Berlin 3.6.2021

BERGISCHE UNIVERSITÄT WUPPERTAL

Was ist Wahrscheinlichkeit?

Was wahr scheint?

***Heute*:**

Verschiedene mathematische Konzeptualisierungen des Wahrscheinlichkeitsbegriffs und Anmerkungen zur wissenschaftlichen Methodik

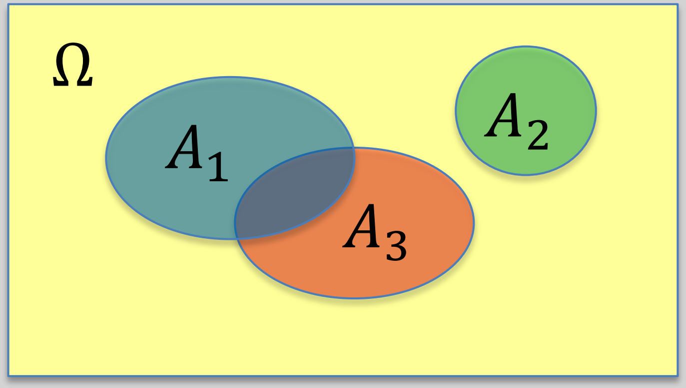
Oliver Passon

Laplace'sche Wahrscheinlichkeit

$$P = \frac{\text{Anzahl der Günstigen}}{\text{Anzahl der Möglichen}}$$

Voraussetzung:
Elementarereignisse, die „gleich **wahrscheinlich**" sind. Damit ist diese Definition jedoch zirkulär.

Die Axiome von Kolmogorov (1933)



$$\xrightarrow{P} \mathbb{R}$$

Andrej N. Kolmogorov
(1903-1987)

Eine Abbildung P von $A_i \subset \Omega$ in die reellen Zahlen heißt Wahrscheinlichkeit, wenn

1. Für alle A_i gilt: $0 \leq P(A_i) \leq 1$
2. Es gilt: $P(\Omega) = 1$
3. Für disjunkte Teilmengen ($A_1 \cap A_2 = \emptyset$) gilt: $P(A_1 \cup A_2) = P(A_1) + P(A_2)$

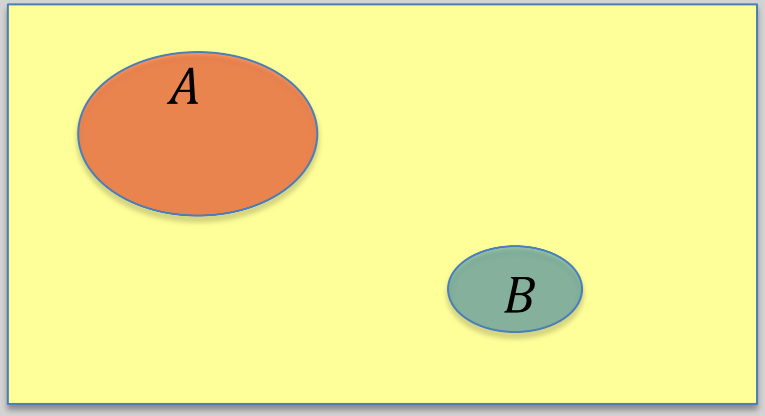
Beispiel: In obigem Diagramm können die Flächeninhalte als Wahrscheinlichkeiten aufgefasst werden.

Man beachte: Die Frage nach der **Bedeutung** von Wahrscheinlichkeit bleibt unbeantwortet!

Aus den Kolmogorov Axiomen folgt:

Die „bedingte Wahrscheinlichkeit von A falls B vorliegt“:

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

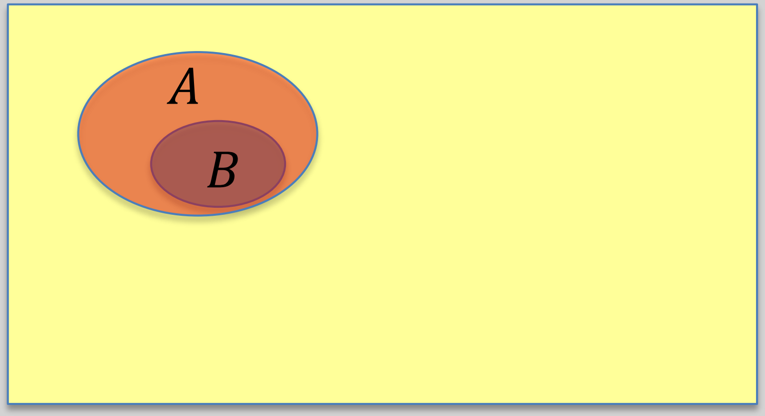


$$A \cap B = \emptyset \qquad \frac{P(A \cap B)}{P(B)} = 0$$

Aus den Kolmogorov Axiomen folgt:

