Wahrscheinlichkeitstheorie und Frequentistische Inferenz

BSc Psychologie WiSe 2024/25

Prof. Dr. Dirk Ostwald

(2) Elementare Wahrscheinlichkeiten

Gemeinsame Wahrscheinlichkeiten

Bedingte Wahrscheinlichkeiten

Unabhängige Ereignisse

Selbstkontrollfragen

Gemeinsame Wahrscheinlichkeiten

Bedingte Wahrscheinlichkeiten

Unabhängige Ereignisse

Selbstkontrollfragen

Definition (Gemeinsame Wahrscheinlichkeit)

$(\Omega, \mathcal{A}, \mathbb{P})$ sei ein Wahrscheinlichkeitsraum und es seien $A, B \in \mathcal{A}$. Dann heißt

$$\mathbb{P}(A \cap B) \tag{1}$$

die *gemeinsame Wahrscheinlichkeit von A und B*.

Bemerkungen

- Zur Abgrenzung nennt man $\mathbb{P}(A)$ auch manchmal auch *totale Wahrscheinlichkeit von A*.
- Mit dem Theorem zur Abgeschlossenheit von σ-Algebren bezüglich Durchschnitten ist gesichert, dass $A \cap B \in \mathcal{A}$, woraus die Existenz von $\mathbb{P}(A \cap B) \in [0, 1]$ folgt.
- Intuitiv entspricht $\mathbb{P}(A \cap B)$ der Wahrscheinlichkeit, dass A und B gleichzeitig eintreten.
- In der Mechanik des W-Raummodells ist $\mathbb{P}(A \cap B)$ die Wahrscheinlichkeit, dass in einem Durchgang des Zufallsvorgang ein ω mit sowohl $\omega \in A$ und $\omega \in B$ realisiert wird.

© 2025 Dirk Ostwald CC BY 4.0

Gemeinsame Wahrscheinlichkeiten

Beispiel

- Wir betrachten das Wahrscheinlichkeitsraummodells des Werfens eines fairen Würfels.
- Wir betrachten die Ereignisse

$$A := \{2, 4, 6\} \qquad \text{Es fällt eine gerade Augenzahl}$$
$$B := \{4, 5, 6\} \qquad \text{Es fällt eine Augenzahl größer als Drei}$$

- Dann gilt $A \cap B = \{4, 6\}$.
- Die Interpretation von $A \cap B = \{4, 6\}$ ist

  “Es fällt eine gerade Augenzahl und diese Augenzahl ist größer als Drei.”

  Mit $\mathbb{P}(\{4\}) = \mathbb{P}(\{6\}) = \frac{1}{6}$ und der σ-Additivität von $\mathbb{P}$ ergibt sich

$$\begin{aligned}
\mathbb{P}(A \cap B) &= \mathbb{P}(\{2, 4, 6\} \cap \{4, 5, 6\}) \\
&= \mathbb{P}(\{4, 6\}) \\
&= \mathbb{P}(\{4\} \cup \{6\}) \\
&= \mathbb{P}(\{4\}) + \mathbb{P}(\{6\}) \\
&= \frac{1}{6} + \frac{1}{6} \\
&= \frac{1}{3}.
\end{aligned} \tag{2}$$

Gemeinsame Wahrscheinlichkeiten

Interpretation von $\mathbb{P}(A \cup B)$

- $\mathbb{P}(A \cap B)$ und $\mathbb{P}(A \cup B)$ sollten nicht verwechselt werden.
- $\cup$ entspricht dem *inklusiven oder*, also *und/oder*.
- Δ entspricht dem *exklusiven oder*, also *entweder..., oder ..., aber nicht beides*.
- $A \cup B$ entspricht also dem Ereignis, dass A und/oder B eintritt.
- $\omega \in A \cup B$ ist also schon erfüllt, wenn "nur" $\omega \in A$ oder "nur" $\omega \in B$ gilt.
- Für obiges Beispiel gilt

$$A \cup B = \{2, 4, 6\} \cup \{4, 5, 6\} = \{2, 4, 5, 6\} \tag{3}$$

mit der Interpretation

"Es fällt eine gerade Augenzahl oder es fällt eine Augenzahl größer als Drei oder es fällt eine gerade Augenzahl und diese Augenzahl ist größer als Drei".

und der Wahrscheinlichkeit

$$\mathbb{P}(\{2, 4, 5, 6\}) = 4/6 = 2/3. \tag{4}$$

Theorem (Weitere Eigenschaften von Wahrscheinlichkeiten)

$(\Omega, \mathcal{A}, \mathbb{P})$ sei ein Wahrscheinlichkeitsraum und es seien $A, B \in \mathcal{A}$ Ereignisse. Dann gelten

1. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.
2. $A \subset B \Rightarrow \mathbb{P}(A) \le \mathbb{P}(B)$.
3. $\mathbb{P}(A \cap B^c) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$
4. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

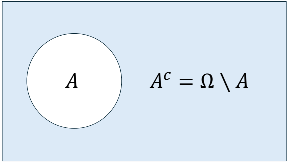
Bemerkungen

- Die Eigenschaften sind in der Analyse probabilistischer Modelle oft nützlich.

