Stochastik

Dies ist eine Live-Transkription der Vorlesung Stochastik, gehalten von **Prof. Dr. Peter Pickl** im **Sommersemester 2024** an der Uni Tübingen. Dieser Mitschrieb ist fehlerbehaftet (`https://github.com/thomasxhua/stochastik-pickl-ss24-skript/`).

Inhaltsverzeichnis

[15.04.2024]

In der Stochastik geht es um die Modellierung von Experimenten, deren Ausgang vom Zufall abhängt.

I Endliche Wahrscheinlichkeitsräume

I.1 Ergebnismenge, Ereignismenge

I.1.1 Definition (Ergebnismenge)

Die Menge Ω, welche die möglichen Ausgänge eines Zufallexperiementes beschreibt, dennen wir *Grundmenge* oder *Ergebnismenge.*

I.1.2 Definition (Ereignismenge)

Die Potenzmenge $\mathcal{P}(\Omega)$, d.h. die Menge aller Teilmengen Ω, nennen wir *Ereignismenge.*

I.1.3 Beispiel

1) Wir werfen einen Würfel.
    - $\Omega = \{1, 2, \dots, 6\}$,
    - $\mathcal{P}(\Omega) = \{\emptyset, \Omega, \{1\}, \dots, \{1, 1\}, \dots\}$,
2) Glücksrad: $\Omega = [0, 2\pi[$ beschreibt die möglichen Winkel eines Glückradspiels. $\mathcal{P}(\Omega)$ ist klar (keine geeignete Ereignismenge, siehe Kapitel 2).

I.2 Wahrscheinlichkeitsmaß

Das Wahrscheinlichkeitsmaß wird auf der Ereignismenge definiert. Grund: Für überabzählbare Mengen (Glücksrad z.B.) haben einzelne Ausgänge häufig Wahrscheinlichkeit 0, obwohl global gesehen existiert ein sinnvolles Wahrscheinlichkeitsmaß (siehe Kapitel 2).

Die Wahrscheinlichkeit quantifiziert die Plausibilität der entsprechenden Ereignisse. Sie gibt die relative Häufigkeit an, wie oft ein bestimmtes Ereignis nach sehr häufigen Wiederholen unter identischen Umständen eintritt.

I.2.1 Definition (Wahrscheinlichkeitsmaß)

Eine Abbildung $\mathbb{P} \colon \mathcal{P}(\Omega) \to \mathbb{R}$ nennt man *Wahrscheinlichkeitsmaß* $:\Leftrightarrow$

Ka) $\mathbb{P}(A) \geq 0, \ \forall A \subset \Omega$,

Kb) $\mathbb{P}(\Omega) = 1$,

Kc) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B), \ \forall A, B \subset \Omega, \ A \cap B = \emptyset$.

I.2.2 Bemerkung

Die Axiome a)-c) nennt man *Axiome von Kolmogorov* (werden in 2 ebenfalls leicht angepasst).

I.2.3 Satz

Für jedes Wahrscheinlichkeitsmaß $\mathbb{P}\colon \mathcal{P}(\Omega) \to \mathbb{R}$ für beliebiges Ω gelten:

a) $\mathbb{P}(A^C) = 1 - \mathbb{P}(A),\ \forall A \subset \Omega$.

b) $\mathbb{P}(A) \leq 1$.

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

I.2.4 Beweis

(weggelassen)

I.2.5 Bemerkung

Über die Wahrscheinlichkeiten der Elementarereignisse (d.h. der einelementigen Ereignisse), wird das Wahrscheinlichkeitsmaß eindeutig festgelegt.

Betrachte $\mathbb{P}(A)$ für $A = \{w_1, w_2, \ldots, w_k\} \subset \Omega$. Durch mehrmaliges Anwenden von **K**c) erhält man

$$\mathbb{P}(A) = \sum_{i=1}^{k} \mathbb{P}(\{w_k\}).$$

I.2.6 Definition (Wahrscheinlichkeitsraum)

Das Paar $(\Omega, \mathbb{P})$ nennt man auch *Wahrscheinlichkeitsraum* (auch $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$).

——————— [18.04.2024] ———————

I.2.7 Bemerkung

Wie findet man nun das richtige Wahrscheinlichkeitsmaß, d.h. jenes, welches zu meinem Experiment passt?

1) Ausprobieren (siehe unten, "Statistik").

2) Analyse der physikalischen Eigenschaften. Praktikabel, falls Symmetrie in den relevanten physikalischen Eigenschaften herrscht:

I.2.8 Laplace-Annahme (Indifferenzprinzip)

Falls es keinen Grund zur Annahme gibt, dass die verschiedenen Ausgänge des Experiments im Wesentlichen zu unterscheiden sind, nehmen wir an, dass die Wahrscheinlichkeiten aller Elementarerignisse gleich sind.

I.2.9 Folgerung

Sei Ω ein Ergebnisraum. Unter der Laplace-Annahme gilt:

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

I.2.10 Beweis

$$1 \overset{\mathbf{K}}{=} \mathbb{P}(\Omega) = \mathbb{P}(\bigcup_{\omega\in\Omega}\{\omega\}) \overset{\mathbf{K}}{=} \sum_{\omega\in\Omega} \mathbb{P}(\{\omega\}) \overset{\mathbf{L}}{=} |\Omega| \cdot \mathbb{P}(\{\omega\})$$

$\Rightarrow \forall \omega \in \Omega$ gilt:

$$\mathbb{P}(\{w\}) = \frac{1}{|\Omega|}.$$

Außerdem:

$$\mathbb{P}(A) = \mathbb{P}(\bigcup_{\omega\in A}\{\omega\}) \overset{\mathbf{K}}{=} \sum_{\omega\in A} \mathbb{P}(\{\omega\}) = |A| \cdot \frac{1}{|\Omega|}.$$

I.2.11 Beispiel

1) Werfen eines ungezinkten Würfels:

$$\mathbb{P}(\{2,4,6\}) = \frac{3}{6} = \frac{1}{2}.$$

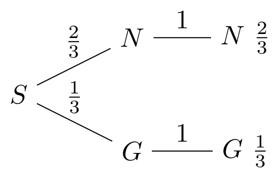
2) Gesamte Augenzahl bei zweimaligen Werfen des Würfels

$$\Omega = \{2,\ldots,12\},$$

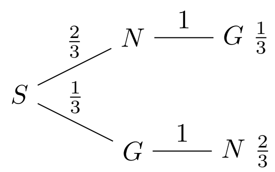
Laplace-Annahme gilt nicht. Die Elemente sind wesentlich verschieden, z.B. 2 hat nur Option $(1,1)$, 7 hat die Optionen $\{(1,6),\ldots\}$.

3) Ziegenproblem: Wir befinden uns in einer Gameshow, dürfen zwischen drei Toren wählen. Hinter einemal Tor ist ein Gewinn, hinter zweien eine Niete. Der Moderator öffnet eines der nicht-gewählten Tore. Hinter diesen ist eine Niete. Er bietet daraufhin an, das Tor zu wechseln. Ist der Wechsel sinnvoll?

   - Problem 1: Spielregeln müssen ergänzt werden. Wie handelt der Moderator? Wir gehen davon aus, dass er oder sie in jedem Fall ein nicht-gewähltes Tor mit Niete öffnet.
   - Problem 2: Man ist geneigt, von einer Laplace-Situation auszugehen. Dies ist falsch, da die Tore durch die Wahl und die Reaktion des Moderators zu unterscheiden sind.



Ohne Wechseln



Mit Wechseln

mit Start (S), Gewinn (G) und Niete (N).

I.3 Kombinatorik

Wie bestimmt man in einer Laplace-Situation $|A|$ und $|\Omega|$?

I.3.1 Beispiele

a) Sei $\Omega = A \times B$, so ist $|\Omega| = |A| \cdot |B|$ Münzwurf, dann Würfel:

$$A = \{\mathrm{K}, \mathrm{Z}\},\ B = \{1, \dots, 6\}.$$

b) Urne mit N durchnummerierten Kugeln. Wir ziehen davon nacheinander ohne Zurücklegen k Kugeln (Reihenfolge wird berücksichtigt):

$$|\Omega| = N(N-1)\dots(N-k+1) = \frac{N!}{(N-k)!}.$$

c) Zahlenlotto "6 aus 49", wie b) ohne Rücksicht auf Reihenfolge:

$$|\Omega| = \frac{N!}{(N-k)!} \cdot \frac{1}{k!}.$$

d) Anzahl der Anagramme von `MISSISSIPPI`.

$$|\Omega| = \frac{11!}{\underbrace{4!}_{\mathtt{I}}\,\underbrace{4!}_{\mathtt{S}}\,\underbrace{2!}_{\mathtt{P}}}.$$

II Allgemeine Wahrscheinlichkeitsräume

II.0 Anpassung des 3. Axioms von Kolmogorov

Wir werden das 3. Axiom von Kolmogorov anpassen:

$$\mathbb{P}(\bigcup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} \mathbb{P}(A_j), \text{ falls } A_j \cap A_k = \emptyset, \ \forall j \neq k.$$

Schwieriger ist die anpassung des Definitionsbereiches von $\mathbb{P}$. Warum ist das nötig? Betrachte das Beispiel "Glücksrad".

[22.04.2024]

II.0.1 Lemma

Es gibt kein translationsinvariantes Wahrscheinlichkeitsmaß auf $\mathbb{P}([0, 2\pi[)$.

II.0.2 Beweis

WA: Es gibt ein solches $\mathbb{P}$.

Wir zerlegen nun $[0, 2\pi[$ in überabzählbar viele abzählbare Teilmengen:

$$a \sim b :\Leftrightarrow a - b \in \mathbb{Q}.$$

Dies definiert Äquivalenzklassen, diese dienen zur Zerlegung von $[0, 2\pi[$:

- $\{0, \frac{1}{2}, \frac{1}{3}, 1, \ldots\}$,
- $\{\sqrt{2}, \sqrt{2} + \frac{1}{2}, \ldots\}$,
- $\{\pi, \pi + \frac{1}{2}, \ldots\}$,
- ...

Wir wählen aus jeder Äquivalenzklasse einen Representanten $\leftarrow \{0, \sqrt{2}, \pi, \ldots\}$. R ist natürlich überabzählbar. Für jedes $q \in [0, 2\pi[\cap \mathbb{Q}$ definieren wir

$$R_q := \{r + q, \ r \in R\} = R + q.$$

Nach Annahme der Translationsinvarianz ist

$$\mathbb{P}(R_q) = \mathbb{P}(R), \ \forall q \in [0, 2\pi[\cap \mathbb{Q}.$$

Es gilt:

$$[0, 2\pi[\ \subseteq \bigcup_{q \in]-2\pi, 2\pi[\cap \mathbb{Q}} R_q \subset \ ]-2\pi, 4\pi[.$$

$\Rightarrow$ da $\mathbb{P}([0,2\pi[) = 1$, gilt, dass

$$\begin{aligned} &\mathbb{P}(\bigcup_{q\in]-2\pi,2\pi[\cap\mathbb{Q}} R_q) \geq 1. \\ \Rightarrow &\sum_{j\in]-2\pi,2\pi[} \mathbb{P}(R) \geq 1 \\ \Rightarrow &\mathbb{P}(R) \neq 0, \end{aligned}$$

aber falls der Inhalt $\mathbb{P}(R) > 0$, folgt, dass das Intervall $]-2\pi,4\pi[$ unendlichen Inhalt hat. Dieses überdeckt jedoch $[0,2\pi[$ dreimal. $\lightning$

II.1 Ereignisraum

Wir schränken den Definitionsbereich von $\mathbb{P}$ ein, um solch problematische Mengen zu umgehen.

II.1.1 σ-Algebra

Sei Ω eine Menge. Eine σ-Algebra bzgl. Ω ist eine Teilmenge von $\mathcal{P}(\Omega)$ mit:

a) $\Omega \in \mathcal{A}$,

b) $A \in \mathcal{A} \Rightarrow A^C \in \mathcal{A}$,

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

II.1.2 Beispiele

a) $\mathcal{P}(\Omega)$, sowie $\{\emptyset, \Omega\}$ ist jeweils σ-Algebra.

b)
- $\Omega = \{1,2,3,4\}$.
- $\mathcal{A} = \{\emptyset, \Omega, \{1,2\}, \{3\}, \{4\}, \{3,4\}, \{1,2,4\}, \{1,2,3\}\}$.

II.1.3 Satz

Sei $\mathcal{A}$ eine σ-Algebra bzgl. Ω. Dann gilt:

a) $\emptyset \in \mathcal{A}$,

b) abzählbare Schritte von Ereignissen sind in $\mathcal{A}$,

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

II.1.4 Beweis

a) $\emptyset = \Omega^C \in \mathcal{A}$,

b) $\bigcap_{n=1}^{\infty} A_n = \left[\bigcup_{n=1}^{\infty} A_n^C\right]^C \in \mathcal{A}$,

c) $A \setminus B = A \cap B^C$.

II.1.5 Satz

Sei Ω eine Menge. Der Schnitt beliebiger σ-Algebren ergibt wieder eine σ-Algebra.

II.1.6 Beweis

Seien $\mathcal{A}_i$ für $i \in \mathcal{I}$ σ-Algebren (bzgl. Ω). Z.z. $\mathcal{A} := \bigcap_{i\in\mathcal{I}} \mathcal{A}_i$ ist σ-Algebra.

a) $\Omega \in \mathcal{A}_i,\ \forall i \in \mathcal{I} \Rightarrow \Omega \in \mathcal{A}$.

b) Sei $E \in \mathcal{A} \Rightarrow E \in \mathcal{A}_i,\ \forall i \in \mathcal{I} \Rightarrow E^C \in \mathcal{A}_i,\ \forall i \in \mathcal{I}$.

c) Seien $(E_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ ($E_n \in \mathcal{A}\ \forall n \in \mathbb{N}$):

$$\begin{aligned}
&\Rightarrow E_n \in A_i,\ \forall n \in \mathbb{N}, \forall i \in \mathcal{I} \\
&\Rightarrow \bigcup_{n=1}^{\infty} E_n \in \mathcal{A}_i,\ \forall i \in \mathcal{I} \text{ (da } \mathcal{A}_i\ \sigma\text{-Algebra)} \\
&\Rightarrow \bigcup_{n=1}^{\infty} E_n \in \mathcal{A}.
\end{aligned}$$

II.1.7 Bemerkung

Vereinigungen von σ-Algebren ergeben *nicht* notwendigerweise eine σ-Algebra.

II.1.8 Beispiel

$\Omega = \{1, 2, 3\}$.

- $\mathcal{A}_1 = \{\Omega, \emptyset, \{1, 2\}, \{3\}\}$,
- $\mathcal{A}_2 = \{\Omega, \emptyset, \{1, 3\}, \{2\}\}$,
- $\mathcal{A}_1 \cup \mathcal{A}_2 = \{\Omega, \emptyset, \{1, 2\}, \{1, 3\}, \{2\}, \{3\}\} \not\ni \{2\} \cup \{3\}$.

