Stochastik Vorlesung 2.7: Satz von Bayes, Paradoxa, Hypergeometrische Verteilung

Peter Nejjar

2. Juli 2025

Einleitung und Motivation

- **Problemstellung:** Wie aktualisieren wir Wahrscheinlichkeiten bei neuen Informationen?
- **Beispiel:** Medizinische Diagnose: Von $\mathbb{P}(\text{positiv}|\text{krank})$ zu $\mathbb{P}(\text{krank}|\text{positiv})$.
- **Bedeutung:** Fundamentales Werkzeug zur Aktualisierung von Wahrscheinlichkeiten basierend auf Evidenz. Grundlage der **Bayesianischen Statistik**.

Gemeinsame und Bedingte Wahrscheinlichkeit

- **Gemeinsame Wahrscheinlichkeit:** $\mathbb{P}(A \cap B)$
  - Wahrscheinlichkeit, dass A **und** B eintreten.
- **Bedingte Wahrscheinlichkeit:**

Definition (Bedingte Wahrscheinlichkeit)

Für $\mathbb{P}(B) > 0$:

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

  - Intuition: Reduktion des Stichprobenraums auf B.
  - Daraus folgt die **Produktregel:**

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

Gesetz der totalen Wahrscheinlichkeit

Theorem (Gesetz der totalen Wahrscheinlichkeit)

Seien $A_1, \ldots, A_n$ *eine* **Partition des Stichprobenraums** Ω*. Dann gilt für ein beliebiges Ereignis* B*:*

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

- **Beweis** $B \cap A_1, B \cap A_2, \ldots, B \cap A_n$ ist eine Partition von B, daher ist $\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B \cap A_i)$. Da nun $\mathbb{P}(B \cap A_i) = \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)$ gilt, folgt die Aussage.

Der Satz von Bayes

Theorem (Satz von Bayes)

Für Ereignisse A und B mit $\mathbb{P}(A) > 0, \mathbb{P}(B) > 0$:

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

Erweiterte Form für Partition $A_1, \ldots, A_n$:

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)}{\sum_{j=1}^{n} \mathbb{P}(B|A_j) \cdot \mathbb{P}(A_j)}$$

Beweis des Satzes von Bayes

Beweis.

Aus der Definition der bedingten Wahrscheinlichkeit:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \quad (1)$$

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \quad (2)$$

Aus (2) folgt $\mathbb{P}(A \cap B) = \mathbb{P}(B|A) \cdot \mathbb{P}(A)$.
Setzen in (1) ein:

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

Der Nenner $\mathbb{P}(B)$ wird durch das Gesetz der totalen Wahrscheinlichkeit ersetzt für die erweiterte Form. □

Terminologie im Satz von Bayes

- $\mathbb{P}(A|B)$: **Posteriori-Wahrscheinlichkeit** (Posterior)
  - Wahrscheinlichkeit der Hypothese A, *nachdem* Evidenz B beobachtet wurde.
- $\mathbb{P}(B|A)$: **Likelihood**
  - Wahrscheinlichkeit, Evidenz B zu beobachten, *wenn* Hypothese A wahr ist.
- $\mathbb{P}(A)$: **Prior-Wahrscheinlichkeit** (Prior)
  - Anfängliche Überzeugung für Hypothese A, *bevor* Evidenz B beobachtet wurde.
- $\mathbb{P}(B)$: **Evidenz** / **Marginalwahrscheinlichkeit der Evidenz**
  - Normalisierungsfaktor, Gesamtwahrscheinlichkeit von B.

Medizinische Diagnose: Szenario

- Seltene Krankheit: **0,1%** der Bevölkerung ($\mathbb{P}(\text{Krank}) = 0.001$).
- Diagnostischer Test:
  - **Sensitivität von 95%:** $\mathbb{P}(\text{Positiv}|\text{Krank}) = 0.95$.
  - **Spezifität von 90%:**
    $\mathbb{P}(\text{Negativ}|\text{Nicht Krank}) = 0.90 \implies \mathbb{P}(\text{Positiv}|\text{Nicht Krank}) = 0.10$.

Frage: Wie hoch ist die Wahrscheinlichkeit, tatsächlich krank zu sein, wenn der Test positiv ausfällt? ($\mathbb{P}(\text{Krank}|\text{Positiv})$)?

