Mathe III

Ralf Herbrich

Eigenschaften von Wahrscheinlichkeitsräumen

Hasso Plattner Institut
Digital Engineering · Universität Potsdam

Überblick

1. Bedingte Wahrscheinlichkeit
2. Satz von Bayes
3. Stochastische Unabhängigkeit
4. Mehrstufige Modelle

Überblick

1. **Bedingte Wahrscheinlichkeit**
2. Satz von Bayes
3. Stochastische Unabhängigkeit
4. Mehrstufige Modelle

Bedingte Wahrscheinlichkeit: Einführendes Beispiel

- **Beispiel (Bedingte Wahrscheinlichkeit).**
  - **Situation**: Für die 3 Erstplatzierten eines Radrennens werden Fahrradklingeln verschenkt. Hierfür dürfen die Gewinner eine Fahrradklingel verdeckt aus einer Urne ziehen. Es gibt eine blaue, eine rote und eine goldene Klingel in der Urne. Wie wahrscheinlich ist es, dass du eine goldene Fahrradklingel ziehst?
    - **Ergebnismenge**: $\Omega = \{\text{blau}, \text{rot}, \text{gold}\}$
    - **Gleichverteilung**: $P_{\text{naive}}(\text{gold}) = \frac{1}{3}$

  - **Neue Situation**: Da du nur Zweitplatzierter warst, darfst du erst als zweites ziehen und siehst, wie der Erstplatzierte vor dir eine rote Klingel zieht. Wie wahrscheinlich ist es jetzt, dass du eine goldene Fahrradklingel ziehst?
    - **Ergebnismenge**: $\Omega = \{\text{blau}, \text{gold}\}$
    - **Gleichverteilung**: $P_{\text{naive}}(\text{gold}) = \frac{1}{2}$

- **Was ist passiert?**

Bedingte Wahrscheinlichkeit: Herleitung

- **Was ist passiert?**
  - Wir wollen die Wahrscheinlichkeit für ein Ereignis $A = \{\text{gold}\}$ bestimmen.
  - Dabei haben wir den Eintritt des Ereignisses $B = \{\text{blau}, \text{gold}\}$ beobachtet.
  - Daraufhin konnten wir die Eintrittswahrscheinlichkeit für $A = \{\text{gold}\}$ neu bewerten.
  - Diese neue Wahrscheinlichkeit nennen wir $P(A|B)$.
    - Es kann nur noch ein Ergebnis eintreten, welches in B ist.
    - Es können nur noch Ereignisse eintreten, welche Teilmenge von B sind.
    - B ist nun ein sicheres Ereignis.

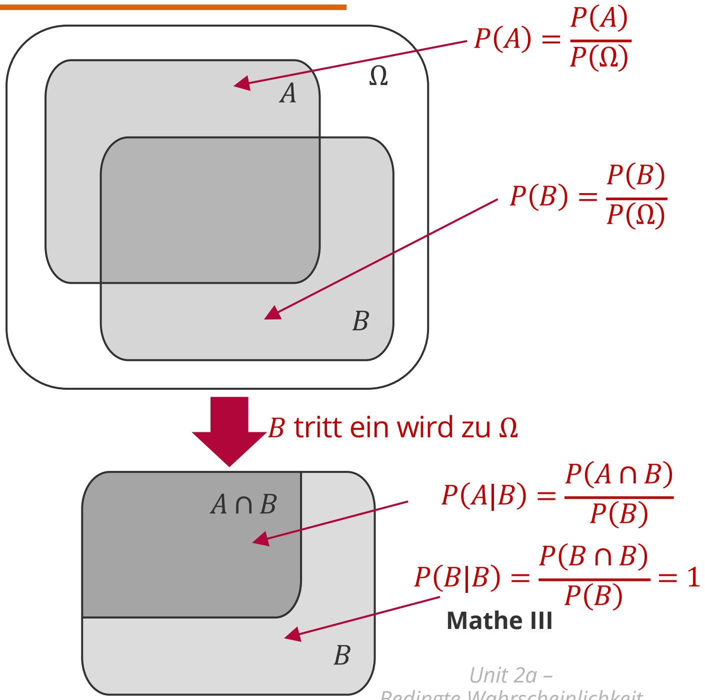
$$P(A|B) = \frac{P(A|B)}{P(B|B)} = \frac{P(A \cap B|B)}{P(B|B)} = \frac{P(A \cap B)}{P(B)}$$

