Wahrscheinlichkeitstheorie und Statistik

2. Woche

Dávid Tóth (BME SZIT)

9., 12. September 2024

Erinnerung

Laplacescher Wahrscheinlichkeitsraum (klasszikus valószínűségi mező)

Ω endlich, Ereignisse: $\mathcal{P}(\Omega)$ =Potenzmenge von Ω,
$A \subset \Omega \Rightarrow \mathbb{P}(A) = |A| \,/\, |\Omega|$

Eigenschaften der Wahrscheinlichkeit (im Laplaschen Raum):

1. Für jedes Ereignis $A \subset \Omega$ gilt $\mathbb{P}(A) \in [0; 1]$.
2. $\mathbb{P}(\Omega) = \dfrac{|\Omega|}{|\Omega|} = 1$, $\mathbb{P}(\emptyset) = \dfrac{|\emptyset|}{|\Omega|} = \dfrac{0}{|\Omega|} = 0$.
3. Falls die Ereignisse $A, B \subset \Omega$ disjunkt sind, dann gilt

$$\mathbb{P}(A \cup B) = \frac{|A \cup B|}{|\Omega|} = \frac{|A| + |B|}{|\Omega|} = \mathbb{P}(A) + \mathbb{P}(B).$$

4. Für jedes Ereignis $A \subset \Omega$ gilt $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$.

Wahrscheinlichkeitsräume

Der Begriff des Laplaceschen Raums ist nicht genügend zu unseren Zielen.

Gründe:

- Im Laplaschen Raum muss die Menge Ω endlich sein, aber es gibt oft unenedlich viele mögliche Eregbnisse.
- Wenn man (z.B.) das Ergebnis eines Wurfes von 2 Würfeln beschreiben möchte, ist es möglich, nur die zwei Augenzahlen (ohne Reihenfolge) anzugeben, aber die verschidene Ereignisse sollten verschiedene Wahrscheinlichkeiten haben.
- Das Verhalten eines gezinkten Würfels kann nicht durch einen Lapalceschen Raum beschrieben werden.

Wahrscheinlichkeitsräume

Wir geben eine abstrakte Definition des Wahrscheinlichkeitsraums an:

- Die Wahrscheinlichkeit wird nicht durch eine definierende Formel, sondern durch ihre *Eigenschaften* angegeben.
- Im Laplaschen Raum sind die Wahrscheinlichkeiten Zahlen, die den Ereignissen zugeordnet werden. Also ist die Wahrscheinlichkeit **eine reelle Funktion**, deren Definitionsbereich die Menge der Ereignisse ist.
- Die Werte dieser Funktion sind im Intervall $[0; 1]$. Wir erwarten im Allgemeinen, dass der Wert 1 die Bestimmheit und der Wert 0 die Unmöglichkeit representieren. ($\Longrightarrow \mathbb{P}(\Omega) = 1$, $\mathbb{P}(\emptyset) = 0$)

Wahrscheinlichkeitsräume

Wir geben eine abstrakte Definition des Wahrscheinlichkeitsraums an:

- Falls zwei Ereignis simultan nicht eintreten kann, dann ist es auch eine naheliegende Erwartung, dass die Wahrscheinlichkeit dafür, dass eines der zwei Eregnisse eintritt, gleich die Summe der einzelnen Wahrscheinlichken ist.
- Mathematisch formuliert:

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

- Im Laplaschen Raum sind diese Eigenschaften *Folgerungen* der definierenden Formel.
- Im Allgemeinfall werden wir diese Eigenschaften *vorschreiben*.
- Um die Definition einfach zu halten, schreiben wir so wenig Eigenschaften vor, wie möglich ist, die andere werden auch hier Folgerungen.

Wahrscheinlichkeitsmass

Wahrscheinlichkeitsmass (valószínűségi mérték):

Ω: eine fixe Ergebnismenge, $\mathcal{F} \subset \mathcal{P}(\Omega)$ eine fixe Menge von Ereignissen (wenn Ω abzählbar ist, dann wird (in diesem Kurs) immer $\mathcal{F} = \mathcal{P}(\Omega)$)

Definition. Eine Funktion $\mathbb{P} : \mathcal{F} \to [0; 1]$ heisst *Wahrscheinlichkeitsmass* falls

- $\mathbb{P}(\Omega) = 1$,
- (σ-Additivität) $A_1, A_2 \cdots \in \mathcal{F}$, $A_i \cap A_j = \emptyset \, (\forall \, i \neq j) \Rightarrow$

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

Wahrscheinlichkeitsraum

Wahrscheinlichkeitsraum (valószínűségi mező):

Ein Wahrscheinlichkeitsraum ist ein Tripel $(\Omega, \mathcal{F}, \mathbb{P})$, wo

- Ω eine beliebige nichtleere Menge ist,
- $\mathcal{F} \subset \mathcal{P}(\Omega)$ eine (fixe) Menge von Ereignissen ist,
- $\mathbb{P} : \mathcal{F} \to [0; 1]$ ein Wahrscheinlichkeitsmass ist.

Eingenschaften des Wahrscheinlichkeitsmasses

Eingenschaften des Wahrscheinlichkeitsmasses:

$$\boxed{\mathbb{P}(\emptyset) = 0}$$

Beweis. Seien $A_1 = \emptyset, A_2 = \emptyset, \ldots, A_i = \emptyset, \ldots$, dann gilt $A_i \cap A_j = \emptyset$, also

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

Aber $\mathbb{P}(\emptyset) \geq 0$, also muss $\mathbb{P}(\emptyset) = 0$ gelten.

Eingenschaften des Wahrscheinlichkeitsmasses

Die Wahrscheinlichkeit der Vereinigungsmenge von zwei disjunkten (sich gegenseitig ausschließenden) Ereignissen ist gleich die Summe der Wahrscheinlichkeiten dieser Ereignisse.

D.h., falls $A, B \in \mathcal{F}$ Ereignisse sind, dann

$$\boxed{A \cap B = \emptyset \Rightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).}$$

(Das Maß ist *additiv*.)

