Stochastik I

Repetitorium

Leonardo Vela, Martin Dattge

13. Januar 2021

Vorab zur Klarstellung

- Alles an klausurrelevantem Stoff an einem Tag durchzukriegen ist unmöglich.
- Wir werden versuchen, so viel Stoff aufzugreifen wie möglich.
- Bitte denkt nicht, dass nur die heute behandelten Sätze/Def etc. in der Klausur kommen werden, wir haben die Klausur bisher selbst noch nicht gesehen....

$\Longrightarrow$ Message: Kein Anspruch auf Vollständigkeit!

Ablauf heute

- Wir wiederholen Kapitel für Kapitel mit Euch.
- Es wird immer wieder Aufgabenblöcke zwischendurch geben, für deren Bearbeitung Ihr 20-30 Minuten Zeit bekommt.

Ablauf heute

- Wir wiederholen Kapitel für Kapitel mit Euch.
- Es wird immer wieder Aufgabenblöcke zwischendurch geben, für deren Bearbeitung Ihr 20-30 Minuten Zeit bekommt.
- Ja, alle Lösungen sind getext und Ja, die Slides werden hochgeladen.

Ablauf heute

- Wir wiederholen Kapitel für Kapitel mit Euch.
- Es wird immer wieder Aufgabenblöcke zwischendurch geben, für deren Bearbeitung Ihr 20-30 Minuten Zeit bekommt.
- Ja, alle Lösungen sind getext und Ja, die Slides werden hochgeladen.
- Heute keine Breakout-Rooms, da ihr es in der Klausur auch alleine hinbekommen müsst, also testet Euch heute selbst wie gut ihr alles schon verstanden habt!

Ablauf heute

- Wir wiederholen Kapitel für Kapitel mit Euch.
- Es wird immer wieder Aufgabenblöcke zwischendurch geben, für deren Bearbeitung Ihr 20-30 Minuten Zeit bekommt.
- Ja, alle Lösungen sind getext und Ja, die Slides werden hochgeladen.
- Heute keine Breakout-Rooms, da ihr es in der Klausur auch alleine hinbekommen müsst, also testet Euch heute selbst wie gut ihr alles schon verstanden habt!
- Bei Fragen während der Bearbeitung der Aufgaben einfach Mikro anmachen, Hand heben oder in den Chat schreiben :)

Kurze Klausurbearbeitungshinweise

- Ihr müsst uneigentliche Integrale z.B. mit MCT “abschneiden“, das heißt es darf NICHT unendlich an einer Stammfunktion auftauchen.
  Alle Sätze/RR, die ihr aus Ana1 kennt sind nämlich nur für stetige Intervalle definiert.

Kurze Klausurbearbeitungshinweise

- Ihr müsst uneigentliche Integrale z.B. mit MCT “abschneiden“, das heißt es darf NICHT unendlich an einer Stammfunktion auftauchen.
  Alle Sätze/RR, die ihr aus Ana1 kennt sind nämlich nur für stetige Intervalle definiert.
- Wohldefiniertheit zeigt man, indem man von einer Funktion den Positiv- und Negativteil separat betrachtet. Wenn beide unendlich sind wäre das sehr schlecht, da dann nicht wohldefiniert...
- Macht Euch deswegen aber nicht verrückt, in (fast) allen Beispielen die Ihr kennt ist der Integrand nichtnegativ und somit ist das Integral wohldefiniert, weil der Negativteil 0 ist.

Kurze Klausurbearbeitungshinweise

- Ihr müsst uneigentliche Integrale z.B. mit MCT "abschneiden", das heißt es darf NICHT unendlich an einer Stammfunktion auftauchen.
  Alle Sätze/RR, die ihr aus Ana1 kennt sind nämlich nur für stetige Intervalle definiert.
- Wohldefiniertheit zeigt man, indem man von einer Funktion den Positiv- und Negativteil separat betrachtet. Wenn beide unendlich sind wäre das sehr schlecht, da dann nicht wohldefiniert...
- Macht Euch deswegen aber nicht verrückt, in (fast) allen Beispielen die Ihr kennt ist der Integrand nichtnegativ und somit ist das Integral wohldefiniert, weil der Negativteil 0 ist.

$\Rightarrow$ Dann geht´s jetzt los mit Kapitel 1 !

σ-Algebra

Definition

Sei Ω nicht leer. $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ heißt σ-Algebra, falls

1. $\Omega \in \mathcal{A}$,
2. $A \in \mathcal{A} \Rightarrow A^C \in \mathcal{A}$, das nennt man auch stabil (oder abgeschlossen) unter Komplementbildung,
3. $A_1, A_2, \ldots \in \mathcal{A} \Rightarrow \bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$, das nennt man auch stabil (oder abgeschlossen) unter abzählbarer Vereinigung.

Elemente von $\mathcal{A}$ heißen **messbare Mengen**. Ist $\mathcal{A} \subseteq \mathcal{B}$ und $\mathcal{A}, \mathcal{B}$ sind σ-Algebren, so nennt man $\mathcal{A}$ Unter-σ-Algebra von $\mathcal{B}$.

Maß

Definition

Für eine σ-Algebra $\mathcal{A}$ heißt $\mu\colon \mathcal{A} \longrightarrow [0, \infty]$ ein **Maß auf** $\mathcal{A}$, falls folgende Eigenschaften gelten:

1. $\mu(\emptyset) = 0$
2. Sind $A_1, A_2, \ldots \in \mathcal{A}$ paarweise disjunkte Mengen, so gilt $\mu(\dot\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mu(A)$. Wir nennen diese Eigenschaft σ-Additivität.

Ein Maß μ heißt **endlich**, falls $\mu(\Omega) < \infty$. μ heißt **Wahrscheinlichkeitsmaß**, falls $\mu(\Omega) = 1$.

Stetigkeit von Maßen

Satz

Sei $(\Omega, \mathcal{A}, \mu)$ *ein Maßraum und* $(A_n)_{n\in\mathbb{N}}$ *eine Folge messbarer Mengen, so gelten:*

i) *Aus* $A_n \uparrow A$ *(d. h.* $A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots, \bigcup_{n=1}^{\infty} A_n = A$*) folgt* $\lim\limits_{n\to\infty} \mu(A_n) = \mu(A)$.

ii) *Aus* μ *endlich und* $A_n \downarrow A$ *(d. h.* $A_1 \supseteq A_2 \supseteq \ldots, \bigcap\limits_{n=1}^{\infty} A_n = A$*) folgt* $\lim\limits_{n\to\infty} \mu(A_n) = \mu(A)$.

Erzeugte σ-Algebra

Definition

Sei $\mathcal{E} \subseteq \mathcal{P}(\Omega)$, so existiert genau eine σ-Algebra $\mathcal{A}$ mit

i) $\mathcal{E} \subseteq \mathcal{A}$

ii) Ist $\mathcal{E} \subseteq \mathcal{B}$ und $\mathcal{B}$ ist eine σ-Algebra, so gilt $\mathcal{A} \subseteq \mathcal{B}$.

Dabei bedeutet ii, dass $\mathcal{A}$ die kleinste σ-Algebra ist, die $\mathcal{E}$ enthält.
Für $\mathcal{E} \subseteq \mathcal{P}(\Omega)$ heißt

$$\sigma(\mathcal{E}) = \bigcap_{\substack{\mathcal{E} \subseteq \mathcal{B}, \\ \mathcal{B}\ \sigma\text{-Alg.}}} \mathcal{B}$$

die von $\mathcal{E}$ **erzeugte σ-Algebra**. Ist $= \sigma(\mathcal{E})$, so nennt man $\mathcal{E}$ einen Erzeuger von .

Hauptsatz

Satz

Ist $\mathcal{E} \subseteq \mathcal{P}(\Omega)$ $\cap$*-stabil, so gilt* $d(\mathcal{E}) = \sigma(\mathcal{E})$.

Hauptsatz

Satz

Ist $\mathcal{E} \subseteq \mathcal{P}(\Omega)$ $\cap$-stabil, so gilt $d(\mathcal{E}) = \sigma(\mathcal{E})$.

- Für einen messbaren Raum $(\Omega, \mathcal{A})$ und zwei (endliche) Maße μ_1, μ_2 ist das insbesondere wichtig für die Menge
$$\mathcal{M} = \{A \in \mathcal{A} \mid \mu_1(A) = \mu_2(A)\}$$
die ein Dynkin-System ist, falls wir zusätzlich $\mu_1(\Omega) = \mu_2(\Omega)$ fordern.
- Finden wir Gleichheit der Maße auf einem $\cap$-stabilen Erzeuger von $\mathcal{A}$ so wissen wir, dass Gleichheit auf ganz $\mathcal{A}$ gilt.

Satz

Es sei $(\Omega, \mathcal{A})$ *ein messbarer Raum,* $\mathcal{E}$ *ein* $\cap$*-stabiler Erzeuger von* $\mathcal{A}$ *und* μ_1, μ_2 *seien Maße auf* $\mathcal{A}$*. Zudem gelten:*

i) *Es gibt eine Folge* $(E_n) \subseteq \mathcal{E}$ *mit* $E_n \uparrow \Omega$*,* $n \to \infty$*, und* $\mu_i(E_n) < \infty$ *für alle* $n \in \mathbb{N}$*,* $i = 1, 2$*.*

ii) $\mu_1(A) = \mu_2(A)$ *für alle* $A \in \mathcal{E}$*.*

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

Semiring

Definition

$\mathcal{S} \subseteq \mathcal{P}(\Omega)$ heißt **Semiring**, falls

i) $\emptyset \in \mathcal{S}$

ii) $A, B \in \mathcal{S} \Rightarrow A \cap B \in \mathcal{S}$, also ist $\mathcal{S}$ $\cap$-stabil

iii) $A, B \in \mathcal{S} \Rightarrow$ es gibt paarweise disjunkte Mengen $C_1, \dots, C_m \in \mathcal{S}$ mit $A \setminus B = \dot{\bigcup}_{k=1}^{m} C_k$.

