

Cheatsheet WuS

Nicolas Wehrli

June 2023

1 Grundbegriffe

1.1 Wahrscheinlichkeitsraum

Axiome von Kolmogorov

Das Tuple $(\Omega, \mathcal{A}, \mathbb{P})$ ist ein **Wahrscheinlichkeitsraum** mit

I. **Grundraum** Ω mit $\Omega \neq \emptyset$, wobei $\omega \in \Omega$ ein Elementarereignis ist.

II. **σ -Algebra** $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ wobei gilt:

1. $\Omega \in \mathcal{A}$
2. $A \in \mathcal{A} \implies A^c \in \mathcal{A}$
3. $A_1, A_2, \dots \in \mathcal{A} \implies \bigcup_i A_i \in \mathcal{A}$

III. **Wahrscheinlichkeitsmass** \mathbb{P} auf (Ω, \mathcal{A}) ist eine Abbildung $\mathbb{P} : \mathcal{A} \mapsto [0, 1]$, wobei gilt:

1. $\mathbb{P}(\Omega) = 1$
2. $A_1, A_2, \dots \in \mathcal{A}, \forall i \neq j : A_i \cap A_j = \emptyset \implies \mathbb{P}(\bigcup_i A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$

De-Morgan

Sei $(A_i)_{i \geq 1}$ eine Folge von beliebigen Mengen. Dann gilt

$$\left(\bigcup_{i=1}^{\infty} A_i \right)^c = \bigcap_{i=1}^{\infty} (A_i)^c$$

Daraus folgt

1. $A_1, A_2, \dots \in \mathcal{A} \implies \bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$
2. $A, B \in \mathcal{A} \implies (A \cup B), (A \cap B) \in \mathcal{A}$

und für $A, B \in \mathcal{A}$

1. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
2. $A \subseteq B \implies \mathbb{P}(A) \leq \mathbb{P}(B)$
3. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

Sei $A_1, A_2, \dots \in \mathcal{A}$, dann gilt:

Union Bound

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

Siebformel

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n \sum_{1 \leq i_1 < \dots < i_k \leq n} \mathbb{P}(A_{i_1} \cap \dots \cap A_{i_k})$$

Atome

Sei Ω nicht leer und diskret. Sei \mathcal{F} eine beliebige σ -Algebra über Ω .

Eine nichtleere Menge $A \in \mathcal{F}$ heisst **atomare** Menge von \mathcal{F} falls für alle $B \in \mathcal{F}$ gilt:

$$B \subseteq A \implies B = \emptyset \vee B = A$$

(Intuitiv: A ist die kleinste nichtleere Menge bezüglich der Inklusion in \mathcal{F})

Die Menge der atomaren Mengen von \mathcal{F} bezeichnen wir mit $\text{Atom}(\mathcal{F})$.

Jedes Element von \mathcal{F} lässt sich als abzählbare Vereinigung von Elementen aus $\text{Atom}(\mathcal{F})$ schreiben.

1.2 Bedingte Wahrscheinlichkeiten

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

Bedingte Wahrscheinlichkeit

Sei $A, B \in \mathcal{A}$ und $\mathbb{P}(B) > 0$, dann ist die **bedingte Wahrscheinlichkeit von A gegeben B**

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

Satz der totalen Wahrscheinlichkeit

Sei $(B_i)_{i \in I}$ eine Partition von Ω . Dann gilt für jedes beliebige $A \in \mathcal{A}$

$$\mathbb{P}(A) = \sum_{i: \mathbb{P}(B_i) > 0} \mathbb{P}(A|B_i) \mathbb{P}(B_i)$$

Satz von Bayes

Aus der Definition der bedingten W'keit folgt sofort die Bayessche Formel, welche den Zusammenhang zwischen $\mathbb{P}(A|B)$ und $\mathbb{P}(B|A)$ beschreibt:

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

Mit dem **Satz der totalen W'keit** können wir $\mathbb{P}(A)$ umschreiben und kommen auf folgende Form:

Sei $(B_i)_{i \in I}$ eine **Partition** von Ω . Dann gilt für jedes beliebige $A \in \mathcal{A}, \mathbb{P}(A) > 0$

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i)}{\sum_{j: \mathbb{P}(B_j) > 0} \mathbb{P}(A|B_j) \cdot \mathbb{P}(B_j)}$$

Intuition Bayessche Statistik

In dieser Form würde man A als das **eingetretene Ereignis** und die B_i als die verschiedenen **Hypothesen** verstehen.