Bedingte Wahrscheinlichkeit

- **Definition (Bedingte Wahrscheinlichkeit)**. Für einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ bezeichnen wir für alle $A, B \in \mathcal{F}$ mit $P(B) > 0$ die Wahrscheinlichkeit $P(A|B)$ als die bedingte Wahrscheinlichkeit von A unter der Bedingung B, wobei:

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

- **Zwei mögliche Interpretationen**:
  - Ist $P(A)$ die Eintrittswahrscheinlichkeit von A vor dem Experiment, so ist $P(A|B)$ die Eintrittswahrscheinlichkeit, nachdem B im Experiment eingetreten ist.
  - Bei häufiger Wiederholung des Zufallsexperiments ist $P(A|B)$ der Anteil der Fälle, in denen A eintritt unter der Gesamtheit der Fälle, in denen B eintritt.



Bedingte Wahrscheinlichkeit

- **Bemerkungen (Bedingte Wahrscheinlichkeit)**.
  1. Die bedingte Wahrscheinlichkeit sagt **nichts** über kausale Zusammenhänge aus.
     - Wenn durch Eintreten eines Ereignisses B das Eintreten eines Ereignisses A wahrscheinlicher geworden ist, so heißt das nicht automatisch, dass es an B lag, wenn A letztendlich eintritt.
     - **Beispiel**: Gegeben wir beobachten, dass viele Personen kurze Hosen & Röcke tragen erhöht das die bedingte Wahrscheinlichkeit, dass es draußen warm ist. Aber die hohen Außentemperaturen sind keine kausale Folge vom Tragen von kurzen Hosen & Röcken!
  2. Die Wahrscheinlichkeitsverteilung P, welche wir vor der Kenntnis jeglicher Ereigniseintritte annehmen, wird auch **a priori Wahrscheinlichkeit** genannt.
  3. Die Wahrscheinlichkeitsverteilung, welche sich nach Neubewertung aufgrund des Eintritts eines Ereignisses ergibt, wird auch **a posteriori Wahrscheinlichkeit** genannt.

Bedingte Wahrscheinlichkeit: Vater-Tochter Dinner Beispiel

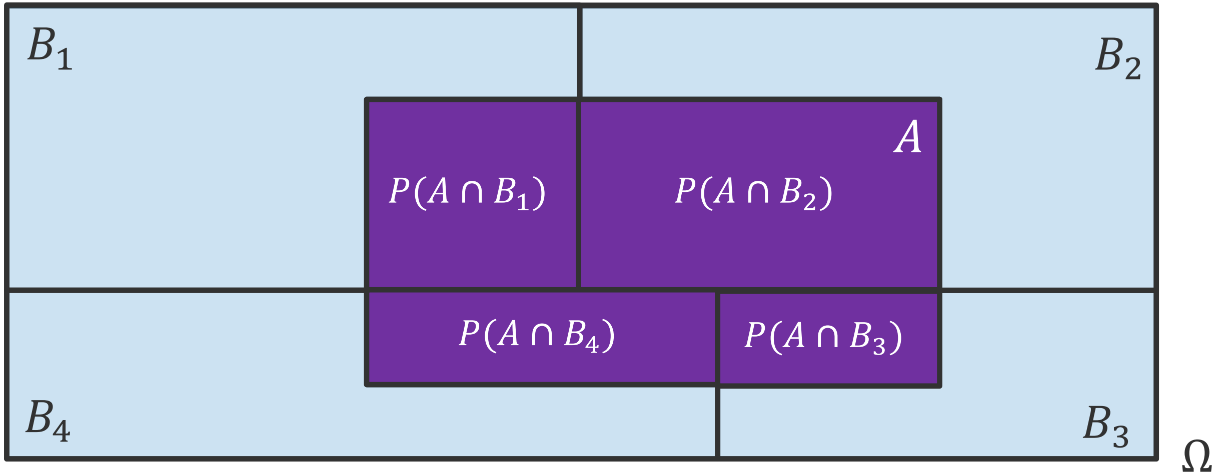
- **Situation**: Eine Firma veranstaltet ein Vater-Tochter Dinner für Angestellte, die mindestens eine Tochter haben. Diese dürfen dann ihre jüngste Tochter mit zum Dinner bringen. Wenn wir wissen, dass Robert eingeladen ist und zwei Kinder hat, was ist die bedingte Wahrscheinlichkeit, dass beides Töchter sind?
  - **Intuitive Antwort**: Wir wissen bereits, dass mindestens ein Kind eine Tochter ist (da Robert eingeladen ist). Daher ist die Wahrscheinlichkeit, dass sein zweites Kind eine Tochter ist 50%.
  - **Richtige Antwort**:
    - $\Omega = \{(\text{Junge}, \text{Junge}), (\text{Junge}, \text{Mädchen}), (\text{Mädchen}, \text{Junge}), (\text{Mädchen}, \text{Mädchen})\}$
    - $P_{\text{naive}}(\omega) = \frac{1}{4}$
    - Robert ist eingeladen: $B = \{(\text{Mädchen}, \text{Mädchen}), (\text{Junge}, \text{Mädchen}), (\text{Mädchen}, \text{Junge})\}$
    - Robert hat zwei Töchter: $A = \{(\text{Mädchen}, \text{Mädchen})\}$
    - $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$