Fortsetzungssatz von Carathéodory

Satz

Sei ein Semiring und $\mu : \mathcal{S} \to [0, \infty]$ *eine Mengenfunktion mit*

- $\mu(\emptyset) = 0,$
- μ *ist* σ-***additiv*** *(d.h. sind* $A_1, A_2, \ldots \in S$ *paarweise disjunkt mit* $A := \dot{\bigcup}_{k=1}^{\infty} A_k \in \mathcal{S}$, *so gilt* $\mu(A) = \sum_{k=1}^{\infty} \mu(A_k)$*).*

Dann existiert ein Maß $\bar{\mu}$ *auf* $\sigma(\mathcal{S})$ *mit* $\mu(A) = \bar{\mu}(A)$ *für alle* $A \in \mathcal{S}$.

Existenz und Eindeutigkeit von Maßen

Satz

Ist $(\Omega, \mathcal{A})$ *ein messbarer Raum,* $\mathcal{E}$ *ein* $\cap$*-stabiler Semiring mit* $\sigma(\mathcal{E}) = \mathcal{A}$*. Sei* $\mu : \mathcal{E} \longrightarrow [0, \infty]$ *mit*

- $\mu(\emptyset) = 0$
- μ *ist* σ*-additiv*
- *es gibt Folge eine* $E_1, E_2, \ldots \in \mathcal{E}$ *mit* $E_n \uparrow \Omega$ *und* $\mu(E_n) < \infty$ *für alle* $n \in \mathbb{N}$*.*

Dann existiert genau ein Maß $\bar{\mu}$ *auf* $\mathcal{A} = \sigma(\mathcal{E})$*, so dass* $\bar{\mu}(A) = \mu(A)$ *für alle* $A \in \mathcal{E}$*.*

Aufgabenblock 1

a) Sei μ ein Maß und $\mathcal{A}$ eine σ-Algebra. Zeige, dass für beliebige $A_1, A_2, \ldots, A_n \in \mathcal{A}$

$$\mu\left(\bigcup_{k=1}^{n} A_k\right) \leq \sum_{k=1}^{n} \mu(A_k)$$

gilt.

b) Zeige, dass der Schniit zweier Dynkin-Systeme $\mathcal{D}_1$ und $\mathcal{D}_2$ wieder ein Dynkin-System ist.

c) Zeige, dass für beliebiges $\mathcal{E} \subseteq \mathcal{P}(\Omega)$

$$d(\sigma(\mathcal{E})) = \sigma(d(\mathcal{E}))$$

gilt.

Lösung Aufgabenblock 1

a) Per Induktion. Für $n = 1$ gilt die Behauptung offensichtlich.

IA: Sei $n = 2$. Dann gilt

$$\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2) - \mu(A_1 \cap A_2) \leq \mu(A_1) + \mu(A_2)$$

IV: Die Behauptung gelte für beliebiges, aber festes $n \in \mathbb{N}$

IS: Definiere $A := \cup_{k=1}^{n} A_k$. Mit dem gleichem Argument wie für $n = 2$ folgt

$$\mu(A \cup A_{n+1}) \leq \mu(A) + \mu(A_{n+1}) \leq \sum_{k=1}^{n+1} \mu(A_k)$$

b) Definition checken:

i) Da $\Omega \in \mathcal{D}_1$ und $\Omega \in \mathcal{D}_2$ folgt $\Omega \in \mathcal{D}_1 \cap \mathcal{D}_2$

ii) Sei $A \in \mathcal{D}_1 \cap \mathcal{D}_2$. Dann muss auch $A \in \mathcal{D}_1$ und $A \in \mathcal{D}_2$ gelten. Da $\mathcal{D}_1, \mathcal{D}_2$ Dynkin-Systeme sind, ist $A^c \in \mathcal{D}_1$ und $A^c \in \mathcal{D}_2$. Damit ist auch $A^c \in \mathcal{D}_1 \cap \mathcal{D}_2$

iii) Seien $A_1, A_2, \ldots \in \mathcal{D}_1 \cap \mathcal{D}_2$ paarweise disjunkt. Dann sind $A_1, A_2, \ldots \in \mathcal{D}_1$ und $A_1, A_2, \ldots \in \mathcal{D}_2$. Damit ist auch $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{D}_1$ und $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{D}_2$. Damit folgt $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{D}_1 \cap \mathcal{D}_2$.

c) Da jede σ-Algebra ein Dynkin-Sytem ist, gilt $d(\sigma(\mathcal{E})) = \sigma(\mathcal{E})$. Mit der Monotonie von d folgt

$$\mathcal{E} \subseteq d(\mathcal{E}) \subseteq \sigma(\mathcal{E}).$$

Mit der Monotonie Eigenschaft und Idempotenz von σ folgt dann

$$\sigma(\mathcal{E}) \subseteq \sigma(d(\mathcal{E})) \subseteq \sigma(\sigma(\mathcal{E})) = \sigma(\mathcal{E}).$$

Definition

Eine Funktion

$$F\colon \mathbb{R} \longrightarrow \mathbb{R}, \quad x \mapsto F(x),$$

welche die Eigenschaften

i) $0 \leq F(x) \leq 1$ für alle $x \in \mathbb{R}$,

ii) F ist monoton steigend,

iii) F ist rechtsstetig,

iv) $\lim\limits_{x \longrightarrow -\infty} F(x) = 0$ und $\lim\limits_{x \longrightarrow +\infty} F(x) = 1$,

erfüllt heißt Verteilungsfunktion.

Verteilungsfunktionen und Wahrscheinlichkeitsmaße

Satz

Für jede Verteilungsfunktion F gibt es genau ein Wahrscheinlichkeitsmaß $\mathbb{P}_F$ auf $\mathcal{B}(\mathbb{R})$ mit $\mathbb{P}_F((-\infty, x]) = F(x)$.

Definition

Sei $\mathbb{P}$ ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Dann heißt die Funktion

$$F\colon \mathbb{R} \to [0,1], \quad x \mapsto \mathbb{P}((-\infty, x])$$

Verteilungsfunktion von $\mathbb{P}$.

- Auf Üb 2 habt ihr gezeigt, dass F aus der Definition eine Verteilungsfunktion ist.
- Jede Verteilungsfunktion korrespondiert also mit genau einem Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Diskrete vs. Stetige Verteilungsfunktionen

Definition

Ist $f : \mathbb{R} \to [0, \infty]$ integrierbar mit $\int_{\mathbb{R}} f(x)\mathrm{d}x = 1$, dann heißt f Dichtefunktion der Verteilungsfunktion

$$F(t) = \int_{-\infty}^{t} f(x)\mathrm{d}x, t \in \mathbb{R}$$

Ist umgekehrt F von obiger Form, so heißt f Dichte von F.

Verteilungsfunktionen mit Dichten nennt man absolutstetig.

Definition

Für $a_1, \ldots, a_n \in \mathbb{R}$, $N \in \mathbb{N}$ oder $N = \infty$ mit $p_1, \cdots, p_n \geq 0$ und $\sum_{k=1}^{N} p_k = 1$ ist

$$F(t) = \sum_{k=1}^{N} p_k \mathbb{1}_{[a_k,\infty)}(t), t \in \mathbb{R}$$

eine Verteilungsfunktion. Die zugehörigen Maße $\mathbb{P}_F$ werden diskrete Verteilungen genannt.

Juhu

Juhu Kapitel 1 ist schon geschafft, kurz durchatmen und glücklich sein. :)
Weiter geht´s mit Nummer 2: Abbildungen zwischen messbaren Räumen

Messbarkeit von Funktionen

Definition

Seien $(\Omega, \mathcal{A}), (\Omega', \mathcal{A}')$ messbare Räume und $f\colon \Omega \to \Omega'$. Dann heißt f $(\mathcal{A} - \mathcal{A}')$-messbar, falls

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

für alle $A' \in \mathcal{A}'$ gilt.

Messbarkeit von Funktionen

Definition

Seien $(\Omega, \mathcal{A}), (\Omega', \mathcal{A}')$ messbare Räume und $f: \Omega \to \Omega'$. Dann heißt f $(\mathcal{A} - \mathcal{A}')$-messbar, falls

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

für alle $A' \in \mathcal{A}'$ gilt.

Satz

Ist $\mathcal{E}'$ ein Erzeuger von $\mathcal{A}'$ und $f: \Omega \to \Omega'$. Dann ist die Messbarkeit von f äquivalent zu

$$A' \in \mathcal{E}' \implies f^{-1}(A') \in \mathcal{A}.$$

- Der obige Satz ist sehr nützlich um die Messbarkeit von Funktionen nachzuweisen.

Ein letztes Mal: Erzeuger von $\mathcal{B}(\mathbb{R})$ bzw. $\mathcal{B}(\bar{\mathbb{R}})$

Die folgenden Mengensysteme sind zum Beispiel Erzeuger von $\mathcal{B}(\mathbb{R})$:

$$\mathcal{E}_1 := \{(-\infty, t) : t \in \mathbb{R}\},$$
$$\mathcal{E}_2 := \{(-\infty, t] : t \in \mathbb{R}\},$$
$$\mathcal{E}_3 := \{(a, b) : a, b \in \mathbb{R}, a < b\},$$
$$\mathcal{E}_4 := \{[a, b] : a, b \in \mathbb{R}, a < b\},$$
$$\mathcal{E}_5 := \{(t, \infty) : t \in \mathbb{R}\},$$
$$\mathcal{E}_6 := \{[t, \infty) : t \in \mathbb{R}\}.$$

Mit Satz 2.1.4. müssen wir um die Messbarkeit einer $\mathbb{R}$-wertigen Funktion nachzuweisen, also nur die Urbilder der Mengen in einem dieser Mengensysteme betrachten.

Definition

Sei $f : \Omega \longrightarrow \Omega'$ für einen messbaren Raum $(\Omega', \mathcal{A}')$. Dann ist

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

eine σ-Algebra und $\mathcal{A}$ ist die kleinste σ-Algebra auf Ω, für die f $(\mathcal{A}, \mathcal{A}')$-messbar ist. Wir nennen die σ-Algebra auch $\sigma(f)$.

Bildmaß oder push-forward

Definition

Sei $f : (\Omega, \mathcal{A}) \to (\Omega', \mathcal{A}')$ messbar und μ ein Maß auf $\mathcal{A}$. Dann ist

$$\mu_f(B) := \mu(f^{-1}(B)), B \in \mathcal{A}'$$

ein Maß auf $\mathcal{A}'$. Dieses Maß heist Bildmaß oder push-forward.

