Wahrscheinlichkeitstheorie

Einleitung

- Wahrscheinlichkeitsräume, $\underbrace{\text{Abbildung}}_{\text{zufällige Variable}}$: Wahrsch.-Raum $\longrightarrow$ Wahrsch.-Raum

- Stochastik: Wahrscheinlichkeitstheorie und Mathematische Statistik

Ziel: Mit math. Methoden Einsichten gewinnen in zufallsbeeinflusste Abläufe

  - Untersuchung von Gesetzmäßigkeiten des Zufalls (Wahrscheinlichkeitstheorie)
  - Schlussfolgerung aus Beobachtungsdaten (mathematische Statistik)

- Formulierung von Modellen für Experimente, die als zufallsbeeinflusst aufgefasst werden

Wahrscheinlichkeitsräume

Definition: (Wahrscheinlichkeitsraum / Axiomensystem von Kolmogorov)

Sei Ω eine nicht leere Menge (von Elementarereignissen). Sei $\mathcal{A} \subset \mathcal{P}(\Omega)$ eine σ-Algebra (Ereignisalgebra). Sei $P: \mathcal{A} \to [0,1]$ gegeben. Dann ist $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum, wenn gilt:

(W1) P ist ein Maß auf $(\Omega, \mathcal{A})$, d.h.: Für alle Folgen (A_n) in $\mathcal{A}$ mit $A_i \cap A_j = \emptyset$ für alle $i,j \in \mathbb{N}$ mit $i \neq j$ gilt:

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} P(A_j) \qquad (\text{„}\sigma\text{-Additivität"})$$

(W2) $P(\Omega) = 1$ („normiert")

Bemerkung:
- .) Ω muss mindestens alle möglichen Beobachtungsergebnisse eines möglichen Zufallsexperiments enthalten. Dann werden die Elemente von Ω auch Elementarereignisse genannt.
- .) $\mathcal{A}$ beschreibt damit eine Menge von Ereignissen. Beim einem Ereignis handelt es sich immer um eine Teilmenge von Ω.
  (Sei $\omega \in \Omega$ das Ergebnis eines Versuchs. Dann ist Ereignis $A \in \mathcal{A}$ eingetreten, wenn $\omega \in A$.
  $\Rightarrow \omega \notin \emptyset$ ($\emptyset$ ist das unmögliche Ereignis)
  $\Rightarrow \omega \in \Omega$ (Ω ist das sichere Ereignis))
- .) P heißt auch Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{A})$

Beispiele: ·) einmaliger Münzwurf: $\Omega_1 = \{0, 1\}$ (Ergebnis entweder Kopf oder Zahl)

·) n-maliger Münzwurf unter Beachtung der Reihenfolge:

$$\Omega_2 = \{0,1\}^n = \{0,1\} \times \ldots \times \{0,1\}$$
$$= \{ (a_1, \ldots, a_n) \mid a_i \in \{0,1\} \}$$

·) n-maliger Münzwurf ohne Beachtung der Reihenfolge:

$\Rightarrow$ Messung der Anzahl 1 und 0 (1 reicht aus)
$\Rightarrow$ maximal n Einsen und minimal Null

$$\Rightarrow \Omega_3 = \{0, 1, \ldots, n\}$$

Ereignis: mindestens einmal 1 werfen

mit Reihenfolge: $A = \{ (a_1, \ldots, a_n) \mid a_i \in \{0,1\}, \sum_{i=1}^{n} a_i > 0 \} \subset \Omega_2$

ohne Reihenfolge: $A = \{1, 2, \ldots, n\} \subset \Omega_3$

<u>Proposition:</u> (Eigenschaften vom Wahrscheinlichkeitsmaß)

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

a) $P(\Omega \setminus A) = 1 - P(A)$ für alle $A \in \mathcal{A}$

b) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ für alle $A, B \in \mathcal{A}$

Beweis: b) $A' := A \setminus B$, $B' := B \setminus A \Rightarrow A' \cap B' = \emptyset$