Beweis. Seien $A_1 = A, A_2 = B, A_3 = \emptyset, \ldots, A_i = \emptyset, \ldots$, dann gilt $A_i \cap A_j = \emptyset$ falls $i \neq j$, also

$$\mathbb{P}(A \cup B) = \mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i) = \mathbb{P}(A) + \mathbb{P}(B) + 0.$$

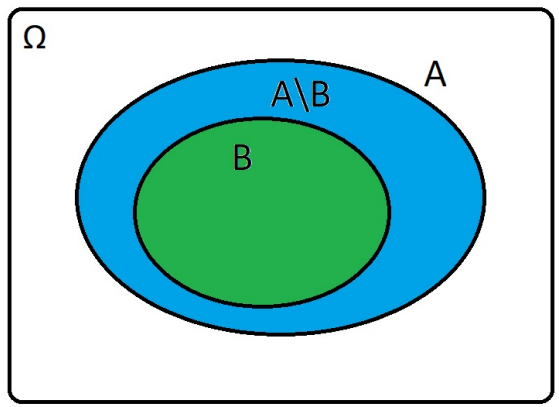
Eingenschaften des Wahrscheinlichkeitsmasses

$A \cap \overline{A} = \emptyset$ gilt, also

$$\boxed{\mathbb{P}(A) + \mathbb{P}(\overline{A}) = \mathbb{P}(A \cup \overline{A}) = \mathbb{P}(\Omega) = 1.}$$

Wenn $B \subset A$, dann $A = B \cup (A \setminus B)$, und $B \cap (A \setminus B) = \emptyset$, also

$$\mathbb{P}(A) = \mathbb{P}(B) + \mathbb{P}(A \setminus B).$$



Aber die Werte von $\mathbb{P}$ sind nichnegativ, und deshalb gilt

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

Siebformel

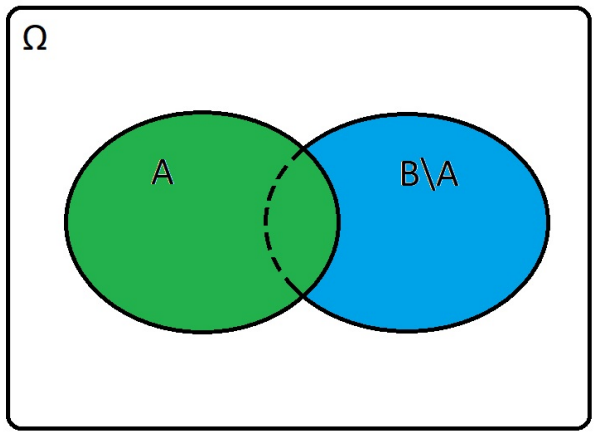
Wie können wir $\mathbb{P}(A \cup B)$ ausdrücken, wenn $A \cap B \neq \emptyset$?

Siebformel (szita-formula, Poincaré-formula) für 2 Ereignisse:

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

Beweis. Das Ereignis $A \cup B$ kann als die Vereinigung zwei disjunkter Ereignisse ausgedrückt werden:

$$A \cup B = A \cup (B \setminus A)$$



Aus der Additivität des Masses folgt $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A)$.

Siebformel

Siebformel (szita-formula, Poincaré-formula) für 2 Ereignisse:

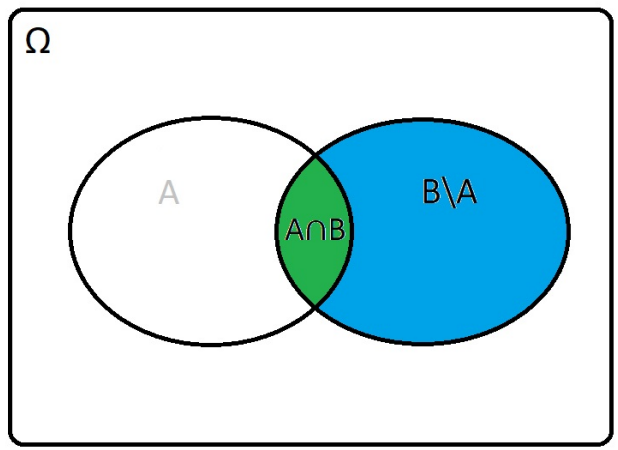
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

Beweis. Zur rechten Seite addieren wir $\mathbb{P}(A \cap B)$ mit positivem und auch negativem Zeichen, so ändert sich der Wert nicht:

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

Endlich ist $B = (B \setminus A) \cup (A \cap B)$ eine disjunkte Vereinigung, also

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



Siebformel

Siebformel für 3 Ereignisse:

$$\begin{aligned}\mathbb{P}(A \cup B \cup C) =& \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) \\ & - \mathbb{P}(A \cap B) - \mathbb{P}(B \cap C) - \mathbb{P}(C \cap A) \\ & + \mathbb{P}(A \cap B \cap C).\end{aligned}$$

Siebformel für n Ereignisse:

$$\mathbb{P}\left(\bigcup_{k=1}^{n} A_n\right) = \sum_{\emptyset \neq I \subset \{1,\ldots,n\}} (-1)^{|I|+1} \cdot \mathbb{P}\left(\bigcap_{i \in I} A_i\right).$$

Boolesche Ungleichung

Boolesche Ungleichung (Boole-egyenlőtlenség): Seien $A_1, \ldots, A_n \in \mathcal{F}$ Ereignisse. Dann gilt

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

Das folgt leicht aus der Siebformel durch vollständige Induktion:

- für $n = 1$ ist die Ungleichung trivial,
- falls die Ungleichung für $n - 1$ gilt, dann

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) &= \mathbb{P}\left(\left(\bigcup_{i=1}^{n-1} A_i\right) \cup A_n\right) \\
&= \mathbb{P}\left(\bigcup_{i=1}^{n-1} A_i\right) + \mathbb{P}(A_n) - \mathbb{P}\left(\left(\bigcup_{i=1}^{n-1} A_i\right) \cap A_n\right) \\
&\leq \mathbb{P}\left(\bigcup_{i=1}^{n-1} A_i\right) + \mathbb{P}(A_n) \leq \sum_{i=1}^{n} \mathbb{P}(A_i).
\end{aligned}$$

Boolesche Ungleichung

Aufgabe. Zeigen wir die Boolsche Ungleichung für abzählbar unendlich viele Ereignisse: seien $A_1, \ldots, A_n, \ldots \in \mathcal{F}$ Ereignisse, dann gilt

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

Konstruktion von Wahrscheinlichkeitsmassen

Bisher sahen wir noch nicht, dass der Laplaschen Raum wirklich ein Wahrscheinlichkeitsraum ist!