Satz der Totalen Wahrscheinlichkeit

- **Satz (Totale Wahrscheinlichkeit).** Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum und $\Omega = \bigcup_{i \in I} B_i$ eine abzählbare Zerlegung von Ω in paarweise disjunkte Ereignisse B_i. Dann gilt für alle $A \in \mathcal{F}$:

$$P(A) = \sum_{i \in I} P(A \cap B_i) = \sum_{i \in I} P(A|B_i) \cdot P(B_i)$$

- **Geometrischer Beweis**



Satz der Totalen Wahrscheinlichkeit

- **Satz (Totale Wahrscheinlichkeit).** Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum und $\Omega = \bigcup_{i \in I} B_i$ eine abzählbare Zerlegung von Ω in paarweise disjunkte Ereignisse B_i. Dann gilt für alle $A \in \mathcal{F}$:

$$P(A) = \sum_{i \in I} P(A \cap B_i) = \sum_{i \in I} P(A|B_i) \cdot P(B_i)$$

- **Algebraischer Beweis**: Wir definieren die Mengen

$$C_i = A \cap B_i$$

Dann gilt per Definition der Mengen C_i, dass sie paarweise disjunkt sind (d.h., $C_i \cap C_j = \emptyset \Leftrightarrow i \neq j$) und dass $A = \bigcup_{i \in I} C_i$. Aus 2. folgt daher

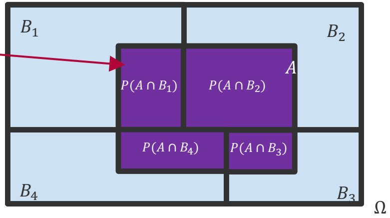
$$P(A) = P\left(\bigcup_{i \in I} C_i\right) = \sum_{i \in I} P(C_i) = \sum_{i \in I} P(A \cap B_i)$$

Der zweite Teil folgt aus der Definition der bedingten Wahrscheinlichkeit

$$P(A|B_i) = \frac{P(A \cap B_i)}{P(B_i)} \Leftrightarrow P(A \cap B_i) = P(A|B_i) \cdot P(B_i)$$

Kolmogorov Axiome

1. Für alle $A \in \mathcal{F}$, $0 \leq P(A) \leq 1$
2. Für alle $A_1, \ldots, A_n \in \mathcal{F}$ mit $A_i \cap A_j = \emptyset \Leftrightarrow i \neq j$ gilt
$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$
3. Es gilt $P(\Omega) = 1$



4-Felder-Tafel

- Zur Berechnung bedingter Wahrscheinlichkeiten kann eine **4-Felder-Tafel** hilfreich sein.
- Betrachte dafür einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ sowie zwei Ereignisse $A, B \in \mathcal{F}$. Dann ist die 4-Felder-Tafel darstellbar als Tabelle von Häufigkeiten oder Wahrscheinlichkeiten.