In der Bayesschen Statistik versucht man die Hypothese zu finden, so dass $\mathbb{P}(B_i|A)$ **maximiert** wird.

(Wurde in der Vorlesung nicht weiter behandelt)

1.3 Unabhängigkeit

Unabhängigkeit von zwei Ereignissen

Zwei Ereignisse $A, B \in \mathcal{A}$ heissen **unabhängig**, wenn

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

- $\mathbb{P}(A) \in \{0, 1\} \implies A$ zu jedem Ereignis unabhängig
- A zu sich selbst unabhängig $\implies \mathbb{P}(A) \in \{0, 1\}$
- A, B unabhängig $\implies A, B^c$ unabhängig

Wenn $\mathbb{P}(A) > 0, \mathbb{P}(B) > 0$ gilt:

$$A, B \text{ unabhängig} \iff \mathbb{P}(A|B) = \mathbb{P}(A) \iff \mathbb{P}(B|A) = \mathbb{P}(B)$$

Wir können die Definition der Unabhängigkeit auf beliebige Mengen von Ereignissen erweitern.

Allgemeine Unabhängigkeit

Eine Kollektion von Ereignissen $(A_i; i \in I)$ heisst (**stochastisch**) **unabhängig**, wenn

$$J \subseteq I \text{ endlich} \implies \mathbb{P}\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} \mathbb{P}(A_i)$$

2 Zufallsvariablen

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

Zufallsvariable

Eine (reellwertige) **Zufallsvariable** auf Ω ist eine messbare Funktion $X : \Omega \rightarrow \mathbb{R}$.

$$X : \Omega \rightarrow \mathbb{R} \text{ messbar} \iff \forall x \in \mathbb{R} : X^{-1}(\{x\}) \in \mathcal{A}$$

Die Eigenschaft **messbar** ist bezüglich dem Wahrscheinlichkeitsmass \mathbb{P} relevant (i.e. dann ist $\mathbb{P}(X = x) := \mathbb{P}(\{\omega \in \Omega \mid X(\omega) = x\})$ wohldefiniert).

Diese Definition von **messbar** ist für diskrete Ω äquivalent zu derjenigen der Vorlesung, die die rechte Seite vom ' \iff ' für alle abgeschlossenen Teilmengen $B \subset \mathbb{R}$ fordert.

Für die Messbarkeit von X ist nur $X(\Omega) \subseteq \mathbb{R}$ entscheidend und jede Teilmenge $A \subseteq X(\Omega)$ ist abzählbar (da Ω abzählbar). Somit kann $X^{-1}(A)$ als abzählbare Vereinigung von $\bigcup_{x \in A} X^{-1}(\{x\})$ geschrieben werden.

($\implies X^{-1}(A) \in \mathcal{A}$ per Def. σ -Algebra)

Verteilungsfunktion

Die **Verteilungsfunktion** ist die Abbildung $F_X : \mathbb{R} \rightarrow [0, 1]$ definiert durch:

$$F_X(t) := \mathbb{P}(X \leq t), \forall t \in \mathbb{R}$$

Die Funktion erfüllt folgende Eigenschaften:

- 1. F_X ist monoton wachsend
- 2. F_X ist rechtsstetig, i.e. $\lim_{h \downarrow 0} F_X(x+h) = F_X(x)$
- 3. $\lim_{x \rightarrow -\infty} F_X(x) = 0$ und $\lim_{x \rightarrow \infty} F_X(x) = 1$
- 4. $\forall a, b \in \mathbb{R}, a < b : \mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$

Linksstetigkeit

Die Verteilungsfunktion ist nicht immer linksstetig.

Sei $F_X(a-) := \lim_{h \downarrow 0} F_X(a-h)$ für $a \in \mathbb{R}$ beliebig.