$X := A \cap B \Rightarrow A', B', X$ paarweise disjunkt

$$\Rightarrow A' \cup B' \cup X = A \cup B$$
$$\Rightarrow P(A) = P(A' \cup X) \overset{(\sigma\text{-a.})}{=} P(A') + P(X)$$
$$\Rightarrow P(B) = P(B' \cup X) = P(B') + P(X)$$
$$\Rightarrow P(A \cup B) = P(A' \cup B' \cup X) \overset{(\sigma)}{=} P(A') + P(B') + P(X)$$
$$= P(A') + P(X) + P(B') + P(X) - P(X)$$
$$= P(A) + P(B) - P(A \cap B)$$

a) $\Omega \setminus A \cap A = \emptyset \Rightarrow P(\Omega) = P(\Omega \setminus A \cup A) = 1$
$$\overset{(\sigma)}{=} P(\Omega \setminus A) + P(A)$$

$$\Rightarrow P(\Omega \setminus A) = 1 - P(A)$$

□

Wichtige Standardfälle: ·) Diskrete Wahrscheinlichkeitsräume:
Ω ist endlich oder abzählbar unendlich
man setzt dann $\mathcal{A} = \mathcal{P}(\Omega)$

$\Longrightarrow$ jedes $A \in \mathcal{A}$ ist durch $A = \bigcup_{\omega \in A} \{\omega\}$ ausdrückbar

$\overset{(\sigma\text{-add})}{\Longrightarrow}$ $\boxed{P(A) = \sum_{\omega \in A} P(\{\omega\})}$ für alle $A \in \mathcal{A}$

$\Longrightarrow$ Wahrscheinlichkeitsmaß P ist bereits eindeutig durch $P(\{\omega\}), \omega \in \Omega$ gegeben

·) kontinuierliche Wahrscheinlichkeitsräume:
meistens $\Omega = \mathbb{R}$, $\mathcal{A} = \mathcal{B} := \mathcal{F}(\{(-\infty, x] \mid x \in \mathbb{R}\})$
eindeutig durch $P((-\infty, x]), x \in \mathbb{R}$ bestimmt

Definition: (Verteilungsfunktion / cumlative distribution function (cdf))

Sei $(\Omega = \mathbb{R}, \mathcal{A} = \mathcal{B}, P)$ ein Wahrscheinlichkeitsraum. Dann definiert man:

$$F : \mathbb{R} \longrightarrow [0,1], \quad F(x) := P((-\infty, x])$$

als die Verteilungsfunktion des Wahrscheinlichkeitsmaßes P. (sie ist eindeutig)

Proposition: (Charakterisierung von Verteilungsfunktionen)

Sei $(\Omega = \mathbb{R}, \mathcal{A} = \mathcal{B}, P)$ der Wahrscheinlichkeitsraum. Sei $F: \mathbb{R} \to [0,1]$ die Verteilungsfunktion von P. Dann gilt:

·) $x_1 < x_2 \Longrightarrow F(x_1) \leq F(x_2)$ für alle $x_1, x_2 \in \mathbb{R}$ (Monotonie)

·) $\lim_{h \to 0^+} F(x+h) = F(x)$ für alle $x \in \mathbb{R}$ (rechtsseitig stetig)

·) $\lim_{x \to \infty} F(x) = 1$, $\lim_{x \to -\infty} F(x) = 0$

Beweis: Übung $\square$

Proposition:

Sei $(\mathbb{R}, \mathcal{B})$ der messbare Raum. Dann existiert zu jeder Funktion $F: \mathbb{R} \to [0,1]$ mit den Eigenschaften, wie in obiger Proposition, ein Wahrscheinlichkeitsmaß P auf $(\mathbb{R}, \mathcal{B})$, sodass F die Verteilungsfunktion von P ist.

Beweis: hier nicht $\square$

Definition: (Verteilungsdichte)

Sei $(\mathbb{R}, \mathcal{B}, P)$ der Wahrscheinlichkeitsraum und $F: \mathbb{R} \to [0,1]$ die Verteilungsfunktion von P. Falls eine (Riemann-)integrierbare Funktion $f: \mathbb{R} \to [0, \infty)$ existiert mit

$$F(x) = \int_{-\infty}^{x} f(t)\, dt$$ für alle $x \in \mathbb{R}$, dann heißt f Verteilungsdichte von P.

Bemerkung: Nicht zu jedem Wahrscheinlichkeitsmaß P existiert Verteilungsdichte. Aber zu jeder integrierbaren Funktion $f: \mathbb{R} \to [0, \infty)$ mit $\int_{\mathbb{R}} f \, d\lambda = 1$ existiert ein Wahrscheinlichkeitsmaß P, sodass f eine Verteilungsdichte von P ist.

