \mathcal{A}$. Borel-Cantelli:

$$\begin{aligned} \sum_{n_0=1}^{\infty} \mathbb{P}\left[|X_{n_0}| \geq \frac{1}{k}\right] < \infty \quad &\Leftrightarrow \quad \mathbb{P}\left(\bigcap_{n_0=1}^{\infty} \bigcup_{n=n_0}^{\infty} \left[|X_n| \geq \frac{1}{k}\right]\right) = 0 \\ &\Leftrightarrow \quad \mathbb{P}\left(\bigcup_{n_0=1}^{\infty} \bigcap_{n=n_0}^{\infty} \left[|X_n| < \frac{1}{k}\right]\right) = 1 \\ &\Leftrightarrow \quad \mathbb{P}(A) = 1 \end{aligned}$$

Bemerkung:

- Unabhängigkeit ist wichtig! Beispiel: $\Omega = [0,1], \mathcal{A} = \mathcal{B}([0,1])$, $\mathbb{P} = \lambda|_{[0,1]}$, $X_n = 1_{[0,a_n)}$ mit $a_n \to 0$ $(n \to \infty)$, $0 < a_n \leq 1$. Dann ist

$$\lim_{n\to\infty} X_n = 0 \qquad \mathbb{P}\left[X_n \geq \frac{1}{k}\right] = a_n$$

2.1.11 Definition

Es sei (X_n) eine Folge von Zufallsgrößen. Ein Ereignis $A \in \mathcal{A}$ heißt ein terminales Ereignis (Restereignis), falls $A \in \sigma(X_n, X_{n+1}, \ldots)$ für alle $n \in \mathbb{N}$.
Hierbei bezeichnet $\sigma(X_n, X_{n+1}, \ldots)$ die σ-Algebra, die durch die Ereignisse der Form $[X_j \in B]$ $(j \geq n, B \in \mathcal{B}(\mathbb{R}))$ erzeugt wird. Vergleiche Skript Maßtheorie, Definition 5.5.

Beispiel:

$$A_r := \lim_{k\to\infty} \sup[X_k > r] = \bigcap_{n=1}^{\infty} \underbrace{\bigcup_{k=n}^{\infty} [X_k > r]}_{\in \sigma(X_n, X_{n+1}, \ldots)}$$

2.1.12 Aufgabe

Sei (X_n) eine Folge von unabhängigen Zufallsgrößen. Zu zeigen:

(i). $A \in \sigma(X_{n+1}, X_{n+2}, \ldots) \Rightarrow A$ ist unabhängig von der σ-Algebra $\sigma(X_1, \ldots, X_n)$

(ii). A unabhängig von $\sigma(X_1, \ldots, X_n)$ für alle $n \geq 1 \Rightarrow A$ ist unabhängig von $\sigma(X_n; n \in \mathbb{N})$.

2.1.13 Satz: 0-1-Gesetz von Kolmogorov

Seien (X_n) unabhängige Zufallsgrößen. Ist A ein terminales Ereignis, so gilt $\mathbb{P}(A) \in \{0,1\}$.

Beweis:

- Sei $A \in \sigma(X_{n+1}, X_{n+2}, \ldots)$ für alle $n \geq 0$. Dann ist A unabhängig von $\sigma(X_1, \ldots, X_n)$ nach 2.1.12(i) für alle $n \in \mathbb{N}$. Nach 2.1.12(ii) ist A unabhängig von $\sigma(X_n; n \in \mathbb{N})$. A ist aber in $\sigma(X_n; n \in \mathbb{N})$ enthalten, also A unabhängig von sich selbst:

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P}(A \cap A) = \mathbb{P}(A) \cdot \mathbb{P}(A) \\ \Rightarrow \mathbb{P}(A) &= 0 \vee \mathbb{P}(A) = 1 \end{aligned}$$

2.1.14 Definition

- Notation: $\mathbb{R}^\infty := \bigtimes_{i=1}^{\infty} \mathbb{R}$,

$$\mathcal{B}^\infty := \sigma(\{B_n \times \mathbb{R} \times \mathbb{R} \times \ldots \subset \mathbb{R}^\infty; B_n \in \mathcal{B}(\mathbb{R}^n)\})$$

  Diese Mengen B bilden eine Algebra (Aufgabe).

- Eine Menge $B \subset \mathbb{R}^\infty$ heißt permutierbar, wenn $t := (t_1, t_2, \ldots) \in B \Rightarrow \tau(t) := (t_{\tau(1)}, t_{\tau(2)}, \ldots) \in B$ für eine beliebige endliche Permutation τ von $\{1, 2, \ldots\}$, d.h. $\tau(j) = j$ bis auf endlich viele $j \in \mathbb{N}$.

- Es sei nun $(X_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsgrößen und $Y := (X_1, X_2, \ldots)$. Die Abbildung $Y : \Omega \to \mathbb{R}^\infty$ ist $\mathcal{B}^\infty$-messbar, d.h. Y ist eine $\mathbb{R}^\infty$-wertige Zufallsgröße, da

$$[Y \in B] = [(X_1, \ldots, X_n) \in B_n] \in \mathcal{A}$$

  (Messbarkeit am Erzeuger). Ist die Menge $B \in \mathcal{B}^\infty$ permutierbar, so heißt auch das Ereignis $A := [Y \in B]$ permutierbar.

- Beispiel: Das Ereignis $[X_n \to 0]$ ist permutierbar.

2.1.15 Satz: 0-1-Gesetz von Hewitt und Savage

Sind die Zufallsgrößen $(X_n)_{n\in\mathbb{N}}$ unabhängig und identisch verteilt und ist A ein permutierbares Ereignis, dann gilt $\mathbb{P}(A) \in \{0,1\}$.

2.1.16 Aufgabe

(Bezeichnungen wie in 2.1.14, Voraussetzungen wie in 2.1.15)

Bezeichne μ die Verteilung von Y, d.h.

$$\mu(B) := \mathbb{P}[Y \in B]$$

für $B \in \mathcal{B}^\infty$. Für $D_n, D \in \mathcal{B}^\infty$ $(n \in \mathbb{N})$ schreiben wir $D_n \to D$, wenn

$$\lim_{n\to\infty} \mu(D_n \Delta D) = 0$$

Dann gilt:

(i). $D_n \to D \Rightarrow \mu(D) = \lim_{n\to\infty} \mu(D_n)$

(ii). $D_n \to D, C_n \to D \Rightarrow D_n \cap C_n \to D$

(iii). Ist $\tau : \mathbb{R}^\infty \to \mathbb{R}^\infty$ eine injektive Abbildung, so ist $\tau(C \Delta D) = \tau(C) \Delta \tau(D)$

(iv). Ist τ eine endliche Permutation, so gilt

$$\forall B \in \mathcal{B}^\infty : \mu(\tau(B)) = \mu(B)$$

Hinweis: Genügt für Erzeuger zu zeigen (siehe 2.1.14) + Unabhängigkeit benutzen

2.1.17 Beweis (Satz 2.1.15)

- Sei $A = [Y \in B]$ ein permutierbares Ereignis, d.h. $Y = (X_1, X_2, \ldots)$, $B \in \mathcal{B}^\infty$ permutierbar. Für $(x_1, x_2, \ldots) \in \mathbb{R}^\infty$ setzen wir

$$\tau_n(x_1, x_2, \ldots) := (x_{n+1}, x_{n+2}, \ldots, x_{2n}, x_1, \ldots, x_n, x_{2n+1}, \ldots)$$

τ_n ist eine endliche Permutation. Aus dem Approximationssatz folgt die Existenz von $B_{k_n} \in \mathcal{B}(\mathbb{R}^{k_n})$ mit

$$D_n := B_{k_n} \times \mathbb{R} \times \mathbb{R} \times \ldots \to B$$

Behauptung: $\tau_{k_n}(D_n) \to B = \tau_{k_n}(B)$. Dies folgt aus

$$\begin{aligned} \mu(\tau_{k_n}(D_n) \Delta\, \tau_{k_n}(B)) &\overset{2.1.16(iii)}{=} \mu(\tau_{k_n}(D_n \Delta B)) \\ &\overset{2.1.16(iv)}{=} \mu(D_n \Delta\, B) \to 0 \end{aligned}$$

Aus 2.1.16(ii) folgt somit

$$\underbrace{D_n \cap \tau_{k_n}(D_n)}_{B_{k_n} \times B_{k_n} \times \mathbb{R} \times \ldots} \to B$$

Man erhält:

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P}[Y \in B] = \mu(B) \\ &= \lim_{n \to \infty} \mu(D_n) = \lim_{n \to \infty} \mathbb{P}[(X_1, \ldots, X_{k_n}) \in B_{k_n}] \end{aligned}$$

und

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P}[Y \in B] = \lim_{n \to \infty} \mathbb{P}[(X_1, \ldots, X_{2k_n}) \in B_{k_n} \times B_{k_n}] \\ &\overset{*}{=} \lim_{n \to \infty} \mathbb{P}[(X_1, \ldots, X_{k_n}) \in B_{k_n}]^2 \end{aligned}$$

also $\mathbb{P}(A) = \mathbb{P}(A)^2$, d.h. $\mathbb{P}(A) \in \{0, 1\}$. Bei $(*)$: Unabhängigkeit und Gleichverteilung der (X_n)

2.1.18 Aufgabe

Sei $(X_n)_{n \in \mathbb{N}}$ eine Folge von unabhängigen Zufallsgrößen.

(i). Die Folge (X_n) konvergiert oder divergiert fast sicher.

(ii). Die Reihe $\sum_{n=1}^\infty X_n$ konvergiert oder divergiert fast sicher.

(iii). Wenn $X_n \to X$ oder $\frac{1}{b_n} \sum_{i=1}^n X_i \to X$ fast sicher mit $b_n \to \infty$, dann ist X fast sicher konstant.

Hinweis: 0-1-Gesetz von Kolmogorov

2.2 Einige Ungleichungen

2.2.1 Satz: Ungleichung von Hájek-Rènyi

Gegeben seien die unabhängigen Zufallsgrößen $X_1, \ldots, X_n \in \mathcal{L}^2(\mathbb{P})$ und n positive Zahlen $r_1 \geq r_2 \geq \ldots \geq r_n$. Wir setzen

$$S_i := \sum_{k=1}^{i} (X_k - \mathbb{E}X_k)$$

für $i \in \{1, \ldots, n\}$. Dann gilt für jedes $m \in \{1, \ldots, n\}$ und jedes $\varepsilon > 0$:

$$\mathbb{P}\left[\max_{m \leq i \leq n} r_i \cdot |S_i| \geq \varepsilon\right] \leq \frac{1}{\varepsilon^2} \cdot \left(r_m^2 \cdot \sum_{j=1}^{m} D^2(X_j) + \sum_{j=m+1}^{n} r_j^2 \cdot D^2(X_j)\right)$$

2.2.2 Folgerung: Ungleichung von Kolmogorov

$$\mathbb{P}\left[\max_{1\leq i\leq n} |S_i| \geq \varepsilon\right] \leq \frac{1}{\varepsilon^2} \cdot \sum_{i=1}^{n} D^2(X_i)$$

Beweis: $r_i = 1, m = 1$ in Satz 2.2.1

2.2.3 Folgerung: Ungleichung von Tschebysheff

$$\mathbb{P}[|X - \mathbb{E}X| \geq \varepsilon] \leq \frac{D^2(X)}{\varepsilon^2}$$

Beweis: $m = n = 1$, $r_1 = 1, X = X_1$ in Satz 2.2.1

Beweis von Satz 2.2.1:

(i). Ohne Beschränkung der Allgemeinheit sei $\mathbb{E}X_j = 0$ für $j = 1, \ldots, n$. Dann ist

$$S_i = \sum_{j=1}^{i} X_j$$

Definiere

$$\begin{aligned} A &:= \left[\max_{m\leq i\leq n} r_i \cdot |S_i| \geq \varepsilon\right] \\ B_i &:= [r_i \cdot |S_i| \geq \varepsilon] \end{aligned}$$

Dann ist $A = \bigcup_{i=m}^{n} B_i$. Setze

$$\begin{aligned} A_m &:= B_m \\ A_{m+1} &:= B_m^c \cap B_{m+1} \\ &\vdots \\ A_n &:= \left(\bigcap_{i=m}^{n-1} B_i^c\right) \cap B_n \end{aligned}$$

Per Konstruktion gilt $A_i \subset B_i$, $\dot\bigcup_{j=m}^{n} A_j = A$. Damit

$$\mathbb{P}(A) = \sum_{j=m}^{n} \mathbb{P}(A_j) \qquad (1)$$

(ii). Sei

$$Z := \sum_{i=m}^{n} (r_i^2 - r_{i+1}^2) \cdot S_i^2$$

wobei $r_{n+1} := 0$. Wegen

$$\mathbb{E}(S_i^2) = D^2(S_i) = \sum_{j=1}^{i} D^2(X_j)$$

folgt

$$\begin{aligned} \mathbb{E}Z &= \sum_{i=m}^{n} \sum_{j=1}^{i} (r_i^2 - r_{i+1}^2) \cdot D^2(X_j) \\ &= \sum_{j=1}^{m} \sum_{i=m}^{n} (r_i^2 - r_{i+1}^2) \cdot D^2(X_j) + \sum_{j=m+1}^{n} \sum_{i=j}^{n} (r_i^2 - r_{i+1}^2) \cdot D^2(X_j) \\ &= r_m^2 \cdot \sum_{j=1}^{m} D^2(X_j) + \sum_{j=m+1}^{n} r_j^2 \cdot D^2(X_j) \end{aligned}$$

$\frac{1}{\varepsilon^2} \cdot \mathbb{E}Z$ ist also die rechte Seite der Ungleichung in Satz 2.2.1.

