Bayesian Inference

Oliver Filla

November 20, 2024

Universität zu Köln

Determinism and its Limits

Why do we need Probabilities?

Determinism

- *Every event is determined by past events.*
- assumed in Classical Mechanics
- contradicted in Quantum Mechanics?

Reality

- limited knowledge, e.g. friction
- limited control over initial state, e.g. in chaotic systems

Example: Throwing dice: Initial angle and velocity are unknown or difficult to control. It is unknown if the dice are fair.

$\Rightarrow$ Thus, we need probabilities to predict events!

Frequentist statistics

Probabilities are determined by distributions of random events.
They are characterized by the frequency of occurring events.

	1	2	3	4	5	6	$\sum$
frequency	10	5	7	6	9	7	44
probability [%]	23	11	16	14	20	16	100

results of dice rolls; expected PMF $p(i) = 1/6 = 16.\bar{6}\,\%$

- Probability Mass Function (PMF) (discrete)
- Probability Density Function (PDF) (continuous)
- Cumulative Distribution Function (CDF) (integrated PDF)

$p(6) = \frac{1}{6} \Rightarrow$ *I expect that every 6th dice roll shows a 6, on average.*

Bayesian statistics

Bayesian statistics describes the credibility of hypotheses.

- hypothesis H: falsifiable assumption that a specific statement is true, e.g. *it will rain today.*
- credibility $C(H) \in [0, 1]$: strength of belief in H
- mathematically, credibilities and probabilities are identical
- difference is conceptual
- $C(H)$ can be updated with new data

$p(6) = \frac{1}{6} \Rightarrow$ *I believe the hypothesis H "the die roll shows a 6" with a credibility of $C(H) = \frac{1}{6}$.*

Conditional Probability

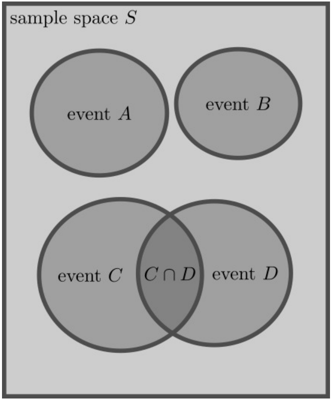
$P(A|B)$ describes the probability of event A with the assumption that event B takes place.

- read: $P(A \text{ given that } B)$
- filters sets of possible events
- joint probability: both A and B:
  $P(A \cap B) = P(A)P(B|A)$

$\Rightarrow$ only when A and B independent:
  $P(A \cap B) = P(A)P(B)$

- joint probability: any A or B:
  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$\Rightarrow$ A and B mutually exclusive:
  $P(A \cup B) = P(A) + P(B)$



Probability & Likelihood

Recap from the talk *The Maximum Likelihood Principle* by Maja Korbmacher.

- Probability P: tool for prediction of events
- Likelihood L: tool for modelling
- equal for given hypothesis H and events E
- but only if no prior knowledge is available
- Bayes' theorem generalizes this, using prior knowledge

$$L(H|E) \stackrel{?}{=} P(E|H) \tag{1}$$

Bayes' Theorem

Bayes' theorem describes the procedure of updating a hypothesis H, given that new evidence E is available. It is the scientific method of changing beliefs.

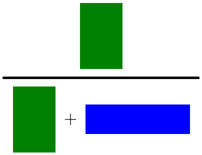
- Hypothesis H
- Evidence E
- follows from joint probability $P(E \cap H)$
- overall likelihood $P(E)$ (*law of total probability*)

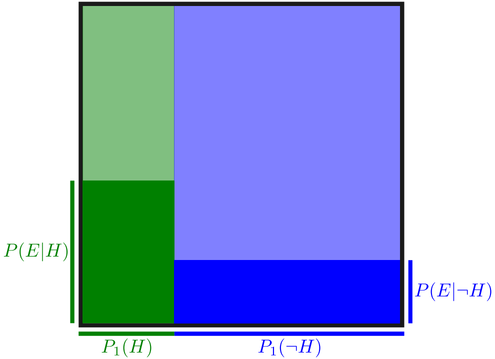
$$
\begin{aligned}
P_2(H|E) &= \frac{P(E|H)P_1(H)}{P(E)} && (2)\\
\Rightarrow L(H|E) &= P(E|H) \cdot \frac{P(H)}{P(E)} && (3)\\
P(E) &= P(E|H)P(H) + P(E|\neg H)P(\neg H) && (4)
\end{aligned}
$$

Bayes' Theorem

Geometric Interpretation

- square: probability 1
- prior $P_1(H)$

$P_2(H|E) =$ 



The black square with side length 1 contains all probabilities.

Updating Credibilities

with Bayes' Theorem

$$P_2(H|E) \quad = \quad \frac{P(E|H)P_1(H)}{P(E)}$$

How to update the credibility of a hypothesis H:

1. assert prior $P_1(H)$
2. evaluate evidence E: $\rightarrow$ likelihood $P(E|H)$
3. calculate overall likelihood $P(E)$
4. calculate posterior probability $P_2(H|E)$

Updating Credibilities

The Posterior Odds ratio

The Posterior Odds ratio compares the posterior probabilities P_2 for two different hypotheses $H_{A,B}$.

- relative probability: compare probabilities
- absolute probabilities have limited use

$\rightarrow$ $\frac{P(A)}{P(B)} = \frac{1}{2}$ tells event B occurs twice as often as event A, for e.g. $P(A) = 0.1, P(B) = 0.2$ or $P(A) = 0.3, P(B) = 0.6$.

- overall likelihood $P(E)$ becomes irrelevant

$$\frac{P_2(H_A|E)}{P_2(H_B|E)} = \frac{P(E|H_A)}{P(E|H_B)} \frac{P_1(H_A)}{P_1(H_B)} \tag{5}$$

Example

rapid antigen tests

Assume a *COVID rapid test* with a sensitivity of 80%, causing 20 % false-negative results, and a specificity of 98%, causing 2% false-positive results.

Source: Robert Koch Institute

Bibliography

[1] Lawrence A. (2019) *Probability in Physics: An Introductory Guide*, Springer Nature Switzerland, Chapters 1 & 7.1 – 7.5, DOI 10.1007/978-3-030-04544-9

[2] Clyde M. *et al.* (2022) *An Introduction to Bayesian Thinking: A Companion to the Statistics with R Course*, Github.com/enstarprise, on 2022-06-15

[3] 3Blue1Brown (2020) *Bayes theorem, the geometry of changing beliefs*, YouTube.com/watch?v=HZGCoVF3YvM, 2024-11-16

[4] Kozyrkov C. (2020) *Are you Bayesian or Frequentist?* YouTube.com/watch?v=GEFxFVESQXc, 2024-11-16

Thank you for your attention!