Spezielle Wahrscheinlichkeitsräume:

•) („Laplace-Experiment") Sei $n \in \mathbb{N}$ und Ω eine beliebige Menge mit $\#\Omega = n$. Dann sei $(\Omega, \mathcal{P}(\Omega), P)$ der Wahrscheinlichkeitsraum.
Es sei:

$$P(\{\omega\}) := \frac{1}{\#\Omega} \quad \text{für alle } \omega \in \Omega$$

$$\Longrightarrow \boxed{P(A) = \sum_{\omega \in A} P(\{\omega\}) = \frac{\#A}{\#\Omega}}$$ P heißt diskrete Gleichverteilung auf Ω

Häufige Methode: $$\underset{\text{Gleichverteilung.}}{(\Omega, \mathcal{P}(\Omega), P)} \xrightarrow[\text{Abbildung}]{\text{messbare}} (\Omega', \mathcal{P}(\Omega'), P')$$

•) $(\Omega = \{0,1\}, \mathcal{P}(\Omega), P)$ mit $P(\{1\}) = p$, $P(\{0\}) = 1-p$
mit $0 \le p \le 1$

•) („n-gliedriges Bernoulli-Schema) $(\Omega = \{0,1\}^n, \mathcal{P}(\{0,1\}^n), P)$ mit $n \in \mathbb{N}$ und

$$P(\{(a_1, \dots, a_n)\}) = \prod_{j=1}^{n} p^{a_j} (1-p)^{1-a_j}$$

•) $\Omega = \{(r_1, \dots, r_n) \mid r_j \in \mathbb{N}_0, \sum_{j=1}^{n} r_j = r\}$ mit $r, n \in \mathbb{N}$
$\mathcal{A} = \mathcal{P}(\Omega)$
(„r Teilchen in n Energiezuständen")

a) Maxwell-Boltzmann-Verteilung: $P_{MB}(\{(r_1, \dots, r_n)\}) = n^{-r} \cdot r! \cdot (r_1! \cdot r_2! \cdots r_n!)^{-1}$

b) Fermi-Dirac-Verteilung: (Voraus. $n \ge r$): $P_{FD}(\{(r_1, \dots, r_n)\}) = \begin{cases} \binom{n}{r}^{-1} & : r_j \in \{0,1\} \\ 0 & : \text{sonst} \end{cases}$

c) Bose-Einstein-Verteilung: $P_{BE}(\{(r_1,\dots,r_n)\}) = \binom{n+r-1}{r}^{-1}$

(auch Gleichverteilung auf Ω)

<u>Konstruktion Maxwell-Boltzmann-Verteilung:</u>

Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum, welcher der Maxwell-Boltzmann-Verteilung genügt. Dann sei $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{P})$ ein (Hilfs-) Wahrscheinlichkeitsraum mit

- $\tilde{\Omega} = \{1,\dots,n\}^r$ (Ordne einem Teilchen ein Energieniveau zu)
- $\tilde{\mathcal{A}} = \mathcal{P}(\tilde{\Omega})$
- $\tilde{P}$ ... diskrete Gleichverteilung auf $\tilde{\Omega}$ (Teilchen verteilen sich gleichwahrscheinlich)

Dann sei weiterhin $Z: \tilde{\Omega} \to \Omega$ messbar, sodass $\tilde{\Omega}$ in Ω überführt wird.

Sei nun $(r_1,\dots,r_n) \in \Omega$ und $\tilde{A} = \{(a_1,\dots,a_r) \in \tilde{\Omega} \mid Z(a_1,\dots,a_r) = (r_1,\dots,r_n)\}$
Dann folgt durch Kombinatorik:

- $\#\tilde{\Omega} = n^r$
- $\#\tilde{A} = \binom{r}{r_1}\binom{r-r_1}{r_2}\cdots\binom{r-r_1-\dots-r_{n-1}}{r_n} = r!\,(r_1!\cdot r_2!\cdot\ldots\cdot r_n!)^{-1}$

$$\Longrightarrow \tilde{P}(\tilde{A}) = \frac{\#\tilde{A}}{\#\tilde{\Omega}} \overset{!}{=} P(\{r_1,\dots,r_n\})$$

<u>Weitere Wahrscheinlichkeitsräume:</u>

- Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum mit $\Omega = \mathbb{N}_0$, $\mathcal{A} = \mathcal{P}(\Omega)$. und (mit $p \in (0,1)$)
$$P(\{k\}) = (1-p)^k p \quad \text{für alle } k \in \mathbb{N}_0$$
(auch „geometrische Verteilung": k... Anzahl der Nullen vor erster Eins)

- („Poisson-Verteilung")
Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum mit $\Omega = \mathbb{N}_0$, $\mathcal{A} = \mathcal{P}(\Omega)$ und $P = \Pi_\lambda$ mit $(\lambda \in (0,\infty))$
$$\Pi_\lambda(\{k\}) = \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{für alle } k \in \mathbb{N}_0$$