(iii). Wir setzen $Y_i := 1_{A_i}$ für $i = m, \dots, n$. Behauptung: Für $m \le i \le j \le n$:

$$\mathbb{E}(Y_i \cdot S_j^2) \ge \frac{\varepsilon^2}{r_i^2} \cdot \mathbb{P}(A_i) \qquad (2)$$

Beweis:

- Sei

$$u_{ij} := S_j - S_i = \sum_{k=i+1}^{j} X_k$$

für $m \le i \le j \le n$. Dann ist $S_j^2 = (S_i + u_{ij})^2$, also

$$\mathbb{E}(Y_i \cdot S_j^2) = \mathbb{E}(S_i^2 Y_i) + \mathbb{E}(Y_i \cdot u_{ij}^2) + 2\mathbb{E}(Y_i \cdot S_i \cdot u_{ij}) \qquad (3)$$

Aufgabe: Die Zufallsgrößen $Y_i \cdot S_i$ und u_{ij} sind unabhängig, also $\mathbb{E}(Y_i \cdot S_i \cdot u_{ij}) = 0$ wegen $\mathbb{E}(u_{ij}) = 0$. Aus Positivität der Summanden:

$$\begin{aligned} \mathbb{E}(Y_i \cdot S_j^2) &\ge \mathbb{E}(Y_i \cdot S_i^2) = \int_{A_i} S_i^2 \, d\mathbb{P} \\ &\overset{A_i \subset B_i}{\ge} \frac{\varepsilon^2}{r_i^2} \cdot \mathbb{P}(A_i) \end{aligned}$$

(iv). Aus $Z \ge 0$ und $1 \ge 1_A = \sum_{j=m}^n 1_{A_j}$ folgt:

$$\begin{aligned} Z &\ge \sum_{i=m}^{n} \underbrace{1_{A_i}}_{Y_i} \cdot Z \\ \Rightarrow \mathbb{E}Z &\ge \mathbb{E}\left(\sum_{i=m}^{n} Y_i \cdot Z \right) \\ &= \sum_{i=m}^{n} \sum_{j=m}^{n} (r_j^2 - r_{j+1}^2) \cdot \mathbb{E}(Y_i \cdot S_j^2) \\ &\overset{i \ge m}{\ge} \sum_{i=m}^{n} \sum_{j=i}^{n} (r_j^2 - r_{j+1}^2) \cdot \mathbb{E}(Y_i \cdot S_j^2) \\ &\overset{(2)}{\ge} \sum_{i=m}^{n} \sum_{j=i}^{n} (r_j^2 - r_{j+1}^2) \cdot \frac{\varepsilon^2}{r_i^2} \cdot \mathbb{P}(A_i) \\ &= \varepsilon^2 \cdot \sum_{i=m}^{n} \frac{\mathbb{P}(A_i)}{r_i^2} \cdot \sum_{j=i}^{n} (r_j^2 - r_{j+1}^2) \overset{(1)}{=} \varepsilon^2 \cdot \mathbb{P}(A) \end{aligned}$$

2.2.4 ?

2.2.5 Satz: Bernstein

Für jede stetige Funktion $f : [0,1] \to \mathbb{R}$ gilt

$$f(x) = \lim_{n \to \infty} \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \cdot \binom{n}{k} \cdot x^k \cdot (1-x)^{n-k}$$

für $x \in [0,1]$, wobei die Konvergenz gleichmäßig ist.

Beispiel: Für $n = 1$: $f(0) + (f(1) - f(0)) \cdot x$

2.2.6 Folgerung: Satz von Weierstraß

Jede stetige Funktion $f : [0, 1] \to \mathbb{R}$ lässt sich gleichmäßig durch Polynome approximieren.

Beweis: (von Satz 2.2.5)

- Für beliebige $x \in [0, 1]$ und $n \in \mathbb{N}$ sei Y_n^x eine Zufallsgröße so, dass $n \cdot Y_n^x$ binomialverteilt ist mit den Parametern n und x. Dann gilt:
$$\mathbb{E}(Y_n^x) = x \qquad D^2(Y_n^x) = \frac{(1-x) \cdot x}{n}$$
Damit
$$\mathbb{E}(f(Y_n^x)) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \cdot \binom{n}{k} \cdot x^k \cdot (1-x)^{n-k}$$
Deshalb brauchen wir nur zu zeigen, dass
$$\lim_{n\to\infty} \mathbb{E}(f(Y_n^x) - f(x)) = 0$$
gleichmäßig auf $[0, 1]$. Nach Tschebysheff-Ungleichung:
$$\begin{aligned} \mathbb{P}[|Y_n^x - x| > n^{-\frac{1}{4}}] &\leq \sqrt{n} \cdot D^2(Y_n^x) = \frac{x \cdot (1-x)}{\sqrt{n}} \\ &\leq \frac{1}{\sqrt{n}} \end{aligned}$$
Sei K eine Konstante mit $|f| \leq K$ und bezeichne
$$A := [|Y_n^x - x| > n^{-\frac{1}{4}}]$$
Dann ist $\mathbb{P}(A) \leq \frac{1}{\sqrt{n}}$ und
$$\begin{aligned} \mathbb{E}(|f(Y_n^x) - f(x)|) &= \int_\Omega |f(Y_n^x(w)) - f(x)|\, d\mathbb{P}(w) \\ &= \int_A |f(Y_n^x(w)) - f(x)|\, d\mathbb{P}(w) + \int_{\Omega\setminus A} |f(Y_n^x(w)) - f(x)|\, d\mathbb{P}(w) \\ &\leq \frac{2K}{\sqrt{n}} + \sup_{|y-z|\leq n^{-\frac{1}{4}}} |f(y) - f(z)| \\ &\to 0 \qquad (n \to \infty) \end{aligned}$$
da f gleichmäßig stetig.

2.2.7 Definition

- Es sei X eine Zufallsgröße. Eine reelle Zahl $m = m(X)$ heißt Median (Zentralwert) von X, wenn
$$\mathbb{P}[X \leq m] \geq \frac{1}{2} \leq \mathbb{P}[X \geq m]$$
gilt. Im Allgemeinen ist der Median nicht eindeutig, z.B. für $\mathbb{P}[X = 0] = \mathbb{P}[X = 1] = \frac{1}{2}$ sind alle $m \in [0, 1]$ Mediane.
- Jede Zufallsgröße X besitzt mindestens einen Median, denn
$$m := \inf_{x\in\mathbb{R}} \{x \in \mathbb{R} : \mathbb{P}[X \leq x] \geq \frac{1}{2}\}$$
Dann gilt $\mathbb{P}[X \leq m - \varepsilon] < \frac{1}{2}$ für alle $\varepsilon > 0$, also $\mathbb{P}[X \leq m] \leq \frac{1}{2}$.
- X heißt symmetrisch, wenn $\mathbb{P}[X < t] = \mathbb{P}[X > -t]$ für alle $t \in \mathbb{R}$. Dann ist 0 ein Median, $\mathbb{E}X = 0$ falls $X \in \mathcal{L}^1(\mathbb{P})$. Summen von unabhängigen, symmetrischen Verteilungen sind symmetrisch.

2.2.8 Satz: Ungleichung von Lévy

Es seien $X_1, \ldots, X_n$ unabhängige Zufallsgrößen und

$$S_j := \sum_{i=1}^{j} X_i$$

Für jedes $\varepsilon > 0$ gilt:

$$\begin{aligned}
\mathbb{P}\left[\max_{1\le j\le n}(S_j - m(S_j - S_n)) \ge \varepsilon\right] &\le 2\cdot\mathbb{P}[S_n \ge \varepsilon] \\
\mathbb{P}\left[\max_{1\le j\le n}|S_j - m(S_j - S_n)| \ge \varepsilon\right] &\le 2\cdot\mathbb{P}[|S_n| \ge \varepsilon]
\end{aligned}$$

wobei $m(Y)$ einen beliebigen Median von Y bezeichnet.

Beweis:

(i). Sei $S_0 := 0$, $T(w) :=$ die kleinste Zahl $j \in \{1, \ldots, n\}$ mit

$$S_j(w) - m(S_j - S_n) \ge \varepsilon$$

falls eine solche Zahl j existiert, sonst $T(w) := n + 1$. Definiere

$$B_j := [m(S_j - S_n) \ge S_j - S_n]$$

für $1 \le j \le n$. Dann gilt $\mathbb{P}(B_j) \ge \frac{1}{2}$, $[T = j] \in \sigma(X_1, \ldots, X_j)$ für $j = 1, \ldots, n$, $B_j \in \sigma(X_{j+1}, \ldots, X_n)$. Es folgt, dass $[T = j]$ und B_j unabhängig nach 2.1.12. Wegen

$$[S_n \ge \varepsilon] \supset \dot{\bigcup_{j=1}^{n}}(B_j \cap [T = j])$$

gilt

$$\begin{aligned}
\mathbb{P}[S_n \ge \varepsilon] &\ge \sum_{j=1}^{n}\mathbb{P}(B_j \cap [T = j]) \\
&= \sum_{j=1}^{n}\mathbb{P}(B_j)\cdot\mathbb{P}[T = j] \\
&\ge \frac{1}{2}\cdot\mathbb{P}[1 \le T \le n]
\end{aligned}$$

(ii). Aufgabe (Hinweis: in (i) $-X_j$ anstelle von X_j)

2.2.9 Folgerung

Es seien $X_1, \ldots, X_n$ unabhängige symmetrische Zufallsgrößen,

$$S_j := \sum_{i=1}^{j} X_i$$

Für jedes $\varepsilon > 0$ gilt:

$$\begin{aligned}
\mathbb{P}\left[\max_{1\le j\le n} S_j \ge \varepsilon\right] &\le 2\mathbb{P}[S_n \ge \varepsilon] \\
\mathbb{P}\left[\max_{1\le j\le n} |S_j| \ge \varepsilon\right] &\le 2\mathbb{P}[|S_n| \ge \varepsilon]
\end{aligned}$$

Beweis: Folgt aus 2.2.8 wegen $m(S_j - S_n) = 0$ für $j = 1, \ldots, n$ (Symmetrie)

2.2.10 Aufgabe

Seien $A_i \in \mathcal{A}$ für $i = 1, \ldots, n$ mit $n \geq 2$. Dann gilt:

$$-\left(1 - \frac{1}{n}\right)^n \leq \mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) - \prod_{i=1}^{n} \mathbb{P}(A_i) \leq \frac{1}{n^{\frac{1}{n-1}}} \cdot \left(1 - \frac{1}{n}\right)$$

2.2.11 Aufgabe

2.2.12 Satz: Jensensche Ungleichung

Es seien $I \subseteq \mathbb{R}$ ein Intervall, $f : I \to \mathbb{R}$ eine konvexe Funktion und $X : \Omega \to I$ mit $X \in \mathcal{L}^1(\lambda^1)$. Dann gilt $\mathbb{E}(f^-(X)) < \infty$ und

$$f(\mathbb{E}X) \leq \mathbb{E}(f(X)) := -\mathbb{E}(f^-(X)) + \mathbb{E}(f^+(X))$$

Ist f im Punkt $\mathbb{E}X$ streng konvex, dann gilt die Gleichheit genau dann wenn X fast sicher konstant ist.