Dann gilt:

$$\mathbb{P}(X = a) = F_X(a) - F_X(a-)$$

Intuitiv folgt daraus

- Wenn F_X in einem Punkt $a \in \mathbb{R}$ nicht stetig ist, dann ist die "Sprunghöhe" $F_X(a) - F_X(a-)$ gleich der Wahrscheinlichkeit $\mathbb{P}(X = a)$.
- Falls F_X stetig in einem Punkt $a \in \mathbb{R}$, dann gilt $\mathbb{P}(X = a) = 0$.

Unabhängigkeit von Zufallsvariablen

Seien X_1, \dots, X_n Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Dann heissen X_1, \dots, X_n **unabhängig**, falls

$$\forall x_1, \dots, x_n \in \mathbb{R} :$$

$$\mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n) = \mathbb{P}(X_1 \leq x_1) \cdot \dots \cdot \mathbb{P}(X_n \leq x_n).$$

Unendlich viele Bernoulli-Experimente

TBD

2.1 Diskrete Zufallsvariablen

Sei $A \in \mathcal{F}$ ein Ereignis.

Wir sagen A tritt **fast sicher (f.s.)** ein, falls $\mathbb{P}(A) = 1$.

Seien $X, Y : \Omega \rightarrow \mathbb{R}$ Zufallsvariablen:

$$X \leq Y \text{ f.s.} \iff \mathbb{P}(X \leq Y) = 1$$

Diskrete Zufallsvariable

Eine Zufallsvariable $X : \Omega \rightarrow \mathbb{R}$ heisst **diskret**, falls eine endliche oder abzählbare Menge $W \subset \mathbb{R}$ existiert, sodass

$$\mathbb{P}(X \in W) = 1$$

Falls Ω endlich oder abzählbar ist, dann ist X immer diskret.

Die **Verteilungsfunktion** einer diskreten ZV X :

$$F_X(x) = \mathbb{P}(X \leq x) = \sum_{y \in W} p(y) \cdot \mathbb{1}_{y \leq x}$$

Die **Gewichtsfunktion** einer diskreten ZV X :

$$\forall x \in X(\Omega) : p(x) = \mathbb{P}(X = x) \text{ wobei } \sum_{x \in X(\Omega)} p(x) = 1$$

2.2 Diskrete Verteilungen

Bernoulli-Verteilung ($X \sim \text{Ber}(p)$):

$X(\Omega) = \{0, 1\}$ und die Gewichtsfunktion ist definiert durch

$$p(1) := \mathbb{P}(X = 1) = p \text{ und } p(0) := \mathbb{P}(X = 0) = 1 - p.$$

Binomialverteilung ($X \sim \text{Bin}(n, p)$):

Wiederholung von n unabhängigen Bernoulli-Experimenten mit gleichem Parameter p .

$$p(k) := \mathbb{P}(X = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \quad \forall k \in \{0, 1, \dots, n\}$$

Geometrische Verteilung ($X \sim \text{Geo}(p)$):

Warten auf den 1-ten Erfolg.

$$p(k) := \mathbb{P}(X = k) = (1-p)^{k-1} \cdot p \quad \forall k \in \mathbb{N} \setminus \{0\}$$

Poisson-Verteilung ($X \sim \text{Poisson}(\lambda)$):

Grenzwert der Binomialverteilung für grosse n und kleine p .

$$p(k) := \mathbb{P}(X = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \quad \forall k \in \mathbb{N}_0, \lambda > 0$$

- 1. Für $X_n \sim \text{Bin}(n, \frac{\lambda}{n})$ gilt $\lim_{n \rightarrow \infty} \mathbb{P}(X_n = k) = \mathbb{P}(Y = k)$ wobei $Y \sim \text{Poisson}(\lambda)$.
- 2. Seien $X_1 \sim \text{Poisson}(\lambda_1)$ und $X_2 \sim \text{Poisson}(\lambda_2)$ unabhängig. Dann gilt $(X_1 + X_2) \sim \text{Poisson}(\lambda_1 + \lambda_2)$.

2.3 Stetige Zufallsvariablen

Stetige Zufallsvariablen, Dichte

Eine Zufallsvariable $X : \Omega \rightarrow \mathbb{R}$ heisst **stetig**, wenn ihre Verteilungsfunktion F_X wie folgt geschrieben werden kann

$$F_X(a) = \int_{-\infty}^a f(x) dx = \text{für alle } a \in \mathbb{R}.$$

wobei $f : \mathbb{R} \rightarrow \mathbb{R}^+$ eine nicht-negative Funktion ist. f wird dann als **Dichte** von X benannt.