	A	$\overline{A}$	Summe
B			
$\overline{B}$			
Summe			

Häufigkeits 4-Felder-Tafel

	A	$\overline{A}$	Summe
B	$\lvert A \cap B\rvert$	$\lvert\overline{A} \cap B\rvert$	$\lvert B\rvert$
$\overline{B}$	$\lvert A \cap \overline{B}\rvert$	$\lvert\overline{A} \cap \overline{B}\rvert$	$\lvert\overline{B}\rvert$
Summe	$\lvert A\rvert$	$\lvert\overline{A}\rvert$	$\lvert\Omega\rvert$

Wahrscheinlichkeits 4-Felder-Tafel

	A	$\overline{A}$	Summe
B	$P(A \cap B)$	$P(\overline{A} \cap B)$	$P(B)$
$\overline{B}$	$P(A \cap \overline{B})$	$P(\overline{A} \cap \overline{B})$	$P(\overline{B})$
Summe	$P(A)$	$P(\overline{A})$	$P(\Omega)$

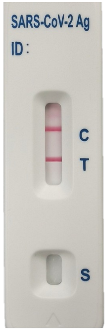
- **Definition (Randverteilung)**. Die in der 4-Felder-Tafel notierten Anteile, welche nur eine der beiden Dimensionen beschreiben (d.h., $P(A), P(\overline{A}), P(B)$ und $P(\overline{B})$), werden Randverteilungen genannt.

4-Felder-Tafel

- **Beispiel (4-Felder-Tafel).** Betrachte eine Population, von welcher 0.15 % mit COVID infiziert sind. Betrachte außerdem einen COVID Test mit Sensitivität von 98% (d.h., $P(\text{Test ist positiv}|\text{Patient hat COVID}) = 0.98$) und Spezifität von 99% (d.h., $P(\text{Test ist negativ}|\text{Patient hat kein COVID}) = 0.99$). Welche Werte enthält die entsprechende 4-Felder-Tafel für ein zufällig gewähltes Individuum?

	Hat COVID	Hat kein COVID	Summe
Test positiv	0.00147	0.009985	0.011455
Test negativ	0.00003	0.988515	0.988545
Summe	0.0015	0.9985	1

  - $\alpha = P(\text{Test positiv} \cap \text{hat COVID}) = P(\text{hat COVID}) \cdot P(\text{Test positiv}|\text{hat COVID})$
  - $\beta = P(\text{Test negativ} \cap \text{hat kein COVID}) = P(\text{hat kein COVID}) \cdot P(\text{Test negativ}|\text{hat kein COVID})$



Überblick

1. Bedingte Wahrscheinlichkeit
2. **Satz von Bayes**
3. Stochastische Unabhängigkeit
4. Mehrstufige Modelle

Satz von Bayes

- **Satz (von Bayes)**. Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum. Dann gilt für alle $A, B \in \mathcal{F}$ mit $P(A) > 0$ und $P(B) > 0$, dass

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

- **Beweis**. Folgt aus der Definition von bedingter Wahrscheinlichkeit und dem „multiplizieren-mit-1" Trick

$$P(A \cap B) \cdot \frac{P(B)}{P(B)} = P(A \cap B) \cdot \frac{P(A)}{P(A)}$$ =1 (per Annahme ist $P(A) > 0$ and $P(B) > 0$)

$$P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$ (per Definition von bedingter Wahrscheinlichkeit)

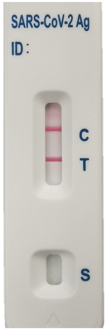
$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

- **Bemerkungen (Satz von Bayes)**. Mit dem Satz von Bayes können wir $P(A|B)$ aus $P(B|A)$ berechnen und umgekehrt.