Beweis:

(i). Sei
$$h(t) := f(\mathbb{E}X) + a \cdot (t - \mathbb{E}X)$$
eine Gerade durch den Punkt $(\mathbb{E}X, f(\mathbb{E}X))$ mit $h(t) \leq f(t)$ für alle $t \in I$. Damit $f^-(X) \leq h^-(X)$ und somit
$$\mathbb{E}(f^-(X)) \leq \mathbb{E}(h^-(X)) < \infty$$
also ist $\mathbb{E}(f(X))$ definiert und
$$\mathbb{E}(f(X)) \geq \mathbb{E}(h(X)) = f(\mathbb{E}X)$$

(ii). Im Fall der Gleichheit gilt
$$\begin{aligned} & \mathbb{E}(f(x) - h(x)) = 0 \\ \Leftrightarrow \quad & f(x) - h(x) = 0 \text{ fast sicher} \\ \Leftrightarrow \quad & X = \mathbb{E}X \text{ fast sicher} \end{aligned}$$
wegen $f(x) - h(x) > 0$ für $X(w) \neq \mathbb{E}X$.

Bemerkung: Bereits in Teil 1 (Maßtheorie) in §13 für positive Funktionen bewiesen.

2.2.13 Definition

X sei eine Zufallsvariable mit Werten $S := \{s_1, \ldots, s_n\}$ und $p(s) := \mathbb{P}[X = s]$ für $s \in S$. Unter der Entropie von X versteht man die Zahl

$$H(X) := \mathbb{E}(-\log_2 p(X)) = -\sum_{s \in S} p(s) \cdot \log_2 p(s)$$

wobei $0 \cdot \log_2 0 := 0$ gesetzt wird.

2.2.14 Aufgabe

Seien X, Y wie in 2.2.13. Zu zeigen:

(i). $0 \leq H(X) \leq \log_2 n$

(ii). $0 = H(X) \Leftrightarrow X = c$ fast sicher, $H(X) = \log_2 n \Leftrightarrow X$ ist gleichmäßig verteilt (iii). $H(X,Y) \leq H(X) + H(Y)$, Gleichheit $\Leftrightarrow$ X und Y unabhängig

(iv). Mit $q(s) := \mathbb{P}[Y = s]$ gilt:

$$\sum_{s \in S} p(s) \cdot \log_2 \left(\frac{p(s)}{q(s)} \right) \geq 0$$

mit Y mit Werten in S. Sogenannter Kullback-Leibler-Abstand zwischen p und q (nicht symmetrisch).

Bemerkung: Für $a \geq 0, b > 0$ definieren wir

$$\begin{aligned} 0 \cdot \log \frac{0}{a} &:= 0 \\ 0 \cdot \log \frac{a}{0} &:= 0 \\ b \cdot \log \frac{b}{0} &:= \infty \\ b \cdot \log \frac{0}{b} &:= -\infty \end{aligned}$$

Hinweis: Jensensche Ungleichung mit $f = -\log_2$.

2.3 Konvergenz von Zufallsgrößen

2.3.1 Definition

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Maßraum. Eine Folge $(X_n)_{n \in \mathbb{N}}$ heißt

(i). fast sicher konvergent gegen eine Zufallsgröße X, wenn $\mathbb{P}[\lim_{n\to\infty} X_n = X] = 1$, d.h.

$$\lim_{n \to \infty} X_n(w) = X(w)$$

für $\mathbb{P}$-fast alle $w \in \Omega$.

(ii). konvergent in Wahrscheinlichkeit gegen die Zufallsgröße X, wenn für alle $\varepsilon > 0$:

$$\lim_{n \to \infty} \mathbb{P}[|X_n - X| > \varepsilon] = 0$$

(iii). konvergent in Verteilung gegen die Zufallsgröße X, wenn für jeden Stetigkeitspunkt x von F gilt:

$$\lim_{n \to \infty} F_n(x) = F(x)$$

2.3.2 Anmerkungen

(i). $[w \in \Omega : \lim X_n = X(w)]$ ist ein Ereignis, siehe 2.1.10.

(ii). Eindeutigkeit des Grenzwertes:

(1) $X_n \to X$, $X_n \to Y$ fast sicher $\Rightarrow X = Y$ fast sicher

(2) Konvergiert $(X_n)_{n \in \mathbb{N}}$ in Wahrscheinlichkeit gegen X und auch gegen Y so gilt $X = Y$ fast sicher. Das folgt aus

$$\begin{aligned} \mathbb{P}[|X - Y| > \varepsilon] &\leq \mathbb{P}\left[|X_n - X| > \frac{\varepsilon}{2}\right] + \mathbb{P}\left[|X_n - Y| > \frac{\varepsilon}{2}\right] \\ &\to 0 \qquad (n \to \infty) \\ \Rightarrow \mathbb{P}[|X - Y| > \varepsilon] &= 0 \end{aligned}$$

für alle $\varepsilon > 0$. Also $X = Y$ fast sicher wegen

$$[X = Y] = \bigcap_{k=1}^{\infty} \left[|X - Y| \leq \frac{1}{k} \right]$$

(3) Für Konvergenz in Verteilung ist Eindeutigkeit von X im allgemeinen nicht gegeben, siehe Beispiel 2.3.5(ii).

2.3.3 Aufgabe

Es seien F eine stetige Verteilungsfunktion und $(F_n)_{n\in\mathbb{N}}$ eine Folge von Verteilungsfunktionen so, dass für alle $x \in \mathbb{R}$ gilt:

$$\lim_{n\to\infty} F_n(x) = F(x)$$

Dann liegt gleichmäßige Konvergenz vor.

2.3.4 Satz

$X_n \to X$ fast sicher $\Rightarrow X_n \to X$ in Wahrscheinlichkeit $\Rightarrow X_n \to X$ in Verteilung

Beweis:

(i). Annahme: $X_n \to X$ fast sicher. Sei $\varepsilon > 0$. Setze

$$A := \left[\lim_{n\to\infty} X_n = X\right] \in \mathcal{A}$$

Dann $\mathbb{P}(A) = 1$ nach Voraussetzung. Seien

$$B_n := [\forall m \geq n : |X_m - X| \leq \varepsilon]$$

Dann $B_n \in \mathcal{A}$ für alle $n \in \mathbb{N}$, $B_n \subset B_{n+1}$. Es gilt

$$\bigcup_{n=1}^{\infty} B_n \supset A$$

Damit:

$$\begin{aligned} 1 \geq \mathbb{P}\left(\bigcup_{n=1}^{\infty} B_n\right) &\geq \mathbb{P}(A) = 1 \\ \Rightarrow \mathbb{P}\left(\bigcup_{n=1}^{\infty} B_n\right) &= 1 \end{aligned}$$

Wegen $B_m = \bigcup_{n=1}^m B_n$ gilt

$$\lim_{m\to\infty} B_m = \bigcup_{n=1}^{\infty} B_n$$

und damit

$$\lim_{n\to\infty} \mathbb{P}(B_n) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} B_n\right) = 1$$

Sei $A_n := [|X_n - X| \leq \varepsilon]$. Dann $A_n \supset B_n$ für $n \in \mathbb{N}$. Mit $\mathbb{P}(A_n) \geq \mathbb{P}(B_n)$ folgt

$$\lim_{n\to\infty} \mathbb{P}(A_n) = \lim_{n\to\infty} \mathbb{P}[|X_n - X| \leq \varepsilon] = 1$$

(ii). Annahme: $X_n \to X$ in Wahrscheinlichkeit. Seien $\varepsilon > 0$,

$$A_n := [|X_n - X| \leq \varepsilon]$$

Dann gilt nach Voraussetzung $\lim_{n\to\infty} \mathbb{P}(A_n) = 1$. Es folgt:

$$\begin{aligned} F_n(t) &= \mathbb{P}[X_n < t] \\ &= \mathbb{P}([X_n < t] \cap A_n) + \mathbb{P}([X_n < t] \cap A_n^c) \\ &\leq \mathbb{P}([X_n < t] \cap A_n) + \mathbb{P}(A_n^c) \\ &= \mathbb{P}[X_n < t, |X_n - X| \leq \varepsilon] + \mathbb{P}(A_n^c) \\ &\leq \mathbb{P}[X_n < t, X - X_n \leq \varepsilon] + \mathbb{P}(A_n^c) \\ &= \mathbb{P}[X < \varepsilon + t] + \mathbb{P}(A_n^c) \to F(t+\varepsilon)\,(n \to \infty) \end{aligned}$$

d.h.

$$\lim_{n\to\infty} \sup F_n(t) \le F(t+\varepsilon)$$

Mit 1.1.12 folgt wie oben:

$$\begin{aligned} F_n(t) &\ge \mathbb{P}([X_n < t] \cap A_n) \ge \dots \\ &\ge \mathbb{P}([X < t-\varepsilon] \cap A_n) \\ &\overset{1.1.12}{\to} F(t-\varepsilon) = \mathbb{P}[X < t-\varepsilon] \qquad (n\to\infty) \\ \Rightarrow \lim_{n\to\infty} \inf F_n(t) &\ge F(t-\varepsilon) \end{aligned}$$

Damit für alle $\varepsilon > 0$:

$$F(t-\varepsilon) \le \lim_{n\to\infty} \inf F_n(t) \le \lim_{n\to\infty} \sup F_n(t) \le F(t+\varepsilon)$$

Sei t eine Stetigkeitsstelle von F. Dann für $\varepsilon \to 0$:

$$\begin{aligned} F(t) &\le \lim_{n\to\infty} \inf F_n(t) \le \lim_{n\to\infty} \sup F_n(t) \le F(t) \\ \Rightarrow \lim_{n\to\infty} F_n(t) &= F(t) \end{aligned}$$

2.3.5 Beispiele

(i). Sei $\Omega = [0,1]$, $\mathcal{A} = \mathcal{B}[0,1], \mathbb{P} = \lambda|_{\mathcal{A}}$ und für $n \in \mathbb{N}$ definiere

$$X_n^i := 1_{[\frac{i-1}{n}, \frac{i}{n}]} \qquad i = 1, \dots, n$$

Definiere Folge $(X_n)_n$ durch $X_1^1, X_2^1, X_2^2, X_3^1, X_3^2, X_3^3, X_4^1, \dots$. Dann

$$\mathbb{P}[X_n^i > \varepsilon] = \lambda\left(\left(\frac{i-1}{n}, \frac{i}{n}\right)\right) = \frac{1}{n} \to 0 \, (n \to \infty)$$

für $\varepsilon \in (0,1)$, also

$$\lim_{n\to\infty} \mathbb{P}[|X_n - 0| > \varepsilon] = 0$$

d.h. $X_n \to 0$ in Wahrscheinlichkeit. Aber: Für kein $w \in \Omega$ liegt punktweise Konvergenz vor.

(ii). $\Omega = [0,1], \mathcal{A} = \mathcal{B}[0,1], \mathbb{P} = \lambda|_{\mathcal{A}}$,

$$X_{2n+1} := 1_{[0,\frac{1}{2}]} \qquad X_{2n} := 1_{[\frac{1}{2},1]}$$

Dann

$$F_n(x) = \begin{cases} 0 & x \le 0 \\ \frac{1}{2} & 0 < x \le 1 \\ 1 & x > 1 \end{cases}$$

Für $F(x) := F_n(x)$ für $x \in \mathbb{R}$ folgt $X_n \to X_i$ in Verteilung für $i = 1, 2$.

2.3.6 Satz

Für eine Folge $(X_n)_{n\in\mathbb{N}}$ von Zufallsgrößen sind die folgenden Aussagen äquivalent:

(i). Es gibt eine Zufallsgröße X so, dass $X_n \to X$ in Wahrscheinlichkeit.

(ii). $\forall \varepsilon > 0$:

$$\lim_{n\to\infty} \sup_{m\ge n} \mathbb{P}[|X_m - X_n| > \varepsilon] = 0$$

(iii). Es gibt eine Zufallsgröße X so, dass jede Teilfolge von $(X_n)_{n\in\mathbb{N}}$ eine Teilfolge besitzt, die fast sicher gegen X konvergiert.