Medizinische Diagnose: Berechnung

- **Priori:** $\mathbb{P}(\text{Krank}) = 0.001$, $\mathbb{P}(\text{Nicht Krank}) = 0.999$
- **Likelihoods:** $\mathbb{P}(\text{Positiv}|\text{Krank}) = 0.95$, $\mathbb{P}(\text{Positiv}|\text{Nicht Krank}) = 0.10$
- **Evidenz ($\mathbb{P}$(Positiv)) - Totale Wahrscheinlichkeit:**

$$\begin{aligned}
\mathbb{P}(\text{Positiv}) &= \mathbb{P}(\text{Positiv}|\text{Krank})\mathbb{P}(\text{Krank}) \\
&\quad + \mathbb{P}(\text{Positiv}|\text{Nicht Krank})\mathbb{P}(\text{Nicht Krank}) \\
&= (0.95 \cdot 0.001) + (0.10 \cdot 0.999) \\
&= 0.00095 + 0.0999 = 0.10085
\end{aligned}$$

- **Bayes-Anwendung ($\mathbb{P}$(Krank|Positiv)):**

$$\begin{aligned}
\mathbb{P}(\text{Krank}|\text{Positiv}) &= \frac{\mathbb{P}(\text{Positiv}|\text{Krank}) \cdot \mathbb{P}(\text{Krank})}{\mathbb{P}(\text{Positiv})} \\
&= \frac{0.00095}{0.10085} \approx 0.00942
\end{aligned}$$

Medizinische Diagnose: Interpretation

- Ergebnis: $\mathbb{P}(\text{Krank}|\text{Positiv}) \approx \mathbf{0,94\%}$
- **Fazit:** Obwohl der Test 95% sensitiv ist, ist die Wahrscheinlichkeit bei einem positiven Test krank zu sein, sehr gering.
- **Grund:** Die extrem niedrige Prävalenz (Prior-Wahrscheinlichkeit) der Krankheit.

Das Ziegenproblem: Szenario

- 3 Türen: 1 Auto, 2 Ziegen.
- Sie wählen eine Tür (z.B. Tür 1).
- Moderator (weiß, wo Auto ist) öffnet *eine andere* Tür mit einer Ziege (z.B. Tür 3).
- Moderator fragt: Bei ursprünglicher Wahl bleiben oder wechseln?

Frage: Ist es von Vorteil zu wechseln, zu bleiben oder ist es egal?

Das Ziegenproblem: Analyse

- **Intuition:** Oft 50/50 angenommen, da zwei Türen übrig. FALSCH!
- **Ereignisse:**
  - C_i: Auto hinter Tür i. ($\mathbb{P}(C_1) = \mathbb{P}(C_2) = \mathbb{P}(C_3) = 1/3$)
  - W_1: Sie wählen Tür 1.
  - M_3: Moderator öffnet Tür 3 (Ziege).
- Wir suchen $\mathbb{P}(C_1|M_3)$ (Bleiben) und $\mathbb{P}(C_2|M_3)$ (Wechseln).

Das Ziegenproblem: Berechnung

- **Likelihoods (wenn Sie Tür 1 gewählt haben):**
  - $\mathbb{P}(M_3|C_1) = 1/2$ (Mod. wählt zufällig zwischen Tür 2 & 3)
  - $\mathbb{P}(M_3|C_2) = 1$ (Mod. MUSS Tür 3 öffnen)
  - $\mathbb{P}(M_3|C_3) = 0$ (Mod. darf nicht Auto öffnen)
- **Evidenz $\mathbb{P}(M_3)$ (Gesetz der totalen Wahrscheinlichkeit):**

$$\begin{aligned}\mathbb{P}(M_3) &= \mathbb{P}(M_3|C_1)\mathbb{P}(C_1) + \mathbb{P}(M_3|C_2)\mathbb{P}(C_2) + \mathbb{P}(M_3|C_3)\mathbb{P}(C_3) \\ &= (1/2 \cdot 1/3) + (1 \cdot 1/3) + (0 \cdot 1/3) = 1/6 + 1/3 = 1/2\end{aligned}$$

- **Posteriori Wahrscheinlichkeit:**
  - **Bleiben ($\mathbb{P}(C_1|M_3)$):** $\frac{\mathbb{P}(M_3|C_1)\cdot\mathbb{P}(C_1)}{\mathbb{P}(M_3)} = \frac{1/2\cdot 1/3}{1/2} = \mathbf{1/3}$
  - **Wechseln ($\mathbb{P}(C_2|M_3)$):** $\frac{\mathbb{P}(M_3|C_2)\cdot\mathbb{P}(C_2)}{\mathbb{P}(M_3)} = \frac{1\cdot 1/3}{1/2} = \mathbf{2/3}$

Das Ziegenproblem: Schlussfolgerung

- Es ist **doppelt so wahrscheinlich**, das Auto zu gewinnen, wenn Sie Ihre Wahl ändern!
- Die Information, die der Moderator durch das Öffnen einer Ziegentür gibt, konzentriert die ursprüngliche $2/3$-Wahrscheinlichkeit (dass das Auto nicht hinter Ihrer Tür ist) auf die verbleibende geschlossene Tür.