II.1.9 Definition (Erzeugte σ-Algebra)

Sei Ω eine Menge, $\mathcal{E} \subset \mathcal{P}(\Omega)$. Die σ-Algebra $\sigma(\mathcal{E})$ definiert durch

$$\sigma(\mathcal{E}) := \bigcap_{\mathcal{A} \text{ ist } \sigma\text{-Alg},\ \mathcal{E}\subset\mathcal{A}} \mathcal{A}$$

nennen wir die *von $\mathcal{E}$ erzeugte σ-Algebra.*

II.1.10 Bemerkung

$\sigma(\mathcal{E})$ ist für nichtleere Ω immer wohldefiniert und wegen Satz σ-Algebra.

II.1.11 Korollar

$\sigma(\mathcal{E})$ ist die kleinste σ-Algebra, die $\mathcal{E}$ enthält, d.h.

a) $\mathcal{E} \in \sigma(\mathcal{E})$,

b) $\forall \widetilde{\mathcal{A}}$ σ-Algebra it $\mathcal{E} \subset \widetilde{\mathcal{A}}$, gilt $\sigma(\mathcal{E}) \subset \widetilde{\mathcal{A}}$.

II.1.12 Beweis

a) $\mathcal{E}$ ist in allen σ-Algebren enthalten, über die der Schnitt gebildet wird.

b) $\widetilde{\mathcal{A}}$ ist ein Kandidat für $\mathcal{A}$ in der Definition. Es wird also auch über $\widetilde{\mathcal{A}}$ der Schnitt gebildet $\Rightarrow \widetilde{A} \supset \sigma(\mathcal{E})$.

II.1.13 Beispiel

$\Omega = \{1,2,3\}, \mathcal{E} = \{\{1,2\}\}$. $\mathcal{P}(\Omega)$ ist σ-Algebra, enthält $\mathcal{E}$ $\mathcal{A}_1$ vom Beispiel oben ebenso. Weitere Kandidaten:

$$\mathcal{B} = \{\Omega, \emptyset, \{1,2\}, \{3\}, \{1\}, \{2\}, \ldots\}.$$

[25.04.2024]

II.1.14 Defintion (Borel-σ-Algebra)

Sei $\Omega = \mathbb{R}$. Die von den offenen Teilmengen von $\mathbb{R}$ erzeugte σ-Algebra heißt *Borel-σ-Algebra*. (alternative Definiton später)

II.2 Wahrscheinlichkeitsmaß

II.2.1 Definition (Wahrscheinlichkeitsmaß)

Sei Ω eine Menge, $\mathcal{A}$ σ-Algebra. Eine Abbildung $\mathbb{P}\colon \mathcal{A} \to \mathbb{R}$ heißt *Wahrscheinlichkeitsmaß* $:\Leftrightarrow$

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

b) $\mathbb{P}(A) \geq 0,\ \forall A \in \mathcal{A}$,

c) $\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$, falls $A_n \in \mathcal{A},\ A_n \cap A_m = \emptyset\ \forall n \neq m$.

II.2.2 Bemerkung

Satz vom Gegenereignis, $\mathbb{P}(\emptyset) = 0$ und $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ gilt weiterhin.

II.2.3 Definition und Satz (Bedingte Wahrscheinlichkeit)

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, d.h. Ω Menge, $\mathcal{A}$ zugehörige σ-Algebra, $\mathbb{P}\colon \mathring{A} \to \mathbb{R}$ Wahrscheinlichkeitsmaß.

Sei $A \in \mathcal{A}$ mit $\mathbb{P}(A) \neq 0$. Dann ist das auf A *bedingte Wahrscheinlichkeitsmaß* definiert durch

$$\mathbb{P}_A(B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}.$$

II.2.4 Beweis

(dass dies ein Wahrscheinlichkeitsmaß ist):

a) $\mathbb{P}_A(\Omega) = \frac{\mathbb{P}(A\cap\Omega)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A)}{\mathbb{P}(A)} = 1.$

b) Zähler ≥ 0, Nenner $> 0 \Rightarrow$ Behauptung.

c)

$$\begin{aligned}
\mathbb{P}_A(\bigcup_{n=1}^{\infty} A_n) &= \frac{\mathbb{P}(A \ \cap \ \bigcup_{n=1}^{\infty} A_n)}{\mathbb{P}(A)} \ (A_n \cap A_m = \emptyset, \ n \neq m) \\
&= \frac{\mathbb{P}(\bigcup_{n=1}^{\infty} \overbrace{(A_n \cap A)}^{\text{paarweise disjunkt}})}{\mathbb{P}(A)} \\
&= \frac{\sum_{n=1}^{\infty} \mathbb{P}(A_n \cap A)}{\mathbb{P}(A)} \\
&= \sum_{n=1}^{\infty} \mathbb{P}_A(A_n).
\end{aligned}$$

II.2.5 Definition (Unabhängigkeit)

Zwei Ereignisse A, B heißen (stochastisch) *unabhängig* $:\Leftrightarrow$

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

II.2.6 Bemerkung

A unabhängig von $B \Rightarrow \mathbb{P}_A(B) = \mathbb{P}(B)$ (falls $\mathbb{P}(A) \neq 0$).

II.2.7 Beispiel

$\Omega = \{1, \ldots, 6\}$.

a) $A = \{3, 4, 5, 6\}, \ B = \{2, 4, 6\}$ sind unabhängig.

b) $\widetilde{A} = \{4, 5, 6\}$ und B wie oben sind nicht unabhängig.

II.2.8 Definition (Limes von Ereignissen)

Seien $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}$, $(B_n)_{n\in\mathbb{N}} \subset \mathcal{A}$. Wir nehmen an:

- $A_n \subset A_{n+1}, \ \forall n \in \mathbb{N}$,
- $B_n \supset B_{n+1}, \ \forall n \in \mathbb{N}$.

Dann ist

$$\begin{aligned}
\lim_{n\to\infty} A_n &:= \bigcup_{n=1}^{\infty} A_n, \\
\lim_{n\to\infty} B_n &:= \bigcap_{n=1}^{\infty} B_n.
\end{aligned}$$

II.2.9 Korrolar

$\lim_{n\to\infty} A_n$ und $\lim_{n\to\infty} B_n$ sind Ereignisse, falls $A_n,\ B_n$ Ereignisse sind.

II.2.10 Beweis

Definition der σ-Algebra, bzw. Satz gleich darunter.

II.2.11 Definition ($\limsup$, $\liminf$)

Sei $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ Dann ist

$$\begin{aligned}\limsup_{n\to\infty} A_n &:= \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n, \\ \liminf_{n\to\infty} A_n &:= \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n.\end{aligned}$$

II.2.12 Korollar

Auch $\limsup$ und $\liminf$ sind Ereignisse (falls $A_n \in \mathcal{A}\ \forall n \in \mathbb{N}$).

II.2.13 Satz (σ-Stetigkeit des Wahrscheinlichkeitsmaßes)

Sei $(A_n)_{n\in\mathbb{N}}, (B_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ eine abfallende bzw. ansteigende Folge von Ereignissen ($A_n \subset A_{n+1},\ B_n \supset B_{n+1},\ \forall n$). Dann ist

a) $\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(\lim_{n\to\infty} A_n)$,

b) $\lim_{n\to\infty} \mathbb{P}(B_n) = \mathbb{P}(\lim_{n\to\infty} B_n)$.

II.2.14 Beweis

Definiere $C_1 := A_1$. $C_2 := A_1 \setminus A_1, \ldots, C_n = A_n \setminus A_{n-1}$. Es gilt:

a) $\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} C_n$.

b) $C_n \cap C_m = \emptyset,\ \forall n \neq m$,

c) $A_n = \bigcup_{k=1}^{n} C_k$.

a)

$$\begin{aligned}\Rightarrow \mathbb{P}(\lim_{n\to\infty} A_n) &\overset{\text{a)}}{=} \mathbb{P}(\bigcup_{n=1}^{\infty} C_n) \\ &\overset{\text{Kc)}}{=} \sum_{n=1}^{\infty} \mathbb{P}(C_n) \\ &= \lim_{n\to\infty} \sigma_{k=1}^{n} \mathbb{P}(C_k) \\ &\overset{\text{Kc)}}{=} \lim_{n\to\infty} \mathbb{P}(\bigcup_{k=1}^{n} C_k) \\ &\overset{\text{c)}}{=} \mathbb{P}(A_n).\end{aligned}$$

$\square$

b)

$$\begin{aligned}\mathbb{P}(\lim_{n\to\infty} B_n) &= 1 - \mathbb{P}(\lim_{n\to\infty} B_n^C) \\ &\overset{\text{Fall a)}}{=} 1 - \lim_{n\to\infty} \mathbb{P}(B_n^C) \\ &= \lim_{n\to\infty} \mathbb{P}(B_n).\end{aligned}$$

[29.04.2024]

II.3 Borelsche σ-Algebra

II.3.1 Erinnerung

$\mathcal{B}^{\mathbb{R}}$ ist die σ-Algebra erzeugt aus allen offenen Teilmengen von $\mathbb{R}$.

II.3.2 Bemerkung

Diese Definition lässt sich auf Grundmengen verallgemeinern, in denen es ein "Konzept" von offenen Teilmengen gibt (topologische Räume, z.B. metrische Räume). Wir werden alternative Definitionen von $\mathcal{B}$ betrachten.

II.3.3 Korollar

a) Sei $\mathcal{E} \subset \mathcal{P}(\Omega)$, $\mathcal{A} \subset \mathcal{P}(\Omega)$ σ-Algebra. $\mathcal{E} \subset \mathcal{A} \Rightarrow \sigma(\mathcal{E}) \subset \mathcal{A}$.

b) Seien $\mathcal{E}, \mathcal{F} \subset \mathcal{P}(\Omega)$. Falls $\mathcal{E} \subset \sigma(\mathcal{F})$ und $\mathcal{F} \subset \sigma(\mathcal{E}) \Rightarrow \sigma(\mathcal{E}) = \sigma(\mathcal{F})$.

II.3.4 Beweis

a)

$$\sigma(\mathcal{E}) = \bigcap_{\mathcal{B}\ \sigma\text{-Algebra},\ \mathcal{E}\subset\mathcal{B}} \mathcal{B}$$

$\rightsquigarrow$ $\mathcal{A}$ ist einer der Kandidaten, über die geschnitten wird.

b)

$$\begin{aligned}\mathcal{E} \subset \sigma(\mathcal{F}) &\overset{\text{a)}}{\Rightarrow} \sigma(\mathcal{E}) \subset \sigma(\mathcal{F}), \\ \mathcal{F} \subset \sigma(\mathcal{E}) &\overset{\text{a)}}{\Rightarrow} \sigma(\mathcal{F}) \subset \sigma(\mathcal{E}).\end{aligned}$$

II.3.5 Satz

Die Borelsche σ-Algebra $\mathcal{B}^{\mathbb{R}}$ ist die aus den offenen Intervallen erzeugte σ-Algebra.

II.3.6 Beweis

Z.z.: Jede offene Teilmenge $\mathbb{R}$ liegt in der aus den offenen Intervallen erzeugten σ-Algebra.

Sei $A \subset \mathbb{R}$ offen. Für jedes $q \in \mathbb{Q} \cap A$ sei

$$r_q = \sup\{\varepsilon \in \mathbb{R}\colon\ ]q-\varepsilon,\ q+\varepsilon[\ \subset A\}.$$

Das Intervall $]q-r_q,\ q+r_q[$ ist $\subset A$ $(*)$. (WA: $]q-r_q,\ q+r_q[\ \not\subset A \Rightarrow \exists x \in\ ]q-r_q,\ q+r_q[$ mit $x \notin A$. Wähle $\varepsilon = \frac{r_q+|q-x|}{2} \Rightarrow x \in\ ]q-r_q,\ q+r_q[$, aber $\varepsilon < r_q$ $\lightning$.)

Betrachte $B = \bigcup_{q\in\mathbb{Q}\cap A}\ ]q-r_q,\ q+r_q[$. Z.z. $A = B$ (da B abzählbare Vereinigung offener Intervalle).

- $B \subset A$, da all die Teilintervalle $]q-r_q,\ q+r_q[\ \subset A$ $(*)$.
- $A \subset B$. Sei $y \in A$. Z.z. $y \in B$. Wegen Offenheit von A $\exists \delta_y$, sodass $]y-\delta_y,\ y+\delta_y[\ \subset A$. Wähle $q_y \in \mathbb{Q}\cap\ ]y-\frac{\delta_y}{2}, y+\frac{\delta_y}{2}[$. $r_{q_y} \geq \frac{\delta_y}{2} \Rightarrow y \in\ ]q_y - r_{q_y},\ q_y + r_{q_y}[$.

II.3.7 Satz

Die Borel-σ-Algebra ist genau die σ-Algebra erzeugt aus den Intervallen $]-\infty,\ a]$ mit $a \in \mathbb{R}$.

II.3.8 Beweis

$\mathcal{E} :=$ Menge der offenen Intervalle, $\mathcal{F} := \{]-\infty,\ a],\ a \in \mathbb{R}\}$.

a) Z.z. $\mathcal{E} \subset \sigma(\mathcal{F})$.

   Sei $a < b \in \mathbb{R} \cup \{\pm\infty\}$. Z.z. $]a,b[\ \in \sigma(\mathcal{F})$.

   - 1. Schritt: "$a = -\infty$".

     $$]-\infty,\ b[\ = \underbrace{\bigcup_{n\in\mathbb{N}} \underbrace{\left]-\infty,\ b-\frac{1}{n}\right]}_{\in \mathcal{F}}}_{\in \sigma(\mathcal{F})}.$$

     $\forall x \in\ ]-\infty,\ b[$, d.h. $\forall x < b\ \exists n \in \mathbb{N}$, sodass $x < b - \frac{1}{n} \Rightarrow x \in \bigcup_{n\in\mathbb{N}} \left]-\infty,\ b-\frac{1}{n}\right]$. $\Rightarrow\ ]-\infty,\ b[\ \subset \bigcup_{n\in\mathbb{N}} \left]-\infty,\ b-\frac{1}{n}\right]$. $\forall n \in \mathbb{N}$ ist $\left]-\infty, b-\frac{1}{n}\right]\ \subset\ ]-\infty, b[\ \Rightarrow$ "$\supset$".
   - 2. Schritt:

     $$a \in \mathbb{R}\colon\ ]a,b[\ = \underbrace{]-\infty,b[}_{\in \sigma(\mathcal{F})\ (1.\ \text{Schritt})} \setminus \overbrace{]-\infty,a]}^{\in \mathcal{F} \subset \sigma(\mathcal{F})}$$

     nach Satz ist letzteres in $\sigma(\mathcal{F})$.

b) Z.z. $\forall a \in \mathbb{R}$ ist $]-\infty,\ a]\ \in \sigma(\mathcal{E})$.

$$]-\infty,a]\ =\ (]a,\ +\infty[)^C = (\bigcup_{n\in\mathbb{N}}\ ]a,\ n[)^C.$$

II.3.9 Satz

$\mathcal{B}^{\mathbb{R}}$ ist die σ-Algebra erzeugt aus $\{]-\infty, a] \text{ mit } a \in \mathbb{Q}\}$.

II.3.10 Beweis

Sei $\mathcal{G} = \{]-\infty, a] \colon a \in \mathbb{Q}\}$. Offensichtlich ist $\mathcal{G} \subset \mathcal{F}$.

$$\mathcal{G} \subset \mathcal{F} \subset \sigma(\mathcal{F}).$$

Z.z. $\mathcal{F} \subset \sigma(\mathcal{G})$. Sei $]-\infty, a]$ $a \in \mathbb{R}$.

$$]-\infty, a] \ = \underbrace{\bigcap_{q\in\mathbb{Q},\ q\geq a}]-\infty,\ q]}_{\in\,\sigma(\mathcal{G})}.$$

$]-\infty, a] \ \subset \,]-\infty, q], \forall q \geq a \Rightarrow \,]-\infty, a] \subset \bigcup_{q\in\mathbb{Q},\ q\geq a}]-\infty, q]$ ($\Rightarrow$ "$\subset$").