Die „bedingte Wahrscheinlichkeit von A falls B vorliegt“:

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



$A \cap B = \emptyset$ $\quad \dfrac{P(A \cap B)}{P(B)} = 0$

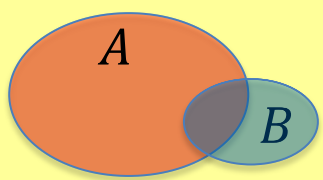
$B \subset A$

$A \cap B = B$ $\quad \dfrac{P(A \cap B)}{P(B)} = \dfrac{P(B)}{P(B)} = 1$

Aus den Kolmogorov Axiomen folgt:

Die „bedingte Wahrscheinlichkeit von A falls B vorliegt“:

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



Hier: $P(A|B) \approx 0{,}4$

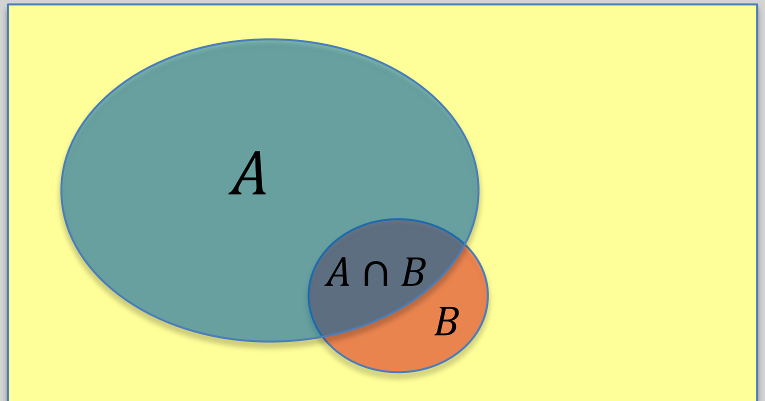
$A \cap B = \emptyset$ $\qquad \dfrac{P(A \cap B)}{P(B)} = 0$

$B \subset A$

$A \cap B = B$ $\qquad \dfrac{P(A \cap B)}{P(B)} = \dfrac{P(B)}{P(B)} = 1$

$$0 \leq \frac{P(A \cap B)}{P(B)} \leq 1$$

Wichtig: $P(A|B) \neq P(B|A)$



$$P(A|B) = \frac{P(A \cap B)}{P(B)} \approx 0{,}5$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \approx 0{,}1$$

Hier gilt offensichtlich: $P(A|B) > P(B|A)$

Aber wieviel größer genau?

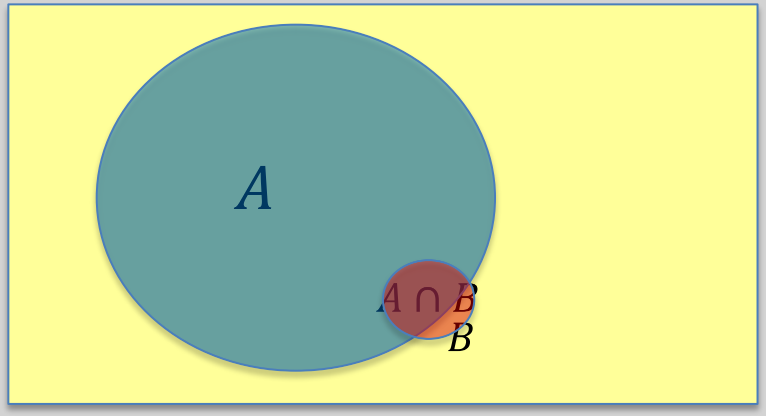
$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}$$

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$ „Satz von Bayes“

Beispiele für $P(A|B) \neq P(B|A)$

Ereignis A: Person ist weiblich
Ereignis B: Person arbeitet in einem Pflegeberuf

Dann gilt $P(A|B) \approx 0{,}85$ („in der Pflege ist der Frauenanteil hoch“)
Aber: $P(B|A) \approx 0{,}05$ („die Pflege ist nur ein Berufsfeld von vielen“)