- Seien $a, b \in \mathbb{R}$ mit $a < b$ und $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum mit $\Omega = (a,b)$, $\mathcal{A} = \mathcal{B}$ und $P = U_{(a,b)}$. Sei $f: (a,b) \to [0,\infty)$ die Verteilungsdichte von $U_{(a,b)}$:
$$f(t) = (b-a)^{-1}\, 1_{(a,b)}(t)$$
$$\Longrightarrow U_{(a,b)}(A) = \int_A f(t)\, d\lambda(t)$$

·) („Gauß – Verteilung“ / Normalverteilung)

Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum mit $\Omega = \mathbb{R}$, $\mathcal{A} = \mathcal{B}$.
und $P = \mathcal{N}_{\mu,\sigma^2}$, wobei $\mu \in \mathbb{R}$, $\sigma \in (0,\infty)$. Sei $f: \mathbb{R} \to [0,\infty)$
die Verteilungsdichte von $\mathcal{N}_{\mu,\sigma^2}$:

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right)$$

$$\Longrightarrow \mathcal{N}_{\mu,\sigma^2}(A) = \int_A f(t)\, d\lambda(t)$$

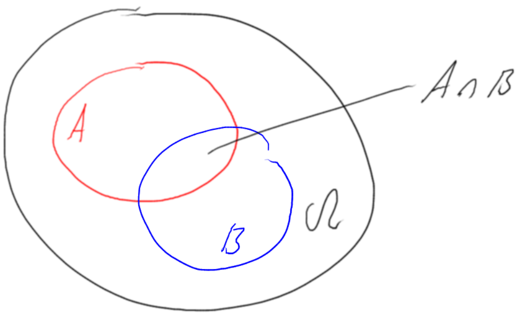
Dann heißt $\mathcal{N}_{0,1}$ auch die Standardnormalverteilung. μ heißt Erwartungswert
und σ^2 heißt Variant.

·) („Cauchy – Verteilung“)

Bedingte Wahrscheinlichkeiten

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

Idee: „B ist eingetreten $\Longrightarrow$ $P(A)$ ändert sich"
Ist B eingetreten, so ist B in diesem Moment das sichere Ereignis.
$\Longrightarrow$ Normierung auf $P(B)$ und Betrachtung von $P(A \cap B)$



B ist eingetreten

$\Longrightarrow$ $\tilde{\Omega} = B$ mit $\tilde{A} = A \cap B$

$$\Longrightarrow \tilde{P}(A \cap B) = \frac{P(A \cap B)}{P(B)}$$

Definition: (Bedingte Wahrscheinlichkeit)

Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum. Sei $A, B \in \mathcal{A}$ und $P(B) > 0$. Dann definiert man die bedingte Wahrscheinlichkeit von A unter der Bedingung B als:

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

Proposition: (Eigenschaften)

Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und sei $A, B \in \mathcal{A}$. Dann gilt:

a) $P(\Omega \mid B) = 1$
b) $P(A \mid B) \in [0, 1]$
c) für alle Folgen (A_n) in $\mathcal{A}$ mit $A_n \cap A_m = \emptyset$, $n, m \in \mathbb{N}$, $n \neq m$ gilt:

$$P\left(\bigcup_{n \in \mathbb{N}} A_n \,\middle|\, B\right) = \sum_{n=1}^{\infty} P(A_n \mid B)$$

Beweis: a) klar.
b) $0 \leq P(A \cap B) \leq P(B) \leq 1$
c)
$$P\left(\bigcup_{n \in \mathbb{N}} A_n \,\middle|\, B\right) = P(B)^{-1} P\left(\bigcup_{n \in \mathbb{N}} A_n \cap B\right)$$
$$(\sigma\text{-add.} + \text{p. disj.}) = P(B)^{-1} \left[\sum_{n=1}^{\infty} P(A_n \cap B)\right]$$
$$= \sum_{n=1}^{\infty} P(A_n \mid B)$$

$\square$

Bemerkung: ·) Damit ist $(\Omega, \mathcal{A}, P(\cdot \mid B))$ für $P(B) > 0$ ein Wahrscheinlichkeitsraum.