Beweis:

- $(i) \Rightarrow (ii)$

  Annahme: Es gibt X so, dass $X_n \to X$ in Wahrscheinlichkeit, d.h. $\forall \varepsilon > 0 \forall \delta > 0 \exists N_\delta(\varepsilon) \in \mathbb{N}$:
$$\mathbb{P}[|X_n - X| > \delta] < \varepsilon$$
  für alle $n \geq N_\delta(\varepsilon)$. Sei $\varepsilon > 0, \delta := \frac{\varepsilon}{2}$. Für $m \geq n \geq N_{\frac{\varepsilon}{2}}(\varepsilon)$ gilt dann:
$$\begin{aligned} \mathbb{P}[|X_m - X_n| > \varepsilon] &\leq \mathbb{P}\left[|X_m - X| > \frac{\varepsilon}{2}\right] + \mathbb{P}\left[|X_n - X| > \frac{\varepsilon}{2}\right] \\ &< 2\varepsilon \\ \Rightarrow \sup_{m \geq n} \mathbb{P}[|X_m - X_n| > \varepsilon] &\leq 2\varepsilon \\ \Rightarrow \lim_{n \to \infty} \sup_{m \geq n} \mathbb{P}[|X_m - X_n| > \varepsilon] &= 0 \end{aligned}$$

- $(ii) \Rightarrow (iii)$

  Gilt (ii), so existiert für jede Zahl $k \geq 1$ ein n_k mit
$$\mathbb{P}\left[|X_n - X_m| > \frac{1}{2^k}\right] < \frac{1}{2^k}$$
  für $n \geq m \geq n_k$. Ohne Beschränkung der Allgemeinheit sei $n_1 < n_2 < \ldots$. Wir setzen
$$\begin{aligned} X'_k &:= X_{n_k} \\ A_k &:= \left[|X'_{k+1} - X'_k| > \frac{1}{2^k}\right] \end{aligned}$$
  Dann gilt $\mathbb{P}(A_k) < \frac{1}{2^k}$, also
$$\sum_{k=1}^{\infty} \mathbb{P}(A_k) < \infty$$
  Mit Borel-Cantelli folgt, dass $\mathbb{P}[\limsup A_k] = 0$, d.h. für fast alle w existiert ein $k_0(w)$ mit
$$|X'_{k+1}(w) - X'_k(w)| \leq \frac{1}{2^k}$$
  für $k \geq k_0(w)$. Für $n \geq k_0(w)$ folgt:
$$\begin{aligned} \sup_{m \geq n} |X'_m(w) - X'_n(w)| &\leq \sum_{k=n}^{\infty} |X'_{k+1}(w) - X'_k(w)| \\ &\leq \sum_{k=n}^{\infty} \frac{1}{2^k} = \frac{1}{2^{n-1}} \to 0 \qquad (n \to \infty) \end{aligned}$$
  Also ist $(X'_k(w))_k$ fast sicher eine Cauchy-Folge, damit konvergiert $X_{n_k} = X'_k$ fast sicher gegen eine Zufallsgröße X. Wegen 2.3.4 gilt auch $X_{n_k} \to X$ in Wahrscheinlichkeit,
$$\begin{aligned} \mathbb{P}[|X_k - X| > \varepsilon] &\leq \mathbb{P}\left[|X_k - X_{n_k}| > \frac{\varepsilon}{2}\right] + \mathbb{P}\left[|X_{n_k} - X| > \frac{\varepsilon}{2}\right] \\ &\to 0 \qquad (k \to \infty) \end{aligned}$$
  d.h. $X_k \to X$ in Wahrscheinlichkeit. Noch zu zeigen: Eine beliebige, fast sicher konvergente Teilfolge X_{n_k} konvergiert gegen X. Gilt nämlich $X_{n_k} \to Y$ fast sicher, dann auch $X_{n_k} \to Y$ in Wahrscheinlichkeit, also $X = Y$ fast sicher wegen Eindeutigkeit des Grenzwertes (2.3.2(ii)).

- $(iii) \Rightarrow (i)$

  Konvergiert X_n nicht in Wahrscheinlichkeit gegen X, so existieren $\varepsilon > 0, \delta > 0$ und eine Teilfolge (X_{n_k}) mit
  $$\mathbb{P}[|X_{n_k} - X| > \varepsilon] > \delta$$
  Dann konvergiert keine Teilfolge von X_{n_k} gegen X in Wahrscheinlichkeit und folglich auch nicht sicher. Widerspruch!

2.3.7 Satz

Für eine Folge $(X_n)_{n\in\mathbb{N}}$ von Zufallsgrößen sind die folgenden Aussagen äquivalent:

(i). (X_n) konvergiert fast sicher gegen eine Zufallsgröße X

(ii). $\forall k \in \mathbb{N} : \mathbb{P}[|X_n - X| > \frac{1}{k}, \text{unendlich oft }] = 0$

(iii). Für alle $\varepsilon > 0$ gilt
$$\lim_{n\to\infty} \mathbb{P}\left[\sup_{m\geq n} |X_m - X| > \varepsilon\right] = 0$$

(iv). Für alle $\varepsilon > 0$ gilt
$$\lim_{n\to\infty} \mathbb{P}\left[\sup_{m\geq n} |X_m - X_n| > \varepsilon\right] = 0$$

Beweis:

- $(i) \Leftrightarrow (ii)$

  Ohne Beschränkung der Allgemeinheit: $X = 0$. Setze
  $$\begin{aligned} A_k &:= \left[|X_n| > \frac{1}{k}, \text{unendlich oft}\right] \\ A &:= \bigcup_{k=1}^{\infty} A_k \end{aligned}$$
  Dann gilt $A^c = [\lim X_n = 0]$. Denn: Wenn $w \notin A$, dann gilt für alle $k \in \mathbb{N}$ bis auf endlich viele n: $|X_n(w)| < \frac{1}{k}$, also $\lim X_n = 0$. Umgekehrt, wenn $\lim X_n = 0$, dann $w \notin A_k$ für alle $k \in \mathbb{N}$, also $w \notin A$. Damit:
  $$\begin{aligned} \mathbb{P}[\lim X_n = 0] = 1 &\Leftrightarrow \mathbb{P}(A^c) = 1 \Leftrightarrow \mathbb{P}(A) = 0 \\ &\Leftrightarrow \forall k \in \mathbb{N} : \mathbb{P}(A_k) = 0 \end{aligned}$$

- $(ii) \Leftrightarrow (iii)$

  Ohne Beschränkung der Allgemeinheit: $X = 0$. Setze wie oben
  $$A_k := \left[|X_n| > \frac{1}{k}, \text{unendlich oft}\right]$$
  Dann
  $$A_k = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \left[|X_m| > \frac{1}{k}\right]$$
  Die Äquivalenz von (ii),(iii) folgt aus
  $$\begin{aligned} \mathbb{P}(A_k) &= \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{m=n}^{\infty} \left[|X_m| > \frac{1}{k}\right]\right) \\ &= \lim_{n\to\infty} \mathbb{P}\left[\sup_{m\geq n} |X_m| > \frac{1}{k}\right] \end{aligned}$$

- $(iii) \Rightarrow (iv)$

  Die Implikation folgt aus

$$
\begin{aligned}
\mathbb{P}\left[\sup_{m\geq n} |X_m - X_n| > \varepsilon\right] &\leq \mathbb{P}\left[\sup_{m\geq n} |X_m - X| > \frac{\varepsilon}{2}\right] + \mathbb{P}\left[|X - X_n| > \frac{\varepsilon}{2}\right] \\
&\to 0
\end{aligned}
$$

- $(iv) \Rightarrow (iii)$

  Ist (iv) erfüllt, so gilt

$$
\begin{aligned}
\sup_{m\geq n} \mathbb{P}[|X_m - X_n| > \varepsilon] &\leq \mathbb{P}\left[\sup_{m\geq n} |X_m - X_n| > \varepsilon\right] \\
&\to 0 \qquad (n\to\infty)
\end{aligned}
$$

  Nach 2.3.6 existiert eine Zufallsgröße X mit $X_n \to X$ in Wahrscheinlichkeit. Damit folgt wegen

$$
\begin{aligned}
\mathbb{P}\left[\sup_{m\geq n} |X_m - X| > \varepsilon\right] &\leq \mathbb{P}\left[\sup_{m\geq n} |X_m - X_n| > \frac{\varepsilon}{2}\right] + \mathbb{P}\left[|X_n - X| > \frac{\varepsilon}{2}\right] \\
&\to 0 \qquad (n\to\infty)
\end{aligned}
$$

  die Behauptung.

2.3.8 Satz: Lévy

Für eine Folge $(X_n)_{n\in\mathbb{N}}$ von unabhängigen Zufallsgrößen konvergiert die Folge

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

genau dann fast sicher, wenn sie in Wahrscheinlichkeit konvergiert.

Beweis:

(i). Nehmen wir an, dass (S_n) in Wahrscheinlichkeit konvergiert. Sei

$$S_{h,n} := S_n - S_h = \sum_{i=h+1}^{n} X_i$$

für $1 \leq h \leq n$. Wegen 2.3.6 existiert für jedes $\varepsilon \in \left(0, \frac{1}{2}\right)$ ein h_0, sodass $\mathbb{P}[|S_{h,n}| \geq \varepsilon] < \varepsilon$ für $n > h \geq h_0$. Dann gilt $|m(S_{h,n})| < \varepsilon$. Folgt aus Aufgabe: Aus $\mathbb{P}[|X| \geq c] < \frac{1}{2}$ folgt $|m(X)| < c$.

Damit:

$$
\begin{aligned}
\mathbb{P}\left[\max_{h<n\leq k} |S_{h,n}| \geq 2\varepsilon\right] &\leq \mathbb{P}\left[\max_{h<n\leq k} |S_{h,n} - m(S_{h,n} - S_{h,k})| \geq \varepsilon\right] \\
&\overset{2.2.8(ii)}{\leq} 2\mathbb{P}[|S_{h,k}| \geq \varepsilon] < 2\varepsilon \\
\Rightarrow \mathbb{P}\left[\sup_{n>h} |S_{h,n}| \geq 2\varepsilon\right] &\leq 2\varepsilon
\end{aligned}
$$

für $k > h \geq h_0$. Nach 2.3.7 konvergiert (S_n) fast sicher gegen eine Zufallsgröße S.

(ii). Fast sichere Konvergenz $\Rightarrow$ Konvergenz in Wahrscheinlichkeit

2.3.9 Aufgabe

Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsgrößen. Dann gilt:

(i). Ist c eine Konstante, so gilt $X_n \to c$ in Wahrscheinlichkeit $\Leftrightarrow X_n \to c$ in Verteilung

(ii). Ist $(X_n)_{n\in\mathbb{N}}$ unabhängig, dann sind $\limsup X_n$ und $\liminf X_n$ konstant.

(iii). Ist $(X_n)_{n\in\mathbb{N}}$ unabhängig, $X_n \to X$ in Wahrscheinlichkeit, dann ist X fast sicher konstant.

(iv). Ist $(X_n)_{n\in\mathbb{N}}$ unabhängig, identisch verteilt und nicht konstant, dann $\mathbb{P}[X_n \text{ konvergiert }] = 0$.

2.3.10 Satz

Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von unabhängigen Zufallsgrößen,

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

Wenn $\sum_{n=1}^{\infty} D^2(X_n) < \infty$ und $\mathbb{E}X_n = 0$ für alle $n \in \mathbb{N}$, dann konvergiert die Reihe $\sum_{n=1}^{\infty} X_n$ fast sicher.

Beweis:

- $\mathbb{E}(S_m - S_n) = 0$ für $n < m$. Mit Tschebysheff-Ungleichung folgt:

$$\begin{aligned} \mathbb{P}[|S_m - S_n| \geq \varepsilon] &\leq \frac{1}{\varepsilon^2} \sum_{i=n+1}^{m} D^2(X_i) \\ &\to 0 \qquad (n, m \to \infty) \end{aligned}$$

  Mit 2.3.6: $(S_m)_m$ konvergent in Wahrscheinlichkeit. Nach Satz von Lévy (2.3.8) folgt fast sichere Konvergenz.

2.3.11 Satz: Jegorov

Es seien Y, $(Y_n)_{n\in\mathbb{N}}$ beliebige Zufallsgrößen, $A \in \mathcal{A}$ mit $0 < \mathbb{P}(A) \leq 1$ und für $w \in A$ gelte

$$\lim_{n\to\infty} Y_n(w) = Y(w)$$

Dann existiert für jedes $\varepsilon > 0$ eine Menge $F \in \mathcal{A}$, sodass $F \subset A$, $\mathbb{P}(F) > \mathbb{P}(A) - \varepsilon$ und die Folge $(Y_n)_{n\in\mathbb{N}}$ konvergiert auf F gleichmäßig gegen Y.