"$\supset$": Dazu "$\subset$" für die Komplete. Sei $x \notin]-\infty,\ a] \Rightarrow x > a$. Wähle $q \in \,]a, x[\, \cap \mathbb{Q} \Rightarrow x \notin \,]-\infty,\ q] \ \Rightarrow x \notin \bigcup_{q\in\mathbb{Q},\ q\geq a}]-\infty, q]$.

[02.05.2024]

II.4 Eindeutigkeitssatz

Ziel ist es zu zeigen, dass unter gewissen Bedingungen das Wahrscheinlichkeitsmaß eindeutig festgelegt ist, falls es auf einem Erzeuger gegeben ist.

Wir benutzen das "Prinzip der guten Mengen".

II.4.1 Satz (Prinzip der guten Mengen)

Falls eine Eingenschaft für $\mathcal{E} \subset \mathcal{P}(\Omega)$ gilt und die Menge, auf der die Eigenschaft gilt eine σ-Algebra ist, so gilt die Eigenschaft auf ganz $\sigma(\mathcal{E})$.

II.4.2 Beweis

Sei $\mathcal{A}$ die Menge, für die die Eigenschaft gilt. Da $\mathcal{E} \subset \mathcal{A}$, $\mathcal{A}$ ist σ-Algebra nach Voraussetzung,

$$\sigma(\mathcal{E}) = \bigcup_{\mathcal{B} \text{ ist } \sigma\text{-Algebra},\ \mathcal{E}\subset\mathcal{B}} \mathcal{B}.$$

$\mathcal{A}$ ist eines der Kandidaten für $\mathcal{B} \Rightarrow \sigma(\mathcal{E}) \subset \mathcal{A}$.

II.4.3 Beweisstrategie für den Eindeutigkeitssatz

Seien $\mathbb{P}, \mathbb{Q} \colon \sigma(\mathcal{E}) \to \mathbb{R}$ Wahrscheinlichkeitsmaße. Wir möchten zeigen, dass unter gewissen Bedingungen die Menge $\mathcal{G}$ definiert durch

$$A \in \mathcal{G} \Leftrightarrow \mathbb{P}(A) = \mathbb{Q}(A)$$

eine σ-Algebra ist.

Nach dem Prinzip der guten Mengen ist dann

$$\mathbb{P}(A) = \mathbb{Q}(A),\ \forall A \in \sigma(\mathcal{E}).$$

Wir beweisen dies in zwei Schritten:

1) Die Menge $\mathcal{G}$ ist ein Dynkin-System.

2) Unter gewissen Bedingungen ist jedes Dynkin-System eine σ-Algebra.

II.4.4 Definition (Dynkin-System)

Eine Teilmenge $\mathcal{D}$ heißt *Dynkin-System* :$\Leftrightarrow$

a) $\Omega \in \mathcal{D}$,

b) $A \in \mathcal{D} \Rightarrow A^C \in \mathcal{D}$,

c) $(A_n)_{n\in\mathbb{N}} \subset \mathcal{D}$, A_n paarweise disjunkt $\Rightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{D}$.

II.4.5 Satz

Seien $\mathbb{P}, \mathbb{Q}$ Wahrscheinlichkeitsmaße auf $\mathcal{A}$. $\mathcal{G} := \{A \subset \mathcal{A} \colon \mathbb{P}(A) = \mathbb{Q}(A)\}$ ist ein Dynkin-System.

II.4.6 Beweis

a) $\mathbb{P}(\Omega) = 1 = \mathbb{Q}(\Omega) \Rightarrow \Omega \in \mathcal{G}$.

b) Sei $A \in \mathcal{G}$, d.h. $\mathbb{P}(A) = \mathbb{Q}(A) \Rightarrow \mathbb{P}(A^C) = 1 - \mathbb{P}(A) = 1 - \mathbb{Q}(A) = \mathbb{Q}(A^C)$, d.h. $A^C \in \mathcal{G}$.

c) Sei $(A_n)_{n\in\mathbb{N}} \subset \mathcal{G}$ paarweise disjunkt, d.h. $\mathbb{P}(A_n) = \mathbb{Q}(A_n),\ \forall n \in \mathbb{N}$

$$\Rightarrow \mathbb{P}(\bigcup_{n\in\mathbb{N}} A_n) \overset{\mathbf{Kc)}}{=} \Sigma_{n\in\mathbb{N}}\mathbb{P}(A_n) \overset{\overleftarrow{\mathbf{Kc)}}}{=} \mathbb{Q}(\bigcup_{n\in\mathbb{N}} A_n).$$

II.4.7 Definition (Schnittstabilität)

Eine Menge $\mathcal{E} \subset \mathcal{P}(\Omega)$ heißt *schnittstabil* :$\Leftrightarrow A \cap B \in \mathcal{E},\ \forall A, B \in \mathcal{E}$.

II.4.8 Satz

Jedes schnittstabile Dynkin-System ist eine σ-Algebra.

II.4.9 Beweis

Axioma a),b) sind identisch (σ-Algebra und Dynkin-System). Es bleibt c) zu zeigen. Sei dazu $\mathcal{A}$ ein schnittstabiles Dynkin-System. Sei $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ beliebig. Z.z. $\bigcup_{n\in\mathbb{N}} A_n \subset \mathcal{A}$.

Sei $A_1, A_2 \in \mathcal{A} \Rightarrow A_1 \cap A_2 \in \mathcal{A}$ (schnittstabil) $\Rightarrow (A_1 \cap A_2)^C \in \mathcal{A}$. Da $\mathcal{A}$ schnittstabil ist, ist

$$\underbrace{(A_1 \cap A_2)^C \cap A_1}_{A_1 \setminus A_2} \in \mathcal{A},\ A_1 \cup A_2 = A_2 \cup (A_1 \setminus A_2) \in \mathcal{A} \text{ (disjunkt)}.$$

$$\bigcup_{n\in\mathbb{N}} A_n = A_1 \cup (A_2 \setminus A_1) \cup ((A_3 \setminus A_1) \setminus A_2) \cup \dots$$

$\square$

II.4.10 Satz

Ein von $\mathcal{E} \subset \mathcal{P}(\Omega)$ erzeugtes Dynkin-System $\delta(\mathcal{E})$ ist bereits schnittstabil, falls $\mathcal{E}$ schnittstabil ist.

II.4.11 Beweis

Sei $\mathcal{E} \subset \mathcal{P}(\Omega)$ schnittstabil, $E \in \mathcal{E}$ beliebig. Sei

$$\mathcal{A}_E := \{A \in \mathbb{P}(\Omega) \colon E \cap A \in \delta(\mathcal{E})\}.$$

Wir zeigen nun, dass $\mathcal{A}_E$ ein Dynkin-System ist.

a) $\Omega \in \mathcal{A}_E$: $E \cap \Omega = E \in \mathcal{E} \subset \delta(\mathcal{E})$.

b) Sei $A \in \mathcal{A}_E \Rightarrow E \cap A \in \delta(\mathcal{E})$:

$$A^C \cap E = (A \cup E^C)^C = ((A \cap E) \cup E^C)^C.$$

c) Sei $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}_E$ paarweise disjunkt. Z.z. $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}_E$. $\forall n \in \mathbb{N}$ ist $A_n \cap E \in \delta(\mathcal{E})$

$$E \cap \bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} E \cap A_n \ (*).$$

$E \cap A_n$ sind paarweise disjunkt, da die A_n paarweise disjunkt $\Rightarrow (*) \in \delta(\mathcal{E}) \Rightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}_E$.

[06.05.2024]

Sei nun $\mathcal{A}_B := \{A \in \delta(\mathcal{E}) \colon A \cap B \in \delta(\mathcal{E})\}$ für $B \in \delta(\mathcal{E})$.

1) $\mathcal{E} \subset \mathcal{A}_B$ wegen des vorigen Schrittes.

2) a) $\Omega \in \mathcal{A}_B$, da $\Omega \cap B = B \in \delta(\mathcal{E})$.
   b) Sei $A \in \mathcal{A}_B$, d.h. $A \cap B \in \delta(\mathcal{E}) \Rightarrow \underbrace{B \setminus (A \cap B)}_{B \cap A^C \in \delta(\mathcal{E})} \in \delta(\mathcal{E})$.

3) Sei $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}_B$ paarweise disjunkt.

$$\left(\dot{\bigcup}_{n\in\mathbb{N}} A_n\right) \cap B = \dot{\bigcup}_{n\in\mathbb{N}} \underbrace{(A_n \cap B)}_{\ni \delta(\mathcal{E}) \text{ n.V.}} \in \delta(\mathcal{E}) \overset{\text{P.d.g.M.}}{\Rightarrow} \mathcal{A}_B \supset \delta(\varepsilon) \text{ (außerdem } \mathcal{A}_B \subset \delta(\mathcal{E})\text{), d.h. } \mathcal{A}_B = \delta(\mathcal{E}).$$

$\Rightarrow \forall A, B \in \delta(\mathcal{E})$ gilt $A \cap B \in \delta(\mathcal{E})$.

II.4.12 Korollar

Sei $\mathcal{E} \subset \mathcal{P}(\Omega)$ schnittstabil $\Rightarrow \delta(\mathcal{E}) = \delta(\mathcal{E})$.

II.4.13 Beweis

- "$\subset$": $\delta(\mathcal{E})$ enthält $\mathcal{E}$ und ist ein Dynkin-System $\overset{\text{P.d.g.M.}}{\Rightarrow} \delta(\mathcal{E}) \subset \delta(\mathcal{E})$.
- "$\supset$": $\delta(\mathcal{E})$ enthält $\mathcal{E}$. $\delta(\mathcal{E})$ ist eine σ-Algebra aufgrund der letzten beiden Sätze $\overset{\text{P.d.g.M.}}{\Rightarrow} \sigma(\mathcal{E}) \subset \delta(\mathcal{E})$.

II.4.14 Satz (Eindeutigkeitssatz)

Sei $\mathcal{E} \subset \mathcal{P}(\Omega)$ schinttstabil. Dann ist $\mathbb{P}\colon \sigma(\mathcal{E}) \to \mathbb{R}$ eindeutig durch $\mathbb{P}$ eingeschränkt auf $\mathcal{E}$ definiert (falls existent).

II.4.15 Beweis

Seien $\mathbb{P}, \mathbb{Q}\colon \sigma(\mathcal{E} \to \mathbb{R}$ Wahrscheinlichkeitsmaße mit $\mathbb{P}(A) = \mathbb{Q}(A)$, $\forall A \in \mathcal{E}$. $\{A \in \sigma(\mathcal{E})\colon \mathbb{P}(A) = \mathbb{Q}(A)\} = \mathcal{A}$ enthält $\mathcal{E}$ und ist ein Dynkin-System $\Rightarrow \mathcal{A} \supset \delta(\mathcal{E}) = \sigma(\mathcal{E})$.

II.4.16 Beispiel

a) Durch $\mathbb{P}(]a, b[) = \frac{b-a}{2\pi}$ wird für unser Glücksradspiel das Wahrscheinlichkeitsmaß auf der entsprechenden Borel-σ-Algebra eindeutig.

b) Sei $f\colon \mathbb{R} \to \mathbb{R}_0^+$ integrierbar mit

$$\int_{-\infty}^{\infty} f(t)dt = 1.$$

Dann ist durch $\mathbb{P}(]-\infty, a]) = \int_{-\infty}^{a} f(t)dt$ das Wahrscheinlichkeitsmaß eindeutig.

II.4.17 Bemerkung

Für die obigen Beispiele kann die Existenz eines Wahrscheinlichkeitsmaßes mit den genannten Eigenschaften gezeigt werden. (siehe Maßtheorie)

III Zufallsvariable

Es liegt nahe, die Menge Ω nach z.B. $\mathbb{R}$ abzubilden, um die Ausgänge des Experimentes zu quantifizieren.

III.0.1 Definition (Zufallsvariable im Diskreten)

Eine Abbildung $X\colon \Omega \to \vartheta$ (falls $|\Omega| < \infty$, $(\Omega, \mathbb{P})$ Wahrscheinlichkeitsraum) nennt man *Zufallsvariable*.

III.0.2 Beispiel

a) Würfel $\Omega = \{\dot{1}, \dot{2}, \ldots, \dot{6}\}$, $\vartheta = \{1, 2, \ldots, 6\}$. X bildet auf die Augenzahl ab.

b) Glücksspiel: Man gewinnt 3 Euro, falls Augenzahl durch 3 teilbar, ansonsten verliert man 1 Euro:
$$y\colon \Omega \to \{-1, 3\}.$$

III.0.3 Bemerkung

Es ist für jede Zufallsvariable in natürlicher Weise ein Wahrscheinlichkeitsmaß auf ϑ definiert:

$$\mathbb{Q}(A) = \mathbb{P}(X^{-1}(A)).$$

III.0.4 Definition (Zufallsvariable)

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, $(\vartheta, \mathcal{B})$ ein sogennanter Maßraum, d.h. $\mathcal{B}$ ist σ-Algebra bzgl. ϑ. Dann nennt man jede Abbildung $X\colon \Omega \to \vartheta$ mit der Eigenschaft

$$(*)\ X^{-1}(B) \in \mathcal{A},\ \forall B \in \mathcal{B}$$

eine *Zufallsvariable*.

III.0.5 Bemerkung

Die Eigenschaft $(*)$ der Abbildung X nennt man $\mathcal{A}$-$\mathcal{B}$ *messbar*.

III.0.6 Satz

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ Wahrscheinlichkeitsraum, ϑ eine Menge mit σ-Algebra $\mathcal{B}$. Dann ist auf B durch

$$\mathbb{Q}(A) = \mathbb{P}(X^{-1}(\mathcal{A}))$$

ein Wahrscheinclihkeitsmaß definiert. D.h. falls $X\colon \Omega \to \vartheta$ $\mathcal{A}$-$\mathcal{B}$-messbar $\Rightarrow$ $(\vartheta, \mathcal{B}, \mathbb{P} \cdot X^{-1})$ ist Wahrscheinlichkeitsraum.

[13.05.2024]

III.0.7 Satz

X Zufallsvariable $\Rightarrow \mathbb{Q} := \mathbb{P} \cdot X^{-1}$ ist Wahrscheinlichkeitsmaß.

III.0.8 Erinnerung

$X^{-1} := \{\omega \in \Omega \colon X(\omega) \in A\},\ X: \Omega \to \vartheta.$

III.0.9 Proposition

Sei $X: \Omega \to \vartheta$ eine Abbildung. $A, B \subset \vartheta$. Dann gilt:

a) $X^{-1}(\vartheta) = \Omega$,

b) $X^{-1}(A^C) = [X^{-1}(A)]^C$,

c) $X^{-1}(A) \cup X^{-1}(B) = X^{-1}(A \cup B)$,

d) $X^{-1}(A) \cap X^{-1}(B) = X^{-1}(A \cap B)$,

e) $X^{-1}(A) \cap X^{-1}(B) = \emptyset$, falls $A \cap B = \emptyset$.