Erinnerung: Laplacescher Wahrscheinlichkeitsraum

Ω: eine endliche Menge, $\mathcal{F} = \mathcal{P}(\Omega)$,

$$A \in \mathcal{F} \Rightarrow \mathbb{P}(A) = \frac{|A|}{|\Omega|}$$

Behauptung: $\mathbb{P}$ ist ein Washrscheinlichkeitsmass (und dann $(\Omega, \mathcal{F}, \mathbb{P})$ ist ein Wahrscheinlichkeistraum). Nämlich erfüllt $\mathbb{P}$ die definierende Eigenschaften des Wahrscheinlichkeitsmasses:

- $\mathbb{P}(\Omega) = 1$,
- $A_1, A_2 \cdots \in \mathcal{F}$, $A_i \cap A_j = \emptyset \, (\forall \, i \neq j) \Rightarrow$

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

(Beweis: Aufgabe)

Konstruktion von Wahrscheinlichkeitsmassen

Im Laplaschen Raum hat jedes Ergebnis gleiche Wahrscheinlichkeit:

$$\omega \in \Omega \Rightarrow \mathbb{P}(\{\omega\}) = \frac{|\{\omega\}|}{|\Omega|} = \frac{1}{|\Omega|}.$$

Anders gesagt: jedes Ergebnis hat ein Gewicht $1/|\Omega|$.

Diese Bemerkung führt zur Verallgemeinerung dieses Modells:

- sei Ω eine endliche Menge,
- ordnen wir jedem Element $\omega \in \Omega$ ein nichtnegatives Gewicht p_ω zu,
- die $\sum_{\omega \in \Omega} p_\omega = 1$ erfüllen.

Konstruktion von Wahrscheinlichkeitsmassen

Behauptung. In der obigen Situation gibt die Funktion

$$\mathbb{P}(A) = \sum_{\omega \in A} p_\omega \qquad (A \subset \Omega)$$

immer ein Wahrscheinlichkeitsmass. (Beweis: Aufgabe)

Bemerkung. Falls man $\mathbb{P}(\{\omega\}) = p_\omega$ erwartet, dann bestimmen die Gewichte p_ω das Wahrscheinlichkeitsmass $\mathbb{P}$ eindeutig:

$$A = \bigcup_{\omega \in A} \{\omega\}$$

ist eine disjunkte Vereinigung, also muss wegen der Additivität des Masses

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{\omega \in A} \{\omega\}\right) = \sum_{\omega \in A} \mathbb{P}\left(\{\omega\}\right) = \sum_{\omega \in A} p_\omega$$

gelten.

Wahrscheinlichkeitsmass

Beispiel. Zwei reguläre Münzen werden geworfen. Die Ergebnisse können als geordnete Paare betrachtet werden, aber wir haben auch die folgende Möglichkeit:

$$\Omega = \{2K, 2Z, KZ\},$$

wo $2K$, $2Z$ bzw. KZ bedeutet, dass wir 2 Köpfe, 2 Zahlen bzw. einen Kopf und eine Zahl bekommen. Aber in diesem Fall wird das Modell nur dann realistisch, wenn wir

$$\mathbb{P}(2K) = \mathbb{P}(2Z) = \frac{1}{4}, \qquad \text{und} \qquad \mathbb{P}(KZ) = \frac{1}{2}$$

erfordern.

Aus der obigen Behauptung und Bemerkung folgt, dass die letzten drei Gleichungen ein Wahrscheinlichkeitsmass eindeutig bestimmen.

Wahrscheinlichkeitsmass

Man kann diese Technik anwenden, wenn Ω endlich oder sogar abzählbar unendlich ist (im zweiten Fall betrachtet mann auch unendliche Summen). In diesem Fall sagen wir, dass der Wahrscheinlichkeitsraum *diskret* ist.

Wir werden später auch solche Beispiele sehen, wo diese Eigenschaft nicht erfüllt ist. Dann kann die Konstruktion eines Masses viel komplizierter sein. Unsere Definition des Masses funktioniert auch im Allgemeinfall, aber die technischen Details der Konstuktionen werden später hier nicht angegeben.

Das Monty-Hall-Paradoxon

Betrachten wir das folgende Spiel:

- Es gibt drei geschlossene Türen. Hinter einer Tür gibt es einen Gewinn, hiter den anderen zwei Türen gibt es nichts.
- Wir sollen eine Tür wählen. Falls der Gewinn hinter dieser Tür ist, wir gewinnen, sonst nicht.
- Aber bevor die gewählte Tür geöffnet wird, eine der zwei anderen Türen wird zuerst geöffnet, eine solche, die keinen Gewinn verbergt.
- Danach können wir entscheiden, ob wir die Tür wechseln möchten, also können wir noch einmal aus der zwei geschlossenen Türen wählen.
- Was ist die gute Strategie? Sollten wir wächseln oder nicht? Ist es vielleicht egal?

Das Monty-Hall-Paradoxon

- Man könnte folgenderweise argumentieren: es bleibt zwei Türen und wir wissen nichts über den Platz des Gewinns, es ist egal, welche Tür man wählt, die Wahrscheinlichkeit ist $1/2$.
- Aber diese Argumetation ist nicht korrekt. Eine Tür wird vor unserer Wahl geöffnet, und diese Tür hängt von unserer ersten Wahl und das gibt uns extra Information.
- Um diese Information zu erleuchten, verändern wir zuerst das Spiel: nehmen wir an, dass es 1000 Türen gibt und genau eine von diesen einen Gewinn deckt. Wir wählen eine, und dann werden 998 aus den anderen 999 Türen geöffnet, die keinen Gewinn verbergen.
- Danach können wir zwischen den zwei geschlossenen Türen wählen. Würden wir wechseln?