Folgerung: (Weitere Eigenschaften)

Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $A, B \in \mathcal{A}$, sodass $P(A), P(B) > 0$. Dann gilt:

a) $$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

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

c) Seien $A_1, \ldots, A_n \in \mathcal{A}$, $n \in \mathbb{N}$ und $P(\bigcap_{j=1}^{n-1} A_j) > 0$, dann gilt:

$$P(\bigcap_{j=1}^{n} A_j) = \prod_{j=1}^{n} P(A_j \mid \bigcap_{k=1}^{j-1} A_k) \quad (\text{mit } P(A_1 | \emptyset) = P(A_1))$$

$$= P(A_1) \cdot P(A_2 | A_1) \cdot P(A_3 | A_1 \cap A_2) \cdot \ldots \cdot P(A_n | A_1 \cap A_2 \cap \ldots \cap A_{n-1})$$

(auch „Pfadregel in Wahrscheinlichkeitsbäumen")

Proposition: (Formel der totalen Wahrscheinlichkeit)

Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $B_1, \ldots, B_n \in \mathcal{A}, n \in \mathbb{N}$. eine Partition von Ω (d.h. $B_i \neq \emptyset$, B_i sind paarweise disjunkt, $\bigcup_{j=1}^{n} B_j = \Omega$) mit $P(B_i) > 0$ für alle $i = 1, \ldots, n$. Dann gilt für alle $A \in \mathcal{A}$:

$$P(A) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)$$

insbesondere falls $P(A) > 0$: $$P(B_j|A) = \frac{P(A|B_j) \cdot P(B_j)}{P(A)}$$

$$\Longrightarrow \quad P(B_j|A) = \frac{P(A|B_j) \cdot P(B_j)}{\sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)} \quad \text{für } j = 1, \ldots, n$$

Beweis: $$P(A) = P(A|\Omega) = P(A \mid \bigcup_{j=1}^{n} B_j) = P(\bigcup_{j=1}^{n} B_j \mid A) \cdot \frac{P(A)}{P(\Omega)}$$

$$= \sum_{j=1}^{n} P(B_j|A) \cdot P(A) = \sum_{j=1}^{n} P(A|B_j) \cdot \frac{P(B_j)}{P(A)} \cdot P(A)$$

$$= \sum_{j=1}^{n} P(A|B_j) \cdot P(B_j)$$

□

Bemerkung: ·) Aussage lässt sich auch für abzählbar viele B_i formulieren

Stochastische Unabhängigkeit von Ereignissen

Idee: Seien $(\Omega, \mathcal{A}, P)$, $A, B \in \mathcal{A}$ wie oben.
Dann soll A unabh. von B sein, wenn $P(A) = P(A|B) = P(A|\Omega\setminus B)$
$\Longrightarrow$ $P(A \cap B) = P(A) \cdot P(B)$ (Produkt geht auch für $B = \emptyset$ oder $B = \Omega$)

Definition: (stochastische Unabhängigkeit)

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

a) Zwei Ereignisse $A, B \in \mathcal{A}$ heißen (stochastisch) unabhängig, wenn

$$P(A \cap B) = P(A) \cdot P(B)$$

b) Die Ereignisse $A_1, \dots, A_n \in \mathcal{A}$, $n \in \mathbb{N}$ heißen vollständig (stochastisch) unabhängig, wenn für alle $I \subset \{1, \dots, n\}$, $\# I \geq 2$ gilt:

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i)$$

c) Eine Folge $(A_k)_{k \in \mathbb{N}}$ in $\mathcal{A}$ heißt unabhängig, wenn für alle $n \in \mathbb{N}$ gilt, dass $A_1, \dots, A_n$ vollständig unabhängig sind.

Bemerkung:
- Begriffe „unvereinbar" ($A \cap B = \emptyset$) und „unabhängig" sind nicht gleich
- Unabhängigkeit hängt nicht nur von A, B, sondern auch von P ab
- (b) kann im Allgemeinen nicht reduziert werden

Proposition:

Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $A, B, A_1, \dots, A_n \in \mathcal{A}$, $n \in \mathbb{N}$. Dann gilt:

a) A und $\emptyset$ sind unabhängig
b) A und Ω sind unabhängig
c) A und B unabhängig $\Longleftrightarrow$ $\Omega\setminus A$ und B unabhängig $\Longleftrightarrow$ A und $\Omega\setminus B$ unabhängig $\Longleftrightarrow$ $\Omega\setminus A$ und $\Omega\setminus B$ unabhängig
d) Wenn $A \cap B = \emptyset$, dann gilt:
A, B unabhängig $\Longleftrightarrow$ $P(A) = 0$ oder $P(B) = 0$
e) $A_1, \dots, A_n$ vollständig unabhängig $\Longrightarrow$ $A_1, \dots, A_m$ vollständig unabhängig für alle $m < n$

<u>Theorem:</u> (Lemma von Borel-Cantelli)

Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum. Sei $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$. Sei weiterhin:

$$A := \limsup_{n\to\infty} A_n := \bigcap_{n=1}^{\infty} \bigcup_{j=n}^{\infty} A_j$$