III.0.10 Beweis

a) $\forall \omega \in \Omega$ gilt $X(\omega) \in \vartheta$.

b) $\omega \in X^{-1}(A^C) \Leftrightarrow X(\omega) \in A^C \Leftrightarrow X(\omega) \notin A \Leftrightarrow \omega \notin X^{-1}(A) \Leftrightarrow w \in [X^{-1}(A)]^C$.

c) $\omega \in X^{-1}(A) \cup X^{-1}(B) \Leftrightarrow \omega \in X^{-1}(A)$ oder $\omega \in X^{-1}(B) \Leftrightarrow X(\omega) \in A$ oder $X(\omega) \in B \Leftrightarrow \omega \in X^{-1}(A \cup B)$. Geht ebenso für beliebige, insbesondere abzählbare Vereinigungen.

d) folgt aus mehrfachem Anwenden von b), und c).

e) folgt aus d), da $X^{-1}(\emptyset) = \emptyset$.

III.0.11 Beweis des Satzes

a) $\mathbb{Q}(\vartheta) := \mathbb{P}(X^{-1} = \mathbb{P}(\Omega) = 1.$

b) Positivität folgt aus b) für $\mathbb{P}$.

c) $\mathbb{Q}(\bigcup_{j=1}^{\infty} A_j)$ für A_j paarweise disjunkt.

$$= \mathbb{P}(X^{-1}(\bigcup_{j=1}^{\infty} A_j)) = \mathbb{P}(\bigcup_{j=1}^{\infty} \underbrace{X^{-1}(A_j)}_{\text{paarweise disjunkt e)}}) = \sum_{j=1}^{\infty} \mathbb{P}(X^{-1}(A_j)) = \sum_{j=1}^{\infty} \mathbb{Q}(A_j).$$

III.0.12 Satz

Sei $X: \Omega \to \vartheta$ messbar ($\mathcal{A} \to \mathcal{B}$), dann sind folgenden Mengen σ-Algebren:

a) $X^{-1}(\mathcal{B}) := \{X^{-1}(B) \colon B \in \mathcal{B}\}$,

b) $X_*(\mathcal{A}) := \{B \in \mathcal{B} \colon \exists A \in \mathcal{A} \colon X^{-1}(B) = A\}$.

a) bzgl. Ω, b) bzgl. ϑ.

III.0.13 Beweis

a) a) $\Omega \in X^{-1}(B)$, da $\Omega = X^{-1}(\vartheta)$.

   b) Sei $A \in X^{-1}(B) \Rightarrow \exists B$ mit $X^{-1}(B) = A \Rightarrow X^{-1}(B^C) = A^C \in X^{-1}(\mathcal{B})$.

   c) Seien $(A_n)_{n\in\mathbb{N}} \subset X^{-1}(\mathcal{B})$

$$\begin{aligned}
&\Rightarrow \forall n \in \mathbb{N}\ \exists B_n \in \mathcal{B} \text{ mit } A_n = X^{-1}(B_n) \\
&\Rightarrow \bigcup_{n\in\mathbb{N}} B_n \in B \text{ und } X^{-1}(\bigcup_{n\in\mathbb{N}} B_n) = \bigcup_{n\in\mathbb{N}} X^{-1}(B_n) \\
&\Rightarrow \bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} X^{-1}(B_n) = X^{-1}(\underbrace{\bigcup_{n\in\mathbb{N}} B_n}_{\in \mathcal{B}}) \\
&\Rightarrow \bigcup_{n\in\mathbb{N}} A_n \in X^{-1}(\mathcal{B}).
\end{aligned}$$

b) ähnlich.

III.0.14 Bemerkung

Die Messbarkeit ging in den Beweis gar nicht ein. Für jede Abbildung $X \colon \Omega \to \vartheta$ gilt, dass für eine σ-Algebra $\mathcal{B}$ bzgl. $X^{-1}(\mathcal{B}$ eine σ-Algebra bzgl. Ω ist.

III.0.15 Korollar

Sei $X \colon \Omega \to \vartheta$ eine Abbildung, $\mathcal{B}$ eine von $\mathcal{E}$ erzeugte σ-Algebra bzgl. ϑ. Dann ist X $\mathcal{A}$-$\mathcal{B}$-messbar bereits., wenn $X^{-1}(E) \in \mathcal{A}$, $\forall E \in \mathcal{E}$ $(*)$.

III.0.16 Beweis

Wir betrachten $X_*(\mathcal{A})$. Falls $(*)$ gilt, ist $\mathcal{E} \subset X_*(\mathcal{A})$. $X_*(\mathcal{A})$ ist σ-Algebra $\Rightarrow B \subset X_*(\mathcal{A}) \Rightarrow X$ ist $\mathcal{A}$-$\mathcal{B}$-messbar.

III.0.17 Beispiel

Jede stetige Funktion $f \colon \mathbb{R} \to \mathbb{R}$ ist $B_\mathbb{R}$-$B_\mathbb{R}$-messbar. Grund: Da f stetig, sind die Urbilder offener Mengen offen. Die offenen Mengen sind ein möglicher Erzeuger von $B_\mathbb{R}$, nach Korollar muss man nur diese untersuchen.

III.0.18 Satz

Seien $\Omega, \vartheta, \Lambda$ Mengen, $\mathcal{A}, \mathcal{B}, \mathcal{C}$ σ-Algebren bzgl. $\Omega, \vartheta, \Lambda$. Sei $X : \Omega \to \vartheta$ $mathcal A$-$\mathcal{B}$-messbar, $Y : \vartheta \to \Lambda$ $\mathcal{B}$-$\mathcal{C}$ messbar. Dann ist $Z := Y \cdot X : \Omega \to \Lambda$ $\mathcal{A}$-$\mathcal{C}$-messbar.

III.0.19 Beweis

Z.z.: $Z^{-1}(A) \in \mathcal{A}\ \forall A \in \mathcal{C}$.

$$Z^{-1}(A) = \underbrace{X^{-1}(\overbrace{Y^{-1}(A)}^{\in \mathcal{B}\text{, da } Y \text{ messbar}})}_{\in \mathcal{A}}.$$

III.0.20 Satz

Seien $X, Y : \Omega \to \mathbb{R}$ $\mathcal{A}$-$\mathcal{B}_\mathbb{R}$ messbar. Dann ist $X + Y$ ebenfalls $\mathcal{A}$-$\mathcal{B}_\mathbb{R}$ messbar.

III.0.21 Beweis

Wegen des Korollars gilt es zu zeigen:

$$(X+Y)^{-1}(]-\infty, a[) \in \mathcal{A},\ \forall a \in \mathbb{R}.$$

Betrachte dazu:

$$\underbrace{\bigcup_{b \in \mathbb{Q}} \underbrace{X^{-1}(\overbrace{]-\infty, b[}^{\in \mathcal{B}_\mathbb{R}})}_{\in \mathcal{A}} \cap \underbrace{Y^{-1}(\overbrace{]-\infty, a-b[}^{\in \mathcal{B}_\mathbb{R}})}_{\in \mathcal{A}}}_{\in \mathcal{A}}.$$

Wir zeigen nun, dass

$$(X+Y)^{-1}(]-\infty, a[) = \bigcup_{b \in \mathbb{Q}} X^{-1}(]-\infty, b[) \ \cap \ Y^{-1}(]-\infty, a-b).$$

- "$\subset$": Sei $\omega \in (X+Y)^{-1}(]-\infty, a[) \Leftrightarrow (X+Y)(\omega) < a$. Sei $b \in \mathbb{Q}$ mit $b \in]X(\omega), X(\omega) + \varepsilon 2[$. $\Rightarrow X(\omega) < b$ und $Y(\omega) + b < X(\omega) + Y(\omega) + \frac{\varepsilon}{2} < a \Rightarrow Y(\omega) < a - b$.
- "$\supset$": Sei $\omega \in \ldots \Rightarrow \exists b$, sodass $X(\omega) < b$ und $Y(\omega) < a - b \Rightarrow (X+Y)(\omega) < a$.

[16.05.2024]

III.0.22 Schreibweise

Die durch Zufallsvariablen bestimmten Ereignisse schreibt man oft in kurzer Form:

- $X \leq 5$ beschreibt $\{\omega \in \Omega \colon X(\omega) \leq 5\}$.
- $X \in B$ für $B \in \mathcal{B}$, $\{\omega \in \Omega \colon X(\omega) \in B\}$.
- $X \leq Y \cdot 2 + 7$, $\{\omega \in \Omega \colon X(\omega) \in 2Y(\omega) + 7\}$.
- $\mathbb{P}(X \leq 5) = \mathbb{P}(\{\omega \ldots\})$.

III.1 Unabhängigkeit

III.1.1 Definition (Unabhängigkeit von Zufallsvariablen)

Zwei reelle Zufallsvariablen X, Y heißen *unabhängig* $:\Leftrightarrow$

$$\forall B_1, B_2 \in \mathcal{B}_\mathbb{R} \colon X \in B_1 \text{ ist unabhängig von } Y \in B_2.$$

III.1.2 Bemerkung

Es reicht, wenn $B_1, B_2 \in \mathcal{E}$ für einen Erzeuger $\mathcal{E}$ von $\mathcal{B}_\mathbb{R}$ unabhängige Ereignisse $X \in B_1$ und $Y \in B_2$ ergeben.

III.1.3 Beispiel

Seien $X, Y\colon \Omega \to \{0,1\}$ $\mathcal{A}$-$\mathcal{B}_{\mathbb{R}}$-messbar.

x/y	0	1	
0	$p_x \cdot p_y$	✓	p_x
1	✓	✓	q_x
	p_y	q_y	

A unabhängig von $B \Rightarrow A$ unabhängig von B^C.

III.1.4 Definition (paarweise Unabhängigkeit)

Sei $(A_i)_{i\in\mathcal{I}}$ Ereignisse.

a) Die $(A_i)_{i\in\mathcal{I}}$ heißen *paarweise unabhängig* $:\Leftrightarrow A_j$ unabhängig von A_k, $\forall j \neq k \in \mathcal{I}$,

b) die $(A_i)_{i\in\mathcal{I}}$ heißen *unabhängig* $:\Leftrightarrow \forall \mathcal{J} \subset \mathcal{I}$ ist $\mathbb{P}(\bigcap_{j\in\mathcal{J}} A_j) = \prod_{j\in\mathcal{J}} \mathbb{P}(A_j)$ (J ist abzählbar).

III.1.5 Definition (paarweise Unabhängigkeit von Zufallsvariablen)

Seien $(X_i)_{i\in\mathcal{I}}$ Zufallsvariablen.

a) Die $(X_i)_{i\in\mathcal{I}}$ heißen *paarweise unabhängig* $:\Leftrightarrow X_j, X_k$ unabhängig $\forall j \neq k \in \mathcal{I}$,

b) die $(X_i)_{i\in\mathcal{I}}$ heißen *unabhängig* $:\Leftrightarrow \forall \mathcal{J} \subset \mathcal{I}$ abzählbar und $\forall (B_j)_{j\in\mathcal{J}} \subset \mathcal{B}_{\mathbb{R}}$ gilt $\mathbb{P}(\bigcap_{j\in\mathcal{J}} X_j \in B_j) = \prod_{i\in\mathcal{J}} \mathbb{P}(X_j \in B_j)$.

III.1.6 Bemerkung

Bei den Anwendungen werden wir meist Situationen betrachten, bei denen a priori klar ist, dass die Zufallsvariablen unabhängig sind (*kausale Unabhängigkeit*).

III.1.7 Beispiel

Seien $X, Y\colon \Omega \to \{0,1\}$ unabhängige Zufallsvariablen. Sei

$$Z := \begin{cases} Z(\omega) = 1 & \text{falls } X(\omega) + Y(\omega) \text{ ungerade,} \\ Z(\omega) = 0 & \text{sonst.} \end{cases}$$

Annahme: $\mathbb{P}(X = 1) = \frac{1}{2},\ \mathbb{P}(Y = 1) = \frac{1}{2}$.

- $\mathbb{P}(Z = 1) = \frac{1}{2}$,
- $\mathbb{P}(X = 1 \text{ und } Z = 1) = \mathbb{P}(X = 1 \text{ und } Y = 0) = \frac{1}{4}$.

Nach obiger Vierfeldertafel reicht dies für die Unabhängigkeit von X und Z. Die Unabhängigkeit von Y und Z ist analog. Für Unabhängigkeit müsste (u.a.) gelten

$$0 = \mathbb{P}(X = 1 \text{ und } Y = 1 \text{ und } Z = 1) \stackrel{?}{=} \mathbb{P}(X = 1) \cdot \mathbb{P}(Y = 1) \cdot \mathbb{P}(Z = 1) = \frac{1}{8} \text{ ↯}.$$

III.2 Verteilungsfunktion

Wir betrachten beliebige reelle Zufallsvariablen, d.h.

$$(\Omega, \mathcal{A}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \mathbb{Q} = \mathbb{P} \cdot X^{-1}).$$

Wegen des Eindeutigkeitssatzes ist dieses $\mathbb{Q}$ durch $\mathbb{P}(X \leq a)$ $\forall a \in \mathbb{R}$ eindeutig! Letzteres ist eine Funktion in der Variable a. Diese nennt man *Verteilungsfunktion*.

III.2.1 Definition (Verteilungsfunktion)

$X : \Omega \to \mathbb{R}$ Zufallsvariablen.

$$V_X : \mathbb{R} \to [0, 1]$$

gegeben durch $V_X(a) = \mathbb{P}(X \leq a)$ nennt man die zu X gehörige *Verteilungsfunktion*.

III.2.2 Satz

Jede Verteilungsfunktion V_X hat folgende Eigenschaften.

a) $\lim_{a\to\infty} V_x(a) = 1,\ \lim_{a\to-\infty} V_x(a) = 0$,

b) V_X ist monoton wachsend,

c) V_X ist rechtsseitig stetig,

d) V_X hat höchstens abzählbar viele Sprungstellen.

III.2.3 Beispiel

a) X ist Augenzahl eines fairen Würfels.

b) $\mathbb{P} \circ X^{-1}$ habe die Dichte

$$\rho(x) = \begin{cases} 1 & \text{falls } x \in \ [-\frac{1}{2}, \frac{1}{2}], \\ 0 & \text{sonst.} \end{cases}$$

III.2.4 Definition (Wahrscheinlichkeitsdichte)

Sei X Zufallsvariable, $B \in \mathcal{B}$. Dann nennt man $\rho \colon \mathbb{R} \to \mathbb{R}$ mit

$$\mathbb{P}(X \in B) = \int_B \rho(x) dx$$

eine *Wahrscheinlichkeitsdichte* von X.