- Wir nutzen also die Messbarkeit der Abbildung f, um ein Maß μ auf $\mathcal{A}$ auf ein Maß μ_f auf $\mathcal{A}'$ rüberzuschieben.
- Wer sich an die Definition der Verteilung von Zufallsvariablen erinnert, hat jetzt hoffentlich nochmal den AHA-Effekt bekommen.

Messbare numerische Funktionen

Definition

Für einen messbaren Raum $(\Omega, \mathcal{A})$ heißt $f : \Omega \to \bar{\mathbb{R}}$ messbare numerische Funktion, falls f $(\mathcal{A}, \mathcal{B}(\bar{\mathbb{R}}))$-messbar ist.

Satz

Sind $f, g : \Omega \to \bar{\mathbb{R}}$ Funktionen von einem messbaren Raum $(\Omega, \mathcal{A})$ nach $(\mathbb{R}, \mathcal{B}(\bar{\mathbb{R}}))$, so sind die Mengen bzw. Funktionen

$$\{f < g\}, \{f \leq g\}, \{f = g\}, \{f \neq g\}$$

$$f + g, \alpha \cdot f \, für \alpha \in \mathbb{R}, f \cdot g, f \wedge g, f \vee g, |f|$$

alle messbar.

Folgen messbarer numerischer Funktionen

Satz

Ist außerdem $(f_n)_{n\in\mathbb{N}}$ *eine Folge messbarer numerischer Funktionen mit* $f_n\colon \Omega \longrightarrow \bar{\mathbb{R}}$ *für alle* $n \in \mathbb{N}$*. Dann sind die für* $\omega \in \Omega$ *punktweise definierten Funktionen*

$$g_1(\omega) := \inf_{n\in\mathbb{N}} f_n(\omega), \quad g_2(\omega) := \sup_{n\in\mathbb{N}} f_n(\omega),$$
$$g_3(\omega) := \liminf_{n\in\mathbb{N}} f_n(\omega), \quad g_4(\omega) := \limsup_{n\in\mathbb{N}} f_n(\omega),$$

messbar. Falls für jedes $\omega \in \Omega$ *der Grenzwert*

$$g(\omega) := \lim_{n\longrightarrow\infty} f_n(\omega)$$

in $\bar{\mathbb{R}}$ *existiert, ist auch* g *eine messbare numerische Funktion .*

$\Longrightarrow$ Message: Gefühlt alles ist messbar!

Aufgabenblock 2

a) Zeige: Jede stetige Abbildung ist $(\mathcal{B}(\bar{\mathbb{R}}), \mathcal{B}(\bar{\mathbb{R}}))$-messbar.

b) Zeige: Ist $h : \Omega \longrightarrow \Omega'$ $(\mathcal{A}, \mathcal{A}')$-messbar und $k : \Omega' \longrightarrow \Omega''$ $(\mathcal{A}', \mathcal{A}'')$-messbar, so ist $k \circ h : \Omega \longrightarrow \Omega''$ $(\mathcal{A}, \mathcal{A}'')$-messbar.

c) Zeige: Sind $f, g : \Omega \to \bar{\mathbb{R}}$ $(\mathcal{A}, \mathcal{B}(\bar{\mathbb{R}})$-messbar, so ist $\{f < g\}$ messbar.

Lösung Aufgabenblock 2

a) Mit Proposition 2.1.4 reicht es, die Messbarkeit auf einem Erzeuger nachzuweisen. Wir wählen $\mathcal{B}(\mathbb{R}) = \sigma(\{O : O \subseteq \mathbb{R} \text{ offen }\})$. Nun ist die Definition der Stetigkeit nach Ana 1 aber äquivalent zu Urbilder offener Mengen sind offen. Damit ist die Aussage gezeigt.

b) Weil k $(\mathcal{A}', \mathcal{A}'')$-messbar ist, gilt $k^{-1}(A'') \in \mathcal{A}' \forall A'' \in \mathcal{A}''$.
Weil h $(\mathcal{A}, \mathcal{A}')$-messbar ist, gilt $h^{-1}(B) \in \mathcal{A} \forall \mathcal{B} \in \mathcal{A}'$.
Also folgt $\forall A'' \in \mathcal{A}''$

$$(k \circ h)^{-1}(A'') = h^{-1}(k^{-1}(A'')) \in \mathcal{A}$$

und somit die Behauptung.

Lösung Aufgabenblock 2

c)

$$\{f < g\} = \bigcup_{t \in \mathbb{Q}} \{f < t < g\} = \bigcup_{t \in \mathbb{Q}} \{f < t\} \cap \{t < g\}.$$

Die erste Gleichheit gilt, da $\mathbb{Q}$ dicht in $\mathbb{R}$ liegt.
Die Behauptung folgt, da $\mathbb{Q}$ abzählbar ist und f, g messbar nach Voraussetzung, also ist die rechte Seite in $\mathcal{A}$ und damit auch die Linke.

Die Gebetsmühle der Integrationstheorie

- Wir leiten zuerst das Lebesgue-Integral für einfache Funktionen her.
- Für nichtnegative, messbare Funktionen führen wir das Lebesgue-Integral über die Supremumsdarstellung ein.
- Nützlich: Durch den Satz der monotonen Konvergenz einfacher Funktionen können wir das Integral einer nichtnegativen, messbaren Funktion f auch als Grenzwert einer Folge von Integralen von $f_n \in \mathcal{E}^+$ schreiben, sofern $f_n \uparrow f$ gilt.
- Über die Zerlegung $f = f^+ - f^-$ kommen wir dann im letzten Schritt zum Lebesgue-Integral für numerische, messbare Funktionen.

Das Lebesgue-Integral: Einfache Funktionen

Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $\mathcal{E} := \{f : f \text{ einfache Funktion}\}$ bzw. $\mathcal{E}^+ := \{f : f \text{ einfache positive Funktion}\}$. Hierbei heißt eine Funktion einfach, falls $\alpha_1, \ldots, \alpha_n \in \overline{\mathbb{R}}$ und $A_1, \ldots, A_n$ existieren, sodass die A_k paarweise disjunkt sind mit $\bigcup_{k=1}^n A_k = \Omega$ und

$$f(\omega) = \sum_{k=1}^{n} \alpha_k \mathbb{1}_{A_k}(\omega), \quad \forall \omega \in \Omega.$$

Für ein $f \in \mathcal{E}^+$ ist dann das Lebesgue-Integral definiert durch

$$\int_\Omega f \, \mathrm{d}\mu := \sum_{k=1}^{n} \alpha_k \mu(A_k).$$

Das Lebesgue-Integral: Nichtnegative, messbare Funktionen

Für eine nichtnegative, messbare Funktion $f : (\Omega, \mathcal{A}) \to (\bar{\mathbb{R}}, \mathcal{B}(\bar{\mathbb{R}}))$ definieren wir dann (etwas unhandlich, dafür aber direkt wohldefiniert) das Lebesgue-Integral durch

$$\int_\Omega f \, \mathrm{d}\mu := \sup \left\{ \int_\Omega g \, \mathrm{d}\mu : g \text{ einfach}, 0 \le g \le f \right\}.$$

Der folgende Satz vereinfacht weitere Beweise durch eine äquivalente Schreibweise.

Satz (Monotone Konvergenz für einfache Funktionen)

Sei $(f_n) \subseteq \mathcal{E}^+$ mit $f_n \uparrow f$, $n \to \infty$, für eine nichtnegative messbare numerische Funktion f. Dann gilt

$$\lim_{n\to\infty} \int_\Omega f_n \, \mathrm{d}\mu = \int_\Omega f \, \mathrm{d}\mu,$$

wobei in der Gleichheit $+\infty = +\infty$ möglich ist.

Dies ist besonders nützlich, da nach 3.1.6 für jede für eine nichtnegative messbare numerische Funktion eine solche Folge in $\mathcal{E}^+$ existiert.

Das Lebesgue-Integral: Integrierbare Funktionen

Sei $f : (\Omega, \mathcal{A}) \to (\bar{\mathbb{R}}, \mathcal{B}(\bar{\mathbb{R}}))$ eine messbare numerische Funktion. Dann sind nach der Vorlesung $f^+ := \max\{f, 0\}, f^- := -\min\{f, 0\}$ messbare und nichtnegative Funktionen. Falls

$$\int_\Omega f^+ \,\mathrm{d}\mu < \infty, \quad \int_\Omega f^- \,\mathrm{d}\mu < \infty,$$

gilt, heißt f integrierbar und das Lebesgue-Integral für eine integrierbare Funktion f ist definiert durch

$$\int_\Omega f \,\mathrm{d}\mu := \int_\Omega f^+ \,\mathrm{d}\mu - \int_\Omega f^- \,\mathrm{d}\mu.$$

Wir schränken uns hierbei auf integrierbare Funktionen ein, um den Fall "$\infty - \infty$" zu vermeiden.