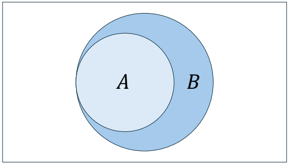
© 2025 Dirk Ostwald CC BY 4.0

Gemeinsame Wahrscheinlichkeiten

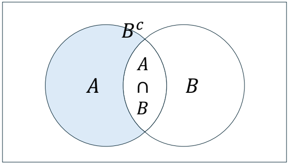
$$\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$$



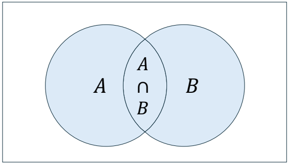
$$A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$$



$$\mathbb{P}(A \cap B^c) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$$



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

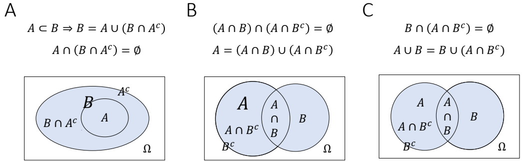


Die hellblau unterlegten Flächen entsprechen jeweils den Wahrscheinlichkeiten von Interesse.

Gemeinsame Wahrscheinlichkeiten

Beweis

Die zweite, dritte, und vierte Aussage dieses Theorems basieren auf elementaren mengentheoretischen Aussagen und der σ-Addivität von $\mathbb{P}$. Wir wollen diese elementaren mengentheoretischen Aussagen hier nicht beweisen, sondern verweisen jeweils auf die Intuition, die durch die Venn-Diagramme in untenstehender Abbildung vermittelt wird.



Zu 1.: Wir halten zunächst fest, dass aus $A^c := \Omega \setminus A$ folgt, dass $A^c \cup A = \Omega$ und dass $A^c \cap A = \emptyset$. Mit der Nomiertheit und der σ-Additivität von $\mathbb{P}$ folgt dann

$$\mathbb{P}(\Omega) = 1 \Leftrightarrow \mathbb{P}(A^c \cup A) = 1 \Leftrightarrow \mathbb{P}(A^c) + \mathbb{P}(A) = 1 \Leftrightarrow \mathbb{P}(A^c) = 1 - \mathbb{P}(A). \tag{5}$$

Zu 2.: Zunächst gilt (vgl. Abbildung A)

$$A \subset B \Rightarrow B = A \cup (B \cap A^c) \text{ mit } A \cap (B \cap A^c) = \emptyset. \tag{6}$$

Mit der σ-Additivät von $\mathbb{P}$ folgt dann aber

$$\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \cap A^c). \tag{7}$$

Mit $\mathbb{P}(B \cap A^c) \geq 0$ folgt dann $\mathbb{P}(A) \leq \mathbb{P}(B)$.

Gemeinsame Wahrscheinlichkeiten

Beweis (fortgeführt)

Man beachte, dass bei $A \subset B$, aber $\mathbb{P}(B \cap A^c) = 0$ gilt, dass $\mathbb{P}(A) = \mathbb{P}(B)$. Ein Beispiel für diesen Umstand ergibt sich in dem Wahrscheinlichkeitsraum mit überabzählbarem Ergebnisraum $\Omega := \mathbb{R}$ für die Ereignisse $A := [0, 1[$ und $B := [0, 1]$ mit $\mathbb{P}(B) > 0$. Dann gilt $1 \in B$ aber $1 \notin A$, also $A \subset B$. Gleichzeitig gilt aber, wie wir im Kontext kontinuierlicher Zufallsvariablen sehen werden, $\mathbb{P}(B \cap A^c) = \mathbb{P}(1) = 0$ und damit $\mathbb{P}(A) = \mathbb{P}(B)$.

Zu 3.: Zunächst gilt (vgl. Abbildung B)

$$(A \cap B) \cap (A \cap B^c) = \emptyset \text{ und } A = (A \cap B) \cup (A \cap B^c). \tag{8}$$

Mit der σ-Addivität von $\mathbb{P}$ folgt dann

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c) \Leftrightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A) - \mathbb{P}(A \cap B^c) \tag{9}$$

Zu 4.: Zunächst gilt (vgl. Abbildung C)

$$B \cap (A \cap B^c) = \emptyset \text{ und } A \cup B = B \cup (A \cap B^c). \tag{10}$$

Mit der σ-Addivität von $\mathbb{P}$ folgt dann

$$\mathbb{P}(A \cup B) = \mathbb{P}(B) + \mathbb{P}(A \cap B^c). \tag{11}$$

Mit 3. folgt dann

$$\mathbb{P}(A \cup B) = \mathbb{P}(B) + \mathbb{P}(A) - \mathbb{P}(A \cap B) \tag{12}$$