[27.05.2024]

III.2.5 Beweis

a) $\lim_{a\to\infty} V_X(a) = \lim_{a\to\infty} \mathbb{P}(X \leq a)$. Betrachte Folge $a_n = n$.

$$\lim_{n\to\infty} \{X \leq a_n\} = \{X \in \mathbb{R}\}.$$

$\lim_{n\to\infty} \mathbb{P}(X \leq a_n) = \mathbb{P}(\lim_{n\to\infty} X \leq a_n) = 1$. Dies reicht aus wegen b). Der zweite Teil geht analog. Betrachte $\{X \leq a_n\}$ für $a_n = -n$.

$$\lim_{n\to\infty} \mathbb{P}(X \leq a_n) = \mathbb{P}(\lim_{n\to\infty} X \leq a_n) = \mathbb{P}(\emptyset) = 0.$$

b) Sei $a \leq b$.

$$\begin{aligned} V_X(b) &= \mathbb{P}(X \leq b) = \mathbb{P}(X \leq a \cap X \in\,]a,b]) \\ &= \underbrace{\mathbb{P}(X \leq a)}_{V_X(a)} + \underbrace{\mathbb{P}(X \in\,]a,b])}_{\geq 0} \end{aligned}$$

$\Rightarrow V_X(b) \geq V_X(a)$.

c) Sei $(h_n)_{n\in\mathbb{N}}$ eine positive Nullfolge. (Rechtsseitig stetig: Für alle solchen $(h_n)_{n\to\mathbb{N}}$ gilt: $\lim_{n\to\infty} V_X(a+h_n) = V_X(a)$.) Sei zunächst h_n monoton fallend.

$$\begin{aligned} \lim_{n\to\infty} V_X(a+h_n) &= \lim_{n\to\infty} \mathbb{P}(X \leq a + h_n) \underbrace{=}_{\text{monoton fallend}} \mathbb{P}(\lim_{n\to\infty} X \leq a + h_n) \\ &= \mathbb{P}(\bigcap_{n\to\mathbb{N}} X \leq a + h_n) = \mathbb{P}(X \leq a). \end{aligned}$$

Für allgemeines h_n betrachte $a_n := \sup_{k\geq n} h_k$, die a_n sind fallend und $a_n \geq h_n\ \forall n \in \mathbb{N}$.

$$\lim_{n\to\infty} V_X(a+a_n) = V_X$$
$$V_X(a+a_n) \geq V_X(a+h_n) \geq V_X(a), \text{ wegen Monotonie von } V.$$

d) Sei $n \in \mathbb{N}$. Frage: Wie viele Sprungstellen hat V_X mit Höhe $geq\frac{1}{n}$? Wegen a), Antwort: höchstens n.

Die Menge aller Sprungstellen ist:

$$\bigcup_{n\to\mathbb{N}} \{x \in \mathbb{R}\colon \text{Sprunghöhe ist} \geq \frac{1}{n}\}.$$

Diese ist als abzählbare Vereinigung abzählbarer Mengen abzählbar.

III.2.6 Beispiel

Wir werfen zunächst eine Münze, danach werfen wir einen Pfeil auf eine Dartscheibe:

$$\Omega = \{K, Z\} \times \{\omega \in B_0(1)\}.$$

Sei X Zufallsvariable definiert durch

$$X((a,\omega)) := \begin{cases} \frac{1}{2} & \text{falls } a = K, \\ |\omega| & \text{falls } a = Z. \end{cases}$$

$b < \frac{1}{2}$:

$$V_X(b) = \mathbb{P}(X \leq b) = \mathbb{P}(a = Z \text{ und } |\omega| \leq b) = \frac{1}{2} \cdot b^2.$$

$b = \frac{1}{2}$:

$$\begin{aligned} V_X(\frac{1}{2}) &= \mathbb{P}(a = K \text{ oder } (a = Z \text{ und } |\omega| \leq \frac{1}{2})) \\ &= \frac{1}{2} + \frac{1}{2}\frac{1}{4} = \frac{5}{8}. \end{aligned}$$

$b > \frac{1}{2}$:

$$V_X(b) = \mathbb{P}(a = K \text{ oder } (a = Z \text{ und } |\omega| \leq b)) = \frac{1}{2} + \frac{1}{2} \cdot b^2.$$

III.3 Besondere Verteilungen

III.3.1 Binomialverteilung

n-facher Münzwurf (Unabhängigkeit vorausgesetzt). Man untersucht die Häufigkeit von "Kopf" (bzw. Zahl).

$$\mathbb{P}(k\text{-mal Kopf}) = \binom{n}{k} p^k q^{n-k}$$

mit p =Wahrscheinlichkeit für "Kopf" bei einem Wurf, $q = 1 - p$.

III.3.2 Definition (binomialverteilt)

Sei X eine Zufallsgröße $\Omega \to \{0, 1, \ldots, n\}$. X heißt *binomialverteilt*, falls

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

mit Parameter p. $\mathbb{P}$ erfüllt die Axiome von Kolmogoroff (siehe Satz):

a) Elementarwahrscheinlichkeit nicht negativ:

$$\sum_{k=0}^{n} \mathbb{P}(X = k) \stackrel{?}{=} 1 = (p + q)^n = \sum_{k=0}^{n} p^k q^{n-k} \binom{n}{k}.$$

III.3.3 Poisson-Verteilung

Wir betrachten obigen Fall für sehr kleines p, jedoch n so groß, dass wir "Treffer" erwarten können. Wir setzen $p \cdot n = \lambda$ und betrachten sehr kleine p bzw. große n.

Wir nutzen folgende Formel:

$$\begin{aligned} (1 - \frac{\lambda}{n})^n &\stackrel{?}{\approx} e^{-\lambda}, \text{ für } n \gg 1 \\ \Leftrightarrow n \cdot \log(1 - \frac{\lambda}{n}) &\stackrel{?}{\approx} -\lambda \end{aligned}$$

Lineare Approximation: $f(x+\varepsilon) \approx f(x) + f'(x) \cdot \varepsilon$,

$$\Rightarrow n \cdot 1 \cdot \frac{-\lambda}{n} = -\lambda.$$

$$\begin{aligned}
\mathbb{P}(k \text{ Treffer}) &= \binom{n}{k} \cdot \left(\frac{\lambda}{n}\right)^k \cdot \left(1 - \frac{\lambda}{n}\right)^{n-k} \\
&\approx \frac{\cancel{n!}}{k! \cancel{(n-k)!}} \cdot \frac{\lambda^k}{\cancel{n^k}} \left(1 - \frac{\lambda}{n}\right)^n \\
&\approx \frac{\lambda^2}{k!} e^{-\lambda}.
\end{aligned}$$

III.3.4 Definition (Poisson-verteilt)

Falls

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda},$$

nennt man X *Poisson-verteilt.* Sie gibt Bernoulli-Situationen, d.h. unabhängige 0-1-Fragen, wie oben, für sehr unwahrscheinliche Ereignisse bei hoher "Strichprobe" wieder. Da

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} a^{\lambda} = 1,$$

wird so in der Tat ein Wahrscheinlichkeitsmaß definiert.

III.3.5 Bemerkung

Namensgebend ist das französische Wort *poisson* (dt. Fisch). Die Verteilung wurde während der Untersuchung des großen Fischsterbens im Jahr 1873 in Irkutsks, Sibirien, entdeckt [1]. Hierbei wurde der Parameter λ als *la grange* (dt. die Scheune) bezeichnet, da er den Erwartungswert und die Varianz beschreibt und somit die Verteilung charakterisiert. Der Parameter k wurde *la place* (dt. die Stelle) genannt, da er die Anzahl der Ereignisse beschreibt.

[03.06.2024]

III.3.6 Definition (normalverteilt)

Eine Zufallsgröße X mit der Wahrscheinlichkeitsdichte

$$\rho(x) := \frac{1}{C} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

für die Parameter $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$ und C, sodass $\int_{-\infty}^{\infty} \rho(x)\, dx = 1$ nennt man *normalverteilt.*

Erinnerung: $\mathbb{P}(X \in [a,b]) = \int_a^b \rho(x)\, dx$.

III.3.7 Bestimmung von C

$$
\begin{aligned}
\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx &= \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}}\, dx \\
&\overset{y\,=\,\frac{x}{\sigma}}{=} \sigma \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\, dy \\
&\overset{(*)}{=} \sigma \cdot \sqrt{2\pi}
\end{aligned}
$$

$(*)$:

$$
\begin{aligned}
\left(\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\, dy\right)^2 &= \int\int e^{-\frac{y^2}{2}}\, dy\; e^{-\frac{z^2}{2}}\, dz \\
&= \int\int e^{-\frac{y^2+z^2}{2}}\, dydz \\
&\overset{\text{Polark.}}{=} \int_0^{\infty} 2r\pi \cdot e^{-\frac{r^2}{2}}\, dr \\
&= 2\pi[\,\underbrace{-e^{-\frac{r^2}{2}}}_{\prime\, =\, re^{\frac{r^2}{2}}}\,]_0^{\infty} \\
&= 2\pi(0-(-1)) = 2\pi \cdot 1
\end{aligned}
$$

IV Invarianzen

IV.1 Erwartungswert

Wir denken an ein Glückspiel mit Geldeinsatz, welches wir häufig und unbhängig wiederholen. Welchen Gewinn/Verlust erwarten wir im Mittel zu erreichen?

IV.1.1 Definition (Erwartungswert diskret)

Sei $X : \Omega \to \mathbb{R}$ eine Zufallsgröße. $|X(\Omega)|$ sei endlich. Dann ist

$$\mathbb{E}(X) := \sum_{a \in X(\Omega)} a \cdot \mathbb{P}(X = a)$$

der *Erwartungswert* von X.

IV.1.2 Bemerkung

Die Definition kann auf abzählbare Mengen direkt erweitert werden. Hier kann es passieren, dass $\mathbb{E}$ nicht existiert. Beipsiel:

$$\mathbb{P}(X = a) = \frac{C}{a^2},\ a \in \mathbb{N}.$$

Für das Beispiel wäre $\mathbb{E}(X) = +\infty$, für andere Beispiele schlägt auch diese "uneigentliche" Definition fehl.

IV.1.3 Satz

Falls Ω endlich ist, gilt:

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega) \cdot \mathbb{P}(\{\omega\}).$$

IV.1.4 Beispiel

Man wirft einen Würfel

$$X(1) = -1,\ X(2) = -1,\ X(3) = -1,\ X(4) = -1,\ X(5) = 0,\ X(6) = 3.$$

Also ist

$$\mathbb{E}(x) = -1 \cdot \frac{1}{6} + \ldots 3 \cdot \frac{1}{6} = (-1) \cdot (\frac{4}{6}) + 0 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6}$$

IV.1.5 Beweis

Distributivgesetz:

$$\mathbb{E} = -\int_{-\infty}^{0} V_X(a)\, da + \int_{0}^{\infty} 1 - V_X(a)\, da.$$

IV.1.6 Bemerkung

Letztere Formel lässt uns den Begriff Erwartungswert verallgemeiner. Sowohl der positive als auch der negative Teil existieren immer (zumindest uneigentlich). Einziges mögliches Problem: "$+\infty + -\infty$".

Eine andere Möglichkeit besteht mithilfe des Lebesgue-Integrals:

$$\mathbb{E}(x) = \int_{-\infty}^{\infty} x \, d\widetilde{\mathbb{P}}.$$

Man integriert bzgl. des Maßes $\widetilde{\mathbb{P}} = \mathbb{P} \cdot X^{-1}$.

IV.1.7 Satz

Für stetig verteilte Zufallsvariablen X, d.h. solch, die eine Dichte besitzen, ist

$$\mathbb{E}(X) = \int x\rho(x) \, dx.$$

IV.1.8 Beweis

Lassen wir weg.

IV.1.9 Satz

Der Erwartungswert ist linear, d.h. gegeben $X, Y : \Omega \to \mathbb{R}$ Zufallsvariablen, $a \in \mathbb{R}$:

$$\mathbb{E}(aX + Y) = a\mathbb{E}(X) + \mathbb{E}(Y),$$

falls $\mathbb{E}(X)$ und $\mathbb{E}(Y)$ existieren.

IV.1.10 Beweis

a) diskret:

$$\begin{aligned}
\mathbb{E}(X + aY) &= \sum_{z \in (X+aY)(\Omega)} z\mathbb{P}(X + aY = z) \\
(\text{falls } |\Omega| < \infty) \quad &= \sum_{\omega \in \Omega} (X + aY)(\omega) \cdot \mathbb{P}(\{\omega\}) \\
&= \sum_{\omega \in \Omega} (X(\omega) + aY(\omega))(\omega) \cdot \mathbb{P}(\{\omega\}) \\
&= \mathbb{E}(X) + a\mathbb{E}(Y)
\end{aligned}$$

b) Im stetigen Fall kann man ρ als Treppenstufenfunktion approximieren und die Linearität folgt mit a).

IV.1.11 Bemerkung

Die Formel $\mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$ gilt im Allgemeinen <u>nicht</u>. Beispiel: $X = 1$, falls $\omega =$ Zahl, $X = 0$ sonst, $Y = X$.

IV.1.12 Definition (fast sicher)

Ein Ereignis A heißt "*fast sicher*", falls $\mathbb{P}(A) = 1$.

IV.1.13 Satz

a) Falls $X = Y$ fast sicher, so ist $\mathbb{E}(X) = \mathbb{E}(Y)$.

b) Sei $X \equiv \mu \in \mathbb{R}$, so ist $\mathbb{E}(X) = \mu$.

IV.1.14 Beweis

a) $\mathbb{E}(X - Y)$ ist fast sicher $= 0 \Rightarrow \mathbb{E}(X - Y) = 0$,

b) klar.

[06.06.2024]

Verteilung mehrerer Zufallsvariablen

Peter Pickl

Mathematisches Institut, Universität Tübingen

5. Juni 2024

Mehrere Zufallsvariablen

- Gegeben mehrere Zufallsvariablen X_i mit $i \in I$ für eine Indexmenge I.
- Wir möchten die Wahrscheinlichkeiten von Ereignissen behandeln, die sich durch die X_i ausdrücken lassen
- Zum Beispiel $\cap_{i\in\mathbb{N}} X_i \in B_i$, wobei die B_i Elemente der Borel'schen σ-Algebra sind.
- So lange wir nur Verknüpfungen zwischen den Ereignissen zulassen, die nach den Axiomen der σ-Algebra legitim sind, erhalten wir so jeweils ein Ereignis.
- Legitim sind natürlich abzählbare Schnitte und Vereinigung, Differenz- und Komplementbildung.