Transformationssatz

Satz (Abstrakter Transformationssatz)

Seien $(\Omega, \mathcal{A})$, $(\Omega', \mathcal{A}')$ *messbare Räume,* μ *ein Maß auf* $\mathcal{A}$, $f\colon \Omega \longrightarrow \Omega'$ *messbar,* $g\colon \Omega' \longrightarrow \overline{\mathbb{R}}$ *messbar. Dann ist* g μ_f*-integrierbar genau dann, wenn* $g \circ f$ μ*-integrierbar ist. Ist eine dieser Aussagen erfüllt, so gilt ebenfalls die Transformationsformel*

$$\int_{\Omega'} g \mathrm{d}\mu_f = \int_{\Omega} g \circ f \mathrm{d}\mu.$$

Konvergenzsätze

Satz (Monotone Konvergenz Theorem (MCT))

Seien $f, f_1, f_2, \ldots : \Omega \to \overline{\mathbb{R}}$ *messbar und es gelte* $0 \leq f_1 \leq f_2 \leq \ldots \leq f$ *sowie* $f = \lim_{n\to\infty} f_n$ μ*-f.ü. Dann gilt*

$$\lim_{n\to\infty} \int_\Omega f_n \, \mathrm{d}\mu = \int_\Omega f \, \mathrm{d}\mu,$$

$+\infty = +\infty$ *ist dabei möglich.*

Konvergenzsätze

Satz (Monotone Konvergenz Theorem (MCT))

Seien $f, f_1, f_2, \ldots : \Omega \to \overline{\mathbb{R}}$ *messbar und es gelte* $0 \leq f_1 \leq f_2 \leq \ldots \leq f$ *sowie* $f = \lim_{n\to\infty} f_n$ μ*-f.ü. Dann gilt*

$$\lim_{n\to\infty} \int_\Omega f_n \, \mathrm{d}\mu = \int_\Omega f \, \mathrm{d}\mu,$$

$+\infty = +\infty$ *ist dabei möglich.*

Satz (Dominierte Konvergenz Theorem (DCT))

Seien $f, f_1, f_2, \ldots : \Omega \to \overline{\mathbb{R}}$ *messbar. Gilt* $\lim_{n\to\infty} f_n = f$ μ*-fast überall und* $|f_n| \leq g$ μ*-fast überall für alle* $n \in \mathbb{N}$*, für eine beliebige* μ*-integrierbare nichtnegative messbare numerische Funktion* g*.*
Dann sind $f, f_1, f_2, \ldots$ μ*-integrierbar und*

$$\lim_{n\to\infty} \int_\Omega f_n \, \mathrm{d}\mu = \int_\Omega f \, \mathrm{d}\mu.$$

Die Funktion g *spielt keine große Rolle (sie muss nur existieren) und wird Majorante für die Folge* (f_n) *genannt.*

Satz (Hölder-Ungleichung)

Seien $p, q > 1$ *mit* $\frac{1}{p} + \frac{1}{q} = 1$. *Dann gilt*

$$\int_\Omega |fg| \mathrm{d}\mu \leq \left(\int_\Omega |f|^p \mathrm{d}\mu \right)^{\frac{1}{p}} \left(\int_\Omega |g|^q \mathrm{d}\mu \right)^{\frac{1}{q}}.$$

Integralabschätzungen

Satz (Hölder-Ungleichung)

Seien $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt

$$\int_\Omega |fg| \mathrm{d}\mu \leq \left(\int_\Omega |f|^p \mathrm{d}\mu \right)^{\frac{1}{p}} \left(\int_\Omega |g|^q \mathrm{d}\mu \right)^{\frac{1}{q}}.$$

Satz (Minkowski-Ungleichung)

Sei $p \geq 1$, so gilt

$$\left(\int_\Omega |f+g|^p \mathrm{d}\mu \right)^{1/p} \leq \left(\int_\Omega |f|^p \mathrm{d}\mu \right)^{1/p} + \left(\int_\Omega |g|^p \mathrm{d}\mu \right)^{1/p}.$$

Beide Seiten können den Wert $+\infty$ annehmen.

Produktmaße und Fubini

Satz (Existenz des Produktmaßes)

Sind μ_1, μ_2 σ-endliche Maße auf $\mathcal{A}_1, \mathcal{A}_2$, so existiert ein eindeutiges Produktmaß $\mu_1 \otimes \mu_2$ auf $\mathcal{A}_1 \otimes \mathcal{A}_2$ mit

$$\mu_1 \otimes \mu_2(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2)$$

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

Produktmaße und Fubini

Satz (Existenz des Produktmaßes)

Sind μ_1, μ_2 σ-endliche Maße auf $\mathcal{A}_1, \mathcal{A}_2$, so existiert ein eindeutiges Produktmaß $\mu_1 \otimes \mu_2$ auf $\mathcal{A}_1 \otimes \mathcal{A}_2$ mit

$$\mu_1 \otimes \mu_2(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2)$$

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

Satz (Satz von Fubini)

Seien $(\Omega_1, \mathcal{A}_1, \mu_1), (\Omega_2, \mathcal{A}_2, \mu_2)$ σ-endliche Maßräume und $f : \Omega_1 \times \Omega_2 \longrightarrow \overline{\mathbb{R}}$ sei $(\mathcal{A}_1 \otimes \mathcal{A}_2, \mathcal{B}(\overline{\mathbb{R}}))$-messbar. Dann gilt (u.a.)

$$\begin{aligned}\int_{\Omega_1 \times \Omega_2} f \,\mathrm{d}\mu_1 \otimes \mu_2 &= \int_{\Omega_1} \left(\int_{\Omega_2} f_{\omega_1}(\omega_2) \mathrm{d}\mu_2(\omega_2) \right) \mathrm{d}\mu_1(\omega_1) \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} f_{\omega_2}(\omega_1) \mathrm{d}\mu_1(\omega_1) \right) \mathrm{d}\mu_2(\omega_2)\end{aligned}$$

Aufgabenblock 3

a) Sei $f : \mathbb{R} \to \mathbb{R}, x \mapsto (-x^2 + c)\mathbf{1}_{[-1,1]}(x)$. Finde eine Folge von einfachen Funktionen f_n für die $f_n \uparrow f$ gilt. Berechne $\int_{-1}^{1} f_n \mathrm{d}\lambda$ und dann $\int_{-1}^{1} f \mathrm{d}\lambda$

b) Sei X eine nichtnegative Zufallsvariable auf $(\Omega, \mathcal{A})$ mit $X \sim F$. Zeige, dass

$$\int_0^\infty (1 - F(t))\mathrm{d}t = \mathbb{E}[X]$$

gilt. *Hinweis: Es gilt* $\mathbb{P}(X \leq t) = F(t)$ *und* $\mathbb{P}(A) = \mathbb{E}[\mathbf{1}_A]$ *für* $A \in \mathcal{A}$.

Lösung Aufgabenblock 3

a) Wir wählen die Darstellung (eine von vielen möglichen)

$$f_n : \mathbb{R} \longrightarrow \mathbb{R}, x \mapsto \mathbf{1}_0(x) + \sum_{k=1}^{2^n} \mathbf{1}_{[\frac{-k}{2^n}, \frac{-(k-1)}{2^n}) \cup (\frac{k-1}{2^n}, \frac{k}{2^n}]}(x) \cdot (-\left(\frac{k}{2^n}\right)^2 + 1)$$

Es gilt $f_n \in \mathcal{E}^+$ für alle $n \in \mathbb{N}$ und $f_n \uparrow f$.

$$\begin{aligned} \int_{\mathbb{R}} f_n \, \mathrm{d}\lambda &= \lambda(\{0\}) + \sum_{k=1}^{2^n} \lambda([\frac{-k}{2^n}, \frac{-(k-1)}{2^n}) \cup (\frac{k-1}{2^n}, \frac{k}{2^n}]) \cdot (-\left(\frac{k}{2^n}\right)^2 + 1) \\ &= \sum_{k=1}^{2^n} \frac{2}{2^n}(-\left(\frac{k}{2^n}\right)^2 + 1) \end{aligned}$$

Lösung Aufgabenblock 3

$$= \sum_{k=1}^{2^n} \frac{2}{2^n}\left(-\left(\frac{k}{2^n}\right)^2 + 1\right)$$

$$= \sum_{k=1}^{2^n} \frac{-2k^2}{2^{3n}} + \sum_{k=1}^{2^n} \frac{2}{2^n}$$

$$= \frac{-2}{2^{3n}} \sum_{k=1}^{2^n} k^2 + 2$$

$$= \frac{-2}{2^{3n}} \frac{2^n(2^n - 1)(2 \cdot 2^n - 2)}{6} + 2$$

$$= \frac{-2}{2^{3n}} \frac{2^n(2^n - 1)(2 \cdot 2^n - 2)}{6} + 2$$

$$
\begin{aligned}
\lim_{n\to\infty} \int_{\mathbb{R}} f_n \,\mathrm{d}\lambda &= \lim_{n\to\infty} \frac{-2}{2^{3n}} \frac{2^n(2^n-1)(2\cdot 2^n - 2)}{6} + 2 \\
&= \lim_{n\to\infty} \frac{-2}{2^{3n}} \frac{2\cdot 2^{3n} - 6\cdot 2^{2n} + 2\cdot 2^n}{6} + 2 \\
&= -2\cdot \frac{2}{6} + 2 = \frac{-4}{6} + 2 = \frac{4}{3} \\
&= \int_{\mathbb{R}} f \,\mathrm{d}\lambda
\end{aligned}
$$

b)

$$
\begin{aligned}
\int_0^\infty (1 - F(t))\mathrm{d}t &= \int_0^\infty (1 - \mathbb{P}(X \le t))\mathrm{d}t \\
&= \int_0^\infty \mathbb{P}(X > t)\mathrm{d}t \\
&= \int_0^\infty \mathbb{E}[\mathbf{1}_{[t,\infty)}(X)]\mathrm{d}t \\
&= \int_0^\infty \int_\Omega \mathbf{1}_{[t,\infty)}(X(\omega))\ \mathrm{d}\mathbb{P}\ \mathrm{d}t \\
&= \int_0^\infty \int_\Omega \mathbf{1}_{[0,X(\omega))}(t)\ \mathrm{d}\mathbb{P}\ \mathrm{d}t \\
&= \int_\Omega \int_0^\infty \mathbf{1}_{[0,X(\omega))}(t)\ \mathrm{d}t\ \mathrm{d}\mathbb{P} \\
&= \mathbb{E}[\int_0^X 1\ \mathrm{d}t] = \mathbb{E}[X]
\end{aligned}
$$

Bleibt noch zu zeigen, dass wir Fubini benutzen können. Dafür müssen wir zeigen, dass die Abbildung

$$(\omega, t) \mapsto \mathbf{1}_{[0, X(\omega))}(t)$$

$\mathcal{A} \otimes \mathcal{B}(\mathbb{R})$-messbar ist. Für Indikatorfunktionen reicht es aus die Messbarkeit der Indikatormenge zu zeigen. Dafür verweisen wir auf die Lösung von Aufgabe 1 auf Übungsblatt 10.