Dann gilt:

$$\boxed{\begin{array}{l} a) \ \sum_{n=1}^{\infty} P(A_n) < \infty \implies P(A) = 0 \\ b) \ \sum_{n=1}^{\infty} P(A_n) = \infty \text{ und } (A_n) \text{ ist unabhängig} \implies P(A) = 1 \end{array}}$$

Beweis: a) $A \subset \bigcup_{k=n}^{\infty} A_k$ für alle $n \in \mathbb{N}$

$$\implies P(A) \leq P\left(\bigcup_{k=n}^{\infty} A_k\right) \leq \sum_{k=n}^{\infty} P(A_k) \quad \text{für alle } n\in\mathbb{N}$$
$$\implies P(A) \leq \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0, \text{ wenn } \sum_{k=1}^{\infty} P(A_k) < \infty$$

b) $$P(\Omega \setminus A) = P\left(\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^c\right) \leq \sum_{n=1}^{\infty} P\left(\bigcap_{k=n}^{\infty} A_k^c\right)$$
$$= \sum_{n=1}^{\infty} \lim_{m\to\infty} P\left(\bigcap_{k=n}^{m} A_k^c\right) = \sum_{n=1}^{\infty} \lim_{m\to\infty} \prod_{k=n}^{m} (1 - P(A_k))$$
$$\leq \sum_{n=1}^{\infty} \lim_{m\to\infty} \exp\left(-\sum_{k=n}^{m} P(A_k)\right) = 0$$

$\square$

Beispiel: Sei $(\Omega, \mathcal{A}, P)$ mit $\Omega = \{1,\dots,6\}^{\mathbb{N}} = \{(a_n)_{n\in\mathbb{N}} \text{ in } \{1,\dots,6\}\}$
und $\mathcal{A} := \sigma(\{B_1 \times \dots \times B_n \times \Omega \mid B_1,\dots,B_n \in \mathcal{P}(\{1,\dots,6\}), n\in\mathbb{N}\})$
mit $P(B_1 \times \dots \times B_n \times \Omega) = \dfrac{\prod \# B_n}{\#\{1,\dots,6\}^n}$

Betrachte jetzt Folge $(A_n)_{n\in\mathbb{N}}$ mit $A_n = \{1,\dots,6\}^{n-1} \times \{6\} \times \Omega$

$$\implies P(A_n) = \frac{1}{6} \implies \sum_{n=1}^{\infty} P(A_n) = \infty$$

$\overset{\text{(Borel-Cant.)}}{\implies}$ $P(\limsup_{n\to\infty} A_n) = 1$ „es fällt unendlich oft eine 6“

<u>Zufällige Variablen</u>

<u>messbare Abbildungen zwischen Wahrscheinlichkeitsräumen:</u>

Beispiele:
- Zweimal würfeln, X... Summe der Augenzahlen
- Herleitung der Maxwell-Boltzmann-Verteilung
- Koordinatentransformation (kartesisch in Polar)

Definition: (Messbarkeit und induziertes Wahrscheinlichkeitsmaß)

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

a) Eine Abbildung $X : \Omega \to \Omega'$ heißt $(\mathcal{A}, \mathcal{A}')$-messbar oder auch zufällige Variable, wenn gilt: $X^{-1}(A) \in \mathcal{A}$ für alle $A \in \mathcal{A}'$

b) Sei X eine $(\mathcal{A}, \mathcal{A}')$-messbare Abbildung. Dann heißt das Wahrscheinlichkeitsmaß P_X auf $(\Omega', \mathcal{A}')$ mit $P_X(A) = P(X^{-1}(A))$ für $A \in \mathcal{A}'$ das von X induzierte Wahrscheinlichkeitsmaß.

Lemma: (induziertes Wahrscheinlichkeitsmaß ist wohldefiniert)

Seien $(\Omega, \mathcal{A}, P)$ und $(\Omega', \mathcal{A}')$ und X, P_X wie oben. Dann ist P_X ein Wahrscheinlichkeitsmaß auf $(\Omega', \mathcal{A}')$.