Wichtiges aktuelles Beispiel:

Ereignis „+“ : Person ist positiv auf COVID 19 getestet
Ereignis „inf“ : Person ist mit COVID 19 infiziert

Dann gilt $P(+|\mathrm{inf}) \approx 0{,}95$ („Infizierte werden sehr sicher erkannt“)

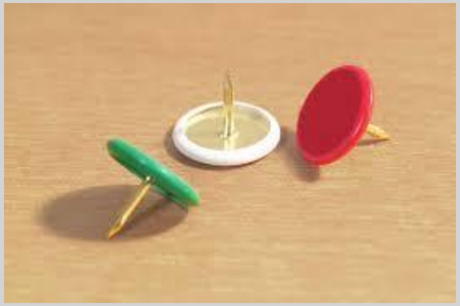
Aber: $P(\mathrm{inf}|+) = \frac{P(+|\mathrm{inf})P(\mathrm{inf})}{P(+)}$ hängt von $P(\mathrm{inf})$ ab – der sog. „Vortestwahrscheinlichkeit“.

Ganz vergessen: Wir wollten doch wissen, was Wahrscheinlichkeit ***bedeutet***!

Der „Marktführer“: Die frequentistische Interpretation von Wkt.

$$P(A) \approx \frac{\text{Häufigkeit von } A \text{ bei } n \text{ Versuchen}}{n}$$

(**Idee**: Anzahl n der Versuche möglichst groß machen!)



Ludwig von Mises
(1881-1973)

Vorteil: Diese Interpretation setzt die Wahrscheinlichkeit nicht bereits voraus (und kann leicht ermittelt werden)!
Nachteil: Dieser Wahrscheinlichkeitsbegriff ist nur auf wiederholbare Vorgänge anwendbar. Die Wahrscheinlichkeit eines Einzelereignisses oder einer Hypothese hat in dieser Konzeption keinen Sinn.

Im alltäglichen Leben (und in der Wissenschaft) interessiert man sich aber **sehr** häufig für die Wahrscheinlichkeit von nicht-wiederholbaren Vorgängen („Hypothesen“).

Beispiele:

- Kommt Albert am nächsten Wochenende auf die Party?
- Bestehe ich diese Prüfung?
- Gibt es eine von Menschen verursachte Klimakrise?
- …

Um hier über „Wahrscheinlichkeit“ reden zu können, braucht man einen anderen Wahrscheinlichkeitsbegriff

Subjektivistischer Wahrscheinlichkeitsbegriff („Bayes'sche Statistik“)

$P(A) =$ Grad der Überzeugung, dass A eintritt bzw. zutrifft

klingt komisch…
Vorteil: $P(A)$ ist nicht an wiederholbaren Vorgang geknüpft.
Aber: Ohne jede Vorkenntnis kann der Wert $P(A)$ tatsächlich nur geraten werden (und verschiedene Personen mögen unterschiedliche „Überzeugungen“ haben).

Aber sobald neue Informationen vorliegen („Daten“ D) kann die Berechnung dieser Wahrscheinlichkeit als Problem der bedingten Wahrscheinlichkeit aufgefasst werden: $P(A|D)$

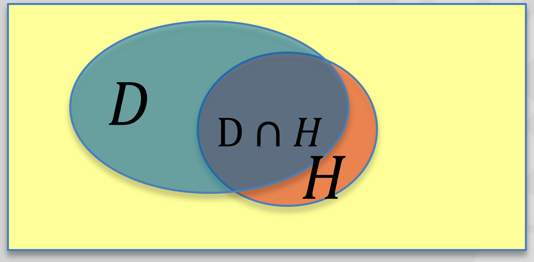
These: Der Satz von Bayes beschreibt dabei bestimmte Aspekte des (nicht nur) wissenschaftlichen Vorgehens

ursprüngliche Überzeugung	zusätzliche Informationen	→	neue Überzeugung

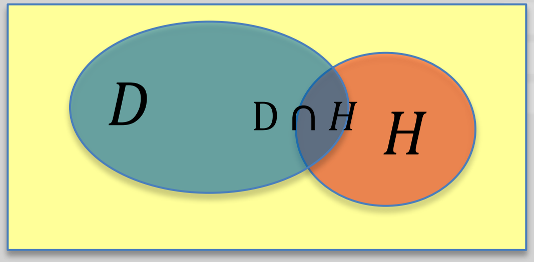
$$P(\text{Hypothese}) \cdot \frac{P(\text{Daten}|\text{Hypothese})}{P(\text{Daten})} = P(\text{Hypothese}|\text{Daten})$$

