Wahrscheinlichkeitstheorie und Frequentistische Inferenz

BSc Psychologie WiSe 2021/22

Prof. Dr. Dirk Ostwald

(3) Elementare Wahrscheinlichkeiten

Repräsentation zentraler Eigenschaften

Modellierung

Modell

Wahrscheinlichkeitstheorie

Probabilistisches Modell

$(\Omega, \mathcal{A}, \mathbb{P}), X : \Omega \to \mathbb{R}$

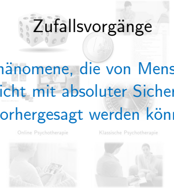
Wahrscheinlichkeitsrechnung

$\mathbb{P}_X(S) = \mathbb{P}(X^{-1}(S))$

Realität

Zufallsvorgänge

Phänomene, die von Menschen nicht mit absoluter Sicherheit vorhergesagt werden können.



Vorhersagen

Quantifizierung von Unsicherheit

Jede Augenzahl kommt im Mittel gleich häufig vor.
Ich denke, jede Augenzahl hat die gleiche Wahrscheinlichkeit.

Modellierung

Modell

Wahrscheinlichkeitstheorie

$\Omega := \{1,2,3,4,5,6\}$

$\mathcal{A} := \mathcal{P}(\Omega)$

$\mathbb{P}(\{\omega\}) := 1/6$

Realität

Zufallsvorgang

Werfen eines fairen Würfels

Vorhersagen

Wenn ich **nicht weiß**, ob eine Augenzahl größer als Drei gefallen ist,
dann die die Wahrscheinlichkeit, dass die Augenzahl gerade ist $\mathbf{1}/\mathbf{2}$.

Wenn ich **weiß**, dass eine Augenzahl größer als Drei gefallen ist,
dann ist die Wahrscheinlichkeit dafür, dass die Augenzahl gerade ist $\mathbf{2}/3$.

Interpretation

Gemeinsame Wahrscheinlichkeiten

Weitere Eigenschaften

Unabhängige Ereignisse

Bedingte Wahrscheinlichkeiten

Selbstkontrollfragen

Interpretation

Gemeinsame Wahrscheinlichkeiten

Weitere Eigenschaften

Unabhängige Ereignisse

Bedingte Wahrscheinlichkeiten

Selbstkontrollfragen

Wiederholung der Definition

Es seien Ω eine Ereignismenge und $\mathcal{A}$ eine σ-Algebra auf Ω. Eine Abbildung

$$\mathbb{P} : \mathcal{A} \to [0,1], A \mapsto \mathbb{P}(A) \tag{1}$$

mit den Eigenschaften

- $\mathbb{P}(A) \geq 0$ für alle $A \in \mathcal{A}$ (Nicht-Negativität),
- $\mathbb{P}(\Omega) = 1$ (Normiertheit),
- $\mathbb{P}\left(\cup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$ für paarweise disjunkte $A_i \in \mathcal{A}$ (σ-Additivität)

heißt ein *Wahrscheinlichkeitsmaß* oder einfach *Wahrscheinlichkeit*. Der Funktionswert $\mathbb{P}(A)$ heißt *Wahrscheinlichkeit des Ereignisses* $A \in \mathcal{A}$.

In der W-Theorie sind Wahrscheinlichkeiten anhand ihrer mathematischen Eigenschaften definiert.

Die Interpretation des mathematischen Wahrscheinlichkeitsbegriffs $\mathbb{P}(A)$ ist nicht eindeutig.

Es gibt mindestens zwei unterschiedliche Interpretationen:

Nach der **Frequentistischen Interpretation** ist $\mathbb{P}(A)$ die idealisierte relative Häufigkeit, mit der das Ereignis A unter den gleichen äußeren Bedingungen einzutreten pflegt. Zum Beispiel ist die frequentistische Interpretation von $\mathbb{P}(\{6\})$ im Modell des Werfen eines Würfels "Wenn man einen Würfel unendlich oft werfen würde und die relative Häufigkeit des Elementareignisses $\{6\}$ bestimmen würde, dann wäre diese relative Häufigkeit gleich $\mathbb{P}(\{6\})$."

Nach der **Bayesianischen Interpretation** ist $\mathbb{P}(A)$ der Grad der Sicherheit, den eine Beobachter:in aufgrund ihrer subjektiven Einschätzung der Lage dem Eintreten des Ereignisses A zumisst. Zum Beispiel ist die Bayesianische Interpretation von $\mathbb{P}(\{6\})$ im Modell des Werfen eines Würfels "Basierend auf meinen Erfahrung mit dem Werfen eines Würfels bin ich mir zu $\mathbb{P}(\{6\}) \cdot 100$ Prozent sicher, dass der Würfel beim nächsten Wurf eine 6 zeigt."