Gemeinsame Wahrscheinlichkeiten

Bedingte Wahrscheinlichkeiten

Unabhängige Ereignisse

Selbstkontrollfragen

Definition (Bedingte Wahrscheinlichkeit)

$(\Omega, \mathcal{A}, \mathbb{P})$ sei ein Wahrscheinlichkeitsraum und $A, B \in \mathcal{A}$ seien Ereignisse mit $\mathbb{P}(B) > 0$. Die *bedingte Wahrscheinlichkeit des Ereignisses A gegeben das Ereignis B* ist definiert als

$$\mathbb{P}(A|B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}. \tag{13}$$

Bemerkungen

- $\mathbb{P}(A|B)$ ist die mit $1/\mathbb{P}(B)$ skalierte gemeinsame Wahrscheinlichkeit $\mathbb{P}(A \cap B)$.
- Die Festlegung der gemeinsamen Wahrscheinlichkeit $\mathbb{P}(A \cap B)$ legt $\mathbb{P}(A|B)$ schon fest.
- Üblicherweise gilt $\mathbb{P}(A|B) \neq \mathbb{P}(B|A)$, z.B.

$$\mathbb{P}(\text{Tod}|\text{Erhängen}) \neq \mathbb{P}(\text{Erhängen}|\text{Tod}). \tag{14}$$

- Eine Verallgemeinerung für $\mathbb{P}(B) = 0$ ist möglich, aber technisch aufwändig.

© 2025 Dirk Ostwald CC BY 4.0

Bedingte Wahrscheinlichkeiten

Werfen eines Würfels

Wir betrachten erneut das Modell $(\Omega, \mathcal{A}, \mathbb{P})$ des fairen Würfels. Wir wollen die bedingte Wahrscheinlichkeit für das Ereignis "Es fällt eine gerade Augenzahl" gegeben das Ereignis "Es fällt eine Zahl größer als Drei" berechnen. Wir haben oben bereits gesehen, dass das Ereignis "Es fällt eine gerade Augenzahl" der Teilmenge $A := \{2, 4, 6\}$ von Ω entspricht, und dass das Ereignis "Es fällt eine Zahl größer als Drei" der Teilmenge $B := \{4, 5, 6\}$ von Ω entspricht. Weiterhin haben wir gesehen, dass unter der Annahme, dass der modellierte Würfel fair ist, gilt, dass

$$\mathbb{P}(\{2, 4, 6\}) = \mathbb{P}(\{2\}) + \mathbb{P}(\{4\}) + \mathbb{P}(\{6\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2} \tag{15}$$

$$\mathbb{P}(\{4, 5, 6\}) = \mathbb{P}(\{4\}) + \mathbb{P}(\{5\}) + \mathbb{P}(\{6\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}. \tag{16}$$

Schließlich hatten wir auch gesehen, dass das Ereignis $A \cap B$, also das Ereignis "Es fällt eine gerade Augenzahl, die größer als Drei ist", die Wahrscheinlichkeit

$$\mathbb{P}(A \cap B) = \mathbb{P}(\{2, 4, 6\} \cap \{4, 5, 6\}) = \mathbb{P}(\{4, 6\}) = \mathbb{P}(\{4\}) + \mathbb{P}(\{6\}) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3} \tag{17}$$

hat. Nach Definition der bedingten Wahrscheinlichkeit ergibt sich dann direkt

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\{4, 6\})}{\mathbb{P}(\{4, 5, 6\})} = \frac{\mathbb{P}(\{4\} \cup \{6\})}{\mathbb{P}(\{4\} \cup \{5\} \cup \{6\})} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{1}{3} \cdot \frac{2}{1} = \frac{2}{3}. \tag{18}$$

Wenn man also weiß, dass eine Augenzahl größer als Drei gefallen ist, ist die Wahrscheinlichkeit, dass es sich dabei um eine gerade Augenzahl handelt 2/3. Wenn man ersteres nicht weiß, ist die Wahrscheinlichkeit für das Fallen einer geraden Augenzahl (nur) 1/2. Dies Beispiel verdeutlicht insbesondere, dass das Bedingen auf einem Ereignis der Inkorporation von neuem Wissen in wahrscheinlichkeitstheoretische Modellen entspricht.