2.3.12 Folgerung

Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von unabhängigen Zufallsgrößen mit $|X_n| \leq c$ $(n \in \mathbb{N})$ für ein $c \in \mathbb{R}$,

$$S_n := \sum_{k=1}^{n} X_k$$

Wenn $(S_n)_{n\in\mathbb{N}}$ fast sicher gegen eine Zufallsgröße S konvergiert, dann existiert ein $d > 0$, sodass die Wahrscheinlichkeit des Ereignisses

$$E := \bigcap_{n=1}^{\infty} \{w : |S_n(w)| \leq d\}$$

positiv ist.

Beweis:

- Es existiert $K \in \mathbb{R}$:
$$\mathbb{P}[|S| \leq K] > 0$$
(sonst wäre $|S| = \infty$ fast sicher). Setze $A := [|S| \leq k]$. Wegen 2.3.11 existiert $F \subset A$ mit $\mathbb{P}(F) > 0$, sodass $S_n \to S$ gleichmäßig auf F, d.h. es existiert ein $n_0 \in \mathbb{N}$:
$$|S_n(w) - S(w)| \leq 1$$
für $n \geq n_0, w \in F$. Damit
$$|S_n(w)| \leq |S(w)| + 1 \leq K + 1 \qquad (n \geq n_0, w \in F)$$
Andererseits ist
$$|S_n| = \left|\sum_{i=1}^{n} X_i\right| \leq n \cdot c < n_0 \cdot c$$
für $n < n_0$. Also gilt für alle $n \in \mathbb{N}, w \in F$:
$$|S_n(w)| \leq d := \max\{K + 1, n_0 \cdot c\}$$

2.3.13 Satz

Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von unabhängigen Zufallsgrößen mit $|X_n| \leq c$ für ein $c \in \mathbb{R}$. Die Reihe $\sum_{n=1}^{\infty} X_n$ konvergiert genau dann fast sicher, wenn
$$(1)\ \sum_{n=1}^{\infty} \mathbb{E}(X_n) < \infty \qquad \text{und} \qquad (2)\ \sum_{n=1}^{\infty} D^2(X_n) < \infty$$

Bemerkung: Aus $|X_n| \leq c$ folgt $X_n \in \mathcal{L}^p(\mathbb{P})$.

Beweis:

(i). Nehmen wir an, dass (1) und (2) gelten und sei $Y_n := X_n - \mathbb{E}X_n$. Dann ist $\mathbb{E}Y_n = 0$, $D^2(Y_n) = D^2(X_n)$. Aus 2.3.10 folgt, dass $\sum_{n=1}^{\infty} Y_n$ fast sicher konvergiert, also konvergiert auch $\sum_{n=1}^{\infty} X_n$ fast sicher.

(ii). Sei $\sum_{n=1}^{\infty} X_n$ fast sicher.

- Spezialfall: $\mathbb{E}X_n = 0$ für alle $n \in \mathbb{N}$, dann ist (1) erfüllt. Sei $S_0 := 0$, E und d wie in 2.3.12. Setze
$$E_n := \bigcap_{i=0}^{n} [|S_i| \leq d] \qquad (n \in \mathbb{N})$$
Dann $(E_n)_n \downarrow E$, $E = \bigcap_{n=0}^{\infty} E_n$. Sei $F_n := E_{n-1} \backslash E_n$ $(n \in \mathbb{N})$,
$$a_n := \int_{E_n} S_n^2 d\,\mathbb{P}$$
Dann
$$\begin{aligned} a_n - a_{n-1} &= \left(\int_{E_{n-1}} S_n^2\, d\mathbb{P} - \int_{F_n} S_n^2\, d\mathbb{P}\right) - \int_{E_{n-1}} S_{n-1}^2\, d\mathbb{P} \\ &= \int_{E_{n-1}} X_n^2\, d\mathbb{P} + 2\int_{E_{n-1}} X_n \cdot S_{n-1}\, d\mathbb{P} - \int_{F_n} S_n^2\, d\mathbb{P} \end{aligned}$$

Vorlesung: "Stochastik" von Zoltan Sasvári, Technische Universität Dresden, Sommersemester 2010

Es gilt:

$$\begin{aligned}
\int_{E_{n-1}} X_n^2 \, d\mathbb{P} &= \mathbb{E}\left(1_{E_{n-1}} \cdot X_n^2\right) \overset{\text{unabh}}{=} \mathbb{E}\left(1_{E_{n-1}}\right) \cdot \mathbb{E}(X_n^2) \\
&\overset{\mathbb{E}X_n=0}{=} \mathbb{P}(E_{n-1}) \cdot D^2(X_n) \geq \mathbb{P}(E) \cdot D^2(X_n) \\
\int_{E_{n-1}} X_n \cdot S_{n-1} \, d\mathbb{P} &= \mathbb{E}\left(1_{E_{n-1}} \cdot X_n \cdot S_{n-1}\right) \\
&\overset{\text{unabh}}{=} \underbrace{\mathbb{E}X_n}_{0} \cdot \mathbb{E}(1_{E_{n-1}} \cdot S_{n-1}) = 0 \\
\int_{F_n} S_n^2 \, d\mathbb{P} &\leq (d+c)^2 \cdot \mathbb{P}(F_n)
\end{aligned}$$

da $|S_n| = |S_{n-1} + X_n| \leq d + c$ auf F_n. Wir erhalten:

$$\begin{aligned}
a_n - a_{n-1} &\geq \mathbb{P}(E) \cdot D^2(X_n) - (d+c)^2 \cdot \mathbb{P}(F_n) \\
\Rightarrow \underbrace{\sum_{n=1}^{k} a_n - a_{n-1}}_{=a_k} &\geq \mathbb{P}(E) \cdot \sum_{n=1}^{k} D^2(X_n) - (d+c)^2 \cdot \underbrace{\sum_{n=1}^{k} \mathbb{P}(F_n)}_{\leq 1}
\end{aligned}$$

Andererseits ist

$$a_k \leq \mathbb{P}(E_k) \cdot d^2 \leq d^2$$

Damit

$$\begin{aligned}
\mathbb{P}(E) \cdot \sum_{n=1}^{k} D^2(X_n) &\leq d^2 + (d+c)^2 \\
\Rightarrow \sum_{n=1}^{\infty} D^2(X_n) &< \infty
\end{aligned}$$

- Allgemeiner Fall: Wir betrachten den Wahrscheinlichkeitsraum $(\Omega \times \Omega, \mathcal{A} \otimes \mathcal{A}, \mathbb{P} \times \mathbb{P})$ und die Zufallsgrößen

  $$Z_n(w_1, w_2) := X_n(w_1) - X_n(w_2) \qquad (w_1, w_2 \in \Omega)$$

  $E \subset \Omega$ sei das Ereignis, dass $\sum_{n=1}^{\infty} X_n$ konvergiert. Dann ist $\mathbb{P}(E) = 1$, also

  $$(\mathbb{P} \times \mathbb{P})(E \times E) = \mathbb{P}(E) \cdot \mathbb{P}(E) = 1$$

  Die Reihe $\sum_{n=1}^{\infty} Z_n$ konvergiert auf $E \times E$, d.h. diese Reihe konvergiert fast sicher bzgl. $\mathbb{P} \times \mathbb{P}$. Weiterhin gilt $\mathbb{E}Z_n = 0$ mit Satz von Fubini und $|Z_n| \leq 2c, D^2(Z_n) = 2D^2(X_n)$. Mit Spezialfall folgt (2).
  Noch zu zeigen: (1). Setze $Y_n := X_n - \mathbb{E}X_n$. Dann $\mathbb{E}Y_n = 0$ und

  $$\sum_{n=1}^{\infty} D^2(Y_n) = \sum_{n=1}^{\infty} D^2(X_n) < \infty$$

  wegen (2). Aus 2.3.10 folgt: Die Reihe $\sum_{n=1}^{\infty} Y_n$ konvergiert fast sicher. Damit (1), da $\mathbb{E}X_n = X_n - Y_n$.

2.3.14 Definition

Zwei Folgen $(Y_n)_{n\in\mathbb{N}}$ und $(Z_n)_{n\in\mathbb{N}}$ von Zufallsgrößen heißen äquivalent

$$:\Leftrightarrow \sum_{n=1}^{\infty} \mathbb{P}[Y_n \neq Z_n] < \infty$$

Aus Borel-Cantelli folgt dann:

$$\mathbb{P}[Y_n \neq Z_n, \text{unendlich oft}] = 0$$

Damit: $\sum_{n=1}^{\infty} Y_n$ konvergiert fast sicher $:\Leftrightarrow \sum_{n=1}^{\infty} Z_n$ konvergiert fast sicher.

2.3.15 Satz: Dreireihensatz von Kolmogorov

Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von unabhängigen Zufallsgrößen. Die Reihe $\sum_{n=1}^{\infty} X_n$ konvergiert genau dann fast sicher, wenn

(i). $\sum_{n=1}^{\infty} \mathbb{P}[|X_n| > 1] < \infty$

(ii). $\sum_{n=1}^{\infty} \mathbb{E}(X_n') < \infty$

(iii). $\sum_{n=1}^{\infty} D^2(X_n') < \infty$

wobei

$$X_n'(w) := \begin{cases} X_n(w) & |X_n(w)| \leq 1 \\ 0 & \text{sonst} \end{cases}$$

Beweis:

- (ii),(iii) $\Rightarrow$ $\sum_{n=1}^{\infty} X_n'$ konvergiert fast sicher nach 2.3.13. Aus (i) folgt, dass die Folgen $(X_n)_{n\in\mathbb{N}}$, $(X_n')_{n\in\mathbb{N}}$ äquivalent sind, damit $\sum_{n=1}^{\infty} X_n$ fast sicher konvergent.
- Sei jetzt $\sum_{n=1}^{\infty} X_n$ fast sicher konvergent. Dann gilt $X_n \to 0$ fast sicher, also

  $$\mathbb{P}[|X_n| > 1, \text{unendlich oft}] = 0$$

  und mit Borel-Cantelli folgt (i). Außerdem folgt, dass $(X_n)_n$, $(X_n')_n$ äquivalent sind, somit

  $$\sum_{n=1}^{\infty} X_n' < \infty$$

  und nach 2.3.13 gilt (ii),(iii).

2.4 Starkes und schwaches Gesetz der großen Zahlen

2.4.1 Definition

- Eine Folge $(X_n)_{n\in\mathbb{N}}$ integrierbarer Zufallsgrößen heißt dem starken Gesetz der großen Zahlen genügend, wenn gilt:

  $$\lim_{n\to\infty} \frac{1}{n} \cdot \sum_{i=1}^{n} (X_i - \mathbb{E}X_i) = 0 \, \text{fast sicher}$$

- Eine Folge $(X_n)_{n\in\mathbb{N}}$ heißt dem schwachen Gesetz der großen Zahlen genügend, wenn für alle $\varepsilon > 0$:

  $$\lim_{n\to\infty} \mathbb{P}\left[\left|\frac{1}{n} \cdot \sum_{i=1}^{n} (X_i - \mathbb{E}X_i)\right| > \varepsilon\right] = 0$$

  d.h. wenn Konvergenz in Wahrscheinlichkeit vorliegt.