Das Geburtstagsparadoxon: Problemstellung

- **Frage:** Wie viele Personen müssen in einem Raum sein, damit die Wahrscheinlichkeit, dass *mindestens zwei* von ihnen am selben Tag Geburtstag haben, größer als 50% ist?
- **Intuition:** Oft wird eine hohe Zahl (z.B. 100) erwartet, da es 365 Tage gibt.
- **Tatsächliche Antwort** ist aber viel kleiner

Das Geburtstagsparadoxon: Lösungsweg

- Berechnung über das **Komplementärereignis**:
  - Wahrscheinlichkeit, dass *keine zwei* Personen am selben Tag Geburtstag haben.
  - Dann: $\mathbb{P}(\text{Match}) = 1 - \mathbb{P}(\text{kein Match})$.
- Annahme: 365 Tage im Jahr (keine Schaltjahre).
- Für n Personen:
  - 1. Person: 365 Möglichkeiten
  - 2. Person: 364 Möglichkeiten (nicht am gleichen Tag wie 1.)
  - ...
  - n. Person: $(365 - n + 1)$ Möglichkeiten

Das Geburtstagsparadoxon: Formeln & Ergebnis

- Anzahl der Möglichkeiten für unterschiedliche Geburtstage:

$$N_{\text{unterschiedlich}} = 365 \cdot 364 \cdot \ldots \cdot (365 - n + 1) = \frac{365!}{(365 - n)!}$$

- Gesamtzahl der möglichen Geburtstagskombinationen:

$$N_{\text{gesamt}} = 365^n$$

- Wahrscheinlichkeit für *kein Match*:

$$\mathbb{P}(\text{kein Match}) = \frac{N_{\text{unterschiedlich}}}{N_{\text{gesamt}}} = \prod_{i=0}^{n-1} \left(1 - \frac{i}{365}\right)$$

- Wahrscheinlichkeit für *mindestens ein Match*:

$$\mathbb{P}(\text{Match}) = 1 - \mathbb{P}(\text{kein Match})$$

Ergebnis: Z.B. für $n = \mathbf{23}$ Personen ist $\mathbb{P}(\text{Match}) \approx 0.507 > 50\%$.

Die Hypergeometrische Verteilung

- Diskrete Wahrscheinlichkeitsverteilung.
- Beschreibt Wahrscheinlichkeit einer Anzahl von Erfolgen in einer Stichprobe, die **ohne Zurücklegen** aus einer endlichen Population gezogen wird.
- Population besteht aus zwei Arten von Elementen (Erfolg, Misserfolg).

Hypergeometrische Verteilung: Formel

Gegeben:

- N: Gesamtgröße der Population
- K: Anzahl der Erfolge in der Population
- $N - K$: Anzahl der Misserfolge in der Population
- n: Größe der gezogenen Stichprobe
- k: Anzahl der Erfolge in der Stichprobe, die wir beobachten wollen

Wahrscheinlichkeitsverteilung/Wahrscheinlichkeitsmassenfunktion (engl. :PMF): Eine Zufallsvariable X ist hypergeometrisch verteilt mit den obigen Parametern, wenn gilt

$$\mathbb{P}(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

- $\binom{a}{b}$: Binomialkoeffizient ("a über b").
- Mögliche Werte für k: $\max(0, n - (N - K)) \leq k \leq \min(n, K)$.

Herleitung der hypergeometrischen Verteilung

Sei $\Omega = \{A \subseteq \{1, \ldots, N\} : |A| = n\}$ versehen mit der Potenzmenge als σ-Algebra, und der Gleichverteilung als Wahrscheinlichkeitsmaß. (In der Vorlesung wurde an der Tafel der Fall $n = 6, K = 6, N = 49$ (d.h. Lotto 6 aus 49) besprochen). Sei nun A_K eine $K-$elementige Teilmenge von $\{1, \ldots, N\}$. Z.B. $K = 6$ und $A_K =$ Gewinnzahlen im Lotto. Die Zufallsvariable

$$\begin{aligned} X &: \Omega \to \{0, \ldots, K\} \\ X(\omega) &= |\omega \cap A_K| \end{aligned}$$

Man kann dann nachweisen, dass $\mathbb{P}_X$ gerade die hypergeometrische Verteilung mit Parametern n, K, N ist.