Wert eines BDI-II Items

Wir betrachten erneut das Beispiel für den Wert einer Patient:in für das Item "Traurigkeit" der deutschen Version des BDI-II. Wir haben oben den möglichen Itemantworten (0), (1), (2), (3) die Wahrscheinlichkeiten

$$\mathbb{P}(\{0\}) := \frac{2}{10}, \quad \mathbb{P}(\{1\}) := \frac{3}{10}, \quad \mathbb{P}(\{2\}) := \frac{4}{10}, \quad \mathbb{P}(\{3\}) := \frac{1}{10}, \tag{19}$$

zugeordnet. Wir nehmen nun an, dass wir von einer Patient:in wissen, dass sie einen Wert größer als (1) gewählt hat und fragen nach den Wahrscheinlichkeiten, dass die Patient:in den Wert (2) oder (3) gewählt hat. Dazu betrachten wir zunächst die Wahrscheinlichkeit des Ereignisses "Die Patient:in hat einen Wert größer als (1) gewählt". Offenbar entspricht dies der Teilmenge $B := \{2,3\}$ von Ω. Mit der σ-Additivität ergibt sich für die Wahrscheinlichkeit für dieses Ereignis wie oben gesehen $\mathbb{P}(\{2,3\}) = 1/2$. Wir betrachten also die Ereignisse

Verbale Beschreibung	Mengenform
Die Patient:in hat einen Wert größer als (1) gewählt	$\omega \in B := \{2,3\}$
Die Patient:in hat Antwort (2) gewählt	$\omega \in A_1 := \{2\}$
Die Patient:in hat Antwort (3) gewählt	$\omega \in A_2 := \{3\}$

Es ergeben sich

$$\mathbb{P}(A_1|B) = \frac{\mathbb{P}(A_1 \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\{2\} \cap \{2,3\})}{\mathbb{P}(\{2,3\})} = \frac{\mathbb{P}(\{2\})}{\mathbb{P}(\{2,3\})} = \frac{4}{10} \cdot \frac{2}{1} = \frac{8}{10} \tag{20}$$

und

$$\mathbb{P}(A_2|B) = \frac{\mathbb{P}(A_2 \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\{3\} \cap \{2,3\})}{\mathbb{P}(\{2,3\})} = \frac{\mathbb{P}(\{3\})}{\mathbb{P}(\{2,3\})} = \frac{1}{10} \cdot \frac{2}{1} = \frac{2}{10}. \tag{21}$$

Man beachte, dass

$$\mathbb{P}(A_1|B) + \mathbb{P}(A_2|B) = \frac{8}{10} + \frac{2}{10} = 1. \tag{22}$$

Bedingte Wahrscheinlichkeiten

Theorem (Bedingte Wahrscheinlichkeit)

Für ein festes $B \in \mathcal{A}$ mit $\mathbb{P}(B) > 0$ sei

$$\mathbb{P}(\cdot|B) : \mathcal{A} \to [0,1], A \mapsto \mathbb{P}(A|B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \tag{23}$$

Dann ist $\mathbb{P}(\cdot|B)$ ein Wahrscheinlichkeitsmaß und heißt die *bedingte Wahrscheinlichkeit gegeben Ereignis* B.

Bemerkungen

- Die Rechenregeln der Wahrscheinlichkeitstheorie gelten für die Ereignisse links des Strichs.
- Im Gegensatz zu $\mathbb{P}(\cdot \cap B)$ definiert $\mathbb{P}(\cdot|B)$ ein Wahrscheinlichkeitsmaß für alle $A \in \mathcal{A}$.
- Beispielsweise gelten für $0 < \mathbb{P}(B) < 1$, dass $\mathbb{P}(\Omega|B) = 1$ aber $\mathbb{P}(\Omega \cap B) = \mathbb{P}(B) < 1$.
- $\mathbb{P}(\cdot|B)$ ist also normalisiert, $\mathbb{P}(\cdot \cap B)$ dagegen nicht.

Beweis

Wir weisen die definierenden Eigenschaften eines Wahrscheinlichkeitsmaßes nach.

(1) $\mathbb{P}(A|B) \geq 0$ für alle $A \in \mathcal{A}$

Nach Definition der bedingten Wahrscheinlichkeit gilt $\mathbb{P}(A|B) = \mathbb{P}(A \cap B)/\mathbb{P}(B)$ mit $\mathbb{P}(B) > 0$. Mit dem Theorem zur Abgeschlossenheit von σ-Algebren bezüglich Durchschnitten und der Definiton von $\mathbb{P}$ gilt $\mathbb{P}(A \cap B) \geq 0$ und damit folgt die Aussage.

(2) $\mathbb{P}(\Omega|B) = 1.$

Es gilt

$$\mathbb{P}(\Omega|B) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1. \tag{24}$$

(3) $\mathbb{P}\left(\cup_{i=1}^{\infty} A_i|B\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i|B)$ für paarweise disjunkte $A_1, A_2, ... \in \mathcal{A}$.