Mittagspause

Puuh jetzt haben wir uns alle mal eine Pause verdient.
Später kommt dann noch Kapitel 4, bis dahin: Guten Hunger ;)

Zufallsvariablen

Definition

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Dann heißt eine $\mathcal{A} - \mathcal{B}(\mathbb{R})$-messbare Abbildung

$$X\colon \Omega \longrightarrow \mathbb{R}, \quad \omega \mapsto X(\omega),$$

Zufallsvariable.

- Eine Zufallsvariable X ordnet also jedem Elementarereignis ω einen Wert/eine Auszahlung in $\mathbb{R}$ zu.
- Wenn $f : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ messbar ist, ist auch $f(X)$ eine Zufallsvariable, da die Verkettung messbarer Abbildungen messbar ist.
- X muss $\mathcal{A} - \mathcal{B}(\mathbb{R})$-messbar sein, damit wir die Wahrscheinlichkeit, dass X Werte in einer Menge $B \in \mathcal{B}(\mathbb{R})$ annimmt, bestimmen können. $\rightsquigarrow$ Verteilung

Verteilungen von Zufallsvariablen

Definition

Sei X eine Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Dann heißt

$$\mathbb{P}_X(B) := \mathbb{P}(X \in B) \overset{\text{Not.}}{=} \mathbb{P}(\{\omega \in \Omega : X(\omega) \in B\}) = \mathbb{P}(X^{-1}(B)), \quad B \in \mathcal{B}(\mathbb{R}),$$

Verteilung von X.

- Hier wird die Messbarkeit von X wichtig, da wir dadurch wissen, dass $X^{-1}(B) \in \mathcal{A}$ für alle $B \in \mathcal{B}(\mathbb{R})$ gilt und somit $\mathbb{P}_X(B)$ für alle $B \in \mathcal{B}(\mathbb{R})$ wohldefiniert ist.
- $\mathbb{P}_X$ ist der push-forward (das Bildmaß) von X.
- Nach der Vorlesung ist $\mathbb{P}_X$ ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.
  $\rightsquigarrow$ Verteilungsfunktionen und Erwartungswerte

Verteilungsfunktionen von Zufallsvariablen

Definition

Sei X eine Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Dann heißt

$$F_X \colon \mathbb{R} \longrightarrow [0,1], \quad t \mapsto F_X(t) := \mathbb{P}(X \leq t) = \mathbb{P}_X((-\infty, t]),$$

Verteilungsfunktion von X.

- Eine Zufallsvariable heißt absolutstetig mit Dichte f, wenn die Verteilungsfunktion $F_X(t)$ von X die Dichte f hat.
- Eine Zufallsvariable heißt diskret,wenn $F_X(t)$ eine diskrete Verteilungsfunktion ist.

Wichtige Rechenobjekte

Definition

Sei X eine Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. So heißen, falls die Integrale wohldefiniert sind:

$$\begin{aligned}
\mathbb{E}[X] &:= \int_\Omega X(\omega)\,\mathrm{d}\mathbb{P}(\omega) \text{ \textit{Erwartungswert} von } X,\\
\mathbb{E}[X^k] &:= \int_\Omega X^k(\omega)\,\mathrm{d}\mathbb{P}(\omega) \text{ \textit{ktes Moment} von } X,\\
\mathbb{V}[X] &:= \mathbb{E}[(X - \mathbb{E}[X])^2] \text{ \textit{Varianz} von } X,\\
\mathbb{E}[\mathrm{e}^{\lambda X}] &:= \int_\Omega \mathrm{e}^{\lambda X(\omega)}\,\mathrm{d}\mathbb{P}(\omega) \text{ \textit{exponentielles Moment} von } X.
\end{aligned}$$

Anwendung des Trafos um rechnen zu können

Satz (Transformationssatz)

Seien $(\Omega, \mathcal{A}), (\Omega', \mathcal{A}')$ messbare Räume, μ ein Maß auf $(\Omega, \mathcal{A})$, $f\colon \Omega \to \Omega'$ und $g\colon \Omega' \to \bar{\mathbb{R}}$. Dann ist g μ_f-integrierbar genau dann, wenn $g \circ f$ μ-integrierbar ist, und falls eine dieser Eigenschaften erfüllt ist gilt

$$\int_\Omega g \circ f \,\mathrm{d}\mu = \int_{\Omega'} g \,\mathrm{d}\mu_f .$$

Also gilt für eine Zufallsvariable X und eine messbare Funktion $g\colon \mathbb{R} \to \mathbb{R}$

$$\mathbb{E}[g(X)] \coloneqq \int_\Omega g(X(\omega)) \,\mathrm{d}\mathbb{P}(\omega) = \int_{\mathbb{R}} g(x) \,\mathrm{d}\mathbb{P}_X(x),$$

wenn $g \circ X$ $\mathbb{P}$-integrierbar oder f $\mathbb{P}_X$-integrierbar ist.

Es reicht also aus die Verteilung der Zufallsvariablen X zu kennen!

Beispiele

$$\begin{aligned} \mathbb{E}[X] &= \int_{\mathbb{R}} x \, \mathrm{d}\mathbb{P}_X(x), \\ \mathbb{V}[X] &= \mathbb{E}[(X - \mathbb{E}[X])^2] = \int_{\mathbb{R}} (x - \mathbb{E}[X])^2 \, \mathrm{d}\mathbb{P}_X(x), \\ \mathbb{E}[\mathrm{e}^{tX}] &= \int_{\mathbb{R}} \mathrm{e}^{tx} \, \mathrm{d}\mathbb{P}_X(x), \end{aligned}$$

falls die Integrale existieren.

- $\mathbb{P}_X$ ist (Danke Trafo!) ein W-Maß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, das heißt wir können jetzt wieder die üblichen Berechnungsregeln definieren, falls $\mathbb{P}_X$ eine Dichte hat oder eine Summe von Dirac-Maßen ist. (siehe nächste Slide)

Berechnungsregeln

Satz

Sei X eine Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ und $g : \mathbb{R} \to \bar{\mathbb{R}}$ messbar, so gelten:

i) *Ist X absolutstetig mit Dichte f, so gilt*

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}} g(x) f(x) dx$$

ii) *Ist X diskret und nimmt die Werte $a_1, \ldots, a_N \in \mathbb{R}$ mit Wahrscheinlichkeiten $p_1, \ldots, p_N$ an, so gilt*

$$\mathbb{E}[g(X)] = \sum_{k=1}^{N} g(a_k) p_k$$

Rechenregeln für den Erwartungswert

Satz

Seien X, Y Zufallsvariablen auf $(\Omega, \mathcal{A}, \mathbb{P})$ mit $\mathbb{E}[|X|], \mathbb{E}[|Y|] < \infty$ und $\alpha, \beta \in \mathbb{R}$, so gelten:

i) $\mathbb{E}[\alpha X + \beta Y] = \alpha \mathbb{E}[X] + \beta \mathbb{E}[Y]$

ii) $X \geq 0 \, \mathbb{P} f.s. \Rightarrow \mathbb{E}[X] \geq 0$ *und* $X \geq Y \, \mathbb{P} f.s. \Rightarrow \mathbb{E}[X] \geq \mathbb{E}[Y]$

iii) *Ist* $X = \alpha \, \mathbb{P} - f.s.$, *d.h.* $\mathbb{P}(X = \alpha) = 1$, *so ist* $\mathbb{E}[X] = \alpha$

iv) $\mathbb{P}(X \in A) = \mathbb{E}[\mathbb{1}_A(X)]$, *insbesondere gilt* $F_X(t) = \mathbb{E}[\mathbb{1}_{(-\infty, t]}(X)], t \in \mathbb{R}$

Momenterzeugende Funktion von Zufallsvariablen

Definition

Sei X eine Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Dann heißt

$$\mathcal{M}_X \colon D \mapsto [0, \infty), \quad t \mapsto \mathcal{M}_X(t) := \mathbb{E}[\mathrm{e}^{tX}]$$

momenterzeugende Funktion von X, wobei

$$D := \{t \in \mathbb{R} : \mathbb{E}[\mathrm{e}^{tX}] < \infty\}.$$

Momenterzeugende Funktion von Zufallsvariablen

Definition

Sei X eine Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Dann heißt

$$\mathcal{M}_X \colon D \mapsto [0, \infty), \quad t \mapsto \mathcal{M}_X(t) := \mathbb{E}[\mathrm{e}^{tX}]$$

momenterzeugende Funktion von X, wobei

$$D := \{t \in \mathbb{R} : \mathbb{E}[\mathrm{e}^{tX}] < \infty\}.$$

- Es gilt immer $D \neq \emptyset$, da $\mathcal{M}_X(0) = \mathbb{E}[\mathrm{e}^{0X}] = \mathbb{E}[1] = 1 < \infty$.
- Falls ein $\epsilon > 0$ existiert, mit $(-\epsilon, \epsilon) \subseteq D$, so gilt nach der Vorlesung

  $$\mathbb{E}[X^k] = \mathcal{M}_X^{(k)}(0)$$

  wobei $\mathcal{M}_X^{(k)}(0)$ die k–te Ableitung an der Stelle 0 ist.
  $\rightsquigarrow$ Deshalb momenterzeugende Funktion.

Markov- und Tschebycheffungleichung

Satz (Markov- und Tschebycheffungleichung)

Sei X eine Zufallsvariable, dann gelten für $a > 0$ folgende Ungleichungen

i) Für $h : \mathbb{R} \longrightarrow (0, \infty)$ wachsend gilt

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[h(X)]}{h(a)} \qquad \text{(Markovungleichung)}$$

ii) Für $h : [0, \infty) \longrightarrow (0, \infty)$ wachsend gilt

$$\mathbb{P}(|X| \geq a) \leq \frac{\mathbb{E}[h(|X|)]}{h(a)} \qquad \text{(Markovungleichung)}$$

iii)

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq a) \leq \frac{\mathbb{V}[X]}{a^2} \qquad \text{(Tschebycheffungleichung)}$$

Aufgabenblock 4

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum auf dem die Zufallsvariablen

$$X \sim \mathrm{Poi}(\lambda) \qquad Y \sim \mathrm{Bin}(n, p) \qquad Z \sim \mathcal{U}([0, 1])$$

mit $\lambda > 0$, $p \in [0, 1]$, $n \in \mathbb{N}$ gegeben seien.

a) Berechne den Erwartungswert der Zufallsvariable $aX^2 + b$, $a, b \in \mathbb{R}$.

b) Berechne die momenterzeugende Funktion und die Varianz von Y.

c) Berechne die Dichtefunktion der Zufallsvariable $-Z$.