Hypergeometrische Verteilung: Anwendungsbeispiel - Qualitätskontrolle

Szenario:

- Lieferung: $N = 100$ Glühbirnen
- Defekt: $K = 10$ Glühbirnen
- Stichprobe: $n = 5$ Glühbirnen, zufällig gleichverteilt gezogen

Frage: Wahrscheinlichkeit, dass genau $k = 2$ defekte Glühbirnen in der Stichprobe sind?

$$\mathbb{P}(X = 2) = \frac{\binom{10}{2}\binom{100-10}{5-2}}{\binom{100}{5}} = \frac{\binom{10}{2}\binom{90}{3}}{\binom{100}{5}}$$

$$\mathbb{P}(X = 2) = \frac{45 \cdot 117480}{75287520} = \frac{5286600}{75287520} \approx \mathbf{0.0702} \quad (\mathbf{7,02\%})$$

Hypergeometrische Verteilung: Anwendungsbereiche

- **Qualitätskontrolle**
- **Kartenspiele** (z.B. Wahrscheinlichkeit, Asse in einem Pokerblatt zu erhalten)
- **Biologie** (z.B. Populationsschätzung durch Fang-Wiederfang)
- **Urnenmodelle** (klassisches Ziehen ohne Zurücklegen)

Lotto 6 aus 49: Grundlagen

- Ziehen **ohne Zurücklegen** aus endlicher Menge $\implies$ Hypergeometrische Verteilung!
- **Gesamtpopulation (N):** 49 Kugeln
- **Anzahl der Erfolgs-Kugeln in Population (K):** 6 Gewinnzahlen
- **Größe der gezogenen Stichprobe (n):** 6 getippte Zahlen auf dem Lottoschein
- **Anzahl der 'Misserfolgs-Zahlen' ($N - K$):** $49 - 6 = 43$ Nieten

Formel:

$$\mathbb{P}(X = k) = \frac{\binom{6}{k}\binom{43}{6-k}}{\binom{49}{6}}$$

Gesamtzahl möglicher 6er-Kombinationen: $\binom{49}{6} = \mathbf{13.983.816}$

Lotto 6 aus 49: Wahrscheinlichkeiten (ohne Superzahl)

- **6 Richtige ($k = 6$):**

$$\mathbb{P}(X=6) = \frac{\binom{6}{6}\binom{43}{0}}{13.983.816} = \frac{1 \cdot 1}{13.983.816} = \mathbf{1} \text{ zu } \mathbf{13.983.816} \approx \mathbf{0,000007}$$

- **5 Richtige ($k = 5$):**

$$\mathbb{P}(X=5) = \frac{\binom{6}{5}\binom{43}{1}}{13.983.816} = \frac{6 \cdot 43}{13.983.816} = \frac{\mathbf{258}}{\mathbf{13.983.816}} \approx \mathbf{0,001845\%}$$

- **4 Richtige ($k = 4$):**

$$\mathbb{P}(X=4) = \frac{\binom{6}{4}\binom{43}{2}}{13.983.816} = \frac{15 \cdot 903}{13.983.816} = \frac{\mathbf{13.545}}{\mathbf{13.983.816}} \approx \mathbf{0,09686\%}$$

- **3 Richtige ($k = 3$):**

$$\mathbb{P}(X=3) = \frac{\binom{6}{3}\binom{43}{3}}{13.983.816} = \frac{20 \cdot 12.341}{13.983.816} = \frac{\mathbf{246.820}}{\mathbf{13.983.816}} \approx \mathbf{1,765\%}$$

Lotto 6 aus 49: Mit Superzahl

- Die Superzahl ist eine Ziffer von 0 bis 9 (10 Möglichkeiten).
- Die Wahrscheinlichkeit, die richtige Superzahl zu haben, ist $\mathbf{1/10}$.
- Superzahl und Gewinnzahlen werden als unabhängig modelliert
- **Jackpot (6 Richtige und Superzahl):**

$$\mathbb{P}(\text{6 Richtige } \cap \text{Superzahl richtig}) = \mathbb{P}(\text{6 Richtige}) \cdot \mathbb{P}(\text{Superzahl richtig})$$

$$\mathbb{P}(\text{6 Richtige } \cap \text{ Superzahl richtig }) = \frac{1}{13.983.816} \cdot \frac{1}{10} = \frac{1}{139.838.160}$$

- Wahrscheinlichkeit für den Hauptgewinn:
  $\mathbf{1}$ zu $\mathbf{140}$ Millionen $\approx \mathbf{0,00000000715}\%$.

Fazit: Lotto-Gewinnwahrscheinlichkeiten sind extrem gering.