Mit dem Theorem zu den Assoziativ- und Distributivgesetzen von Vereinigung und Durchschnitt zunächst, dass

$$\left(\cup_{i=1}^{\infty} A_i\right) \cap B = (A_1 \cup A_2 \cup \cdots) \cap B = (A_1 \cap B) \cup (A_2 \cap B) \cup \cdots = \cup_{i=1}^{\infty} (A_i \cap B). \tag{25}$$

Weiterhin gilt mit $A_i \cap A_j = \emptyset$ und dem Theorem zu den Assoziativ- und Distributivgesetzen von Vereinigung und Durchschnitt, dass auch die $A_i \cap B$ für $i = 1, 2, ...$ paarweise disjunkt sind, denn

$$(A_i \cap B) \cap \left(A_j \cap B\right) = \left(A_i \cap A_j\right) \cap (B \cap B) = \emptyset \cap B = \emptyset. \tag{26}$$

Mit der σ-Additivität von $\mathbb{P}$ folgt damit dann, dass

$$\mathbb{P}\left(\cup_{i=1}^{\infty} A_i|B\right) = \frac{\mathbb{P}\left(\left(\cup_{i=1}^{\infty} A_i\right) \cap B\right)}{\mathbb{P}(B)} = \frac{\mathbb{P}\left(\cup_{i=1}^{\infty} (A_i \cap B)\right)}{\mathbb{P}(B)} = \frac{\sum_{i=1}^{\infty} \mathbb{P}(A_i \cap B)}{\mathbb{P}(B)} = \sum_{i=1}^{\infty} \frac{\mathbb{P}(A_i \cap B)}{\mathbb{P}(B)} = \sum_{i=1}^{\infty} \mathbb{P}(A_i|B). \tag{27}$$

$\square$

Theorem (Gemeinsame und bedingte Wahrscheinlichkeiten)

Es seien $(\Omega, \mathcal{A}, \mathbb{P})$ ein W-Raum und $A, B \in \mathcal{A}$ mit $\mathbb{P}(A) > 0$ und $\mathbb{P}(B) > 0$. Dann gilt

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(A). \tag{28}$$

Beweis

Mit der Definition der jeweiligen bedingten Wahrscheinlichkeit folgen direkt

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \Leftrightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B) \tag{29}$$

und

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} \Leftrightarrow \mathbb{P}(A \cap B) = \mathbb{P}(B|A)\mathbb{P}(A). \tag{30}$$

Bemerkung

- Gemeinsame Wahrscheinlichkeiten können aus bedingten und totalen Wahrscheinlichkeiten berechnet werden.
- Definition von $\mathbb{P}(A)$ und $\mathbb{P}(B|A)$ definiert $\mathbb{P}(A \cap B)$ implizit mit.

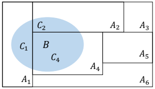
Theorem (Gesetz der totalen Wahrscheinlichkeit)

$(\Omega, \mathcal{A}, \mathbb{P})$ sei ein Wahrscheinlichkeitsraum und $A_1, ..., A_k$ sei eine Partition von Ω. Dann gilt für jedes $B \in \mathcal{A}$, dass

$$\mathbb{P}(B) = \sum_{i=1}^{k} \mathbb{P}(B \cap A_i). \tag{31}$$

<u>Beweis</u>

Für $i = 1, ..., k$ sei $C_i := B \cap A_i$, so dass $\cup_{i=1}^k C_i = B$ und $C_i \cap C_j = \emptyset$ für $1 \leq i, j \leq k, i \neq j$, visuell



Dann gilt mit der σ-Additivität von $\mathbb{P}$, dass

$$\mathbb{P}(B) = \mathbb{P}\left(\cup_{i=1}^k C_i\right) = \sum_{i=1}^{k} \mathbb{P}(C_i) = \sum_{i=1}^{k} \mathbb{P}(B \cap A_i). \tag{32}$$

Bemerkung

- Totale Wahrscheinlichkeiten können aus gemeinsamen Wahrscheinlichkeiten berechnet werden.
- Im Kontext von Zufallsvariablen wird dieser Prozess auch als *Marginalisierung* bezeichnet.

© 2025 Dirk Ostwald CC BY 4.0

Bedingte Wahrscheinlichkeiten

Beispiel

Als Beispiel für die Darstellung der Wahrscheinlichkeit eines Ereignisses mithilfe des Gesetzes des totalen Wahrscheinlichkeit betrachten wir das Wahrscheinlichkeitsraummodell des fairen Würfels. Für den Ergebnisraum $\Omega := \{1,2,3,4,5,6\}$ betrachten wir das Ereignis $B := \{2,4,6\}$ ("Es fällt eine gerade Zahl") mit Wahrscheinlichkeit $\mathbb{P}(B) = \frac{1}{2}$ sowie die Partition $A_1 := \{1,2\}$, $A_2 := \{3\}$, $A_3 := \{4,5\}$, $A_4 := \{6\}$ von Ω. Dann gelten offenbar

$$
\begin{aligned}
\mathbb{P}(B \cap A_1) &= \mathbb{P}(\{2,4,6\} \cap \{1,2\}) &&= \mathbb{P}(\{2\}) &&= \tfrac{1}{6} \\
\mathbb{P}(B \cap A_2) &= \mathbb{P}(\{2,4,6\} \cap \{3\}) &&= \mathbb{P}(\emptyset) &&= 0 \\
\mathbb{P}(B \cap A_3) &= \mathbb{P}(\{2,4,6\} \cap \{4,5\}) &&= \mathbb{P}(\{4\}) &&= \tfrac{1}{6} \\
\mathbb{P}(B \cap A_4) &= \mathbb{P}(\{2,4,6\} \cap \{6\}) &&= \mathbb{P}(\{6\}) &&= \tfrac{1}{6}
\end{aligned}
$$

