

Probability and Bayes

(Conditional versus Marginal)

M Loecher

Bayes Theorem

Bayes Theorem

“Inverting” Conditional Probabilities

$$\begin{aligned}P(A|B) &= \frac{P(A) \cdot P(B|A)}{P(B)} = \frac{P(A)}{P(B)} \cdot P(B|A) \\&= \frac{P(A) \cdot P(B|A)}{P(B \cap A) + P(B \cap \bar{A})} \\&= \frac{P(A) \cdot P(B|A)}{P(A) \cdot P(B|A) + P(\bar{A}) \cdot P(B|\bar{A})} \\&= \frac{P(A) \cdot P(B|A)}{\sum_k P(A_k) \cdot P(B|A_k)}\end{aligned}$$

In the context observations/measurements do we call $P(A)$ the **prior** and $P(A|B)$ the **posterior probability**

Bayes Theorem, Example I

Kahnemann, Thinking, Fast and Slow

A cab was involved in a hit-and-run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:

85% of the cabs in the city are Green and 15% are Blue. A witness identified the cab as Blue. The court tested the reliability of the witness under the circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colors 80% of the time and failed 20% of the time. What is the probability that the cab involved in the accident was Blue rather than Green?

Bayes Theorem, Example II

Cliches/Stereotypes

- ▶ The proportions of green voters that commute to work by bike are similar in Holland and Deutschland. Are the inverse probabilities (i.e. how likely does someone vote for the green party once we know that she/he commutes to work by bike) similar as well ?
- ▶ Assume the proportion of truly right-wing extremists in the USA is about 5%. Let us further assume that 90% of those voted for Trump. How reasonable if the -often heard- "conclusion" that 90% of all Trump voters are right-wing extremists?

Solutions

- ▶ We also need to assume that the proportion of green voters are similar in Holland and Germany! Define events $B = \text{bike to work}$, $G = \text{votes green}$, then

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

We are assuming that $P(B|G)_{\text{holland}} \approx P(B|G)_{\text{germany}}$ and $P(G)_{\text{holland}} \approx P(G)_{\text{germany}}$. So, due to $P(B)_{\text{holland}} > P(B)_{\text{germany}}$ it follows that $P(G|B)_{\text{holland}} < P(G|B)_{\text{germany}}$, inference is less "sharp".

- ▶ Define events $T = \text{votes for Trump}$, $R = \text{right-wing extremists}$. We need to add information on the percentage of non-right wing Trump voters: $P(T|\bar{R}) \approx 0.45$. Then

$$P(R|T) = \frac{P(R)}{P(T)} \cdot P(T|R) = \frac{0.05}{0.05 \cdot 0.9 + 0.95 \cdot 0.45} \cdot 0.9 \approx 0.095$$

So, the -often heard- "conclusion" that 90% of all Trump voters are right-wing extremists is unreasonable.