Lösung Aufgabenblock 4

a) Da $X \sim \text{Poi}(\lambda)$ gilt $p_k = e^{-\lambda}\frac{\lambda^k}{k!}$. Wir berechnen zunächst das zweite Moment von X mithilfe von Indexverschiebungen und Linearität:

$$\begin{aligned}
\mathbb{E}[X^2] &= \sum_{k=0}^{\infty} k^2 \cdot e^{-\lambda}\frac{\lambda^k}{k!} \\
&= \sum_{k=1}^{\infty} k \cdot e^{-\lambda}\frac{\lambda^k}{(k-1)!} \\
&= \sum_{k=0}^{\infty} (k+1) \cdot e^{-\lambda}\frac{\lambda^{(k+1)}}{k!} \\
&= \lambda \sum_{k=0}^{\infty} k \cdot e^{-\lambda}\frac{\lambda^k}{k!} + \lambda \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!}
\end{aligned}$$

Lösung Aufgabenblock 4

$$
\begin{aligned}
&= \lambda \mathbb{E}[X] + \lambda \sum_{k=0}^{\infty} p_k \\
&= \lambda^2 + \lambda,
\end{aligned}
$$

wobei $\mathbb{E}[X] = \lambda$ gerne nochmal gecheckt werden darf wer uns nicht glaubt.
Also gilt für $a, b \in \mathbb{R}$ mit Linearität des Erwartungswerts:

$$\mathbb{E}[aX^2 + b] = a \cdot \mathbb{E}[X^2] + b = a \cdot (\lambda^2 + \lambda) + b.$$

Lösung Aufgabenblock 4

b) Da $Y \sim \mathrm{Bin}(n, p)$ gilt, dass $\mathcal{M}_Y(t)\ \forall t \in \mathbb{R}$ definiert ist, da Y nur Masse auf $\{0, \ldots, n\}$ hat.
Durch den Binomischen Lehrsatz und da für $Y \sim \mathrm{Bin}(n, p)$ $p_k = \binom{n}{k} p^k (1-p)^{n-k}$, erhalten wir:

$$\begin{aligned}\mathcal{M}_Y(t) = \mathbb{E}[e^{tY}] &= \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=0}^{n} \binom{n}{k} (pe^t)^k (1-p)^{n-k} \\ &= (pe^t + (1-p))^n.\end{aligned}$$

Lösung Aufgabenblock 4

Mit dem Satz aus der Vorlesung gilt nun $\mathbb{E}[Y^k] = \mathcal{M}_Y^{(k)}(0)$.

Zweimaliges Ableiten der Momenterzeugenden Funktion gibt:

$$\mathcal{M}_Y^{(1)}(t) = n(pe^t + (1-p))^{n-1} \cdot pe^t$$

$$\mathcal{M}_Y^{(2)}(t) = (n-1)n(pe^t + (1-p))^{n-2} \cdot pe^t \cdot pe^t + n(pe^t + (1-p))^{n-1} \cdot pe^t$$

Lösung Aufgabenblock 4

Mit dem Satz aus der Vorlesung gilt nun $\mathbb{E}[Y^k] = \mathcal{M}_Y^{(k)}(0)$.
Zweimaliges Ableiten der Momenterzeugenden Funktion gibt:

$$\mathcal{M}_Y^{(1)}(t) = n(pe^t + (1-p))^{n-1} \cdot pe^t$$
$$\mathcal{M}_Y^{(2)}(t) = (n-1)n(pe^t + (1-p))^{n-2} \cdot pe^t \cdot pe^t + n(pe^t + (1-p))^{n-1} \cdot pe^t$$

Ausgewertet an der Stelle $t = 0$ gibt:

$$\begin{aligned} \mathbb{E}[Y] &= \mathcal{M}_Y^{(1)}(0) = n(1)^{n-1} \cdot p = np \\ \mathbb{E}[Y^2] &= \mathcal{M}_Y^{(2)}(0) = (n-1)n(1)^{n-2} \cdot p^2 + np \\ &= n^2p^2 - np^2 + np \end{aligned}$$

Also $\mathbb{V}(Y) = \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 = np - np^2 = np(1-p)$.

c) Sei $t \in \mathbb{R}$. Dann gilt da $Z \sim \mathcal{U}([0,1])$

$$\mathbb{P}(-Z \leq t) = \int_{-\infty}^{t} \mathbb{1}_{[0,1]}(-z)\mathrm{d}z$$

Da $0 \leq -z \leq 1 \Leftrightarrow 0 \geq z \geq -1$ gilt

$$\mathbb{P}(-Z \leq t) = \int_{-\infty}^{t} \mathbb{1}_{[0,1]}(-z)\mathrm{d}z = \int_{-\infty}^{t} \mathbb{1}_{[-1,0]}(z)\mathrm{d}z$$

Damit gilt $-Z \sim \mathcal{U}([-1,0])$ mit Dichte $f(z) = \mathbb{1}_{[-1,0]}(z)$.

Zufallsvektoren

Definition

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

$$X\colon \Omega \longrightarrow \mathbb{R}^d, \quad \omega \mapsto X(\omega) = \begin{pmatrix} X_1(\omega) \\ X_2(\omega) \\ \vdots \\ X_d(\omega) \end{pmatrix},$$

heißt Zufallsvektor, falls sie $\mathcal{A}$-$\mathcal{B}(\mathbb{R}^d)$-messbar ist.

- X ist genau dann ein Zufallsvektor, wenn alle $X_i\colon \Omega \longrightarrow \mathbb{R}$ Zufallsvariablen sind, also wenn sie $\mathcal{A}$-$\mathcal{B}(\mathbb{R})$-messbar sind.

Verteilungen von Zufallsvektoren

Definition

Das Maß

$$\mathbb{P}_X(B) := \mathbb{P}(X^{-1}(B)), \quad B \in \mathcal{B}(\mathbb{R}^d),$$

auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ heißt Verteilung des Zufallsvektors X.

- Um die Verteilung der Zufallsvariablen X_i, $i \in \{1, \ldots, d\}$, zu berechnen nutzen wir folgende Gleichung

$$\mathbb{P}(X_i \in B) = \mathbb{P}(X_1 \in \mathbb{R}, X_2 \in \mathbb{R}, \ldots, X_i \in B, \ldots, X_d \in \mathbb{R}) = \mathbb{P}_X(\mathbb{R} \times \ldots \times B \times \ldots \times \mathbb{R}), \quad B \in \mathcal{B}(\mathbb{R}),$$

man nennt $\mathbb{P}_{X_i}$ dann auch eindimensionale Randverteilung von X_i.

Definition

Seien $X_1, ..., X_d$ Zufallsvariablen auf $(\Omega, , \mathbb{P})$.

i) $X_1, ..., X_d$ haben die **gemeinsame Dichte** f, falls die gemeinsame Verteilungsfunktion F absolutstetig ist und Dichte f hat.

ii) $X_1, ..., X_d$ heißen **diskret**, falls $a_1, ..., a_N \in \mathbb{R}^d$ existieren mit

$$\mathbb{P}(X = a_k) = \mathbb{P}(X_1 = a_{k,1}, ..., X_d = a_{k,d}) = p_k$$

und $\sum_{k=1}^{N} p_k = 1$ für ein $N \in \mathbb{N} \cup \{+\infty\}$.

Gemeinsame Dichten

Satz

Seien $X_1, \dots, X_d$ Zufallsvariablen auf $(\Omega, \mathcal{A}, \mathbb{P})$.

(i) *Haben $X_1, \dots, X_d$ die gemeinsame Dichte f, so haben $X_1, \dots, X_d$ Dichten $f_1, \dots, f_d$ und es gilt*

$$f_i(x) = \underbrace{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty}}_{(d-1)\text{-viele}} \underbrace{f(x_1, \dots, x_d)}_{x_i \text{ fest}} \underbrace{x_1 \dots x_d}_{\text{ohne } x_i}, \qquad x \in \mathbb{R},$$

ist eine Dichte von X_i für $i = 1, \dots, d$. In Worten: Ist X absolutstetig, so sind alle X_i absolutstetig und die Dichten der X_i entstehen durch Ausintegrieren aller anderen Variablen.

(ii) *Die Rückrichtung gilt im Allgemeinen nicht. Es gibt also absolutstetige Zufallsvariablen, die keine gemeinsame Dichte haben.*

Unabhängige Zufallsvariablen

Definition

Seien $X_1, \ldots, X_d$ Zufallsvariablen auf $(\Omega, , \mathbb{P})$.

i) $X_1, \ldots, X_d$ heißen **unabhängig**, falls die gemeinsame Verteilungsfunktion in die Randverteilungsfunktionen faktorisiert, d. h.

$$F_X(t_1, \ldots, t_d) = F_{X_1}(t_1) \cdot \ldots \cdot F_{X_d}(t_d), \quad t_i \in \mathbb{R}$$

oder mit Wahrscheinlichkeiten geschrieben

$$\mathbb{P}(X_1 \leq t_1, \ldots, X_d \leq t_d) = \mathbb{P}(X_1 \leq t_1) \cdot \ldots \cdot \mathbb{P}(X_d \leq t_d), \quad t_i \in \mathbb{R}.$$

ii) $X_1, \ldots, X_d$ heißen **abhängig**, falls sie nicht unabhängig sind.

iii) $X_1, \ldots, X_d$ heißen **unabhängig und identisch verteilt (u.i.v.)**, falls sie unabhängig und identisch verteilt $(F_{X_1} = \ldots = F_{X_d})$ sind. Weil die gemeinsame Verteilungsfunktion F bei u.i.v. Zufallsvariablen schon eindeutig durch jede Randverteilungsfunktion festgelegt ist, gibt man oft nur die Verteilung von X_1 an.