Damit ergibt sich dann

$$
\sum_{i=1}^{4} \mathbb{P}(B \cap A_i) = \mathbb{P}(B \cap A_1) + \mathbb{P}(B \cap A_2) + \mathbb{P}(B \cap A_3) + \mathbb{P}(B \cap A_4) = \frac{1}{6} + 0 + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}. \tag{33}
$$

und somit

$$
\mathbb{P}(B) = \sum_{i=1}^{4} \mathbb{P}(B \cap A_i). \tag{34}
$$

Man beachte, dass die Partitionseigenschaft von A_1, A_2, A_3, A_4 essentiell für die Validität des Theorems in diesem Beispiel ist. Wäre z.B. $A_4 = \{4,6\}$ und A_3 und A_4 damit nicht disjunkt, so ergäbe die rechte Seite des Gesetzes von der totalen Wahrscheinlichkeit einen Wert größer $\mathbb{P}(B)$. Ergäbe die Vereinigung der A_1, A_2, A_3, A_4 dagegen nicht Ω, weil zum Beispiel $A_4 = \emptyset$, so ergäbe sich umgekehrt ein Wert kleiner $\mathbb{P}(B)$.

Theorem (Theorem von Bayes)

$(\Omega, \mathcal{A}, \mathbb{P})$ sei ein Wahrscheinlichkeitsraum und $A_1, ..., A_k$ sei eine Partition von Ω mit $\mathbb{P}(A_i) > 0$ für alle $i = 1, ..., k$. Wenn $\mathbb{P}(B) > 0$ gilt, dann gilt für jedes $i = 1, ..., k$, dass

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\sum_{i=1}^k \mathbb{P}(B|A_i)\mathbb{P}(A_i)}. \tag{35}$$

<u>Beweis</u>

Mit der Definition der bedingten Wahrscheinlichkeit und dem Gesetz der totalen Wahrscheinlichkeit gilt

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(A_i \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\sum_{i=1}^k \mathbb{P}(B \cap A_i)}. = \frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\sum_{i=1}^k \mathbb{P}(B|A_i)\mathbb{P}(A_i)}. \tag{36}$$

$\square$

Bedingte Wahrscheinlichkeiten

Bemerkungen

- Das Theorem von Bayes ist eine zu $\mathbb{P}(A_i \cap B)/\mathbb{P}(B)$ alternative Formel, um die bedingte Wahrscheinlichkeit $\mathbb{P}(A_i|B)$ zu berechnen.
- Das Theorem von Bayes gilt unabhängig von der Frequentistischen oder Bayesianischen Interpretation von Wahrscheinlichkeiten.
- Im Rahmen der *Frequentistischen Inferenz* wird das Theorem von Bayes recht selten benutzt; hier steht vor allem die Tatsache $\mathbb{P}(A) = \mathbb{P}(A|B)$ bei Unabhängikeit von A und B im Vordergrund.
- Im Rahmen der *Bayesianischen Inferenz* ist das Theorem von Bayes zentral; hier wird $\mathbb{P}(A_i)$ oft *Prior Wahrscheinlichkeit* und $\mathbb{P}(A_i|B)$ oft *Posterior Wahrscheinlichkeit des Ereignisses* A_i genannt. Wie oben erläutert entspricht $\mathbb{P}(A_i|B)$ dann der aktualisierten Wahrscheinlichkeit von A_i, wenn man um das Eintreten von B weiß.

Bedingte Wahrscheinlichkeiten

Die drei zentralen Rechenregeln der Wahrscheinlichkeitstheorie

Summationsregel (Gesetz von der totalen Wahrscheinlichkeit)

$$\mathbb{P}(B) = \sum_{i=1}^{k} \mathbb{P}(B \cap A_i), \text{ wenn } \Omega = \cup_{i=1}^{k} A_i \text{ für } A_i \cap A_j = \emptyset \tag{37}$$

Multiplikationsregel (Gemeinsame und bedingte Wahrscheinlichkeiten)

$$\mathbb{P}(A \cap B) = \mathbb{P}(B|A)\mathbb{P}(A) \tag{38}$$

Theorem von Bayes

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)} \tag{39}$$

Gemeinsame Wahrscheinlichkeiten

Bedingte Wahrscheinlichkeiten

Unabhängige Ereignisse

Selbstkontrollfragen

Unabhängige Ereignisse

Definition (Unabhängige Ereignisse)

