

Satz von Bayes und Dateninterpretation

Klasse 11

2026-08-19

Grundidee

Oft kennt man die Wahrscheinlichkeiten **vorwärts** — vom Verursacher zur Wirkung, etwa $P_A(B)$ — gesucht ist aber die **Rückrichtung** $P_B(A)$.

Dafür braucht man zwei Formeln. Zuerst den Nenner über alle vollständigen Alternativen $A_1, A_2, \dots$ (**Satz der totalen Wahrscheinlichkeit**):

$$P(B) = P(A_1) \cdot P_{A_1}(B) + P(A_2) \cdot P_{A_2}(B) + \dots$$

Und dann die Umkehrung selbst (**Satz von Bayes**):

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

Vorgehen: Baumdiagramm vorwärts zeichnen $\rightarrow$ alle Pfade zu B addieren, das ergibt $P(B) \rightarrow$ den *einen* gesuchten Pfad durch diese Summe teilen.

Basisrateneffekt

Ist ein Ereignis A sehr selten, so kann $P_B(A)$ klein bleiben, obwohl der „Test“ B sehr zuverlässig ist. Die vielen falschen Treffer aus der riesigen Gegengruppe überwiegen dann die wenigen echten. Man darf $P_A(B)$ und $P_B(A)$ deshalb nie verwechseln.

Korrelation ist keine Kausalität. Treten zwei Merkmale gemeinsam gehäuft auf, kann das an einer gemeinsamen dritten Ursache liegen, an der Richtung des Zusammenhangs oder am Zufall — ein Nachweis von Ursache und Wirkung ist es nicht.

Aufgaben

Aufgabe 1

Ein Betrieb bezieht Bauteile von zwei Lieferanten. 80 % stammen von L1, 20 % von L2. Von den Teilen aus L1 sind 1 % defekt, von denen aus L2 sind 6 % defekt.

- Berechne die Wahrscheinlichkeit, dass ein zufällig entnommenes Teil defekt ist.
- Ein Teil ist defekt. Mit welcher Wahrscheinlichkeit stammt es von L1?

Aufgabe 2

Urne A enthält 3 rote und 7 blaue Kugeln, Urne B enthält 6 rote und 4 blaue. Es wird zuerst per fairem Münzwurf eine Urne gewählt und daraus eine Kugel gezogen.

- Wie wahrscheinlich ist es, eine rote Kugel zu ziehen?
- Die gezogene Kugel ist rot. Mit welcher Wahrscheinlichkeit stammt sie aus Urne A?

Aufgabe 3

Ein Test auf eine Krankheit hat die Sensitivität $P_K(+) = 0,99$ und die Spezifität $P_{\bar{K}}(-) = 0,95$. In der Bevölkerung sind 0,5 % tatsächlich krank.

- Berechne $P(+)$.
- Berechne $P_+(K)$ und deute das Ergebnis in einem Satz.

Aufgabe 4

Die Grafik zeigt 10 000 getestete Personen. 1 % von ihnen ist tatsächlich krank; der Test erkennt 90 % der Kranken und liefert bei 5 % der Gesunden fälschlich ein positives Ergebnis.