2.4.2 Lemma: Kronecker

Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen und $(t_n)_{n\in\mathbb{N}}$ eine monoton wachsende, gegen ∞ strebende Folge positiver Zahlen derart, dass die Reihe

$$\sum_{k=1}^{\infty} \frac{a_k}{t_k} =: b$$

konvergiert. Dann gilt:

$$\lim_{n\to\infty} \frac{1}{t_n} \cdot \sum_{k=1}^{n} a_k = 0$$

Beispiel:

$$a_k = \frac{1}{k}, t_k = k \Rightarrow \lim_{n\to\infty} \frac{1}{n} \cdot \sum_{k=1}^{n} \frac{1}{k} = 0$$

Beweis:

- Definiere $b_0 := 0$,

$$b_n := \sum_{k=1}^{n} \frac{a_k}{t_k}$$

Dann ist $a_n = (b_n - b_{n-1}) \cdot t_n$, also

$$\begin{aligned}
\sum_{k=1}^{n} a_k &= \sum_{k=1}^{n} t_k \cdot (b_k - b_{k-1}) \\
&= t_n \cdot b_n + \sum_{k=1}^{n-1} t_k \cdot b_k - \sum_{k=2}^{n} t_k \cdot b_{k-1} \\
&= t_n \cdot b_n + \sum_{k=1}^{n-1} \cdot (t_k - t_{k+1}) \cdot b_k \\
\Rightarrow \frac{1}{t_n} \sum_{k=1}^{n} a_k &= b_n + \sum_{k=1}^{n-1} \frac{t_k - t_{k+1}}{t_n} \cdot b_k \\
&\to b - b \qquad (n \to \infty)
\end{aligned}$$

Noch zu zeigen ist die Konvergenz des zweiten Terms. Für ein beliebiges $\varepsilon > 0$ wählen wir $n_0 \in \mathbb{N}$ so, dass $|b_n - b| < \varepsilon$ für $n \geq n_0$.

$$\begin{aligned}
\left| \sum_{k=1}^{n} \left(\frac{t_{k+1} - t_k}{t_n} \cdot b_k \right) - b \right| &= \left| -\frac{t_1}{t_n} \cdot b + \sum_{k=1}^{n-1} \frac{t_{k+1} - t_k}{t_n} \cdot (b_k - b) \right| \\
&= \left| -\frac{t_1}{t_n} \cdot b + \sum_{k=1}^{n_0-1} \frac{t_{k+1} - t_k}{t_n} \cdot (b_k - b) + \sum_{k=n_0}^{n-1} \frac{t_{k+1} - t_k}{t_n} \cdot (b_k - b) \right| \\
&\leq \frac{t_1}{t_n} \cdot |b| + \frac{1}{t_n} \cdot \sum_{k=1}^{n_0-1} (t_{k+1} - t_k) \cdot |b_k - b| + \frac{t_n - t_{n_0}}{t_n} \cdot \varepsilon \\
&\to 0 + 0 + \varepsilon \qquad (n \to \infty)
\end{aligned}$$

2.4.3 Satz: Kolmogorov

Es sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von unabhängigen Zufallsgrößen mit $X_n \in L^2(\mathbb{P})$. Gilt

$$\sum_{n=1}^{\infty} \frac{D^2(X_n)}{n^2} < \infty$$

so genügt die Folge dem starken Gesetz der großen Zahlen.

Beweis:

- Sei

$$Y_n := \frac{1}{n} \cdot \sum_{i=1}^{n} (X_i - \mathbb{E}X_i)$$

Nach Ungleichung von Hàjek-Rényi folgt mit $r_i := \frac{1}{i}$:

$$\begin{aligned}
\mathbb{P}\left[\sup_{m\le i\le n}|Y_i|\ge\varepsilon\right] &\le \frac{1}{\varepsilon^2}\cdot\left(\frac{1}{m^2}\sum_{j=1}^{m}D^2(X_j)+\sum_{j=m+1}^{n}\frac{D^2(X_j)}{j^2}\right)\\
\overset{n\to\infty}{\Rightarrow}\mathbb{P}\left[\sup_{i\ge m}|Y_i|\ge\varepsilon\right] &\le \frac{1}{\varepsilon^2}\cdot\left(\frac{1}{m^2}\cdot\sum_{j=1}^{m}D^2(X_j)+\sum_{j=m+1}^{\infty}\frac{D^2(X_j)}{j^2}\right)\\
\overset{2.4.2}{\Rightarrow}\lim_{m\to\infty}\mathbb{P}\left[\sup_{i\ge m}|Y_i|\ge\varepsilon\right] &= 0
\end{aligned}$$

wegen

$$\sum_{j=m+1}^{\infty}\frac{D^2(X_j)}{j^2}\to 0 \qquad (m\to\infty)$$

da die Reihe nach Voraussetzung konvergiert. Nach 2.3.7 folgt nun $Y_n\to 0$ fast sicher.

2.4.4 Beispiel

Wir betrachten eine Folge von Versuchen. Sei $A\in\mathcal{A}$ mit $\mathbb{P}(A)=:p$. Betrachte Folge von Zufallsgrößen

$$X_k=\begin{cases}1 & \text{A im k-ten Versuch eingetreten}\\ 0 & \text{sonst}\end{cases}$$

Dann $(X_k)_k$ unabhängig, $\mathbb{P}[X_k=1]=p,\mathbb{P}[X_k=0]=1-p$, $\mathbb{E}X_k=p, D^2(X_k)=p\cdot(1-p)$. Sei

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

S_n ist die Anzahl des Eintretens von A in n Versuchen (absolute Häufigkeit von A). S_n ist binomialverteilt, $\mathbb{E}S_n=n\cdot p$, $D^2(S_n)=n\cdot p\cdot(1-p)$. $\frac{S_n}{n}$ ist relative Häufigkeit. Aus 2.4.3 folgt:

2.4.5 Satz: Borel

$$\lim_{n\to\infty}\frac{S_n}{n}=p$$

Bemerkung:

- Man kann zeigen (Rényi):

  $$\mathbb{P}\left[\left|\frac{S_n}{n}-p\right|\ge\varepsilon\right]\le 2\cdot\exp\left(-n\cdot\varepsilon^2\cdot\frac{1}{2pq\cdot\left(1+\frac{\varepsilon}{2pq}\right)^2}\right)$$

  wobei $q=1-p$, $0<\varepsilon\le p\cdot q$.

- Alternative Abschätzung mit Tschebysheff-Ungleichung:

  $$\mathbb{P}\left[\left|\frac{S_n}{n}-p\right|\ge\varepsilon\right]\le\frac{p\cdot q}{n\cdot\varepsilon^2}$$

- Beispiel: Für $p=q=\frac{1}{2}$, $\varepsilon=\frac{1}{20}$ erhält man:

  $$\mathbb{P}\left[\left|\frac{S_n}{n}-p\right|\ge\frac{1}{20}\right]\le\frac{1}{100}$$

  für $n\ge 10^4$ (Tschebysheff), $n\ge 1283$ (Rényi).

2.4.6 Satz: Etemadi, 1981

Jede Folge paarweise unabhängiger, integrierbarer Zufallsgrößen, die identisch verteilt sind, genügt dem schwachen Gesetz der großen Zahlen.

Ohne Beweis (siehe Heinz Bauer, Wahrscheinlichkeitstheorie)

2.4.7 Satz: Khinchin

Gilt für eine Folge paarweise unkorrelierter Zufallsgrößen $(X_n)_{n\in\mathbb{N}}$ mit $X_n \in L^2(\mathbb{P})$ die Beziehung

$$\lim_{n\to\infty} \frac{1}{n^2} \cdot \sum_{i=1}^{n} D^2(X_i) = 0$$

so genügt die Folge dem schwachen Gesetz der großen Zahlen. Beweis:

- Setzen wir
$$S_n := \sum_{i=1}^{n} X_i - \mathbb{E}X_i$$
so ist $D^2(S_n) = \sum_{i=1}^{n} D^2(X_i)$. Aus Tschebysheff-Ungleichung und aus der Voraussetzung folgt:
$$\begin{aligned}\mathbb{P}\left[\frac{S_n}{n} \geq \varepsilon\right] &\leq \frac{D^2\left(\frac{S_n}{n}\right)}{\varepsilon^2} \\ &= \frac{1}{n^2 \cdot \varepsilon^2} \cdot \sum_{i=1}^{n} D^2(X_i) \to 0 \qquad (n \to \infty)\end{aligned}$$

2.4.8 Satz

Für jede Zufallsgröße X gilt:

(i).
$$\sum_{n=1}^{\infty} \mathbb{P}[|X| \geq n] \leq \mathbb{E}(|X|) \leq 1 + \sum_{n=1}^{\infty} \mathbb{P}[|X| \geq n]$$

(ii). Git $X(\Omega) \subseteq \mathbb{N}_0$, so erhält man
$$\mathbb{E}X = \sum_{n=1}^{\infty} \mathbb{P}[X \geq n]$$

Beweis:

(i). Für den Beweis können wir annehmen, dass $X \geq 0$. Die Ereignisse
$$A_n := [n \leq X < n+1] \qquad (n \in \mathbb{N}_0)$$
sind paarweise disjunkt und $\Omega = \dot{\bigcup}_{j\in\mathbb{N}_0} A_j$. Damit:
$$\mathbb{E}X = \int_\Omega X \, d\mathbb{P} = \sum_{n=0}^{\infty} \int_{A_n} X \, d\mathbb{P}$$
Aus der Definition von A_n folgt:
$$\begin{aligned} n \cdot \mathbb{P}(A_n) &\leq \int_{A_n} X \, d\mathbb{P} \leq (n+1) \cdot \mathbb{P}(A_n) \\ \Rightarrow \sum_{n=1}^{\infty} n \cdot \mathbb{P}(A_n) &\leq \mathbb{E}X \leq 1 + \sum_{n=1}^{\infty} n \cdot \mathbb{P}(A_n)\end{aligned}$$

Setzen wir $B_n := [X \geq n]$ für $n \in \mathbb{N}_0$, so gilt $B_{n+1} \subset B_n$, $A_n = B_n \backslash B_{n+1}$. Es folgt daher für jedes $N \in \mathbb{N}$:

$$\begin{aligned}
\sum_{n=1}^{N} n \cdot \mathbb{P}(A_n) &= \sum_{n=1}^{N} n \cdot (\mathbb{P}(B_n) - \mathbb{P}(B_{n+1})) \\
&= \sum_{n=1}^{N} n \cdot \mathbb{P}(B_n) - \sum_{n=1}^{N} (n-1) \cdot \mathbb{P}(B_n) - N \cdot \mathbb{P}(B_{N+1}) \\
&= \sum_{n=1}^{N} \mathbb{P}(B_n) - N \cdot \mathbb{P}(B_{N+1}) \\
\Rightarrow \sum_{n=1}^{N} n \cdot \mathbb{P}(A_n) + N \cdot \mathbb{P}(B_{N+1}) &= \sum_{n=1}^{N} \mathbb{P}(B_n)
\end{aligned}$$

Ist $\mathbb{E}X < \infty$:

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{n \in \mathbb{N}_0} B_n\right) &= 0 \\
\Rightarrow 0 &\leq N \cdot \mathbb{P}(B_{N+1}) \leq (N+1) \cdot \mathbb{P}(B_{N+1}) \\
&\leq \int_{B_{N+1}} X \, d\mathbb{P} \to 0 \qquad (N \to \infty) \\
\Rightarrow \sum_{n=1}^{\infty} n \cdot \mathbb{P}(A_n) &= \sum_{n=1}^{\infty} \mathbb{P}(B_n)
\end{aligned}$$

Damit folgt (i).

Ist $\mathbb{E}X = \infty$, dann ist

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$$

und es folgt, dass dann auch

$$\sum_{n=1}^{\infty} \mathbb{P}(B_n) = \infty$$

und damit gilt (i).

(ii). Nimmt X nur Werte in $\mathbb{N}_0$ an, dann ist $A_n = [X = n]$ und somit

$$\mathbb{E}X = \sum_{n=1}^{\infty} n \cdot \mathbb{P}(A_n)$$

Die Gleichheit (ii) folgt somit aus (i).

Bemerkung:

- (ii) folgt auch aus 1.4.18,

$$\mathbb{E}X = \int_0^{\infty} 1 - F(y) \, dy - \int_{-\infty}^{0} F(y) \, dy$$

2.4.9 Satz

Für eine Folge $(X_n)_{n \in \mathbb{N}}$ unabhängiger, identisch verteilter Zufallsgrößen konvergiere

$$\frac{1}{n} \cdot \sum_{i=1}^{n} X_i \to Y \text{ fast sicher}$$

mit Y Zufallsgröße. Dann ist $X_n \in \mathcal{L}^1(\mathbb{P})$ und Y fast sicher konstant, nämlich $Y = \mathbb{E}(X_n)$ fast sicher.

Beweis:

- Setzen wir
$$S_n := \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$$
so konvergiert die Folge $(S_n)_{n\in\mathbb{N}}$ nach Voraussetzung fast sicher gegen Y. Somit konvergiert die Folge
$$\frac{1}{n} \cdot X_n = S_n - \frac{n-1}{n} \cdot S_{n-1}$$
fast sicher gegen 0. Daher können die Ereignisse
$$C_n := [|X_n| \geq n]$$
nur mit Wahrscheinlichkeit 0 für unendlich viele n eintreten, d.h. $\mathbb{P}[\limsup C_n] = 0$. Nach Borel-Cantelli folgt
$$\sum_{n=1}^{\infty} \mathbb{P}(C_n) < \infty$$
da die Ereignisse C_n, C_m unabhängig sind (wegen X_n, X_m unabhängig) für $n \neq m$.
- Da die Zufallsgrößen X_k identisch verteilt sind, gilt $\mathbb{P}(C_n) = \mathbb{P}[|X_k| \geq n]$. Daher konvergiert die Reihe
$$\sum_{n=1}^{\infty} \mathbb{P}[|X_k| \geq n] = \sum_{n=1}^{\infty} \mathbb{P}(C_n) < \infty$$
für alle $k \in \mathbb{N}$. Mit 2.4.8(i) folgt $X_k \in \mathcal{L}^1(\mathbb{P})$.
- Aus dem Satz von Etemadi (2.4.6) folgt nun
$$\lim_{n\to\infty} S_n = d \,\text{fast sicher}$$
mit d als gemeinsamer Erwartungswert aller X_n.