Bedingte Wahrscheinlichkeit

- Um solche Situationen analisieren zu können, müssen wir bestimmen, wie die neue Information unseres Modell verändert.
- Nehmen wir an, dass ein Wahrscheinlichkeitsraum angegeben ist, und wir auch die Infomation erfahren, dass eine Ereignis B eintritt.
- Es wäre nicht praktisch, die Ergebnismenge zu verändern: vielleicht erfahren wir nächstes Mal nichts, vielleicht bekommen wir nächstes Mal andere Information, usw.
- Wir sollten nicht immer ein neues Modell konstruieren, sondern wir sollten mit einem flexiblen Modell arbeiten.

$\Rightarrow$ Wir werden deshalb nur das Wahrscheinlichkeitsmass verändern, so dass die Wahrscheinlichkeit von B 1 wird und alle Ereignisse "außerhalb B" unmöglich wird.
(Wir sind sicher, dass B eintritt!)

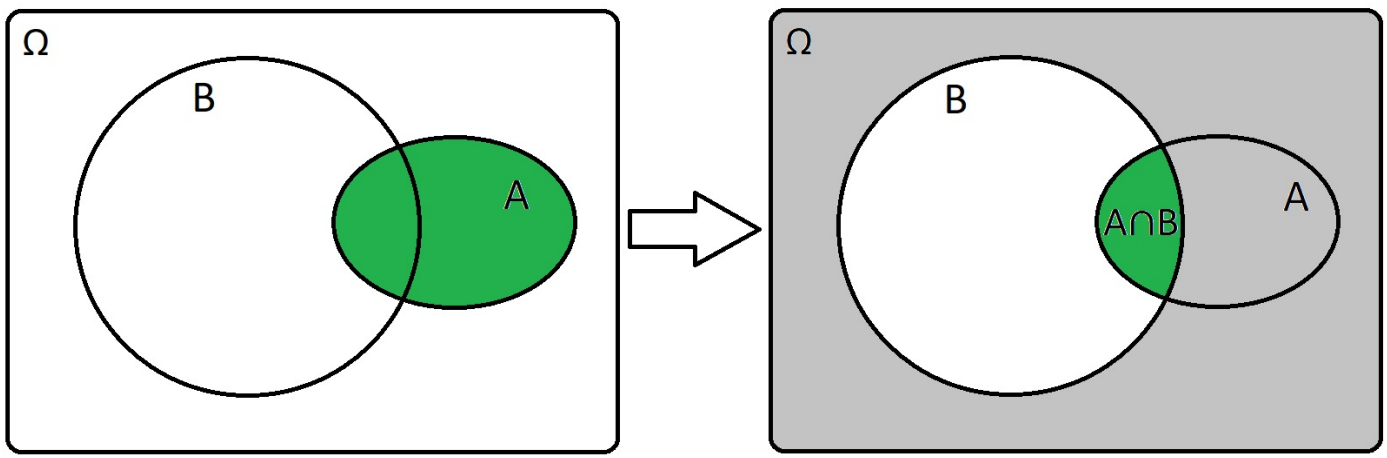
Bedingte Wahrscheinlichkeit

Definition. Seien $A, B \subset \Omega$ zwei Ereignisse mit $\mathbb{P}(B) > 0$. Dann ist *die bedingte Wahrscheinlichkeit* von A unter der Bedingung B (az A esemény B-re vett feltételes valószínűsége) durch

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

definiert.

("Die Wahrscheinlichkeit von A gegeben/angenommen B."
"A valószínűsége, feltéve B.")



Bedingte Wahrscheinlichkeit

Beispiel. Ein Würfel wird geworfen. Sei G das Ereignis, dass wir eine gerade Zahl werfen. Dann gilt

$$\mathbb{P}(\text{das Ergebnis ist größer als } 3 \mid G) = \frac{\mathbb{P}(\{4,6\})}{\mathbb{P}(\{2,4,6\})} = \frac{2}{3}.$$

Weiter gelten (nach unseren Erwartungen)

$$\mathbb{P}(G \mid G) = \frac{\mathbb{P}(G \cap G)}{\mathbb{P}(G)} = \frac{\mathbb{P}(G)}{\mathbb{P}(G)} = 1,$$

$$\mathbb{P}(\{\text{das Ergebnis ist } 1\} \mid G) = \frac{\mathbb{P}(\{1\} \cap G)}{\mathbb{P}(G)} = \frac{\mathbb{P}(\emptyset)}{\mathbb{P}(G)} = 0.$$

Bedingte Wahrscheinlichkeit

Satz. Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Betrachten wir ein Ereignis B mit $\mathbb{P}(B) > 0$, dann definiert die bedingte Wahrscheinlichkeit unter der Bedingung B (also die Funktion $\mathbb{P}_B : \mathcal{F} \to [0, 1]$, die durch

$$\mathbb{P}_B(A) := \mathbb{P}(A \mid B) \qquad (A \in \mathcal{F})$$

angegeben ist) ein Wahrscheinlichkeitsmass, d.h., $(\Omega, \mathcal{F}, \mathbb{P}_B)$ ist ein Wahrscheinlichkeitsraum.

Warum ist dieser Satz nützlich?

Antwort: die Behauptungen, die für ein Wahrscheinlichkeitsmass gelten, gelten auch für die bedingte Wahrscheinlichkeiten.