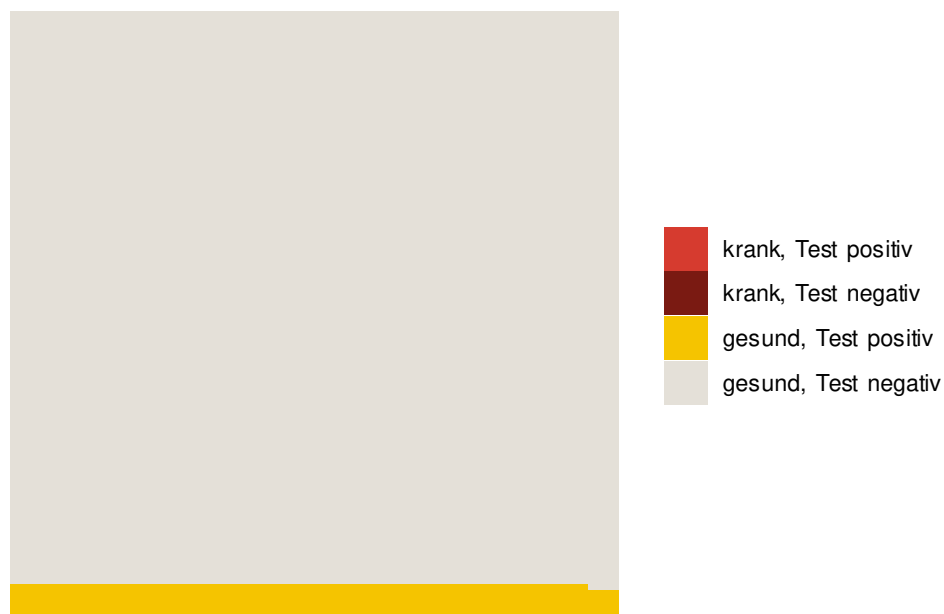


Abbildung 1: 10 000 Personen. Jedes Kästchen ist eine Person.

- Wie viele Personen sind tatsächlich krank, wie viele davon werden erkannt?
- Wie viele Tests fallen insgesamt positiv aus?
- Berechne daraus $P_+(K)$.

- d) Erkläre mit Blick auf die Grafik, warum dieser Wert so klein ist, obwohl der Test 90 % der Kranken erkennt.

Aufgabe 5

Eine Zeitung meldet: „Kinder, die ein eigenes Zimmer haben, schneiden in der Schule deutlich besser ab. Ein eigenes Zimmer verbessert also die Schulleistung.“

- a) Welcher Denkfehler steckt in der Schlussfolgerung?
- b) Nenne eine plausible dritte Größe, die beide Merkmale zugleich beeinflusst.
- c) Formuliere die Meldung so um, dass sie fachlich korrekt ist.

Aufgabe 6

Ein Bauteil wird von drei Werken geliefert: A liefert 50 % (Ausschussquote 1 %), B liefert 30 % (4 %), C liefert 20 % (10 %).

- a) Berechne die Gesamtausschussquote.
- b) Ein defektes Teil wird untersucht. Berechne für jedes Werk die Wahrscheinlichkeit, dass es von dort stammt.
- c) Kontrolliere dein Ergebnis mit einer Eigenschaft, die die drei Werte erfüllen müssen.

Aufgabe 7

Eine Krankheit tritt bei einer von einer Million Personen auf. Ein Test erkennt 99,99 % der Kranken und liefert bei 0,01 % der Gesunden fälschlich ein positives Ergebnis.

- a) Berechne $P(+)$.
- b) Berechne $P_+(K)$.

Aufgabe 8

Fabrik F1 stellt 10^9 Schrauben her, davon ist der Anteil 10^{-6} defekt. Fabrik F2 stellt nur 1000 Schrauben her, davon sind aber 50 % defekt. Alle Schrauben werden vermisch.

Eine zufällig gezogene Schraube ist defekt. Mit welcher Wahrscheinlichkeit stammt sie aus F2?

Lösungen

Lösung 1

- a) Satz der totalen Wahrscheinlichkeit:

$$P(D) = 0,8 \cdot 0,01 + 0,2 \cdot 0,06 = 0,008 + 0,012 = 0,02$$

- b) Satz von Bayes:

$$P_D(L1) = \frac{P(L1) \cdot P_{L1}(D)}{P(D)} = \frac{0,008}{0,02} = 0,4$$

Genau 40 % der defekten Teile stammen von L1 — obwohl L1 die sechsfach bessere Quote hat, weil L1 viermal so viel liefert.

Lösung 2

a) $P(A) = P(B) = 0,5$, $P_A(\text{rot}) = \frac{3}{10}$, $P_B(\text{rot}) = \frac{6}{10}$:

$$P(\text{rot}) = 0,5 \cdot 0,3 + 0,5 \cdot 0,6 = 0,15 + 0,30 = 0,45$$

b)

$$P_{\text{rot}}(A) = \frac{0,15}{0,45} = \frac{1}{3} \approx 0,333$$

Lösung 3

Aus der Spezifität folgt die Falsch-positiv-Rate: $P_{\bar{K}}(+) = 1 - 0,95 = 0,05$.

a)

$$P(+) = 0,005 \cdot 0,99 + 0,995 \cdot 0,05 = 0,00495 + 0,04975 = 0,0547$$

b)

$$P_+(K) = \frac{0,00495}{0,0547} \approx 0,0905$$

Trotz positivem Test ist man nur mit etwa 9 % tatsächlich krank: Weil nur 0,5 % der Bevölkerung krank sind, liefern die vielen Gesunden mit ihren 5 % Fehlalarmen weit mehr positive Tests als die wenigen Kranken.