Gemeinsame Dichten und Unabhängigkeit

Satz

Sind $X_1, \dots, X_d$ Zufallsvariablen mit gemeinsamer Dichte f, dann gilt:

$$X_1, \dots, X_d \text{ sind unabhängig} \quad \Leftrightarrow \quad f(x) = f_1(x_1) \cdot \ldots \cdot f_d(x_d) \quad \text{Lebesgue-fast überall,}$$

wobei $f_1, \dots, f_d$ Dichten von $X_1, \dots, X_d$ sind.

Rechenregeln mit gemeinsamen Dichten

Satz

Sind $X_1, \dots, X_d$ Zufallsvariablen auf $(\Omega, , \mathbb{P})$, so gelten:

i) *Für $A \in \mathcal{B}(\mathbb{R}^d)$ gilt $\mathbb{E}[\mathbf{1}_A(X)] = \mathbb{P}(X \in A)$.*

ii) *Haben $X_1, \dots, X_d$ eine gemeinsame Dichte f, so gilt*

$$\mathbb{E}[g(X_1, \dots, X_d)] = \int_{\mathbb{R}^d} g(x_1, \dots, x_d) f(x_1, \dots, x_d)(x_1, \dots, x_d).$$

iii) *Sind $X_1, \dots, X_d$ diskret und nimmt der Zufallsvektor $X = (X_1, \dots, X_d)$ die Vektoren $a_1, \dots, a_N \in \mathbb{R}^d$ mit Wahrscheinlichkeiten $p_1, \dots, p_N$ an, so gilt*

$$\mathbb{E}[g(X_1, \dots, X_d)] = \sum_{k=1}^{N} p_k g(a_k) = \sum_{k=1}^{N} \mathbb{P}(X = a_k) g(a_k).$$

Wie für $d = 1$ gilt in (ii) und (iii), dass die Erwartungswerte wohldefiniert sind (oder existieren), genau dann, wenn die Integrale bzw. Summen wohldefiniert (oder endlich) sind.

Rechenregeln für unabhängige Zufallsvariablen I

Satz

Sind $X_1, \dots, X_d$ *unabhängige Zufallsvariablen auf* $(\Omega, , \mathbb{P})$, *so gilt*

$$\mathbb{E}[g_1(X_1) \cdot \ldots \cdot g_d(X_d)] = \mathbb{E}[g_1(X_1)] \cdot \ldots \cdot \mathbb{E}[g_d(X_d)]$$

für alle messbaren $g_1, \dots, g_d : \mathbb{R} \longrightarrow \overline{\mathbb{R}}$. *Insbesondere gilt auch*

$$\mathbb{P}(X_1 \in A_1, \dots, X_d \in A_d) = \mathbb{E}[\mathbf{1}_{A_1}(X_1)] \cdot \ldots \cdot \mathbb{E}[\mathbf{1}_{A_d}(X_d)] = \mathbb{P}(X_1 \in A_1) \cdot \ldots \cdot \mathbb{P}(X_d \in A_d)$$

für alle $A_1, \dots, A_d \in \mathcal{B}(\mathbb{R})$.

Rechenregeln für unabhängige Zufallsvariablen II

Satz

Sind $X_1, \dots, X_d$ Zufallsvariablen auf $(\Omega, , \mathbb{P})$, so gelten:

i) *Für $A \in \mathcal{B}(\mathbb{R}^d)$ gilt $\mathbb{E}[\mathbf{1}_A(X)] = \mathbb{P}(X \in A)$.*

ii) *Haben $X_1, \dots, X_d$ eine gemeinsame Dichte f, so gilt*

$$\mathbb{E}[g(X_1, \dots, X_d)] = \int_{\mathbb{R}^d} g(x_1, \dots, x_d) f(x_1, \dots, x_d)(x_1, \dots, x_d).$$

iii) *Sind $X_1, \dots, X_d$ diskret und nimmt der Zufallsvektor $X = (X_1, \dots, X_d)$ die Vektoren $a_1, \dots, a_N \in \mathbb{R}^d$ mit Wahrscheinlichkeiten $p_1, \dots, p_N$ an, so gilt*

$$\mathbb{E}[g(X_1, \dots, X_d)] = \sum_{k=1}^{N} p_k g(a_k) = \sum_{k=1}^{N} \mathbb{P}(X = a_k) g(a_k).$$

Wie für $d = 1$ gilt in (ii) und (iii), dass die Erwartungswerte wohldefiniert sind (oder existieren), genau dann, wenn die Integrale bzw. Summen wohldefiniert (oder endlich) sind.

Faltung diskreter und stetiger Verteilungen

Für diskrete Verteilungen braucht man die Faltung eigentlich überhaupt nicht! Seien X_1, X_2 unabhängige, diskrete Zufallsvariablen. Dann gilt

$$\mathbb{P}(X_1 + X_2 = a) = \sum_{b \in X_2(\Omega)} \mathbb{P}(X_1 = a - b)\mathbb{P}(X_2 = b).$$

Für Summen unabhängiger stetiger Zufallsvariablen braucht man hingegen die Faltung. Diese liefert dann eine Regel für die Dichte der Zufallsvariable $X_1 + X_2$, wenn X_1, X_2 unabhängige, stetige Zufallsvariablen, denn nach der Vorlesung gilt für diese

$$f_{X+Y}(x) = \int_{\mathbb{R}} f_X(x-y) f_Y(y) \,\mathrm{d}y.$$

Aufgabenblock 5

a) Seien $X \sim \mathrm{Poi}(\lambda)$, $Y \sim \mathrm{Poi}(\beta)$ mit $\lambda, \beta > 0$ unabhängige Zufallsvariablen. Zeige, dass $X + Y \sim \mathrm{Poi}(\lambda + \beta)$ gilt.

b) Seien X, Y unabhängige Zufallsvariablen mit $\mathcal{M}_X < \infty, \mathcal{M}_Y(t) < \infty$ für ein $t \in \mathbb{R}$. Zeige, dass $\mathcal{M}_{X+Y}(t) = \mathcal{M}_X(t) \cdot \mathcal{M}_Y(t)$ gilt.

Lösung Aufgabenblock 5

a) Mit der diskreten Faltungsformel

$$
\begin{aligned}
\mathbb{P}(X+Y=n) &= \sum_{k\in\mathbb{N}_0} \mathbb{P}(X=n-k)\cdot\mathbb{P}(Y=k) \\
&= \sum_{k=0}^{n} \mathbb{P}(X=n-k)\cdot\mathbb{P}(Y=k) \\
&= \sum_{k=0}^{n} e^{-\lambda}\frac{\lambda^{(n-k)}}{(n-k)!}\cdot e^{-\beta}\frac{\beta^k}{k!} \\
&= e^{-\lambda}\cdot e^{-\beta}\sum_{k=0}^{n} \frac{\lambda^{(n-k)}}{(n-k)!}\cdot\frac{\beta^k}{k!} \\
&= e^{-(\lambda+\beta)}\sum_{k=0}^{n} \frac{\lambda^{(n-k)}}{(n-k)!}\cdot\frac{\beta^k}{k!}
\end{aligned}
$$

Lösung Aufgabenblock 5

$$
\begin{aligned}
e^{-(\lambda+\beta)} \sum_{k=0}^{n} \frac{\lambda^{(n-k)}}{(n-k)!} \cdot \frac{\beta^k}{k!} &= e^{-(\lambda+\beta)} \frac{n!}{n!} \sum_{k=0}^{n} \frac{\lambda^{(n-k)}}{(n-k)!} \cdot \frac{\beta^k}{k!} \\
&= e^{-(\lambda+\beta)} \frac{1}{n!} \sum_{k=0}^{n} \frac{n!}{(n-k)!k!} \lambda^{(n-k)} \cdot \beta^k \\
&= e^{-(\lambda+\beta)} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} \lambda^{(n-k)} \cdot \beta^k \\
&= e^{-(\lambda+\beta)} \frac{(\lambda+\beta)^n}{n!}
\end{aligned}
$$

Also gilt $\mathbb{P}(X+Y=n) = \mathbb{P}(Z=n)$ für $Z \sim \mathrm{Poi}(\lambda+\beta)$ für alle $n \in \mathbb{N}$ und damit $X+Y \sim \mathrm{Poi}(\lambda+\beta)$.

Lösung Aufgabenblock 5

b) Da die Erwartungswerte unabhängiger Zufallsvariablen faktorisieren gilt für alle $t \in \mathbb{R}$

$$\mathcal{M}_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX} \cdot e^{tY}] = \mathbb{E}[e^{tX}]\mathbb{E}[e^{tY}] = \mathcal{M}_X(t) \cdot \mathcal{M}_Y(t)$$

Definition

Sei X eine Zufallsvariable und $(X_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Dann konvergiert $(X_n)_{n\in\mathbb{N}}$ $(\mathbb{P})$-fast sicher gegen X, falls

$$\mathbb{P}\left(\lim_{n\to\infty} X_n = X\right) = \mathbb{P}\left(\left\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}\right) = 1$$

gilt. Man schreibt

$$X_n \xrightarrow[n\to\infty]{\text{f.s.}} X.$$

- Die fast sichere Konvergenz von Zufallsvariablen ist dazu äquivalent, dass eine $\mathbb{P}$-Nullmenge $N \in \mathcal{A}$ existiert, sodass für alle $\omega \in N^C$

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

gilt.

Konvergenz von Zufallsvariablen

Definition

Sei X eine Zufallsvariable und $(X_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ mit $\mathbb{E}[|X|^p] < \infty, \mathbb{E}[|X_n|^p] < \infty, \forall n \in \mathbb{N}$, für ein $p \geq 1$. Dann konvergiert $(X_n)_{n\in\mathbb{N}}$ im p-ten Mittel (oder in $\mathcal{L}^p$) gegen X, falls

$$\lim_{n\to\infty} \mathbb{E}[|X_n - X|^p] = 0$$

gilt. Man schreibt

$$X_n \xrightarrow[n\to\infty]{\mathcal{L}^p} X.$$

- Ihr kennt die $\mathcal{L}^p$-Konvergenz schon aus dem Kapitel zu $\mathcal{L}^p$-Räumen, nur wird sie hierbei explizit auf Zufallsvariablen angewendet (denkt daran Erwartungswerte sind Integrale!).