Bedingte Wahrscheinlichkeit

Beispiel. Wenn $\mathbb{P}(B) > 0$ gilt, dann haben wir

$$\mathbb{P}_B(\overline{A}) = \boxed{\mathbb{P}(\overline{A} \mid B) = 1 - \mathbb{P}(A \mid B)} = 1 - \mathbb{P}_B(A).$$

Unabhängigkeit von Ereignissen

- Die Vergleichung der Wahrscheinlichkeiten $\mathbb{P}(A)$ und $\mathbb{P}(A \mid B)$ zeigt, wie der Eintritt des Ereignisses B die Wahrscheinlichkeit von A beeinflußt.
- Ein wichtiger Spezialfall ist, wenn sich die Wahrscheinlichkeit von A unter der Bedingung nicht ändert, d.h. wenn

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

  gilt.
- Wenn man diese Gleichung mit $\mathbb{P}(B)$ multipliziert, so ergibt sich

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

  und diese Eigenschaft wird als die Definition der Unabhängigkeit benutzt:

Unabhängigkeit von Ereignissen

Definition. Zwei Ereignisse $A, B \subset \Omega$ sind unabhängig (függetlenek), wenn

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

gilt.

Unabhängigkeit von Ereignissen

Bemerkung.

- Falls $\mathbb{P}(A), \mathbb{P}(B) > 0$, dann gilt

  $$\mathbb{P}(A \mid B) = \mathbb{P}(A) \Leftrightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \Leftrightarrow \mathbb{P}(B \mid A) = \mathbb{P}(B),$$

  also bedeutet jede obige Gleichung, dass A und B unabhängig sind.
- Diese Gleichungen (oder sogar die definierende Gleichung allein) zeigen auch, dass die Unabhängigkeit ein symmetrischer Begriff ist.
- In der Definition nahmen wir nicht an, dass die Ereginisse positive Wahrscheinlichkeiten haben, die definierende Gleichung ist auch dann sinnvoll, falls $\mathbb{P}(A) = 0$ oder $\mathbb{P}(B) = 0$ gilt, und in diesen Fällen sind die Ereignisse automatisch unabhängig. (Warum?)

Unabhängigkeit von Ereignissen

Dieser Begriff sollte dem Phänomen entprechen, wenn zwei Ereignisse einander *in der Wirklichkeit* nicht beeinflussen.

Beispiel. Zwei reguläre Münzen werden geworfen.

- Das Ergebnis des ersten Wurfes beeinflusst das Ergebnis des zweiten Wurfes nicht.
- Bei dem ersten Wurf erhalten wir eine "Zahl" (ungefähr) in der Hälfte der Fälle.
- Bei dem zweiten Wurf erhalten wir eine "Zahl" auch in der Hälfte der Fälle (unabhängig vom Ergebnis des ersten Wurfes).
- Also sollten diese Ereignisse auch in unserem Modell unabhängig sein.

Unabhängigkeit von Ereignissen

Wie sieht diese Situation im mathematischen Modell aus?

$$\Omega = \{(K;K), (K;Z), (Z;K), (Z;Z)\}.$$

Seien

$$\begin{aligned} A =& \{ \text{ das erste Ergebnis ist eine "Zahl" } \}, \\ B =& \{ \text{ das zweite Ergebnis ist eine "Zahl" } \}, \end{aligned}$$

dann gilt $\mathbb{P}(A) = \mathbb{P}(B) = \frac{1}{2}$, und

$$\mathbb{P}(A \cap B) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(A) \cdot \mathbb{P}(B),$$

also sind die Ereignisse A und B auch im Modell (im Sinne der obigen Definition) unabhängig.

Unabhängigkeit von Ereignissen

Bemerkung. Def Begriff der Unabhängigkeit ist mit Wahrscheinlichkeiten definiert, deshalb können solche Fälle auftreten, wenn man so fühlt, dass die Ereignisse einander beeinflussen, und sie dennoch die vorige Gleichung erfüllen.

Beispiel. Zwei Würfel werden geworfen. Seien

$$A = \{\text{die erste Augenzahl ist } 1\}$$
$$B = \{\text{die Augenzahlen sind gleich}\}$$

Man berechnet leicht, dass

$$\mathbb{P}(A) = \frac{1}{6} = \mathbb{P}(B) \qquad \text{und} \qquad \mathbb{P}(A \cap B) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6}$$

gelten, also sind A und B unabhängig.

Unabhängigkeit von Ereignissen

Behauptung. Wenn A und B unabhängig sind, dann sind auch A und $\overline{B}$ unabhängig. Folglich sind auch die Paare $\overline{A}$ und B bzw. $\overline{A}$ und $\overline{B}$ unabhängig.

Beweis.

$$(A \cap B) \cap (A \cap \overline{B}) = \emptyset, \qquad (A \cap B) \cup (A \cap \overline{B}) = A,$$

also

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

Dann (wegen der Unabhängigkeit von A und B)

$$\begin{aligned}\mathbb{P}(A \cap \overline{B}) &= \mathbb{P}(A) - \mathbb{P}(A \cap B) = \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) \\ &= \mathbb{P}(A)(1 - \mathbb{P}(B)) = \mathbb{P}(A)\mathbb{P}(\overline{B}).\end{aligned}$$

Unabhängigkeit von mehreren Ereignissen

Wie könnte die Unabhängigkeit von drei Ereignissen definiert werden?

Beispiel. Eine Münze wird dreimal geworfen.

- Die einzelnen Würfe geben einen Kopf (unabhängig von den anderen Würfen) ungefähr in der Hälfte der Fälle, also erhalten wir 3 Köpfe (ungefähr) in dem Achtel aller Fälle.
- Sei K_i das Ereignis, dass der i-te Wurf einen Kopf gibt. Die obige Behauptung kann dann durch Wahrscheinlichkeiten folgenderweise ausgedrückt werden:

$$\mathbb{P}(K_1 \cap K_2 \cap K_3) = \mathbb{P}(K_1)\mathbb{P}(K_2)\mathbb{P}(K_3).$$

- Diese Gleichung ist eine naheliegende Verallgemeinerung der definierenden Gleichung der Unabhängigkeit.

Unabhängigkeit von mehreren Ereignissen

Falls mehrere Ereignisse unabhängig wäre, dann sollte auch ein beliebig gewähltes Paar von ihnen unabhängig sein.