4-Felder-Tafel und Satz von Bayes

- **Beispiel (4-Felder-Tafel).**

	Hat COVID	Hat kein COVID	Summe
Test positiv	0.00147	0.009985	0.011455
Test negativ	0.00003	0.988515	0.988545
Summe	0.0015	0.9985	1



  - Mit welcher Wahrscheinlichkeit ist ein Individuum infiziert, wenn der Test positiv ist?

$$P(\text{Hat COVID}|\text{Test positiv}) = \frac{P(\text{Hat COVID} \cap \text{Test positiv})}{P(\text{Test positiv})} = \frac{0.00147}{0.011455} \approx 12.8\ \%$$

  - Mit welcher Wahrscheinlichkeit ist ein Individuum infiziert, wenn der Test negativ ist?

$$P(\text{Hat COVID}|\text{Test negativ}) = \frac{P(\text{Hat COVID} \cap \text{Test negativ})}{P(\text{Test negativ})} = \frac{0.00003}{0.988545} \approx 0.003\ \%$$

Satz von Bayes: *False-Positive* Puzzle

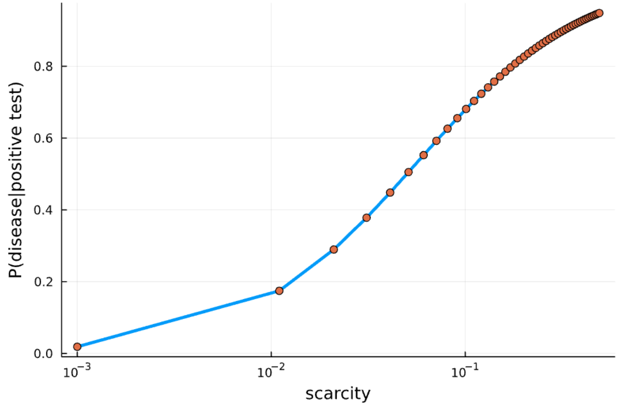
- **Situation**: Ein Test für eine seltene Krankheit hat eine Sensitivität und Spezifität von 95%. Die Krankheit haben nur 0.1% der Bevölkerung. Was ist die Wahrscheinlichkeit, die Krankheit zu haben, wenn der Testausgang positiv ist?
- **Lösung**:

$$A = \text{"Person hat die Krankheit"}$$

$$B = \text{"Test ist positiv"}$$

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

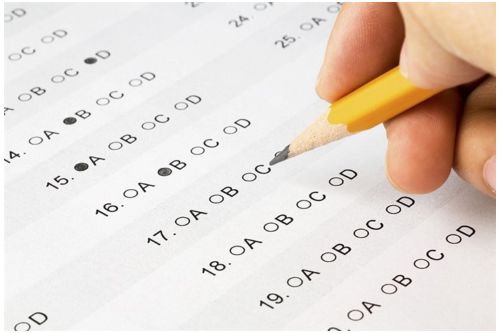
$$P(A|B) = \frac{0.95 \cdot 0.001}{0.95 \cdot 0.001 + 0.05 \cdot 0.999} \approx 0.0187$$



- **Nicht Intuitiv**: 80% der amerikanischen Krankenhausmitarbeiter schätzen die Wahrscheinlichkeit $P(A|B)$ auf 95% (*The Economist* (Februar 20, 1999))!

Satz von Bayes: Klausur Beispiel

- **Situation**: In der Klausur mit einer *multiple choice* Frage mit 4 Antwortmöglichkeiten wissen wir, dass 50% der Studenten die Antwort wissen und 50% der Studenten raten. Wenn die Studenten raten, dann erraten sie mit 25% die richtige Antwort. Wie groß ist die Wahrscheinlichkeit, dass ein Student die Antwort wusste, wenn die Antwort richtig ist?



  - **Lösung**:
    - $A =$ "Antwort richtig"
    - $B =$ "Student weiß die Antwort"

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

$$P(B|A) = \frac{1 \cdot 0.5}{1 \cdot 0.5 + 0.25 \cdot 0.5} = 0.8 = 80\%$$

Viel Spaß bis zur nächsten Vorlesung!