2.4.10 Beispiel

Es sei $(X_n)_{n\in\mathbb{N}}$ eine Folge unabhängiger Zufallsgrößen mit
$$\mathbb{P}[X_n = n] = \mathbb{P}[X_n = -n] = \frac{1}{2n \cdot \log(n+1)}$$
für $n \in \mathbb{N}$,
$$\mathbb{P}[X_n = 0] = 1 - \frac{1}{n \cdot \log(n+1)}$$
Dann genügt (X_n) dem schwachen Gesetz der großen Zahlen, aber nicht dem starken Gesetz der großen Zahlen.

Beweisskizze:

(i). schwaches Gesetz der großen Zahlen: Offensichtlich gilt
$$\mathbb{E}X_n = 0 \qquad D^2(X_n) = \frac{n}{\log(n+1)}$$
Idee: Anwendung des Satz von Khinchin.
$$\frac{1}{n^2} \cdot \sum_{i=1}^{n} D^2(X_i) \leq \frac{1}{\log(n+1)} \to 0 \qquad (n \to \infty)$$

(ii). starkes Gesetz der großen Zahlen: Wähle $A_n := [|X_n| \geq n]$, dann

$$\begin{aligned}\sum_{n=1}^{\infty} \mathbb{P}(A_n) &= \sum_{n=1}^{\infty} \frac{1}{n \cdot \log(n+1)} = \infty \\ \overset{BC}{\Rightarrow} \mathbb{P}\left[\lim_{n\to\infty} \sup A_n\right] &= 1\end{aligned}$$

also $|X_n| \geq n$ für unendlich viele $n \in \mathbb{N}$. Für $w \in \limsup A_n$, $n \geq 2$ für das $|X_n| \geq n$:

$$\begin{aligned}\left|\frac{1}{n} \cdot \sum_{i=1}^{n} X_i - \frac{1}{n-1} \cdot \sum_{i=1}^{n-1} X_i\right| &= \left|\frac{1}{n} \cdot X_n - \left(\frac{1}{n-1} - \frac{1}{n}\right) \cdot \sum_{i=1}^{n-1} X_i\right| \\ &\geq 1 - \frac{1}{n \cdot (n-1)} \cdot \left|\sum_{i=1}^{n-1} X_i\right|\end{aligned}$$

Somit kann $\frac{1}{n}\sum_{i=1}^{n} X_i$ nicht konvergieren. (X_n) genügt nicht dem starken Gesetz der großen Zahlen. (Vergleich Heinz Bauer, S.73,76f)

2.5 Satz von Gliwenko-Cantelli

2.5.1 Definition

Es seien $X_1, \ldots, X_n$ identisch verteilte, unabhängige Zufallsgrößen mit der Verteilungsfunktion F und $w \in \Omega$. Wir ordnen die Zahlen $X_1(w), \ldots, X_n(w)$ um:

$$X_1^*(w) \leq \ldots \leq X_n^*(w)$$

Die empirische Verteilungsfunktion F_n^w wird folgendermaßen definiert:

$$F_n^w(x) := \begin{cases} 0 & x \leq X_1^*(w) \\ \dfrac{k}{n} & X_k^*(w) < x \leq X_{k+1}^*(w) \\ 1 & X_n^*(w) < x \end{cases}$$

(F_n^w ist tatsächlich eine Verteilungsfunktion.)

2.5.2 Satz: Gliwenko-Cantelli

Seien $(X_n)_{n\in\mathbb{N}}$ unabhängige, identisch verteilte Zufallsgrößen mit der Verteilungsfunktion F und $w \in \Omega$. Dann gilt:

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |F_n^w(x) - F(x)| = 0 \text{ fast sicher}$$

2.5.3 Aufgabe

(i). Man berechne die Wahrscheinlichkeit

$$\mathbb{P}\left[F_n^w = \frac{k}{n}\right]$$

für $0 \leq k \leq n$ mit Hilfe der Verteilungsfunktion F und zeige, dass

$$\mathbb{E}(F_n^w(x)) := \int_\Omega F_n^w(x)\, d\mathbb{P}(w) = F(x)$$

(ii). Man zeige:

$$\begin{aligned}\forall x \in \mathbb{R} \quad &: \quad \lim_{n\to\infty} F_n^w(x) = F(x) \text{ fast sicher} \\ \forall x \in \mathbb{R} \quad &: \quad \lim_{n\to\infty} F_n^w(x+0) = F(x+0) \text{ fast sicher}\end{aligned}$$

Hinweis: Betrachte

$$Y_n^x := 1_{[X_n < x]} \qquad\qquad Z_n^x := 1_{[X_n \le x]}$$

Dann folgt unter Verwendung von 2.3:

$$\begin{aligned} F_n^w(x) &= \frac{1}{n} \cdot \sum_{i=1}^n Y_i^x \to F(x) \text{ fast sicher} \\ F_n^w(x+0) &= \frac{1}{n} \cdot \sum_{i=1}^n Z_i^x \to F(x+0) \text{ fast sicher} \end{aligned}$$

(iii). Man beweise Satz 2.5.2.

Es bleibt nur noch zu zeigen, dass aus (i),(ii) die gleichmäßige Konvergenz folgt. Dazu sei für jedes $k \in \mathbb{N}$ und $1 \le j < k$ die kleinste Zahl x_{j_k} mit

$$\frac{j}{k} \le F(x_{j_k} + 0)$$

und es sei $x_{0_k} := -\infty, x_{k_k} = \infty$. Dann ist

$$\begin{aligned} F(x_{j_k}) &\le \frac{j}{k} \\ \Rightarrow F(x_{j+1_k}) - F(x_{j_k} + 0) &\le \frac{1}{k} \end{aligned} \tag{1}$$

Daher gilt für jedes $x \in (x_{j_k}, x_{j+1_k})$:

$$\begin{aligned} F_n^w(x_{j_k} + 0) - F(x_{j_k} + 0) - \frac{1}{k} &\overset{(1)}{\le} F_n^w(x_{j_k} + 0) - F(x_{j+1_k}) \\ &\le F_n^w(x) - F(x) \le F_n^w(x_{j+1_k}) - F(x_{j_k} + 0) \\ &\overset{(1)}{\le} F_n^w(x_{j+1_k}) - F(x_{j+1_k}) + \frac{1}{k} \end{aligned}$$

Daraus folgt:

$$\begin{aligned} \sup_{x \in \mathbb{R}} |F_n^w(x) - F(x)| &\le \max_{1 \le j < k} |F_n^w(x_{j_k}) - F(x_{j_k})| + \max_{0 \le j < k} |F_n^w(x_{j_k} + 0) - F(x_{j_k} + 0)| + \frac{1}{k} \\ &\overset{(ii)}{\to} \frac{1}{k} \text{ fast sicher} \qquad (n \to \infty) \end{aligned}$$

3 Charakteristische Funktionen

3.1 Definition + einige Eigenschaften

3.1.1 Definition

Unter der charakteristischen Funktion einer Zufallsgröße X versteht man die Funktion

$$\begin{aligned} f_X(t) &:= f(t) := \mathbb{E}(\exp(\imath tX)) \\ &= \mathbb{E}(\cos(tX)) + \imath \cdot \mathbb{E}(\sin(tX)) \end{aligned}$$

für $t \in \mathbb{R}$. Ist μ die Verteilung von X, so gilt

$$\begin{aligned} f_X(t) &= \int_{\mathbb{R}} e^{\imath ty} \, d\mu(y) \\ &= \int_{\mathbb{R}} \cos(ty) \, d\mu(y) + \imath \cdot \int_{\mathbb{R}} \sin(ty) \, d\mu(y) \end{aligned}$$

Man nennt f auch die charakteristische Funktion der Verteilung μ. (Vergleich Fourier-Transformierte)

3.1.2 Satz

Jede Verteilung ist durch ihre charakteristische Funktion eindeutig bestimmt.

ohne Beweis (Vgl. 3.3.3)

3.1.3 Satz

Sind X, Y unabhängige Zufallsgrößen, so ist

$$f_{X+Y} = f_X \cdot f_Y$$

Daraus folgt: Das Produkt von charakteristischen Funktionen ist wieder eine charakteristische Funktion.

Beweis:

- Aus X und Y unabhängig folgt $\exp(\imath tX), \exp(\imath tY)$ unabhängig. Damit:

$$\begin{aligned} f_{X+Y}(t) &= \mathbb{E}(\exp(\imath t \cdot (X+Y))) = \mathbb{E}(\exp(\imath tX) \cdot \exp(\imath tY)) \\ &\overset{\text{unabh}}{=} \mathbb{E}(\exp(\imath tX)) \cdot \mathbb{E}(\exp(\imath tY)) = f_X(t) \cdot f_Y(t) \end{aligned}$$

3.1.4 Satz

Ist X eine Zufallsgröße, $Y := a \cdot X + b$ mit $a, b \in \mathbb{R}$, so gilt

$$f_Y(t) = e^{\imath bt} \cdot f_X(a \cdot t)$$

Beweis:

$$\begin{aligned} f_Y(t) &= \mathbb{E}(\exp(\imath tY)) = \mathbb{E}(\exp(\imath t \cdot (aX+b))) = \mathbb{E}(\exp(\imath bt) \cdot \exp(\imath (at)X)) \\ &= e^{\imath bt} \cdot \mathbb{E}(\exp(\imath (at)X)) = e^{\imath bt} \cdot f_X(a \cdot t) \end{aligned}$$

3.1.5 Satz

Für jede charakteristische Funktion f gilt:

(i). $f(0) = 1$

(ii). $f(-t) = \overline{f(t)}$

(iii). $|f(t)| \leq 1$

(iv). f gleichmäßig stetig

Beweis: Aufgabe

Hinweis zu (iv):

$$|f(t+h) - f(t)| = |\mathbb{E}(\exp(\imath tX) \cdot (\exp(\imath hX) - 1))|$$

3.1.6 Lemma

$$\left| e^{\imath z} - \sum_{k=0}^{n} \frac{(\imath z)^k}{k!} \right| \leq \begin{cases} \dfrac{|z|^{n+1}}{(n+1)!} & \operatorname{Im} z \geq 0 \\ \dfrac{|z|^{n+1}}{(n+1)!} \cdot e^{|z|} & \operatorname{Im} z < 0 \end{cases}$$

Beweis: Induktion

(i). $n = 0$: Sei $z = a + \imath \cdot b$ mit $a, b \in \mathbb{R}$ beliebig. Aus

$$e^{\imath z} - 1 = \imath z \cdot \int_0^1 e^{\imath t z}\, dt$$

folgt:

$$\begin{aligned} |e^{\imath z} - 1| &= |z| \cdot \left| \int_0^1 e^{\imath t z}\, dt \right| \leq |z| \cdot \int_0^1 |e^{\imath t z}|\, dt \\ &= |z| \cdot \int_0^1 \underbrace{|e^{\imath a t - t b}|}_{= e^{-tb}}\, dt \\ &\leq \begin{cases} |z| & b \geq 0 \\ |z| \cdot e^{|z|} & b < 0 \end{cases} \end{aligned}$$

(ii). Sei

$$\tau_n(z) := e^{\imath z} - \sum_{k=0}^{n} \frac{(\imath z)^k}{k!}$$

für $n \in \mathbb{N}$. Dann gilt:

$$\tau_{n+1}(z) = \imath z \cdot \int_0^1 \tau_n(tz)\, dt$$

Nun vollständige Induktion.