Seien $A_1, \ldots, A_n \subset \Omega$ Ereignisse. Sie sind *paarweise unabhängig*, falls

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i) \cdot \mathbb{P}(A_j)$$

für alle $1 \leq i, j \leq n$, $i \neq j$ gilt. Folgt es daraus, dass

$$\mathbb{P}(A_1 \cap \cdots \cap A_n) = \mathbb{P}(A_1) \ldots \mathbb{P}(A_n)$$

gilt?

Antwort: es folgt nicht.

Unabhängigkeit von mehreren Ereignissen

Beispiel. Zwei reguläre Münzen werden geworfen. Seien

$$A_1 = \{\text{das erste Ergebnis ist "Kopf", }\}$$
$$A_2 = \{\text{das zweite Ergebnis ist "Kopf"}\},$$
$$A_3 = \{\text{die Anzahl der "Köpfe" ist gerade}\}.$$

Dann sind A_1, A_2 und A_3 paarweise unabhängig, aber

$$\mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3) = \frac{1}{2^3} \neq \frac{1}{4} = \mathbb{P}(A_1 \cap A_2 \cap A_3).$$

Unabhängigkeit von mehreren Ereignissen

Übung. Geben wir solche Ereignisse A, B und C an, für die $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$ gilt, die aber nicht paarweise unabhängig sind.

Unabhängigkeit von mehreren Ereignissen

Um den richtigen Begriff zu erhalten, erfordern wir alle möglichen Gleichungen des obigen Typs:

Definition. Seien $A_1, \ldots, A_n \subset \Omega$ Ereignisse. Sie sind *(gemeinsam) unabhängig* (együttesen függetlenek), wenn

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

für jede Menge $\emptyset \neq I \subset \{1, \ldots, n\}$ gilt.

Unabhängigkeit von mehreren Ereignissen

Beispiel. Eine Münze wird zehnmal geworfen, sei K_i das Ereignis, dass der i-te Wurf einen Kopf gibt. Wir zeigen, dass diese Ereignisse gemeinsam unabhängig sind.

- Die Anzahl der $K - Z$ Sequenzen der Länge 10, also ist die Mächtigkeit der Ergebnismenge Ω 2^{10}.
- Die Ereignisse K_i bestehen aus 2^9 Ergebnissen: aus solchen Sequenzen, deren i-tes Glied ein Kopf ist (und deren anderen Glieder Köpfe oder Zahlen sind).
- Also gilt

$$\mathbb{P}(K_i) = \frac{|K_i|}{|\Omega|} = \frac{2^9}{2^{10}} = \frac{1}{2}$$

für alle $1 \leq i \leq 10$.

Unabhängigkeit von mehreren Ereignissen

Beispiel.

- Wählen wir einige aus den Ereignissen $K_1, \ldots, K_{10}$ aus, sei I die Menge deren Indexe.
- Dann gilt

$$\prod_{i \in I} \mathbb{P}(K_i) = \frac{1}{2^{|I|}}.$$

- $\bigcap_{i \in I} K_i$ besteht aus $2^{10-|I|}$ Ergebnissen.
- Nämlich aus solchen Sequenzen, deren i-tes Glied für alle $i \in I$ ein Kopf ist, und deren anderen $10 - |I|$ Glieder beliebig gewählt werden können.
- D.h.

$$\mathbb{P}\left(\bigcap_{i \in I} K_i\right) = \frac{\left|\bigcap_{i \in I} K_i\right|}{|\Omega|} = \frac{2^{10-|I|}}{2^{10}} = \frac{1}{2^{|I|}} = \prod_{i \in I} \mathbb{P}(K_i),$$

und die Ereignisse $K_1, \ldots, K_{10}$ sind deshalb gemeinsam unabhängig.

Unabhängigkeit von mehreren Ereignissen

Satz. Seien $A_1, \ldots, A_n$ gemeinsam unabhängige Ereignisse, und nehmen wir die Komplemente von einigen von ihnen (die anderen verändern wir nicht). Die so erhaltene Menge von n Ereignissen ist auch gemeinsam unabhängig.

Anwendungen der bedingten Wahrscheinlichkeit

Anwendungen der bedingten Wahrscheinlichkeit:

1. Multiplikationsregel
2. Satz der totalen Wahrscheinlichkeit
3. Bayes-Formel

Multiplikationsregel (szorzási szabály)

Wir behandeln solche Situationen, wo die Eintritte einiger Ereignisse die Berechnung der Wahrscheinlichkeiten von anderen Ereginissen vereinfachen.

Beispiel. Wir ziehen 3 Blätter ohne Zurücklegen aus einem französischen Spielkartenpacket (mit 52 Blätter). Was ist die Wahrscheinlichkeit dafür, dass wir einen König, eine Dame und einen Bube in dieser Reihenfolge ziehen?

Seien

$$K_1 = \{\text{das erste Blatt ist ein König}\},$$
$$D_2 = \{\text{das zweite Blatt ist eine Dame}\},$$
$$B_3 = \{\text{das dritte Blatt ist ein Bube}\}.$$

Wir bestimmen $\mathbb{P}(K_1 \cap D_2 \cap B_3)$.