Beweis:
- $P_X(\mathcal{A}') \subset [0,1]$ klar.
- $P_X(\Omega') = P(X^{-1}(\Omega')) = P(\Omega) = 1$
- $(A_n)_{n \in \mathbb{N}}$ in Ω', sodass alle A_n paarweise disjunkt

$$\Longrightarrow P_X\left(\bigcup_{n\in\mathbb{N}} A_n\right) = P\left(X^{-1}\left(\bigcup_{n\in\mathbb{N}} A_n\right)\right) = P\left(\bigcup_{n\in\mathbb{N}} X^{-1}(A_n)\right) \overset{(\sigma\text{-a.})}{=} \sum_{n=1}^{\infty} P(X^{-1}(A_n)) = \sum_{n=1}^{\infty} P_X(A_n)$$

$\square$

Notation: $P(X \in A) := P_X(A) = P(X^{-1}(A))$

Reellwertige Zufallsvariable (hier auch Zufallsgröße)

- Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und sei $(\mathbb{R}, \mathcal{B})$ der messbare Raum auf $\mathbb{R}$ mit Borel-σ-Algebra. Dann ist eine zufällige Variable $X : \Omega \to \mathbb{R}$ eine reellwertige Zufallsvariable oder auch Zufallsgröße.

Definition: (Verteilungsfunktion einer Zufallsgröße)

Sei X (wie oben) eine Zufallsgröße. Dann ist die Verteilungsfunktion $F_X : \mathbb{R} \to [0,1]$ von X definiert als:

$$F_X(x) := P(X \leq x) \quad \text{für } x \in \mathbb{R}$$

Notationen:
$$\begin{aligned} F_X(x) &= P(X \leq x) = P(\{X \leq x\}) = P(X \in (-\infty, x]) \\ &= P(\{\omega \in \Omega \mid X(\omega) \leq x\}) \\ &= P(X^{-1}((-\infty, x])) = P_X((-\infty, x]) \end{aligned}$$

Stetige Zufallsgrößen

Definition: (stetige Zufallsgröße)

Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $X: \Omega \to \mathbb{R}$ eine Zufallsgröße.
X heißt stetige Zufallsgröße, wenn P_X eine Verteilungsdichte $f_X: \mathbb{R} \to [0, \infty)$ besitzt.

Bemerkung: ·) Damit muss Verteilungsfunktion F_X eine stetige Funktion sein.
·) für alle $B \subset \mathbb{R}$, mit B abzählbar, gilt B ist Nullmenge
$(P_X(B) = P(X \in B) = 0)$

·) $P(a < X \leq b) = P(a \leq X \leq b) = P(a \leq X < b) = P(a < X < b)$
für alle $a, b \in \mathbb{R}$ mit $a < b$

Proposition: (Inversionsverfahren)

Sei $F: \mathbb{R} \to [0,1]$ eine Verteilungsfunktion und $X \sim \mathcal{U}(0,1)$ (d.h. X ist gleichverteilt auf $(0,1)$). Dann besitzt die Zufallsgröße

$$Y := F^{\sim}(X) := \sup \{ x \in \mathbb{R} \mid F(x) < X \} \qquad (\text{„Pseudoinverse"})$$

besitzt die Verteilungsfunktion $F_Y = F$.
Insbesondere gilt also, wenn F streng monoton und stetig ist, dass $Y := F^{-1}(X)$ die Verteilungsfunktion $F_Y = F$ besitzt.

Beweis: (nur Insbesondere)

$$F_Y(x) = P(Y \leq x) = P(F^{-1}(X) \leq x) = P(X \leq F(x)) = F_X(F(x))$$
$$\underset{(F(x) \in [0,1])}{=} F(x) \qquad \square$$

Exponentialverteilung: Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum. Eine Zufallsgröße X auf Ω heißt exponential verteilt mit Parameter $\lambda > 0$, wenn

$$f_X(t) = \lambda e^{-\lambda t} \cdot 1_{[0,\infty)}(t)$$

$$\Rightarrow \quad F_X(t) = (1 - e^{-\lambda t}) \, 1_{[0,\infty)}(t)$$

Notation: $X \sim \mathcal{E}_\lambda$, $P_X = \mathcal{E}_\lambda$

Lemma: (Exponentialverteilung ist gedächtnislos)

Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $X \sim \mathcal{E}_\lambda$, $\lambda > 0$. Dann gilt für alle $s, t > 0$:

$$P(X \geq s+t \mid X > s) = P(X > t)$$

Beweis: klar durch direktes Nachrechnen $\square$

Zufällige Vektoren

Definition: (zufälliger Vektor)

Seien $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $X_1, \dots, X_n$, $n \in \mathbb{N}$ reelle Zufallsgrößen über Ω. Dann heißt $\underline{X} := (X_1, \dots, X_n)$ zufälliger Vektor über $(\Omega, \mathcal{A}, P)$.