Verteilungsfunktion

Es seien nun zwei Zufallsvariablen X, Y gegeben (folgende Betrchtungen gehen für mehr als 2 Zufallsvariablen analog).
Wir möchten das entsprechende Wahrscheinlichkeitsmass für durch X und Y ausdrückbare Ereignisse zusammenfassen.
Dies machen wir wieder mit Hilfe der Verteilungsfunktion

Definition: Seien X, Y Zufallsvariablen. Die Funktion $V_{X,Y} : \mathbb{R}^2 \to [0, 1]$ gegeben durch

$$V_{X,Y}(a, b) := \mathbb{P}\left(X \leq a \text{ und } Y \leq b\right)$$

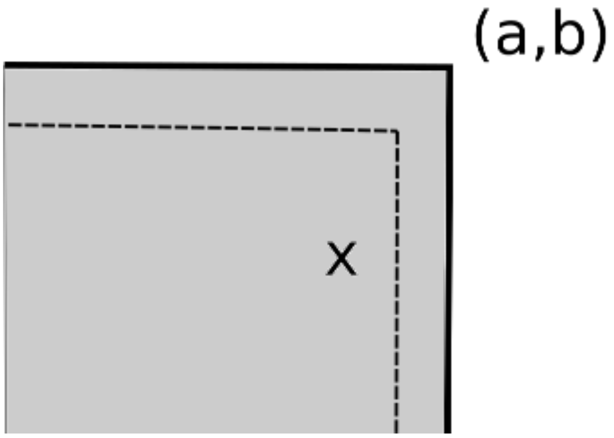
nennt man <u>gemeinsame Verteilungsfunktion von X und Y.</u>

Eindeutigkeit des Masses

Satz: Durch $V_{X,Y}$ sind für beliebige Ereignisse $(X, Y) \in \mathcal{B}(\mathbb{R}^2)$ die Wahrscheinlichkeiten eindeutig festgelegt.
Beweis: Wir benutzen den Eindeutigkeitssatz. Dazu brauchen wir nur zu zeigen, dass die Mengen $\mathcal{E} := \left\{] - \infty, a] \times] - \infty, b] \text{ für } (a, b) \in \mathbb{R}^2\right\}$ die Borel'sche-σ-Algebra erzeugen. Das Mengensystem $\mathcal{E}$ ist nämlich schnittstabil, dadurch greift der Eindeutigkeitssatz!
Durch $\mathcal{E}$ lassen sich zunächst alle Mengen der Form

- $] - \infty, a] \times] - \infty, b[$
- $] - \infty, a[\times] - \infty, b]$
- und $] - \infty, a[\times] - \infty, b[$

erzeugen. Zum Beispiel
$] - \infty, a[\times] - \infty, b[\; = \; \bigcup_{n\in\mathbb{N}}] - \infty, a - \frac{1}{n}] \times] - \infty, b - \frac{1}{n}]$



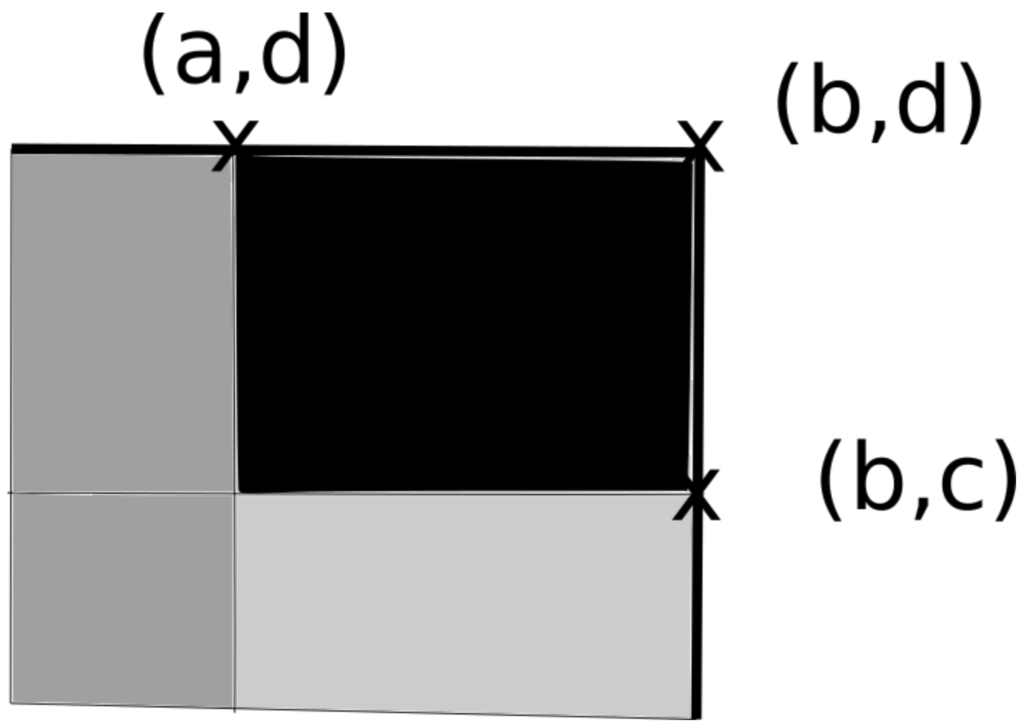
Eindeutigkeit des Masses

Satz: Durch $V_{X,Y}$ sind für beliebige Ereignisse $(X, Y) \in \mathcal{B}(\mathbb{R}^2)$ die Wahrscheinlichkeiten eindeutig festgelegt.
Beweis: Wir benutzen den Eindeutigkeitssatz. Dazu brauchen wir nur zu zeigen, dass die Mengen $\mathcal{E} := \{]-\infty, a]\times]-\infty, b]$ für $(a, b) \in \mathbb{R}^2\}$ die Borel'sche-σ-Algebra erzeugen. Das Mengensystem $\mathcal{E}$ ist nämlich schnittstabil, dadurch greift der Eindeutigkeitssatz!
Durch $\mathcal{E}$ lassen sich zunächst alle Mengen der Form

- $]-\infty, a] \times]-\infty, b[$
- $]-\infty, a[\times]-\infty, b]$
- und $]-\infty, a[\times]-\infty, b[$

erzeugen. Zum Beispiel
$]-\infty, a[\times]-\infty, b[\; = \; \bigcup_{n\in\mathbb{N}}]-\infty, a-\frac{1}{n}] \times]-\infty, b-\frac{1}{n}]$
Es gilt

$$]a, b[\times]c, d[\; =]-\infty, b[\times]-\infty, d[\setminus (]-\infty, a] \times]-\infty, d] \cup]-\infty, b] \times]-\infty, c])$$



Eindeutigkeit des Masses

Satz: Durch $V_{X,Y}$ sind für beliebige Ereignisse $(X, Y) \in \mathcal{B}(\mathbb{R}^2)$ die Wahrscheinlichkeiten eindeutig festgelegt.
Beweis:
Die von $\mathcal{E}$ erzeugte σ-Algebra enthält also die von den offenen Rechtecken erzeugte σ-Algebra.
ŽDass sich durch die offenen Rechtecke alls offenen Mengen erzeugen lassen, geht identisch wie der entsprechende Beweis in $\mathbb{R}$
Wähle eine beliebige offene Teilmenge A von $\mathbb{R}^2$. Für jedes $(a, b) \in \mathbb{Q}^c \cap A$ bilde das größtmögliche offene Rechteck, welches (a, b) enthält.
Die Vereinigung all dieser Rechtecke ist dann gleich der Menge A

Eigenschaften

Definition: Seien X, Y Zufallsvariablen. Die Funktion $V_{X,Y} : \mathbb{R}^2 \to [0,1]$ gegeben durch

$$V_{X,Y}(a,b) := \mathbb{P}\left(X \le a \text{ und } Y \le b\right)$$

nennt man <u>gemeinsame Verteilungsfunktion von X und Y.</u>

Satz: $V_{X,Y}$ hat folgende Eigenschaften:

(a) Für jede Folge $(a_n, b_n)_{n\in\mathbb{N}}$ mit $\lim_{n\to\infty} a_n = -\infty$ oder $\lim_{n\to\infty} b_n = -\infty$ gilt: $\lim_{n\to\infty} V_{X,Y}(a_n, b_n) = 0$.

(b) Für jede Folge $(a_n)_{n\in\mathbb{N}}$ mit $\lim_{n\to\infty} a_n = \infty$ und $\lim_{n\to\infty} b_n = \infty$ gilt: $\lim_{n\to\infty} V_{X,Y}(a_n, b_n) = 1$.

(c) Stetigkeit von rechts-oben: Für jede Folge $(s_n, t_n)_{n\in\mathbb{N}} \subset \mathbb{R}_0^+ \times \mathbb{R}^+$ mit $\lim_{n\to\infty}(s_n, t_n) = 0$ und für jedes $(a,b) \in \mathbb{R}^2$ gilt

$$\lim_{n\to\infty} V_{X,Y}(a + s_n, b + t_n) = V_{X,Y}(a,b)\ .$$

Eigenschaften

(a) Für jede Folge $(a_n, b_n)_{n\in\mathbb{N}}$ mit $\lim_{n\to\infty} a_n = -\infty$ oder $\lim_{n\to\infty} b_n = -\infty$ gilt: $\lim_{n\to\infty} V_{X,Y}(a_n, b_n) = 0$.
Beweis: Es sei o.B.d.A. $\lim_{n\to\infty} a_n = -\infty$. Da das Ereignis "$X \le a_n$ und $Y \le b_n$" eine Teilmenge von $X \le a_n$ ist, gilt

$$V_{X,Y}(a_n, b_n) \le V_X(a_n)$$

Da letztere im Limes $n \to \infty$ Null ergibt, folgt die Aussage ($\lim_{n\to\infty} V_{X,Y}(a_n, b_n) = 0$).

Eigenschaften

(b) Für jede Folge $(a_n)_{n\in\mathbb{N}}$ mit $\lim_{n\to\infty} a_n = \infty$ und $\lim_{n\to\infty} b_n = \infty$ gilt: $\lim_{n\to\infty} V_{X,Y}(a_n, b_n) = 1$.

Beweis: Betrachte $c_n := \inf_{k\geq n} a_n$ und $d_n := \inf_{k\geq n} b_n$. Es gilt offensichtlich, dass $V_{X,Y}(c_n, d_n) \leq V_{X,Y}(a_n, b_n)$.
Es reicht also zu zeigen, dass $\lim_{n\to\infty} V_{X,Y}(c_n, d_n) = 1$.
Die Folgen c_n und d_n sind monoton wachsend und ergeben im Limes ∞.
Daher ist die Folge der Ereignisse "$X \leq c_n$ und $Y \leq d_n$" eine monoton wachsende Folge mit $\lim_{n\to\infty}]-\infty, c_n]\times]-\infty, d_n] = \mathbb{R}^2$.
Somit ist wegen der Stetigkeit von $\mathbb{P}$ der Limes
$\lim_{n\to\infty} V_{X,Y}(c_n, d_n) = \lim_{n\to\infty} \mathbb{P}(X \leq c_n \text{ und } Y \leq d_n) = \mathbb{P}\left(\mathbb{R}^2\right) = 1$.

Eigenschaften

(c)Stetigkeit von rechts-oben: Für jede Folge $(s_n, t_n)_{n\in\mathbb{N}} \subset \mathbb{R}_0^+ \times \mathbb{R}^+$ mit $\lim_{n\to\infty}(s_n, t_n) = 0$ und für jedes $(a, b) \in \mathbb{R}^2$ gilt

$$\lim_{n\to\infty} V_{X,Y}(a + s_n, b + t_n) = V_{X,Y}(a, b) .$$

Beweis: Wie bei der einfachen Verteilungsfunktion können wir uns o.B.d.A auf monoton fallende s_n und t_n einschränken.
Es ist dann $X \leq a + s_n$ und $Y \leq b + t_n$ eine fallende Folge von Ereignissen.
Somit ist
$\lim_{n\to\infty} V_{X,Y}(a + s_n, b + t_n) = \mathbb{P}\left(\bigcap_{n\in\mathbb{N}} X \leq a + s_n \text{ und } Y \leq b + t_n\right) = \mathbb{P}(X \leq a \text{ und } Y \leq b) = V_{X,Y}(a, b)$.

Unabhängige Zufallsgrößen

Im Falle unabhängiger Zufallsgrößen ist die Situation besonders.

- zum Beispiel ist $V_{X,Y}(a,b) = \mathbb{P}(X \leq a \text{ und } Y \leq b) = \mathbb{P}(X \leq a)\mathbb{P}(Y \leq b) = V_X(a)V_Y(b)$.
- Wie sieht es mit der Summe von zwei unabhängigen Zufallsvariablen aus?
- Wir betrachten dazu die Zufallsvariablen X und Y mit den Wahrscheinlichkeitsdichten ρ und η.
- Man erhält wegen der Unabhängigkeit durch das Produkt von ρ und η eine Dichte für (X, Y).
- Es ist also $\mathbb{P}((X,Y) \in A) = \int_A \rho(x)\eta(y)dxdy$.

Verteilungsfunktion

- Insbesondere ist in diesem Fall $V_{X+Y}(a) = \mathbb{P}(X + Y \leq a)$ gegeben durch $\int_{-\infty}^{\infty}\int_{-\infty}^{a-y} \rho(x)\eta(y)dxdy$
- Also ist $V_{X+Y}(a) = \mathbb{P}(X + Y \leq a) = \int V_X(a-y)\eta(y)dy$
- Da die $V_X(a) = \int_{-\infty}^{a} \rho(x)dx$ ist die Dichte immer die Ableitung von V_X.
- $X + Y$ hat daher an der Stelle a die Wahrscheinlichkeitsdichte $\int \rho(a-y)\eta(y)dy$.
- Den letzten Ausdruck nennt man auch die Faltung von ρ mit η
- Schreibweise: $\rho \star \eta := \int \rho(a-y)\eta(y)dy = \int \eta(a-y)\rho(y)dy$

IV.1.15 Satz

a) Falls $X \geq 0$ fast sicher, so ist $\mathbb{E}(x) \geq 0$.

b) Falls $X \leq Y$ fast sicher, so ist $\mathbb{E}(x) \leq \mathbb{E}(Y)$.

c) $\mathbb{E}(|X|) = 0 \Leftrightarrow X = 0$ fast sicher.

d) $\mathbb{E}(|X|) \geq \mathbb{E}(X)$.

IV.1.16 Beweis

a) $\mathbb{P}(X \leq 0) = 0$, $\mathbb{P}(X \leq a) = 0$, $\forall a < 0$. Die Verteilungsfunktion ist 0 $\forall a < 0 \Rightarrow$ der negative Anteil von $\mathbb{E}$ ist $0 \Rightarrow \mathbb{E}(X) \geq 0$.

b) $X \leq Y \Rightarrow Y - X \geq 0 \Rightarrow \mathbb{E}(Y - X) \geq 0 \Rightarrow \mathbb{E}(Y) \geq \mathbb{E}(X)$.

c)
- "$\Leftarrow$": Sei $X = 0$ f.s. $\Leftrightarrow |X| = 0$.

$$V_X(a) = \begin{cases} 0 & \text{für } a < 0 \\ 1 & \text{für } a \geq 0 \end{cases}$$

  $\Rightarrow \mathbb{E}(|x|) = 0$.
- "$\Rightarrow$": Sei $\mathbb{E}(|X|) = 0$. WA: $X = 0$ f.s. gilt nicht $\Rightarrow \exists \varepsilon > 0$ mit $\mathbb{P}(|X| > \varepsilon) > 0$. Falls $\mathbb{P}(|X| < \frac{1}{n}) = 0$, wäre für alle $n \in \mathbb{N}$, hätten wir $\mathbb{P}(\underbrace{\bigcup_{n \in \mathbb{N}} |X| < \frac{1}{n}}_{|X| \neq 0}) = 0$. $\mathbb{P}(|X| > \varepsilon := p > 0 \Rightarrow V_{|X|}(\varepsilon) = 1 - p < 1 \Rightarrow \mathbb{E}(|X|) > 0$ ↯.

d) Es gilt $|X(\omega)| \geq X(\omega)$, $\forall \omega$, mit b) folgt Behauptung.

IV.2 Varianz

Wir denken an ein Glücksspiel. Neben dem Erwartungswert, interessiert uns auch das Risiko.

Das Risiko kann auf unterschiedliche Weise quantifiziert werden. Aufgrund ihrer besonderen Eigenschaften, nehmen wir als Maß für das Risiko die "Varianz".

IV.2.1 Definition (Varianz)

Sei X Zufallsvariable mit $\mathbb{E}(X) =: \mu \in \mathbb{R}$. Dann nennt man

$$\operatorname{Var} X = \mathbb{E}((X - \mu)^2)$$

die *Varianz* von X.