Multiplikationsregel

- Beim ersten Ziehen bekommt man einen König mit Wahrscheinlchkeit $\frac{4}{52} = \frac{1}{13}$.
- Die Berechnung der Wahrscheinlichkeit $\mathbb{P}(D_2)$ ist nicht so einfach, denn die beim zweiten Ziehen gebliebenen Blätter hängen vom ersten Ziehen ab.
- Aber wenn das erste Blatt schon bekannt ist, dann können wir diese Wahrscheinlichkeit leicht bestimmen, denn wir haben in diesem Fall 51 Blätter mit 4 Damen. Also gilt

$$\mathbb{P}(D_2 \mid K_1) = \frac{4}{51}.$$

- Die Definition der bedingten Wahrscheinlichkeit gibt jetzt

$$\mathbb{P}(K_1 \cap D_2) = \mathbb{P}(K_1)\mathbb{P}(D_2 \mid K_1) = \frac{4}{52} \cdot \frac{4}{51}.$$

Multiplikationsregel

- Wir können ähnlich weitergehen: wenn das erste zwei Blätter bekannt sind, dann können wir die Wahrscheinlichkeit von B_3 leicht bestimmen:

$$\mathbb{P}(B_3 \mid K_1 \cap D_2) = \frac{4}{50} = \frac{2}{25}.$$

- Also

$$\begin{aligned}\mathbb{P}((K_1 \cap D_2) \cap B_3) &= \mathbb{P}(K_1 \cap D_2)\mathbb{P}(B_3 \mid K_1 \cap D_2) \\ &= \mathbb{P}(K_1)\mathbb{P}(D_2 \mid K_1)\mathbb{P}(B_3 \mid K_1 \cap D_2) \\ &= \frac{4}{52} \cdot \frac{4}{51} \cdot \frac{4}{50} \approx 0{,}0005.\end{aligned}$$

Multiplikationsregel

Die obige Argumentation kann auch für n Ereignisse durchgeführt werden:

Satz. (Multiplikationsregel) Seien $A_1, \ldots, A_n \subset \Omega$ Ereignisse. Dann gilt

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbb{P}(A_1) \cdot \prod_{i=2}^{n} \mathbb{P}\left(A_i \,\middle|\, \bigcap_{k=1}^{i-1} A_k\right),$$

falls die obigen bedingten Wahscheinlichkeiten definiert sind.

Satz der totalen Wahrscheinlichkeit

Durch die bedingte Wahrscheinlichkeit können verschiedene Fälle getrennt behandelt werden. Nämlich kann die Ergebnismenge in dijsunkte Ergebnisse so zerlegt werden, dass die Berechnung der Wahrscheinlichkeit unter den einzelnen Bedingungen einfacher wird.

Definition. Ein System $A_1, \ldots, A_n \subset \Omega$ paarweise disjunkter Ereignisse (d.h. $A_i \cap A_j = \emptyset$ für alle $1 \leq i, j \leq n$, $i \neq j$) mit $\Omega = \bigcup_{i=1}^{n} A_i$ heißt ein *vollständiges Ereignissystem* oder auch eine *disjunkte Zerlegung von* Ω (teljes eseményrendszer).

Satz der totalen Wahrscheinlichkeit

Satz. (Satz der totalen Wahrscheinlichkeit - teljes valószínűség tétele)
Sei $A_1, \ldots, A_n \subset \Omega$ ein vollständiges Ereignissystem mit $\mathbb{P}(A_i) > 0$.
Dann gilt

$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B|A_i)\mathbb{P}(A_i)$$

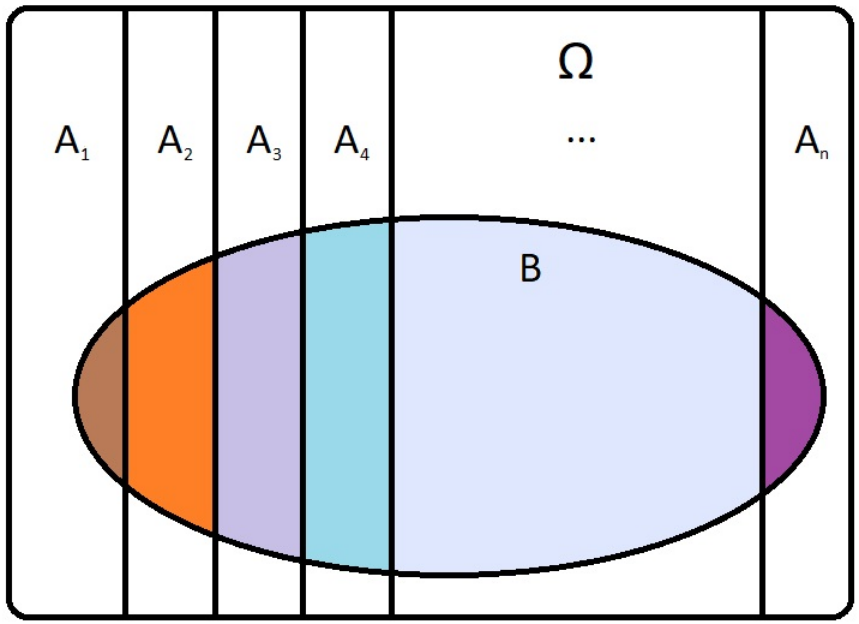
für jedes Ereignis $B \subset \Omega$.

Satz der totalen Wahrscheinlichkeit

Beweis. Das Ereignis B ist eine paarweise disjunkte Vereinigung der Ereignisse $B \cap A_1, \ldots, B \cap A_n$. Folglich gilt

$$\mathbb{P}(B) = \mathbb{P}\left(\cup_{i=1}^{n}(B \cap A_i)\right) = \sum_{i=1}^{n} \mathbb{P}(B \cap A_i) = \sum_{i=1}^{n} \mathbb{P}(B|A_i)\mathbb{P}(A_i),$$

wo wir auch die Definition der bedingten Wahrscheinlichkeit benutzten.



Satz der totalen Wahrscheinlichkeit

Beispiel. (Das Monty-Hall-Paradoxon)

- Es gibt drei geschlossene Türen, hinter einer gibt es einen Gewinn, hiter den anderen gibt es nichts.
- Wir wählen eine Tür, dann wird eine der zwei anderen Türen geöffnet, eine solche, die keinen Gewinn verbergt.
- Danach können wir noch einmal aus der zwei geschlossenen Türen wählen. Was ist die gute Strategie?

Satz der totalen Wahrscheinlichkeit

Beispiel. (Das Monty-Hall-Paradoxon)

- Nehmen wir zuerst an, dass wir unsere erste Wahl **nicht verändern**.
- Sei E das Ereignis, das unsere erste Wahl gut ist, also wählen wir zuerst die Tür, hinter der der Gewinn ist, d.h. $\mathbb{P}(E) = \frac{1}{3}$.
- Da unsere Wahl nicht durch spätere Ereignisse beeinflusst wird, ist die obere Zahl auch die Wahrscheinlichkeit dafür, dass wir am Ende gewinnen.