Konvergenz von Zufallsvariablen

Definition

Sei X eine Zufallsvariable und $(X_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Dann konvergiert $(X_n)_{n\in\mathbb{N}}$ stochastisch (oder in Wahrscheinlichkeit) gegen X, falls für alle $\epsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \epsilon) = 0$$

gilt. Man schreibt

$$X_n \xrightarrow[n\to\infty]{P} X.$$

- Man folgert stochastische Konvergenz häufig direkt aus der Konvergenz im p-ten Mittel mit der Markov-Ungleichung (oder Spezialfall Chebyshev-Ungleichung).

Konvergenz von Zufallsvariablen

Definition

Sei X eine Zufallsvariable und $(X_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Dann konvergiert $(X_n)_{n\in\mathbb{N}}$ in Verteilung gegen X, falls für alle stetigen, beschränkten, reell-wertigen Funktionen f (oder kurz $\forall f \in C_b(\mathbb{R})$)

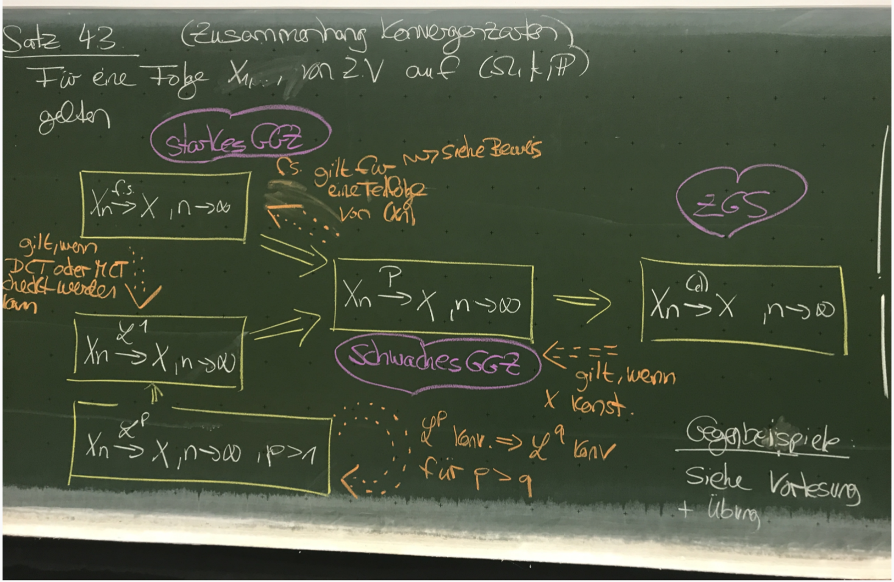
$$\lim_{n\to\infty} \mathbb{E}[f(X_n)] = \mathbb{E}[f(X)]$$

gilt. Man schreibt

$$X_n \xrightarrow[n\to\infty]{(d)} X.$$

- Die Konvergenz in Verteilung ist die schwächste (hier behandelte) Art der Konvergenz von Zufallsvariablen.

Bildchen (Konvergenzdiagramm)



Schwaches GGZ

Satz

i) *Klassische Variante: Sind* $X_1, X_2, \ldots$ *u.i.v. mit* $\mathbb{E}[X_1^2] < \infty$*, so gilt*

$$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow[n\to\infty]{(P)} \mathbb{E}[X_1].$$

ii) *Variante mit schwächeren Annahmen: Sind* $X_1, X_2, \ldots$ *paarweise unkorreliert mit identischem Erwartungswert und*

$$\frac{1}{n^2}\sum_{k=1}^{n} \mathbb{V}(X_k) \xrightarrow[n\to\infty]{} 0,$$

so gilt

$$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow[n\to\infty]{(P)} \mathbb{E}[X_1].$$

Starkes GGZ

Satz

Sind $X_1, X_2, \dots$ u.i.v. mit $\mathbb{E}[|X_1|] < \infty$, so gilt

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{f.s.} \mathbb{E}[X_1].$$

Starkes GGZ

Satz

Sind $X_1, X_2, \ldots$ u.i.v. mit $\mathbb{E}[|X_1|] < \infty$, so gilt

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{f.s.} \mathbb{E}[X_1].$$

- Teil (i) im schwachen GGZ wurde nur aus didaktischen Gründen behandelt, denn diese Aussage folgt direkt aus dem starken Gesetz der großen Zahlen.
- Teil (ii) im schwachen GGZ ist deutlich interessanter, denn unter diesen Annahmen muss das starke GGZ nicht gelten!

Aufgabenblock 6

a) Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein WRaum auf dem die Folgen von Zufallsvariablen $(X_n)_{n\in\mathbb{N}}, (Y_n)_{n\in\mathbb{N}}$ und eine Zufallsvariable Z gegeben seien mit

$$X_n \sim \operatorname{Exp}(n), \quad \mathbb{E}[Y_n] = \frac{1}{\sqrt{n}}, \quad \mathbb{V}[Y_n] = \frac{\sigma^2}{n},$$

für $\sigma > 0$. Zeige $X_n \xrightarrow[n\to\infty]{P} 0$ und $Y_n + Z \xrightarrow[n\to\infty]{P} Z$.

b) Seien $X_1, X_2, \ldots$ paarweise unkorreliert mit identischem Erwartungswert und gelte $\mathbb{V}(X_k) \leq c \, \forall k \in \mathbb{N}, c > 0$.
Zeige oder Widerlege: Das schwache GGZ gilt immer noch!

c) Zeige: Seien $X_1, X_2, \ldots$ u.i.v., so gilt für $A \in \mathcal{B}(\mathbb{R})$

$$\frac{1}{n}\#\{i \leq n : X_i \in A\} \xrightarrow[n\to\infty]{f.s.} \mathbb{P}(X_1 \in A).$$

Lösung Aufgabenblock 6

a) Sei $\epsilon > 0$. Dann gilt für $n \in \mathbb{N}$

$$\begin{aligned}
\mathbb{P}(|X_n - 0| > \epsilon) &= \mathbb{P}(X_n > \epsilon) \\
&= 1 - \mathbb{P}(X_n \leq \epsilon) \\
&= 1 - F_{X_n}(\epsilon) \\
&= 1 - (1 - e^{-n\epsilon}) \\
&= e^{-n\epsilon},
\end{aligned}$$

und somit folgt

$$\lim_{n \longrightarrow \infty} \mathbb{P}(|X_n - 0| > \epsilon) = 0,$$

für alle $\epsilon > 0$, also per Definition

$$X_n \xrightarrow[n \longrightarrow \infty]{P} 0$$

Lösung Aufgabenblock 6

Sei $\epsilon > 0$. Dann gilt für $n \in \mathbb{N}$ mit der Markov-Ungleichung und der Verschiebungsformel

$$\begin{aligned}
\mathbb{P}(|Z + Y_n - Z| > \epsilon) &= \mathbb{P}(|Y_n| > \epsilon) \\
&\leq \frac{\mathbb{E}[|Y_n|^2]}{\epsilon^2} \\
&= \frac{\mathbb{V}(Y_n) + \mathbb{E}[Y_n]^2}{\epsilon^2} \\
&= \frac{\sigma^2 + 1}{n\epsilon^2}
\end{aligned}$$

und somit folgt $\lim_{n \to \infty} \mathbb{P}(|Z + Y_n - Z| > \epsilon) = 0$, also per Definition

$$Z + Y_n \xrightarrow[n \to \infty]{P} Z$$

Lösung Aufgabenblock 6

b) Sei $\epsilon > 0$ beliebig.
O.B.d.A sei $\mathbb{E}[X_1]$ der Erwartungswert aller X_k, sodass
$\mathbb{E}[\frac{1}{n}\sum_{k=1}^n X_k - \mathbb{E}[X_1]] = 0, k = \{1, \ldots, n\}$.
Wie im Beweis des Schwachen GGZ machen wir erst Tschebychef, dann Bienaymé:

$$\begin{aligned}
\mathbb{P}(|\frac{1}{n}\sum_{k=1}^n X_k - \mathbb{E}[X_1]| \geq \epsilon) &= \mathbb{P}(|\sum_{k=1}^n (X_k - \mathbb{E}[X_1])| \geq n\epsilon) \\
&\leq \frac{\mathbb{V}(\sum_{k=1}^n X_k - \mathbb{E}[X_1])}{n^2\epsilon^2} \\
&= \frac{\mathbb{V}(\sum_{k=1}^n X_k)}{n^2\epsilon^2}
\end{aligned}$$

Lösung Aufgabenblock 6

$$= \frac{\sum_{k=1}^{n} \mathbb{V}(X_k)}{n^2\epsilon^2}$$
$$\leq \frac{\sum_{k=1}^{n} c}{n^2\epsilon^2} = \frac{c}{n\epsilon^2} \longrightarrow 0, n \longrightarrow \infty,$$

wobei wir im letzten Schritt die Voraussetzung angewandt haben.
Wir sehen also: Auch unter der Annahme, dass die Varianzen lediglich beschränkt sind, gilt das schwache GGZ noch!

Lösung Aufgabenblock 6

c) Für $i \in \mathbb{N}$ definieren wir

$$Y_i := \mathbb{1}_A(X_i)$$

Die Y_i sind u.i.v nach Korollar 4.2.28 und es gilt

$$\mathbb{E}[Y_1] = \mathbb{E}[\mathbb{1}_A(X_1)] = \mathbb{P}(X_1 \in A) \leq 1 < \infty$$

Also kann das starke Gesetz der großen Zahlen angewandt werden und es gilt

$$\frac{1}{n}\#\{i \leq n : X_i \in A\} = \frac{1}{n}\sum_{i=1}^{n} Y_i \xrightarrow[n\to\infty]{f.s.} \mathbb{E}[Y_1] = \mathbb{P}(X_1 \in A).$$

Die gezeigte Aussage heißt übrigens Empirisches Gesetz der großen Zahlen.

Danke und Viel Erfolg!

- Juhu, Wir haben´s geschafft für heute!
- Wir hoffen, Ihr konntet etwas von heute und der Übung generell mitnehmen und hattet so wie wir auch Spaß dran:)
- Ganz viel Erfolg für die Klausur am Freitag! Daumen für Euch alle sind gedrückt!

Eure Übungskeks Leo und Martin! (Wer wima-memes abonniert hat versteht´s)