IV.2.2 Satz

a) $\operatorname{Var} X \geq 0$ für alle Zufallsvariablen X.

b) $\operatorname{Var} X = 0 \Leftrightarrow X = \mu$ f.s.

c) $\operatorname{Var}(X + a) = \operatorname{Var} X,\ \forall a \in \mathbb{R}$.

d) $\operatorname{Var}(aX) = a^2 \cdot \operatorname{Var} X$.

e) $\operatorname{Var} X = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.

IV.2.3 Beweis

a) Ist eine Folgerung aus a) des letztes Satzes: $(X - \mu)^2 \geq 0$.

b) Ist eine direkte Folgerung asu c) des letzten Satzes, $(X - \mu)^2 = 0 \Leftrightarrow |X - \mu| = 0 \Rightarrow X - \mu = 0$ f.s. $\Leftrightarrow X = \mu$ f.s.

c) $\operatorname{Var}(X + a) = \mathbb{E}[(X + a - \mathbb{E}(X + a))^2] = \mathbb{E}(X + a - \mathbb{E}(X) - a)^2] = \operatorname{Var} X$.

d) $\operatorname{Var}(aX) = \mathbb{E}[(aX - \mathbb{E}(aX)^2] = \mathbb{E}[(aX - a\mathbb{E}(x))^2] = \mathbb{E}[a^2(X - \mathbb{E}(X))^2] = a^2\mathbb{E}[(X - \mathbb{E}(X))^2] = a^2 \operatorname{Var} X$.

e) $\operatorname{Var} X = \mathbb{E}[(X - \mu)^2] = \mathbb{E}(X^2 - 2\mu X + \mu^2) = \mathbb{E}(X^2) - 2\mu \underbrace{\mathbb{E}(X)}_{\mu} + \mu^2 = \mathbb{E}(X^2) - \mu^2$.

IV.2.4 Definition (Standardabweichung)

Sei X eine Zufallsvariable mit $\mathbb{E}(X) \in \mathbb{R}$. Dann nennt man

$$\sigma(x) := \sqrt{\operatorname{Var} X}$$

die *Standardabweichung* von X.

IV.2.5 Bemerkung

σ misst das "Risiko" in den selben Einheiten wie X und μ.

IV.2.6 Beispiele

a) Binomialverteilung: Parameter: p, n:

$$\mathbb{P}(X = k) = \binom{n}{k} p^k q^{n-k} (q = 1 - p).$$

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k=0}^{n} k \binom{n}{k} p^k q^{n-k} = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} p^k q^{n-k} \\
&= \sum_{k=1}^{n} n \cdot \underbrace{\frac{\cdot (n-1)!}{(k-1)!(n-k)!}}_{\sum \binom{n-1}{k-1} p^{k-1} q^{n-1} = 1} p \cdot \underbrace{p^{k-1} q^{n-k}}_{\leftarrow} \\
&= n \cdot p.
\end{aligned}$$

$$\operatorname{Var} X = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

Betrachte zuerst

$$\begin{aligned}
\mathbb{E}(X \cdot (X-1)) &= \sum_{k=0}^{n} k(k-1) \binom{n}{k} p^k q^{n-k} \\
&= \sum_{k=2}^{n} \frac{k \cdot (k-1) \cdot n!}{k!(n-k)!} p^k q^{n-k} \\
&= p^2 \cdot \sum_{k=2}^{n} \frac{n!}{(k-2)!(n-k)!} p^{k-2} q^{n-k} \\
&= p^2 \cdot n(n-1) \cdot \sum_{k=2}^{n} \frac{(n-2)!}{(k-2)!(n-k)!} p^{k-2} q^{n-k}
\end{aligned}$$

Sei $\tilde{k} = k-2$, $\tilde{n} = n-2$.

$$\begin{aligned}
&= p^2 \cdot n(n-1) \cdot \overbrace{\sum_{\tilde{k}=0}^{\tilde{n}} \frac{(\tilde{n})!}{(\tilde{k})!(\tilde{n}-\tilde{k})!} p^{\tilde{k}} q^{\tilde{n}-\tilde{k}}}^{=1} \\
\Rightarrow \mathbb{E}(X(X-1)) &= p^2 n(n-1).
\end{aligned}$$

$$\begin{aligned}
\operatorname{Var} X &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \\
&= \mathbb{E}(X(X-1)) + \mathbb{E}(X) - \mathbb{E}(X)^2 \\
&= p^2 n(n-1) + np - n^2 p^2 \\
&= p^2 n^2 - p^2 n + np - n^2 p^2 \\
&= pn(1-p) \\
&= n \cdot pq.
\end{aligned}$$

b) Normalverteilung: $\varrho(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

$$\begin{aligned}
\mathbb{E}(X) &= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} (x-\mu) e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx + \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \mu e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx \\
&= \underbrace{\left[\frac{1}{\sigma\sqrt{2\pi}} \sigma^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right]_{-\infty}^{\infty}}_{=0} + \mu \cdot \overbrace{1}^{\text{Normiertheit von } \varrho}.
\end{aligned}$$

$$
\begin{aligned}
\operatorname{Var} X = \mathbb{E}((X-\mu)^2) &= \int (x-\mu)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&\overset{x-\mu=y}{=} \int y^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{y^2}{2\sigma^2}} dy \\
&= \int \underbrace{e^{-\frac{y^2}{2\sigma^2}} \left(\frac{-y}{\sigma^2}\right)}_{\text{Stammfunktion } e^{-\frac{y^2}{2\sigma^2}}} \cdot (-y)\sigma^2 \frac{1}{\sigma\sqrt{2\pi}} dy
\end{aligned}
$$

Erinnerung partielle Integration: $\int_a^b fg dx = -\int_a^b F g' dx + [Fg]_a^b$

$$
\begin{aligned}
&= -\int e^{-\frac{y^2}{2\sigma^2}} (-\sigma^2 \frac{1}{\sigma\sqrt{2\pi}}) dy + \underbrace{\left[e^{-\frac{y^2}{2\sigma^2}} (-y)\sigma \frac{1}{\sigma\sqrt{2\pi}}\right]_{-\infty}^{\infty}}_{=0} \\
&= \sigma^2 \underbrace{\int \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2\sigma^2}} dy}_{=1} \\
&= \sigma^2.
\end{aligned}
$$

IV.3 Korrelation

IV.3.1 Definition (Unkorreliertheit)

Zwei Zufallsvariablen X, Y heißen *unkorreliert* $:\Leftrightarrow$

$$
\operatorname{Var}(X+Y) = \operatorname{Var} X + \operatorname{Var} Y.
$$

IV.3.2 Bemerkung

$$
\begin{aligned}
V(X+Y) &= \operatorname{Var} X + \operatorname{Var} Y \\
\Leftrightarrow \mathbb{E}((X+Y)^2) - \mathbb{E}(X+Y)^2 &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 + \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 \\
\Leftrightarrow \mathbb{E}(X^2) + 2\mathbb{E}(XY) + \mathbb{E}(Y^2) - \mathbb{E}(X)^2 - 2\mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(Y)^2 &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 + \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 \\
\Leftrightarrow \mathbb{E}(XY) &= \mathbb{E}(X)\mathbb{E}(Y).
\end{aligned}
$$

Dies ist eine alternative Definition der Unkorreliertheit.

IV.3.3 Satz

X, Y unabhängig $\Rightarrow X, Y$ unkorreliert.

IV.3.4 Beweis

- Diskreter Fall (Ω endlich):

$$\begin{aligned}
\mathbb{E}(XY) &= \sum_{\omega\in\Omega} X(\omega)Y(\omega)\mathbb{P}(\{\omega\}) \\
&= \sum_{x,y} \sum_{\omega \text{ mit } X(\omega)=x, Y(\omega)=y} xy\mathbb{P}(\{\omega\}) \\
&= \sum_{x,y} xy \cdot \mathbb{P}(X = x \text{ und } Y = y) \\
&\overset{\text{Unabh}}{=} \sum_{xy} xy \cdot \mathbb{P}(X = x)\mathbb{P}(Y = y) \\
&= \sum_{x} \mathbb{P}(X = x) \cdot \sum_{y} \mathbb{P}(Y = y) \\
&= \mathbb{E}(X) \cdot \mathbb{E}(Y).
\end{aligned}$$

IV.3.5 Definition (Kovarianz)

Für zwei Zufallsvariablen X, Y nennt man den Ausdruck

$$\operatorname{Cov}(X, Y) := \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

die *Kovarianz* von X, Y.

IV.3.6 Bemerkung

$\operatorname{Cov}(X, X) = \operatorname{Var} X$.

IV.3.7 Satz

a) $\operatorname{Cov}(aX + Y, Z) = a \operatorname{Cov}(X, Z) + \operatorname{Cov}(Y, Z)$.

b) $\operatorname{Cov}(X, Y) = \mathbb{E}([X - \mathbb{E}(X)][Y - \mathbb{E}(Y)])$.

IV.3.8 Beweis

a)

$$\begin{aligned}
\operatorname{Cov}(aX + Y, Z) &= \mathbb{E}((aX + Y)Z) - \mathbb{E}(aX + Y)\mathbb{E}(Z) \\
&= \mathbb{E}(aXZ + YZ) - [a\mathbb{E}(X) + \mathbb{E}(Y)]\mathbb{E}(Z) \\
&= a\mathbb{E}(XZ) + \mathbb{E}(YZ) - a\mathbb{E}(X)\mathbb{E}(Z) - \mathbb{E}(Y)\mathbb{E}(Z) \\
&= a \operatorname{Cov}(X, Z) + \operatorname{Cov}(Y, Z).
\end{aligned}$$

b)

$$\begin{aligned}
\mathbb{E}([X - \mathbb{E}(X)][Y - \mathbb{E}(Y)]) &= \mathbb{E}(XY - X\mathbb{E}(Y) - \mathbb{E}(X)Y + \mathbb{E}(X)\mathbb{E}(Y)) \\
&= \mathbb{E}(XY) = \mathbb{E}\mathbb{E}(Y) - \mathbb{E}(X)\mathbb{E}(Y) + \mathbb{E}(X)\mathbb{E}(Y) \\
&= \operatorname{Cov}(X, Y).
\end{aligned}$$

IV.3.9 Bemerkung

Zur Quantifizierung von Korrelation ist die Kovarianz nur bedingt geeignet. Sie ergibt einen absoluten Wert ohne Vergleichsreferenz. Falls man eine Zufallsgröße z.B. verdoppelt, verdoppelt sich der Wert.

`[ 17.06.2024 ]`

IV.3.10 Definition (Korrelationskoeffizient)

Für zwei Zufallsvariablen X, Y nentnt man

$$\kappa(X,Y) := \frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var} X \cdot \operatorname{Var} Y}}$$

den *Korrelationskoeffizienten.*

IV.3.11 Bemerkung

$|\kappa(X,Y)| = |\kappa(aX,bY)|,\ \forall a,b \in \mathbb{R}$. Vorzeichenwechsel, genau dann wenn a oder b (exklusiv) negativ.

IV.3.12 Satz

$|\kappa(X,Y)| \leq 1$.

IV.3.13 Beweis (Cauchy-Schwarz)

Wir benutzen Cauchy-Schwarz:

$$|\langle X,Y\rangle| \leq \sqrt{\langle X,X\rangle\langle Y,Y\rangle}.$$

Wir beweisen, dass diese für alle positiv semi-definiten, symmetrischen Bilinearformen gilt. Denn die Kovarianz ist eine solche.

- 1. Fall: $\langle x \neq 0$ und $\langle y,y\rangle \neq 0$. Betrachte $\langle ax \pm by, ax \pm by\rangle$ mit a,b so, dass $\langle ax,ax\rangle = \langle by,by\rangle = 1$.

$$\begin{aligned}\langle ax+by, ax+by\rangle &= a^2\langle x,x\rangle + b^2\langle y,y\rangle \pm 2ab\langle x,y\rangle \\ &= 2(1 \pm \langle x,y\rangle ab)\end{aligned}$$

  $\Rightarrow 1 \pm \langle x,y\rangle ab \geq 0 \Rightarrow \pm\langle X,Y\rangle \geq \frac{-1}{ab} \Rightarrow \pm\langle x,y\rangle \leq \frac{1}{ab}$. Da $\langle ax,ax\rangle = 1 \Rightarrow a^2 = \frac{1}{\langle x,x\rangle} \Rightarrow \frac{1}{a} = \sqrt{\langle x,x\rangle},\ \frac{1}{b} = \sqrt{\langle x,x\rangle} \Rightarrow \pm\langle x,y\rangle \leq \sqrt{\langle x,x\rangle\langle y,y\rangle}$.

- $\langle x,x\rangle \neq 0,\ \langle y,y\rangle = 0$.

$$\langle ax \pm by, ax \pm by\rangle = a^2x^2 \pm 2ab\langle x,y\rangle + b^2\langle y,y\rangle$$

  (b,a ist hier ganz allgemein gewählt), $\Rightarrow a^2x^2 \pm 2ab\langle x,y\rangle \leq 0$. Z.z. $\langle x,y\rangle = 0$. WA: $\langle x,y\rangle = \varepsilon > 0$. Wähle $b = 1$ und a hinreichend klein, sowie das negative Vorzeichen $\Rightarrow \lightning$, z.B. $a = \frac{\varepsilon}{\langle x,x\rangle 2}$.

$$a^2x^2 - 2ab\langle x,y\rangle = \frac{\varepsilon^2}{4\langle x,x\rangle} - \frac{2\varepsilon^2}{2\langle x,x\rangle} \leq 0 \ \lightning.$$

- 3. Fall: $\langle x,x\rangle = \langle y,y\rangle = 0$. Betrachte $0 \geq \langle x \pm y, x \pm y\rangle = \langle x,x\rangle + \langle y,y\rangle \pm 2\langle x,y\rangle \Rightarrow \langle x,y\rangle = 0$.

IV.3.14 Beweis (des Satzes)

Für den Vektorraum der beschränkten Zufallsvariablen nehmen wir die Ungleichung von Cauchy-Schwarz her $\Rightarrow$ Behauptung.

IV.3.15 Bemerkung

Falls $\operatorname{Var} X, \operatorname{Var} Y < \infty$, so ist auch $|\operatorname{Cov}(X,Y)| < \infty$. Dies lässt sich mithilfe eines Widersprucharguments zeigen (grob): WA $|\operatorname{Cov}(X,Y)| = \infty \Rightarrow$ durch Abschneiden von X, Y:

$$\widetilde{X}(\omega) = \begin{cases} X(\omega) & \text{falls } X(\omega) \le C, \\ C & \text{sonst,} \end{cases}$$

$\widetilde{Y}(\omega)$ analog. Man findet für jedes noch so große $M \in \mathbb{R}$ ein $C \in \mathbb{R}$, sodass $|\operatorname{Cov}(\widetilde{X}, \widetilde{Y})| \ge M$, aber $|\operatorname{Cov}(\widetilde{X}, \widetilde{Y})| \le \sqrt{\operatorname{Var} \widetilde{X}}\sqrt{\operatorname{Var} \widetilde{Y}} \le \sqrt{\operatorname{Var} X}\sqrt{\operatorname{Var} Y}$. Wähle $M > \sqrt{\operatorname{Var} X \operatorname{Var} Y} \Rightarrow \lightning$.