Lösung 4

a) 1 % von 10 000 sind 100 Kranke. Davon werden 90 % erkannt, also 90 **Personen** (dunkelrot: die 10 übersehenen).

b) Gesunde: $10\,000 - 100 = 9900$. Davon 5 % falsch positiv:

$$0,05 \cdot 9900 = 495$$

Positive Tests insgesamt: $90 + 495 = 585$.

c)

$$P_+(K) = \frac{90}{585} \approx 0,1538$$

d) In der Grafik ist die gelbe Fläche (falsch positiv) mehr als fünfmal so groß wie die hellrote (richtig positiv). Die Gruppe der Gesunden ist so viel größer, dass selbst ihre kleine Fehlerquote von 5 % mehr positive Tests erzeugt als alle Kranken zusammen. Das ist der Basisrateneffekt.

Lösung 5

- a) Aus einer **Korrelation** wird unzulässig auf **Kausalität** geschlossen. Beobachtet wurde nur ein gemeinsames Auftreten, nicht eine Wirkung.
 - b) Plausibel ist das **Einkommen** bzw. die Wohnsituation der Familie: Wer sich eine größere Wohnung leisten kann, kann meist auch Nachhilfe, Bücher und ruhige Lernzeit bereitstellen. Das eigene Zimmer ist dann Begleiterscheinung, nicht Ursache.
 - c) Korrekt wäre etwa: „*Kinder mit eigenem Zimmer schneiden im Mittel besser ab. Ob das Zimmer die Ursache ist oder beide Merkmale auf die wirtschaftliche Lage der Familie zurückgehen, lässt sich aus den Daten nicht entscheiden.*“
-

Lösung 6

- a)
$$P(D) = 0,5 \cdot 0,01 + 0,3 \cdot 0,04 + 0,2 \cdot 0,10 = 0,005 + 0,012 + 0,020 = 0,037$$
- b) Jeweils der eigene Pfad geteilt durch $P(D)$:
$$P_D(A) = \frac{0,005}{0,037} \approx 0,135, \quad P_D(B) = \frac{0,012}{0,037} \approx 0,324, \quad P_D(C) = \frac{0,020}{0,037} \approx 0,541$$
- c) Die drei Werte müssen sich zu 1 addieren, denn das defekte Teil stammt mit Sicherheit aus genau einem der drei Werke:

$$0,135 + 0,324 + 0,541 = 1 \quad \checkmark$$

Werk C liefert nur ein Fünftel der Teile, verursacht aber über die Hälfte des Ausschusses.

Lösung 7

$$P(K) = 10^{-6}, \quad P_K(+) = 0,9999, \quad P_{\bar{K}}(+) = 10^{-4}.$$

- a)
$$P(+) = 10^{-6} \cdot 0,9999 + (1 - 10^{-6}) \cdot 10^{-4} \approx 0,0001010$$
- b)
$$P_+(K) = \frac{10^{-6} \cdot 0,9999}{0,0001010} \approx 0,0099$$

Nur etwa 1 % — bei einem Test, der in beide Richtungen zu 99,99 % richtig liegt. Es ist genau dieselbe Rechnung wie in Aufgabe 3; die Krankheit ist hier nur so selten, dass die Fehlalarme die echten Fälle um das Hundertfache übertreffen.

Lösung 8

Statt mit Anteilen rechnet man hier bequemer mit Stückzahlen.

Defekte aus F1: $10^9 \cdot 10^{-6} = 1000$

Defekte aus F2: $1000 \cdot 0,5 = 500$

$$P_D(\text{F2}) = \frac{500}{1000 + 500} = \frac{500}{1500} = \frac{1}{3}$$

Obwohl F2 millionenfach weniger produziert, stammt jede dritte defekte Schraube von dort. Dasselbe Vorgehen wie in Aufgabe 1 — nur sind die Größenordnungen grotesk.