Interpretation

In Modellen von tatsächlich wiederholbaren Zufallsvorgängen wie dem Werfen eines Würfels ist der Unterschied zwischen Frequentistischer und Bayesianischer Interpretation eher subtil. Es gibt aber viele Zufallsvorgänge, die mit Wahrscheinlichkeiten beschrieben werden, bei denen aufgrund ihrer Einmaligkeit eine frequentistische Interpretation jedoch nicht sinnvoll ist. Zum Beispiel machen Aussagen der Form "Die Wahrscheinlichkeit, dass sich die Temperatur der Erde bis zum Jahr 2100 nur um 2° Celsius erhöht, wenn der CO_2 Ausstoß sofort auf Null gesetzt würde, ist 0.5" nur unter der Bayesianischen Interpretation Sinn.

In Hinblick auf die Definitionen und Rechenregeln für Wahrscheinlichkeiten unterscheiden sich beide Interpretationen allerdings nicht. Sowohl die frequentistische als auch die Bayesianisch geprägte probabilistische Datenanalyse haben mit der Wahrscheinlichkeitstheorie ein identisches mathematisches Bezugssystem.

Wir werden also erst an späterer Stelle wieder auf unterschiedliche Herangehensweisen in der probabilistischen Datenanalyse, die sich aus den unterschiedlichen Interpretation des Wahrscheinlichkeitsbegriffes ergeben, zurückkommen.

Generell kann man vielleicht sagen, dass die Bayesianische Interpretation mathematischer Wahrscheinlichkeiten logisch schlüssiger ist, in vielen wichtigen Anwendungsfeldern wie der Psychologie oder Biomedizin, frequentistisch geprägte Datenanalysen aber weiterhin vorherrschen.

Interpretation

Gemeinsame Wahrscheinlichkeiten

Weitere Eigenschaften

Unabhängige Ereignisse

Bedingte Wahrscheinlichkeiten

Selbstkontrollfragen

Definition (Gemeinsame Wahrscheinlichkeit)

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

$$\mathbb{P}(A \cap B) \tag{2}$$

die *gemeinsame Wahrscheinlichkeit von* A *und* B.

Bemerkungen

- Zur Abgrenzung nennt man $\mathbb{P}(A)$ auch manchmal auch *totale Wahrscheinlichkeit von* A.
- 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.

Gemeinsame Wahrscheinlichkeiten

Beispiel

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

$$
\begin{array}{ll}
A := \{2,4,6\} & \text{Es fällt eine gerade Augenzahl} \\
B := \{4,5,6\} & \text{Es fällt eine Augenzahl größer als Drei}
\end{array}
$$

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

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

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

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

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

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

Interpretation

Gemeinsame Wahrscheinlichkeiten

Weitere Eigenschaften

Unabhängige Ereignisse

Bedingte Wahrscheinlichkeiten

Selbstkontrollfragen

Weitere Eigenschaften

Theorem (Weitere Eigenschaften von Wahrscheinlichkeiten)

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

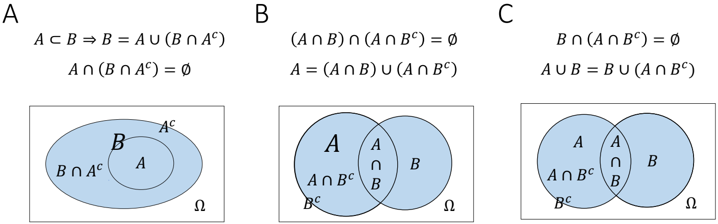
1. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.
2. $A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$.
3. $\mathbb{P}(A \cap B^c) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$
4. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
5. $A \cap B = \emptyset \Rightarrow \mathbb{P}(A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

Bemerkungen

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

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

$$\mathbb{P}(\Omega) = 1 \Leftrightarrow \mathbb{P}(A^c \cup A) = 1 \Leftrightarrow \mathbb{P}(A^c) + \mathbb{P}(A) = 1 \Leftrightarrow \mathbb{P}(A^c) = 1 - \mathbb{P}(A). \tag{6}$$

Weitere Eigenschaften

Beweis (fortgeführt)

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

$$A \subset B \Rightarrow B = A \cup (B \cap A^c) \text{ mit } A \cap (B \cap A^c) = \emptyset. \tag{7}$$

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

$$\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \cap A^c). \tag{8}$$

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

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

$$(A \cap B) \cap (A \cap B^c) = \emptyset \text{ und } A = (A \cap B) \cup (A \cap B^c). \tag{9}$$