Wenn $f : (\mathbb{R}, \mathcal{B}) \rightarrow (\mathbb{R}, \mathcal{B})$ messbar ist, ist die Zufallsvariable X **absolut stetig**.

Intuition: $f(x) dx$ ist die Wahrscheinlichkeit, dass $X \in [x, x + dx]$.

Von F_X zu f :

Sei X eine Zufallsvariable mit stetiger Verteilungsfunktion F_X und F_X stückweise \mathcal{C}^1 , d.h. es gibt $x_0 = -\infty < \dots < x_{n-1} < x_n = +\infty$, sodass F_X auf jedem Intervall (x_i, x_{i+1}) Element von \mathcal{C}^1 ist.

Dann ist X eine stetige Zufallsvariable und die Dichte f kann wie folgt konstruiert werden:

$$\forall x \in (x_i, x_{i+1}) \quad f(x) = F_X'(x).$$

2.4 Stetige Verteilungen

Gleichverteilung ($X \sim \mathcal{U}([a, b])$):

Die Dichte ist auf dem Intervall $[a, b]$ gleich.

$$f_{a,b}(x) = \begin{cases} 0 & x \notin [a, b] \\ \frac{1}{b-a} & x \in [a, b] \end{cases}$$

Exponentialverteilung ($T \sim \text{Exp}(\lambda)$):

Lebensdauer oder Wartezeit eines allg. Ereignisses (Stetiges Äquivalent zur Geometrischen Verteilung).

$$f_\lambda(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

Normalverteilung ($X \sim \mathcal{N}(m, \sigma^2)$):

Häufig verwendete Verteilung. undefiniert für $\sigma = 0$.

$$f_{m,\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

- 1. Seien X_1, \dots, X_n **unabhängige** normalverteilte ZV mit Parametern $(m_1, \sigma_1^2), \dots, (m_n, \sigma_n^2)$, dann ist

$$Z = m_0 + \lambda_1 X_1 + \dots + \lambda_n X_n$$

eine normalverteilte ZV mit Parametern $m = m_0 + \lambda_1 m_1 + \dots + \lambda_n m_n$ und $\sigma^2 = \lambda_1^2 \sigma_1^2 + \dots + \lambda_n^2 \sigma_n^2$.

- 2. Sei $Z \sim \mathcal{N}(0, 1)$ eine **standardnormalverteilte** Zufallsvariable. Dann gilt für $X \sim \mathcal{N}(m, \sigma^2)$

$$X = m + \sigma \cdot Z$$

2.5 Erwartungswert

Erwartungswert - Diskrete ZV

Sei $X : \Omega \rightarrow \mathbb{R}$ eine diskrete Zufallsvariable, $W_X := X(\Omega)$ und $\phi : \mathbb{R} \rightarrow \mathbb{R}$ eine Abbildung. Falls die Summe wohldefiniert ist, gilt:

$$\mathbb{E}(\phi(X)) := \sum_{x \in W_X} \phi(x) \cdot \mathbb{P}(X = x)$$

Wenn $X : \Omega \rightarrow \mathbb{N}_0$, kann man auch den Erwartungswert als

$$\mathbb{E}(X) = \sum_{n=0}^{\infty} \mathbb{P}(X > n)$$

schreiben.

Erwartungswert - Stetige ZV

Sei $X : \Omega \rightarrow \mathbb{R}$ eine stetige Zufallsvariable mit Dichte f . Sei $\phi : \mathbb{R} \rightarrow \mathbb{R}$ eine Abbildung, sodass $\phi(X)$ eine Zufallsvariable ist. Dann gilt

$$\mathbb{E}(\phi(X)) = \int_{-\infty}^{\infty} \phi(x) f(x) dx,$$

solange das Integral wohldefiniert ist.

2.6 Rechnen mit Erwartungswerten

Linearität des Erwartungswertes:

Seien $X, Y : \Omega \rightarrow \mathbb{R}$ ZV mit $\lambda \in \mathbb{R}$, Falls die Erwartungswerte wohldefiniert sind, gilt:

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

Falls X, Y **unabhängig**, dann gilt auch:

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