Satz der totalen Wahrscheinlichkeit

Beispiel. (Das Monty-Hall-Paradoxon)

- Nehmen wir jetzt an, dass wir unsere erste Wahl **verändern**.
- Sei G das Ereignis, das wir am Ende gewinnen.
- Die Ereignisse E und $\overline{E}$ formen ein vollständiges Ereignissystem, also gilt wegen des Stazes der totalen Wahrscheinlichkeit

$$\begin{aligned}\mathbb{P}(G) &= \mathbb{P}(G \mid E)\mathbb{P}(E) + \mathbb{P}(G \mid \overline{E})\mathbb{P}(\overline{E}) \\ &= 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}.\end{aligned}$$

Satz der totalen Wahrscheinlichkeit

Beispiel. (Grippetest)

- Bei einer infizierten Person liefert ein Grippeschnelltest ein positives Ergebnis mit Wahrscheinlichkeit 0,9.
- Allerdings kann der Test auch bei einer gesunden Person mit Wahrscheinlichkeit 0,2 ebenfalls positiv sein.
- Falls nun 1% aller Personen in einer Population tatsächlich infiziert sind, wie groß ist dann die Wahrscheinlichkeit, dass der Test bei einer zufällig gewählten Person ein positives Ergebnis gibt?

Satz der totalen Wahrscheinlichkeit

Betrachten wir die folgende Ereignisse:

$$B = \text{“eine zufällig gewählte Person wird positiv getestet”},$$
$$A_1 = \text{“eine zufällig gewählte Person hat Grippe”},$$
$$A_2 = \text{“eine zufällig gewählte Person hat keine Grippe”}.$$

Die Ereignisse A_1 und A_2 sind disjunkt und

$$A_1 \cup A_2 = \Omega = \text{die ganze Population}.$$

Also bilden A_1 und A_2 ein vollständiges Ereignissystem.

Satz der totalen Wahrscheinlichkeit

Wir wissen, dass $\mathbb{P}(B|A_1) = 0{,}9$ und $\mathbb{P}(B|A_2) = 0{,}2$.

Es ist zusätzlich bekannt, dass 1% aller Personen infiziert sind, also gelten noch $\mathbb{P}(A_1) = 0{,}01$ und $\mathbb{P}(A_2) = 1 - \mathbb{P}(A_1) = 0{,}99$.

Aus dem Satz der totalen Wahrscheinlichkeit folgt

$$\begin{aligned}\mathbb{P}(B) &= \mathbb{P}(B|A_1)\mathbb{P}(A_1) + \mathbb{P}(B|A_2)\mathbb{P}(A_2) \\ &= 0{,}9 \cdot 0{,}01 + 0{,}2 \cdot 0{,}99 \\ &= 0{,}207.\end{aligned}$$

Eine Person wird also mit Wahrscheinlichkeit 0,207 positiv getestet.

Nehmen wir an, dass eine Person, über die es nicht bekannt ist, ob sie gesund oder infiziert ist, positiv getestet wird. Mit welcher Wahrscheinlichkeit ist sie tatsächlich infiziert? $\mathbb{P}(A_1 \mid B) = ?$

Bayes-Formel

Satz. (Einfache Bayes-Formel - egyszerű Bayes-tétel) Seien A und B Ereignisse mit $\mathbb{P}(A), \mathbb{P}(B) > 0$, dann gilt

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

Beweis.

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

Bayes-Formel

Satz. (Bayes-Formel) Seien $A_1, \ldots, A_n \subset \Omega$ Ereignisse mit $\mathbb{P}(A_i) > 0$, die ein vollständiges Ereignissystem bilden. Sei auch $B \subset \Omega$ ein Ereignis mit $\mathbb{P}(B) > 0$, dann gilt für alle $i = 1, \ldots, n$

$$\mathbb{P}(A_i \mid B) = \frac{\mathbb{P}(B \mid A_i)\mathbb{P}(A_i)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B \mid A_i)\mathbb{P}(A_i)}{\sum_{k=1}^{n} \mathbb{P}(B|A_k)\mathbb{P}(A_k)}.$$

Beweis.

Einfache Bayes-Formel + Satz der totalen Wahrscheinlichkeit.

Bayes-Formel

Beispiel. (Grippetest) Eine Person, über die es nicht bekannt ist, ob sie gesund oder infiziert ist, wurde positiv getestet. Mit welcher Wahrscheinlichkeit ist sie tatsächlich infiziert?

$$
\begin{aligned}
B &= \text{“eine zufällig gewählte Person wird positiv getestet”},\\
A_1 &= \text{“eine zufällig gewählte Person hat Grippe”},\\
A_2 &= \text{“eine zufällig gewählte Person hat keine Grippe”}.
\end{aligned}
$$

$$
\mathbb{P}(A_1 \mid B) = \frac{\mathbb{P}(B \mid A_1)\mathbb{P}(A_1)}{\mathbb{P}(B)} = \frac{0{,}9 \cdot 0{,}01}{0{,}207} \approx 0,0435.
$$

Bayes-Formel

Warum bekommen wir dieses überaschende Ergebnis?

$$\mathbb{P}(A_1 \mid B) = \frac{\mathbb{P}(B \mid A_1)\mathbb{P}(A_1)}{\mathbb{P}(B|A_1)\mathbb{P}(A_1) + \mathbb{P}(B|A_2)\mathbb{P}(A_2)}$$

- Das Verhältnis der Personen, die gesund sind, ist zu groß, also ist $\mathbb{P}(A_2)$ viel größer als $\mathbb{P}(A_1)$.
- Ein falsch positives Ergebnis hat auch eine relativ große Wahrscheinlichkeit.
- Die Wahrscheinlichkeit, mit der dieser Test die Krankheit bei einer infizierten Person identifiziert, ist zu klein, um die obigen Werte zu kompensieren.