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

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c) \Leftrightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A) - \mathbb{P}(A \cap B^c) \tag{10}$$

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

$$B \cap (A \cap B^c) = \emptyset \text{ und } A \cup B = B \cup (A \cap B^c). \tag{11}$$

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

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

Mit 3. folgt dann

$$\mathbb{P}(A \cup B) = \mathbb{P}(B) + \mathbb{P}(A) - \mathbb{P}(A \cap B) \tag{13}$$

Zu 5.: Die Aussage folgt direkt aus 4. mit $\mathbb{P}(A \cap B) = \emptyset$ und $\mathbb{P}(\emptyset) = 0$. $\square$

Interpretation

Gemeinsame Wahrscheinlichkeiten

Weitere Eigenschaften

Unabhängige Ereignisse

Bedingte Wahrscheinlichkeiten

Selbstkontrollfragen

Unabhängige Ereignisse

Definition (Unabhängige Ereignisse)

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

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B). \tag{14}$$

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

$$\mathbb{P}\left(\cap_{j\in J} A_j\right) = \prod_{j \in J} \mathbb{P}(A_j). \tag{15}$$

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:

  $\mathbb{P}(A)\mathbb{P}(B) > 0$, aber $\mathbb{P}(A \cap B) = \mathbb{P}(\emptyset) = 0$, also $\mathbb{P}(A \cap B) \neq \mathbb{P}(A)\mathbb{P}(B)$.
- Die Bedingung der beliebigen Untermengen von I sichert die paarweise unabhängig der $A_i, i \in I$.
- Der Sinn der Produkteigenschaft bei Unabhängkeit erschließt sich im Kontext *bedingter Wahrscheinlichkeiten*.

Interpretation

Gemeinsame Wahrscheinlichkeiten

Weitere Eigenschaften

Unabhängige Ereignisse

Bedingte Wahrscheinlichkeiten

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

$$\mathbb{P}(A|B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}. \tag{16}$$

Weiterhin heißt das für ein festes $B \in \mathcal{A}$ mit $\mathbb{P}(B) > 0$ definierte Wahrscheinlichkeitsmaß

$$\mathbb{P}(\ |B) : \mathcal{A} \to [0, 1], A \mapsto \mathbb{P}(A|B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \tag{17}$$

die *bedingte Wahrscheinlichkeit gegeben Ereignis B*.

Bedingte Wahrscheinlichkeiten

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.
- Im Gegensatz zu $\mathbb{P}(\cdot \cap B)$ definiert $\mathbb{P}(\cdot|B)$ ein Wahrscheinlichkeitsmaß für alle $A \in \mathcal{A}$.
- Es gelten also
  - $\mathbb{P}(A|B) \geq 0$ für alle $A \in \mathcal{A}$,
  - $\mathbb{P}(\Omega|B) = 1$ und
  - $\mathbb{P}\left(\cup_{i=1}^{\infty} A_i|B\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i|B)$ für paarweise disjunkte $A_1, A_2, ... \in \mathcal{A}$.
- $\Rightarrow$ Die Rechengeln der Wahrscheinlichkeitstheorie gelten für die Ereignisse links des Strichs.
- Üblicherweise gilt $\mathbb{P}(A|B) \neq \mathbb{P}(B|A)$, z.B.

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

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

Bedingte Wahrscheinlichkeiten

Beispiel

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

$$\mathbb{P}(\{2, 4, 6\}) = \mathbb{P}(\{2\}) + \mathbb{P}(\{4\}) + \mathbb{P}(\{6\}) = 1/6 + 1/6 + 1/6 = 3/6 \tag{18}$$

$$\mathbb{P}(\{4, 5, 6\}) = \mathbb{P}(\{4\}) + \mathbb{P}(\{5\}) + \mathbb{P}(\{6\}) = 1/6 + 1/6 + 1/6 = 3/6. \tag{19}$$

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

$$\mathbb{P}(A \cap B) = \mathbb{P}(\{2, 4, 6\} \cap \{4, 5, 6\}) = \mathbb{P}(\{4, 6\}) = \mathbb{P}(\{4\}) + \mathbb{P}(\{6\}) = 1/6 + 1/6 = 2/6 \tag{20}$$

hat. Nach Definition der bedingten Wahrscheinlichkeit ergibt sich dann direkt

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\{4, 6\})}{\mathbb{P}(\{4, 5, 6\})} = \frac{2/6}{3/6} = 2/6 \cdot 6/3 = 2/3. \tag{21}$$

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.