Beispiele:
- ·) Maxwell-Boltzmann-, Fermi-Dirac-, Bose-Einstein-Verteilung
- ·) n-gliedriges Bernoulli-Schema

Definition: (gemeinsame Verteilungsfunktion)

Seien $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $\underline{X} = (X_1, \dots, X_n)$, $n \in \mathbb{N}$ ein zufälliger Vektor. Unter der gemeinsamen Verteilungsfunktion von $\underline{X}$ versteht man:

$$F_{\underline{X}} : \mathbb{R}^n \to [0,1], \quad F_{\underline{X}}(x) = P(X_1 \leq x_1, \dots, X_n \leq x_n)$$

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

Proposition:

Sei $(\Omega, \mathcal{A}, P)$ wie oben. Seien X, Y zufällige Vektoren auf Ω. Dann gilt:

$$F_X = F_Y \iff P_X = P_Y$$

Beweis: hier nicht. $\square$

Bemerkung: Charakterisierung der Verteilungsfunktion für $n \geq 2$ ist schwieriger.

Definition: (Verteilungsdichte auf $\mathbb{R}^n$)

Eine Funktion $f: \mathbb{R}^n \to [0,\infty)$, $n \in \mathbb{N}$ heißt Dichtefunktion auf $\mathbb{R}^n$, wenn f integrierbar ist und $\int_{\mathbb{R}^n} f(x)\, d\lambda(x) = 1$

Folgerung: Wenn $\underline{X}$ die Verteilungsdichte $f_{\underline{X}}$ besitzt, dann gilt für alle $A \in \mathfrak{B}_n$

$$P_{\underline{X}}(A) = P(\underline{X} \in A) = \int_A f_{\underline{X}}\, d\lambda = \int_{\mathbb{R}^n} 1_A f_{\underline{X}}\, d\lambda$$

Randverteilungen

Zentraler Grenzwertsatz und „Gesetz der Großen Zahlen"

Theorem: (Zentraler Grenzwertsatz)

Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von unabhängig gleichverteilten Zufallsgrößen mit $\mathbb{E}(X_i^2) < \infty$, $var(X_i) = \sigma^2 > 0$, $\mathbb{E}(X_i) = \mu$ für alle $i \in \mathbb{N}$.
Sei dann:

$$\tilde{X}_n := \frac{1}{\sqrt{n\sigma^2}}\left(\sum_{i=1}^{n} X_i - n\mu\right)$$

für alle $n \in \mathbb{N}$. (verschobene und skalierte Summe der Zufallsgrößen, sodass $\mu = 1$, $\sigma^2 = 1$)

Dann gilt für alle $x \in \mathbb{R}$:

$$\lim_{n\to\infty} P(\tilde{X}_n \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{\pi}} e^{-\frac{y^2}{2}}\, dy = \Phi(x)$$

Gilt zusätzlich $\mathbb{E}(|X_i|^3) < \infty$ für alle $i \in \mathbb{N}$, dann

$$\sup_{x\in\mathbb{R}} \left| P(\tilde{X}_n \leq x) - \Phi(x) \right| \leq \frac{8}{10\sigma^3} \cdot \frac{\mathbb{E}(|X_1 - \mathbb{E}X_1|^3)}{\sqrt{n}}$$

mit Schiefe: $\dfrac{\mathbb{E}(|X_1 - \mathbb{E}X_1|^3)}{\sigma^3}$

$\square$

Theorem: (Zentraler Grenzwertsatz von Moivre-Laplace)

Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von unabhängigen gleichverteilten Zufallsgrößen mit $P(X_1 = 1) = 1 - P(X_1 = 0) =: p$ mit $0 < p < 1$. Dann gilt für alle $x \in \mathbb{R}$:

$$\lim_{n\to\infty} P\left([np(1-p)]^{-\frac{1}{2}} \left[\sum_{i=1}^{n} X_i - np \right] \leq x \right) = \Phi(x)$$

$\square$

Theorem: (Gesetz der Großen Zahlen) (auch starkes Gesetz)

Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge unabhängig und gleichverteilter Zufallsgrößen, für die $\mathbb{E}X_i$ existiert $(i \in \mathbb{N})$. Dann gilt:

$$P\left(\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_i = \mathbb{E}X_1 \right) = 1$$

(d.h.: $\frac{1}{n} \sum_{i=1}^{n} X_i \to \mathbb{E}X_1$, $n \to \infty$ fast sicher)

$\square$

<u>Theorem:</u> (Tschebyschevsche Ungleichung)

Sei X eine Zufallsgröße mit $\mathbb{E}X^2 < \infty$. Für alle $c > 0$ gilt

$$P(|X - \mathbb{E}X| \geq c) \leq \frac{\operatorname{Var} X}{c^2}$$

$\square$