IV.3.16 Satz

Sei $\operatorname{Var} X \neq 0$, dann ist

$$\kappa(X, aX + b) = \underbrace{\mathrm{Vz}}_{\text{Vorzeichen}}(a) \in \{\pm 1\}.$$

(Es reicht, $Y = aX + b$ fast sicher, damit $|\kappa(X,Y) = 1|$.)

IV.3.17 Beweis

$$\begin{aligned} \kappa(X, aX + b) &= \frac{\operatorname{Cov}(X, aX + b)}{\sqrt{\operatorname{Var}(X) \operatorname{Var}(aX + b)}} \\ &= \frac{a \operatorname{Cov}(X, X) + \operatorname{Cov}(X, b)}{\sqrt{\operatorname{Var} X \cdot a^2 \operatorname{Var} X}} \\ &= \frac{a \operatorname{Var} X + 0}{|a| \operatorname{Var} X} = \frac{a}{|a|}. \end{aligned}$$

($|\operatorname{Cov}(X,b) \underbrace{\le}_{\text{C.S.}} \sqrt{\operatorname{Var} X \underbrace{\operatorname{Var} b}_{=0}} \Rightarrow \operatorname{Cov}(X,b) = 0$.)

IV.3.18 Satz

Falls $|\kappa(X,Y)| = 1 \Rightarrow Y = aX + b$ fast sicher mit $a, b \in \mathbb{R}$. Es gilt: $a > 0 \Leftrightarrow \kappa(X,Y) > 0$.

IV.3.19 Beispiel

$$\operatorname{Cov}(X,Y) = \operatorname{Cov}(X, Y - aX + aX),$$

wähle a, sodass $\overbrace{Y - aX}^{Z}$ unkorreliert zu X. Da $\operatorname{Cov}(X, Y - aX) = \operatorname{Cov}(X,Y) - a \overbrace{\operatorname{Var} X}^{\neq 0}$, ist dies immer möglich!

$$\kappa(X,Y) = \frac{\operatorname{Cov}(X,Z) + \operatorname{Cov}(X, aX)}{\sqrt{\operatorname{Var} X}\sqrt{\operatorname{Var} Z + aX}} = \frac{a \cdot \operatorname{Var} X}{\sqrt{\operatorname{Var} X (\operatorname{Var} Z + a^2 \operatorname{Var} X)}}.$$

Der größtmögliche Wert wird für $\operatorname{Var} Z = 0$ erreicht, dieses ist genau $1 \Rightarrow \operatorname{Var} Z = 0$. $Z := Y - aX \Rightarrow Y = aX + \underbrace{Z}_{\text{fast sicher konstant}}$.

[20.06.2024]

IV.4 Markov- und Tschebychev-Ungleichung

IV.4.1 Satz (Markov-Ungleichung)

Sei X eine Zufallsgröße, $f\colon \mathbb{R}_0^+ \to \mathbb{R}_0^+$ monoton steigend. Sei $a \in \mathbb{R}^+$ mit $f(a) > 0$. Dann gilt:

$$\mathbb{P}(|X| \geq a) \leq \frac{\mathbb{E}(f(|X|)}{f(a)}.$$

IV.4.2 Beweis

- Sei zunächst $f(x) = x$.

$$\begin{aligned}\mathbb{E}(|X|) &= \mathbb{E}(\mathbb{1}_{|X|\geq a}|X|) + \mathbb{E}(\mathbb{1}_{|X|<a}|X|) \\ &= \mathbb{E}(\underbrace{\mathbb{1}_{|X|\geq a}(|X| - a)}_{\geq 0}) + \mathbb{E}(\underbrace{\mathbb{1}_{|X|\geq a} \cdot a}_{\geq 0}) + \mathbb{E}(\mathbb{1}_{|X|<a}|X|) \\ &\geq \mathbb{E}(\mathbb{1}_{|X|\geq a}) = \mathbb{P}(|X| \geq a) \cdot a.\end{aligned}$$

  $(\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$, da $\mathbb{E}(\mathbb{1}_A) = 0 \cdot \mathbb{P}(\mathbb{1}_A = 0) + \mathbb{P}(\mathbb{1}_A = 1) = \mathbb{P}(A).) \Rightarrow \mathbb{P}(|A| \geq a) \leq \frac{\mathbb{E}(|X|)}{a}$ $(*)$.

- Allgemeiner Fall: Sei $Y := f(|X|),\ b := f(a)$. $(*)$ liefert:

$$\begin{aligned}\mathbb{P}(|Y| \geq b) \leq \frac{\mathbb{E}(|Y|)}{b} &= \frac{\mathbb{E}(Y)}{b} \\ &= \frac{\mathbb{E}(f(|X|))}{f(a)}.\end{aligned}$$

  Z.z. $\mathbb{P}(|X| \geq a) \leq \mathbb{P}(|Y| \geq b) = \mathbb{P}(f(|X|) \geq f(a))$. D.h., z.z.: $|X| \geq a \subset f(|X|) \geq f(a)$. Da f monoton steigent ist, gilt:

$$X \leq Y \Rightarrow f(x) \leq f(y),\ \forall x, y \in \mathbb{R}.$$

  Insbesondere: $|X(\omega)| \geq a \Rightarrow f(|X(\omega)|) \geq f(a)$.

IV.4.3 Beispiel

$\Omega = \{-3, -2, -1, 0, 1, -2, 3\}$, Laplace $X(\omega) = \omega$.

$$\mathbb{P}(|X| \geq a) \leq \frac{\mathbb{E}(|X|}{a} = \frac{2}{a}.$$

Für $a \leq 2$ nicht hilfreich, $a = 3$ ergibt $\mathbb{P}(|X| \geq a) \leq \frac{2}{3}$.

IV.4.4 Korollar (Tschebychev-Ungleichung)

$$\mathbb{P}(|X - \mu| \geq a) \leq \frac{\operatorname{Var} X}{a^2},$$

$\mu = \mathbb{E}(X)$.

IV.4.5 Beweis

Wähle $f(x) = x^2$ und $X \rightsquigarrow X - \mu$.

$$\mathbb{P}(|X - \mu| \geq a) \leq \frac{\mathbb{E}(|X - \mu|^2)}{a^2}.$$

V Gesetze der großen Zahlen

Wir betrachten ein Zufallsexperiment mit zugehöriger Zufallsvariable X, welches wir sehr häufig in unabhängiger Weise wiederholen. Wir machen Aussagen für diese Zufallsvariablen gemeinsam, insbesondere über das gebildete Mittel

$$\overline{X}_n := \frac{1}{n}\sum_{j=1}^{n} x_j.$$

V.1 Schwaches Gesetz der großen Zahlen

V.1.1 Satz (Schwaches Gesetz der großen Zahlen)

Sei Seien $X_1, \ldots, X_n$ identisch verteilte, paarweise unabhängige Zufallsvariablen, je mit Erwartungswert μ und Varianz σ^2. Dann gilt:

$$\forall a > 0 \text{ ist } \mathbb{P}(|\overline{X}_n - \mu| \geq a) \leq \frac{\sigma^2}{na^2}.$$

V.1.2 Beweis

Wegen Tschebychev gilt:

$$\mathbb{P}(|\overline{X}_n - \mu| \geq a) \leq \frac{\mathrm{Var}(\overline{X}_n)}{a^2} \ (*),$$

$$\begin{aligned}
\mathrm{Var}\,\overline{X}_n &= \mathrm{Var}(\frac{1}{n}\sum_{j=1}^{n} X_j) = \frac{1}{n^2}\,\mathrm{Var}(\sum_{j=1}^{n} X_j) \\
&= \frac{1}{n^2}\mathbb{E}((\sum_{j=1}^{n}(X_j - \mu))^2) \\
&\underbrace{=}_{(**)} \frac{1}{n^2}\mathbb{E}(\sum_{j=1}^{n}(X_j - \mu)^2) + \frac{1}{n^2}\mathbb{E}(\sum_{j\neq k}(X_j - \mu)(X_k - \mu)) \\
&= \left[\sum_{j=1}^{n} \mathrm{Var}\,X_j + \sum_{j\neq k} 2\underbrace{\mathrm{Cov}(X_j, X_k)}_{=0}\right]\frac{1}{n^2},
\end{aligned}$$

mit NR $(**)$:

$$\left(\sum_{j=1}^{n} a_j\right)^2 = \sum_{j=1}^{n} a_j \cdot \sum_{k=1}^{n} a_k = \sum_{j=1}^{n}\sum_{k=1}^{k} a_j a_k = \sum_{j=1}^{n} a_j^2 \cdot \sum_{j\neq k} a_j a_k,$$

$$(*) = \frac{1}{n^2}\frac{n\sigma^2}{a^2}.$$

[24.06.2024]

V.1.3 Bemerkung

a) Das obige Gesetz liefert, dass unter den genannten Bedingungen gilt:

$$\forall a > 0 \text{ ist } \lim_{n\to\infty} \mathbb{P}(|\overline{X}_n - \mu| < a) = 1.$$

b) Falls $X_j \to \{0,1\}$, ist $\overline{X}_n$ die relative Häufigkeit von "$X = 1$" und μ die Wahrscheinlichkeit für $X_j = 1$. Das Gesetz sagt etwas darüber aus, wie sich relative Häufigkeit und Wahrscheinlichkeit annähern.

Wir werden später zeigen, dass unter geeigneten Bedingungen $\mathbb{P}(\lim_{n\to\infty} X_n = \mu) = 1$ gilt (starkes Gesetz der großen Zahlen).

V.2 Einschub: Konvergenzbegriffe

V.2.1 Definition (Stochastische Konvergenz)

Sei X eine Zufallsvariable, $(X_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen. Man sagt X_n *konvergiert stochastisch* gegen X

$$:\Leftrightarrow \forall a > 0 \text{ ist } \lim_{n\to\infty} \mathbb{P}(|X_n - X| < a) = 1.$$

V.2.2 Bemerkung

Die Aussage des schwachen Gesetzes (Bemerkung) ist also $\overline{X}_n \xrightarrow[n\to\infty]{\text{stoch.}} \mu$.

V.2.3 Definition (Fast sichere Konvergenz)

Man sagt X_n *konvergiert fast sicher* gegen X

$$:\Leftrightarrow \mathbb{P}(\{\lim_{n\to\infty} X_n(\omega) = X(\omega)\}) = \ 1 (\Leftrightarrow \lim_{n\to\infty} X_n = X).$$

V.2.4 Definition (Konvergenz in Verteilung)

Man sagt X_n *konvergiert* gegen X *in Verteilung*, falls

$$\lim_{n\to\infty} V_{X_n}(t) = V_X(t),$$

$\forall t$, wo V_X stetig ist.

V.2.5 Ausblick

$$\xrightarrow[n\to\infty]{\text{f.s.}} \Rightarrow \xrightarrow[n\to\infty]{\text{stoch.}} \Rightarrow \xrightarrow[n\to\infty]{\text{i.V.}},$$

die Gegenrichtungen gelten nicht.

V.2.6 Satz

Die stochastische Konvergenz impliziert nicht die fast sichere Konvergenz.

V.2.7 Beweis

Gegenbeispiel. $X_n \to \{0,1\}$ von einem Glücksrad.

- $X_1(\omega) = \begin{cases} 1 & \text{falls } \omega \in [0,1] \\ 0 & \text{sonst,} \end{cases}$
- $X_2(\omega) = \begin{cases} 1 & \text{falls } \omega \in [1, 1+\frac{1}{2}] \\ 0 & \text{sonst,} \end{cases}$
- ...
- $X_j(\omega) = \begin{cases} 1 & \text{falls } \omega \in [\sum_{i=1}^{j-1} \frac{1}{i}, \sum_{i=0}^{j} \frac{1}{j}] \\ 0 & \text{sonst,} \end{cases}$

Es gilt $X_n \xrightarrow[n\to\infty]{\text{f.s.}} 0$. $\forall a > 0,\ \mathbb{P}(|X_n - 0| < a) \geq \mathbb{P}(X_n = 0) = \frac{2\pi - \frac{1}{n}}{2\pi} \to_{n\to\infty} 1 \Rightarrow \lim_{n\to\infty} \mathbb{P}(|X_n - 0| < a) = 1,\ \forall a > 0$. Es gilt stochastischce Konvergenz.

Für die fast sichere Konvergenz untersuchen wir $\{\omega \in \Omega\colon \lim_{n\to\infty} X_n(\omega) = 0\}$. Da $\lim_{n\to\infty} \sum_{j=1}^{n} \frac{1}{j} = \infty$ gilt $\forall \omega \in \omega\ X_k(\omega) = 1$ für unendliche viele $k \in \mathbb{N}$, ebenso ist $X_k(\omega) = 0$ für ∞-viele $k \in \mathbb{N}$ $\Rightarrow \lim_{n\to\infty} X_n(\omega)$ existiert nicht $\Rightarrow \mathbb{P}(\lim_{n\to\infty} X_n(\omega) = X) = 0,\ \forall X$ Zufallsvariablen.

V.2.8 Satz

Konvergenz in Verteilung impliziert nicht die stochastische Konvergenz.

V.2.9 Beweis

$X = 1$ mit Whk. $\frac{1}{2}$, $X = -1$ mit Whk. $\frac{1}{2}$. $X_n = -X\ \forall n \in \mathbb{N}$. $|X_n - X| = 2\ \forall \omega \in \Omega$. $\mathbb{P}(|X_n - X| < 1) = 0 \Rightarrow$ stoch. Konvergenz gilt nicht. Es ist $V_X = V_{-X}$. $\lim_{n\to\infty}(X_{X_n}(t) = V_X(t)\ \forall t \in \mathbb{R}(\setminus\{\pm 1\})$.

V.2.10 Satz

Stochastische konvergenz impliziert Konvergenz in Verteilung.

V.2.11 Beweis

Sei $X_n \xrightarrow[n\to\infty]{\text{stoch.}} X$. Betrachte $V_X(t) - V_{X_n}(t)$ für t, sodass V_X stetig an der Stelle t. Z.z. $V_X(t) - V_{X_n}(t) < 2\delta,\ \forall \delta > 0$. Wegen der Stetigkeit gibt es ein $\varepsilon > 0$, sodass $V_X(t+\varepsilon) - V_X(t) < \delta \Rightarrow V_X(t+\varepsilon) - V_X(t) = V_X(t+\varepsilon) - V_X(t) + V_X(t) = V_{X_n}(t)$. Wir betrachten $-V_X(t+\varepsilon) + V_{X_n}(t) = -\mathbb{P}(X \leq t+\varepsilon) + \mathbb{P}(X_n \leq t) \leq \mathbb{P}(X_n \leq t \text{ und } x > t+\varepsilon) \leq \mathbb{P}(|X - X_n| > \varepsilon)$. Wegen stochastischer Konvergenz existiert ein $n \in \mathbb{N}$, sodass $\mathbb{P}(|X_n - X| > \varepsilon) < \delta$. $\Rightarrow V_X(t) - V_{X_n}(t) \geq -\delta - \delta$.

Literaturverzeichnis

[1] P. Pickl. *Stochastik*. Mitschrift der Vorlesung, erstellt von Dominik Bullach und Hannes F. Funk. S. 33. 2015. URL: www.mathematik.uni-muenchen.de/~bohmmech/Teaching/StochastikSoSe17/Stochastik%20Skript.pdf.