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

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B). \tag{40}$$

Eine Menge von Ereignissen $\{A_i | i \in I\} \subset \mathcal{A}$ mit beliebiger Indexmenge I heißt unabhängig, wenn für jede endliche und mindestens zweielementige Untermenge $J \subseteq I$ gilt, dass

$$\mathbb{P}\left(\cap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}(A_j). \tag{41}$$

Bemerkungen

- Die Unabhängigkeit bestimmter Ereignissen kann in der Definition eines probabilistischen Modells vorausgesetzt werden, so zum Beispiel die Unabhängigkeit von Fehlervariablen in statistischen Modellen.
- Die Unabhängigkeit von Ereignissen kann aus der Definition eines probabilistischen Modells folgen.
- Disjunkte Ereignisse mit von Null verschiedener Wahrscheinlichkeit sind nie unabhängig:

$$\mathbb{P}(A)\mathbb{P}(B) > 0, \text{ aber } \mathbb{P}(A \cap B) = \mathbb{P}(\emptyset) = 0, \text{ also } \mathbb{P}(A \cap B) \neq \mathbb{P}(A)\mathbb{P}(B). \tag{42}$$

- Der Sinn der Produkteigenschaft bei Unabhängkeit erschließt sich im Kontext *bedingter Wahrscheinlichkeiten*.
- Die Bedingung der beliebigen Untermengen von I sichert die paarweise Unabhängigkeit der $A_i, i \in I$.

Unabhängige Ereignisse

Beispiele

Unabhängige Ereignisse sind nach Definition immer auch paarweise unabhängig. Beispielsweise gilt für die Menge $\{A_i|i \in \{1,2,3\}\}$ von Ereignissen und die möglichen Untermengen $J_1 := \{1,2\}$, $J_2 := \{1,3\}$, $J_3 := \{2,3\}$, $J := \{1,2,3\}$ der Indexmenge $I = \{1,2,3\}$, dass die Definitionsbedingun der Unabhängigkeit von A_1, A_2, A_3 die Aussagen

$$
\begin{aligned}
\mathbb{P}(A_1 \cap A_2) &= \mathbb{P}(A_1)\,\mathbb{P}(A_2) \\
\mathbb{P}(A_1 \cap A_3) &= \mathbb{P}(A_1)\,\mathbb{P}(A_3) \\
\mathbb{P}(A_2 \cap A_3) &= \mathbb{P}(A_2)\,\mathbb{P}(A_3) \\
\mathbb{P}(A_1 \cap A_2 \cap A_3) &= \mathbb{P}(A_1)\,\mathbb{P}(A_2)\,\mathbb{P}(A_2)
\end{aligned}
\tag{43}
$$

impliziert.

Andersherum kann aus der paarweisen Unabhängigkeit von Ereignissen allerdings nicht auf die Unabhängigkeit von Ereignissen geschlossen werden, wie folgendes Beispiel zeigt (cf. DeGroot and Schervish (2012)): Wir betrachten das Modell des zweifachen Werfens einer fairen Münze mit dem Ergebnisraum

$$\Omega := \{HH, HT, TH, TT\} \tag{44}$$

mit den Elementareigniswahrscheinlichkeiten

$$\mathbb{P}\{HH\} = \mathbb{P}\{HT\} = \mathbb{P}\{TH\} = \mathbb{P}\{TT\} = \frac{1}{4} \tag{45}$$

Unabhängige Ereignisse

Beispiele (fortgeführt)

Dann sind die Ereignisse

Verbale Beschreibung	Mengenform
Im ersten Wurf fällt Kopf	$A_1 := \{HH, HT\}$
Im zweiten Wurf fällt Kopf	$A_2 := \{HH, TH\}$
In beiden Würfen fällt das gleiche Ergebnis	$A_3 := \{HH, TT\}$

zwar paarweise unabhängig, aber nicht unabhängig. Es gelten hier offenbar

$$\mathbb{P}(A_1) = \mathbb{P}(A_2) = \mathbb{P}(A_2) = \frac{1}{2} \tag{46}$$

sowie

$$\begin{aligned}
\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(\{HH\}) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(A_1)\,\mathbb{P}(A_2) \\
\mathbb{P}(A_1 \cap A_3) = \mathbb{P}(\{HH\}) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(A_2)\,\mathbb{P}(A_3) \\
\mathbb{P}(A_2 \cap A_3) = \mathbb{P}(\{HH\}) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(A_1)\,\mathbb{P}(A_3)
\end{aligned} \tag{47}$$

Die Ereignisse sind also paarweise unabhängig. Allerdings gilt auch

$$\mathbb{P}(A_1 \cap A_2 \cap A_3) = \mathbb{P}(\{HH\}) = \frac{1}{4} \neq \frac{1}{8} = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(A_1)\,\mathbb{P}(A_2)\,\mathbb{P}(A_3) \tag{48}$$

und die Ereignisse A_1, A_2, A_3 sind damit nicht unabhängig.