Theorem (Bedingte Wahrscheinlichkeit unter Unabhängigkeit)

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

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A)\mathbb{P}(B)}{\mathbb{P}(B)} = \mathbb{P}(A). \tag{22}$$

Bemerkungen

- Bei Unabhängigkeit von A und B ist es irrelevant, ob auch B eintritt, $\mathbb{P}(A)$ bleibt gleich.
- Stochastische Unabhängigkeit wird also als $\mathbb{P}(A)\mathbb{P}(B)$ modelliert, damit $\mathbb{P}(A|B) = \mathbb{P}(A)$ gilt.

$\Rightarrow$ Die stochastische Unabhängigkeit zweier Ereignisse bedeutet, dass das Wissen um das Eintreten eines der beiden Ereignisse die Wahrscheinlichkeit für das Eintreten des anderen Ereignisses nicht ändert. Andersherum bedeutet stochastische Abhängigkeit, dass das Wissen um das Eintreten eines der beiden Ereignisse die Wahrscheinlichkeit für das Eintreten des anderen Ereignisses verändert, also erhöht oder verringert.

Theorem (Gemeinsame und bedingte Wahrscheinlichkeiten)

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

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(A). \tag{23}$$

Beweis

Mit der Definition der jeweiligen bedingten Wahrscheinlichkeit folgen direkt

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \Leftrightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B) \tag{24}$$

und

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \Leftrightarrow \mathbb{P}(A \cap B) = \mathbb{P}(B|A)\mathbb{P}(A). \tag{25}$$

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.

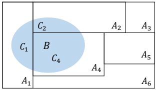
Theorem (Gesetz der totalen Wahrscheinlichkeit)

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

$$\mathbb{P}(B) = \sum_{i=1}^{k} \mathbb{P}(B|A_i)\mathbb{P}(A_i). \tag{26}$$

Beweis

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



Also gilt $\mathbb{P}(B) = \sum_{i=1}^{k} \mathbb{P}(C_i) = \sum_{i=1}^{k} \mathbb{P}(B \cap A_i) = \sum_{i=1}^{k} \mathbb{P}(B|A_i)\mathbb{P}(A_i)$.

$\square$

Bemerkung

- Totale Wahrscheinlichkeiten können aus bedingten und totalen Wahrscheinlichkeiten berechnet werden.
- $\mathbb{P}(B)$ entspricht der gewichteten Summe der bedingten Wahrscheinlichkeiten $\mathbb{P}(B|A_i)$ mit Gewichten $\mathbb{P}(A_i)$.

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

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\sum_{i=1}^k \mathbb{P}(B|A_i)\mathbb{P}(A_i)}. \tag{27}$$

Beweis

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

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(A_i \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\sum_{i=1}^k \mathbb{P}(B|A_i)\mathbb{P}(A_i)}. \tag{28}$$

$\square$

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.

Interpretation

Gemeinsame Wahrscheinlichkeiten

Weitere Eigenschaften

Unabhängige Ereignisse

Bedingte Wahrscheinlichkeiten

Selbstkontrollfragen

Selbstkontrollfragen

1. Erläutern Sie die Frequentistische Interpretation der Wahrscheinlichkeit eines Ereignisses.
2. Erläutern Sie die Bayesianische Interpretation der Wahrscheinlichkeit eines Ereignisses.
3. Geben Sie die Definition der gemeinsamen Wahrscheinlichkeit zweier Ereignisse wieder.
4. Erläutern Sie die intuitive Bedeutung der gemeinsamen Wahrscheinlichkeit zweier Ereignisse.
5. Geben Sie das Theorem zu weiteren Eigenschaften von Wahrscheinlichkeiten wieder.
6. Geben Sie die Definition der Unabhängigkeit zweier Ereignisse wieder.
7. Geben Sie die Definition der bedingten Wahrscheinlichkeit eines Ereignisses und der bedingten Wahrscheinlichkeit wieder.
8. Geben Sie das Theorem zur bedingten Wahrscheinlichkeit unter Unabhängigkeit wieder.
9. Erläutern Sie den Begriff der stochastischen Unabhängigkeit vor dem Hintergrund des Theorems zur bedingten Wahrscheinlichkeit unter Unabhängigkeit.
10. Geben Sie das Theorem zu gemeinsamen und bedingten Wahrscheinlichkeiten wieder.
11. Geben Sie das Gesetz von der totalen Wahrscheinlichkeit wieder.
12. Geben Sie das Theorem von Bayes wieder.
13. Erläutern Sie das Theorem von Bayes im Rahmen der Bayesianischen Inferenz.
14. Beweisen Sie das Theorem von Bayes.