3.1.7 Satz

Es sei f die charakteristische Funktion einer Zufallsgröße X und für ein $k \geq 1$ existiere das Moment M_k. Dann ist f k-mal differenzierbar. Die k-te Ableitung $f^{(k)}$ ist beschränkt, gleichmäßig stetig und

(i).

$$f^{(k)}(t) = \imath^k \cdot \int_{\mathbb{R}} x^k \cdot e^{\imath t x}\, d\mu(x)$$

(ii). $f^{(k)}(0) = i^k \cdot M_k$

(iii). Es gilt:

$$\begin{aligned} f(t) &= \sum_{n=0}^{k} f^{(n)}(0) \cdot \frac{t^n}{n!} + \vartheta(t) \cdot \frac{t^k}{k!} \\ &\overset{(ii)}{=} \sum_{n=0}^{k} i^n \cdot M_n \cdot \frac{t^n}{n!} + \vartheta(t) \cdot \frac{t^k}{k!} \end{aligned}$$

wobei $\vartheta : \mathbb{R} \to \mathbb{C}$ stetig mit $\vartheta(0) = 0$ und

$$|\vartheta(t)| \leq \sup_{0 \leq \eta \leq 1} |f^{(k)}(\eta \cdot t) - f^{(k)}(0)|$$

Beweisskizze:

- Für $h \neq 0$:

$$\begin{aligned} \left| \frac{f(t+h) - f(t)}{h} \right| &= \left| \int_{\mathbb{R}} x \cdot \frac{e^{\imath hx} - 1}{hx} \cdot e^{\imath tx} \, d\mu(x) \right| \\ &\overset{3.1.6}{\leq} \int_{\mathbb{R}} |x| \cdot \frac{|hx|}{|hx|} \, d\mu(x) < \infty \end{aligned}$$

  Aus dominierter Konvergenz folgt mit

$$\frac{e^{\imath hx} - 1}{hx} \to 1$$

  die Behauptung (i) für $k = 0$. Für $k > 1$ vollständige Induktion.

- Aus (i) folgt:

$$|f^{(k)}(t)| \leq \int_{\mathbb{R}} |x|^k \, d\mu(x) < \infty$$

  also $f^{(k)}$ beschränkt. $f^{(k)}$ gleichmäßig stetig, denn

$$\begin{aligned} |f^{(k)}(t+h) - f^{(k)}(t)| &\leq \int_{\mathbb{R}} |x^k| \cdot |e^{\imath hx} - 1| \, d\mu(x) \\ &\to 0 \qquad (h \to 0) \end{aligned}$$

- f ist k-mal differenzierbar, d.h. es gilt die Taylor-Formel:

$$\begin{aligned} f(t) &= \sum_{n=0}^{k} f^{(k)}(0) \cdot \frac{t^n}{n!} + \underbrace{\int_0^t \frac{(t-x)^{k-1}}{(k-1)!} \cdot (f^{(k)}(x) - f^{(k)}(0)) \, dx}_{=:R_k(t)} \\ \overset{t \neq 0}{\Rightarrow} \vartheta(t) &= \frac{R_k(t) \cdot k!}{t^k} \qquad \vartheta(0) := 0 \end{aligned}$$

  Offensichtlich ist $R_k(t)$ stetig (wegen f stetig), also $\vartheta(t)$ stetig für $t \neq 0$. Stetigkeit in 0:

$$\begin{aligned} |R_k(t)| &\leq \sup_{0 \leq \eta \leq 1} |f^{(k)}(\eta \cdot t) - f^{(k)}(0)| \cdot \int_0^{|t|} \frac{(|t| - x)^{k-1}}{(k-1)!} \, dx \\ &= \sup_{0 \leq \eta \leq 1} |f^{(k)}(\eta \cdot t) - f^{(k)}(0)| \cdot \frac{|t|^k}{k!} \end{aligned}$$

3.1.8 Satz

Seien X, $(X_n)_{n\in\mathbb{N}}$ Zufallsgrößen, f, $(f_n)_{n\in\mathbb{N}}$ zugehörige charakteristische Funktionen. Dann gilt:

$$\forall t \in \mathbb{R}: \lim_{n\to\infty} f_n(t) = f(t) \Leftrightarrow \lim_{n\to\infty} X_n = X \text{ in Verteilung}$$

ohne Beweis

3.2 Beispiele

Im Folgenden μ Verteilung, f charakteristische Funktion.

(i). Sei X poisson-verteilt mit Parameter $a > 0$. Dann:

$$\begin{aligned} f(t) &= \mathbb{E}(\exp(\imath tX)) = \sum_{k=0}^{\infty} e^{\imath tk} \cdot \mathbb{P}[X = k] \\ &= \sum_{k=0}^{\infty} e^{\imath tk} \cdot \frac{a^k}{k!} \cdot e^{-a} = e^{-a} \cdot \sum_{k=0}^{\infty} \frac{(a \cdot e^{\imath t})^k}{k!} \\ &= e^{-a} \cdot \exp(a \cdot e^{\imath t}) = \exp(a \cdot (e^{\imath t} - 1)) \end{aligned}$$

Folgerung: X, Y unabhängig, poisson-verteilt mit Parameter a bzw. b. Dann $X + Y$ poisson-verteilt mit Parameter $a + b$.

(ii). X gleichmäßig verteilt in $(-a, a)$, d.h. Dichte

$$p(x) = \frac{1}{2a} \cdot 1_{(-a,a)}(x)$$

Dann:

$$\begin{aligned} f(t) &= \int_{-a}^{a} \frac{e^{\imath tx}}{2a}\, dx = \int_{-a}^{a} \frac{\cos(tx)}{2a} + \imath \cdot \int_{-a}^{a} \frac{\sin(tx)}{2a}\, dx \\ &= \frac{1}{2a} \cdot \left(\left[\frac{\sin(tx)}{t} \right]_{-a}^{a} + \imath \cdot \left[\frac{-\cos(tx)}{t} \right]_{-a}^{a} \right) \\ &= \begin{cases} \dfrac{\sin(at)}{at} & t \neq 0 \\ 1 & t = 0 \end{cases} \end{aligned}$$

(iii). X binomial-verteilt mit Parametern n,p mit $n \in \mathbb{N}$, $p \in (0, 1)$.

$$\begin{aligned} f(t) &= \mathbb{E}(\exp(\imath tX)) = \sum_{k=0}^{n} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \cdot \underbrace{e^{\imath tk}}_{(\exp(\imath t))^k} \\ &= (p \cdot e^{\imath t} + (1-p))^n \end{aligned}$$

Folgerung: X binomial-verteilt mit n,p, Y binomial-verteilt mit m, p, X und Y unabhängig $\Rightarrow X + Y$ binomial-verteilt mit Parametern $n + m$, p

(iv). X sei normalverteilt mit Parametern a und σ. Dann

$$f(t) = \exp\left(\imath at - \frac{\sigma^2 \cdot t^2}{2}\right)$$

Folgerung: X normalverteilt mit a_1, σ_1, Y normalverteilt mit a_2, σ_2, X und Y unabhängig $\Rightarrow X + Y$ normalverteilt mit $a_1 + a_2, \sqrt{\sigma_1^2 + \sigma_2^2}$.

(v). X besitze die Dichte

$$p(x) = \frac{1}{2} \cdot e^{-|x|}$$

p ist tatsächlich eine Dichte, denn

$$\int_{\mathbb{R}} p(x)\,dx = \frac{1}{2} \cdot 2 \cdot \int_0^\infty e^{-|x|}\,dx = 1$$

Dann:

$$\begin{aligned} f(t) &= \frac{1}{2} \cdot \int_{\mathbb{R}} e^{\imath t x} \cdot e^{-|x|}\,\lambda(dx) \\ &= \ldots = \frac{1}{1+t^2} \end{aligned}$$

(vi). X sei Cauchy-verteilt, d.h. besitze Dichte

$$p(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}$$

Mit Residuensatz folgt:

$$f(t) = e^{-|t|}$$

3.3 Umkehrfunktionen

Im Folgenden: f charakteristische Funktion der Verteilung μ

3.3.1 Aufgabe

(i). Sei $-\infty < a < b < \infty$, $x \in \mathbb{R}$. Bestimme

$$\lim_{T\to\infty} \frac{1}{2\pi} \cdot \int_{-T}^{T} \frac{\sin((x-a)\cdot t)}{t} - \frac{\sin((x-b)\cdot t)}{t}\,dt$$

Hinweis: Aus 1.2.6 folgt Grenzwert

$$G(x) := \begin{cases} 1 & x \in (a,b) \\ \frac{1}{2} & x \in \{a,b\} \\ 0 & x \notin [a,b] \end{cases}$$

(ii). Sei $-\infty < a < b < \infty$. Dann:

$$\left| \int_a^b \frac{\sin t}{t}\,dt \right| < 2\pi$$

Hinweis:

$$c_n := \int_{n\cdot\pi}^{(n+1)\cdot\pi} \frac{\sin t}{t}\,dt = (-1)^n \cdot \int_0^\pi \frac{\sin t}{n\cdot\pi + t}\,dt$$

Dann $|c_n| \geq |c_{n+1}|$, Leibniz (?):

$$0 \leq \int_0^T \frac{\sin t}{t}\,dt \leq c_0 < \pi \qquad (T > 0)$$

3.3.2 Satz

Sei $-\infty < a < b < \infty$. Es gilt:

$$\lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-\imath a t} - e^{-\imath b t}}{\imath t} \cdot f(t)\,dt = \mu((a,b)) + \frac{1}{2} \cdot (\mu(\{a\}) + \mu(\{b\}))$$

ohne Beweis

3.3.3 Folgerung

Jede Verteilung ist durch ihre charakteristische Funktion eindeutig bestimmt. (Vgl. 3.1.2)

3.3.4 Satz

Für $a \in \mathbb{R}$ gilt

$$\lim_{T\to\infty} \frac{1}{2T} \cdot \int_{-T}^{T} e^{-\imath at} \cdot f(t)\, dt = \mu(\{a\})$$

Beweis:

$$\begin{aligned}
\frac{1}{2T} \cdot \int_{-T}^{T} e^{-\imath at} \cdot f(t)\, dt &= \frac{1}{2T} \cdot \int_{-T}^{T} \int_{\mathbb{R}} e^{-\imath at} \cdot e^{\imath tx}\, d\mu(x)\, dt \\
&\overset{SvF}{=} \int_{\mathbb{R}} \frac{1}{2T} \int_{-T}^{T} e^{\imath t(x-a)}\, dt\, d\mu(x) \\
&= \mu(\{a\}) + \int_{\mathbb{R}\setminus\{a\}} \frac{\sin(T \cdot (x-a))}{T \cdot (x-a)}\, d\mu(x)
\end{aligned}$$

3.3.5 Satz

Gilt $f \in L^1(\mathbb{P})$, so ist μ absolut stetig und für die zugehörige Dichte gilt

$$p(x) = \frac{1}{2\pi} \cdot \int_{-\infty}^{\infty} e^{-\imath tx} \cdot f(t)\, dt$$

p ist beschränkt und gleichmäßig stetig.

4

Zentraler Grenzwertsatz

4.1 Satz von Moivre und Laplace

$$\varphi(x) := \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{x} \exp\left(-\frac{y^2}{2}\right) dy$$

Seien $(X_n)_{n\in\mathbb{N}} \subset L^2(\mathbb{P})$ unabhängig, identisch verteilt mit $D^2(X_i) = \sigma^2 > 0$. Setze

$$S_n := \frac{1}{\sigma \cdot \sqrt{n}} \cdot \sum_{i=1}^{n} X_i - \mathbb{E}X_i$$

Dann $\mathbb{E}(S_n) = 0, D^2(S_n) = 1$.

4.1.1 Aufgabe

Sei $(r_n)_{n\in\mathbb{N}}$ eine Nullfolge in $\mathbb{R}$. Dann gilt:

$$\lim_{n\to\infty} \left(1 + \frac{r + r_n}{n}\right)^n = e^r$$

4.1.2 Satz: de Moivre-Laplace

Sei $(X_n)_{n\in\mathbb{N}} \subset L^2(\mathbb{P})$ identisch verteilt, unabhängig mit $D^2(X_j) =: \sigma^2$. Dann gilt:

$$\forall x \in \mathbb{R} : \lim_{n\to\infty} \mathbb{P}[S_n < x] = \varphi(x)$$

Bemerkung zum Beweis:

- Betrachte

  $$f_{S_n}(t) = f\left(\frac{t}{\sigma \cdot \sqrt{n}}\right)^n \to \exp\left(-\frac{t^2}{2}\right)$$

  mit f als charakteristische Funktion von X_1. Dazu

  $$f(s) = 1 - \frac{1}{2}\sigma^2 \cdot s^2 + R_2(s)$$

  Mit 4.1.1 folgt dann die Behauptung.