Unabhängige Ereignisse

Theorem (Bedingte Wahrscheinlichkeit unter Unabhängigkeit)

$(\Omega, \mathcal{A}, \mathbb{P})$ sei ein Wahrscheinlichkeitsraum und $A, B \in \mathcal{A}$ seien unabhängige Ereignisse mit $\mathbb{P}(B) > 0$. Dann gilt

$$\mathbb{P}(A|B) = \mathbb{P}(A). \tag{49}$$

<u>Beweis</u>

Unter den Voraussetzungen des Theorems gilt

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A)\mathbb{P}(B)}{\mathbb{P}(B)} = \mathbb{P}(A). \tag{50}$$

$\square$

Bemerkungen

- Bei Unabhängigkeit von A und B ist es irrelevant, ob auch B eintritt, $\mathbb{P}(A)$ bleibt gleich.
- Unabhängigkeit wird also als $\mathbb{P}(A)\mathbb{P}(B)$ modelliert, damit $\mathbb{P}(A|B) = \mathbb{P}(A)$ gilt.
- Die Unabhängigkeit zweier Ereignisse bedeutet also, dass das Wissen um das Eintreten eines der beiden Ereignisse die Wahrscheinlichkeit für das Eintreten des anderen Ereignisses nicht ändert. Andersherum bedeutet Abhängigkeit, dass das Wissen um das Eintreten eines der beiden Ereignisse die Wahrscheinlichkeit für das Eintreten des anderen Ereignisses verändert, also erhöht oder verringert.

Gemeinsame Wahrscheinlichkeiten

Bedingte Wahrscheinlichkeiten

Unabhängige Ereignisse

Selbstkontrollfragen

Selbstkontrollfragen

1. Geben Sie die Definition der gemeinsamen Wahrscheinlichkeit zweier Ereignisse wieder.
2. Erläutern Sie die intuitive Bedeutung der gemeinsamen Wahrscheinlichkeit zweier Ereignisse.
3. Geben Sie das Theorem zu weiteren Eigenschaften von Wahrscheinlichkeiten wieder.
4. Geben Sie die Definition der bedingten Wahrscheinlichkeit eines Ereignisses wieder.
5. Geben Sie das Theorem zur bedingten Wahrscheinlichkeit wieder.
6. Geben Sie das Theorem zu gemeinsamen und bedingten Wahrscheinlichkeiten wieder.
7. Geben Sie das Gesetz der totalen Wahrscheinlichkeit wieder.
8. Geben Sie das Theorem von Bayes wieder.
9. Geben Sie den Beweis des Theorems von Bayes wieder.
10. Geben Sie die Definition der Unabhängigkeit zweier Ereignisse wieder.
11. Geben Sie das Theorem zur bedingten Wahrscheinlichkeit unter Unabhängigkeit wieder.
12. Geben Sie den Beweis des Theorems zur bedingten Wahrscheinlichkeit unter Unabhängigkeit wieder.
13. Erläutern Sie das Theorem zur bedingten Wahrscheinlichkeit unter Unabhängigkeit.

Selbstkontrollfragen - Lösungen

1. Siehe Definition (Gemeinsame Wahrscheinlichkeit).
2. Siehe Bemerkungen zu Definition (Gemeinsame Wahrscheinlichkeit).
3. Siehe Theorem (Weitere Eigenschaften von Wahrscheinlichkeiten).
4. Siehe Definition (Bedingte Wahrscheinlichkeit). Die bedingte Wahrscheinlichkeit eines Ereignisses gegebem ein anderers Ereignis ist ein Wert in $[0, 1]$.
5. Siehe Theorem (Bedingte Wahrscheinlichkeit). Die bedingte Wahrscheinlichkeit gegeben ein Ereignis ist ein Wahrscheinlichkeitsmaß.
6. Siehe Theorem (Gemeinsame und bedingte Wahrscheinlichkeit).
7. Siehe Theorem (Gesetz der totalen Wahrscheinlichkeit).
8. Siehe Theorem (Theorem von Bayes).
9. Siehe Beweis zum Theorem (Theorem von Bayes).
10. Siehe Definition (Unabhängige Ereignisse).
11. Siehe Theorem (Bedingte Wahrscheinlichkeit unter Unabhängigkeit).
12. Siehe Beweis zum Theorem (Bedingte Wahrscheinlichkeit unter Unabhängigkeit).
13. Siehe Bemerkungen zum Theorem (Bedingte Wahrscheinlichkeit unter Unabhängigkeit).

Referenzen

DeGroot, Morris H., and Mark J. Schervish. 2012. *Probability and Statistics*. 4th ed. Boston: Addison-Wesley.