„update-Faktor"

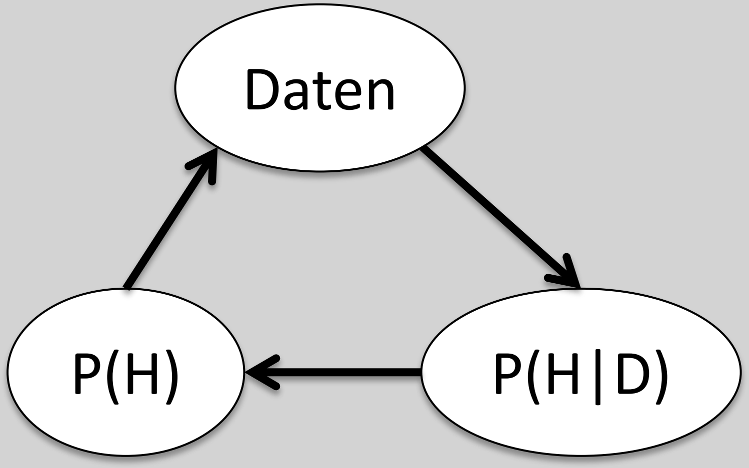
- Er ist > 1 wenn das Zutreffen der Hypothese das Auftreten der Daten wahrscheinlicher macht.



- Er ist < 1, wenn das Zutreffen der Hypothese das Auftreten der Daten unwahrscheinlicher macht…



Dieser Prozess hört (im Prinzip) nie auf:



Auf diese Weise werden sich durch mehr Daten bzw. Informationen (i.d.R.) verschiedene Überzeugungen einander annähern.

- Aber: Daten können auch verschieden interpretiert (oder angezweifelt) werden
- Andere Probleme der „Konsensfindung“:
  - $P(H) = 0$ dann folgt: $P(\mathrm{H}|\mathrm{D}) = 0 \cdot \frac{P(D|H)}{P(D)} = P(H)$

$\rightarrow$ Keine Informationen/Daten können diese Überzeugung verändern!

- $P(H) = 1$ („Hypothese wahr“) dann folgt:

$$P(\mathrm{H}|\mathrm{D}) = 1\cdot\frac{P(D|H)}{P(D)} = 1\cdot\frac{P(\cancel{D|}H)}{P(\cancel{D|}H)\underbrace{P(H)}_{=1}+\cancel{P(D|\bar{H})\underbrace{P(\bar{H})}_{=0}}} = P(\mathrm{H})$$

Auch in diesem Fall können die neuen Daten/Ergebnisse etc. die vormalige Überzeugung nicht ändern! Solche Überzeugungen sind also (Fakten-) resistent. Beispiel: Verschwörungsmythen!

Wissenschaftliches Vorgehen zeichnet sich dadurch aus, dass Überzeugungen nicht absolut sind.

Ebenfalls häufig: Zwei verschiedene Hypothesen (H_0 und H_1) konkurrieren. Neue Daten sollen helfen, zwischen ihnen zu unterscheiden:

$$\frac{P(H_1|D) = \frac{P(D|H_1)P(H_1)}{\cancel{P(D)}}}{P(H_0|D) = \frac{P(D|H_0)P(H_0)}{\cancel{P(D)}}} = \boxed{\frac{P(H_1|D)}{P(H_0|D)} = \frac{P(H_1)}{P(H_0)} \cdot \frac{P(D|H_1)}{P(D|H_0)}}$$

„Bayes Faktor“ (BF)

BF$>$ 1 falls die Daten besser zu H_1 als zu H_0 passen
BF$<$ 1 falls die Daten besser zu H_0 als zu H_1 passen
BF$\approx$ 1 falls die Daten gleich gut zu H_0 und H_1 passen

Auch hier: Ob sich meine Überzeugung ändert hängt auch von den ursprünglichen „Glaubensgraden“ ab.

THE SCIENCES

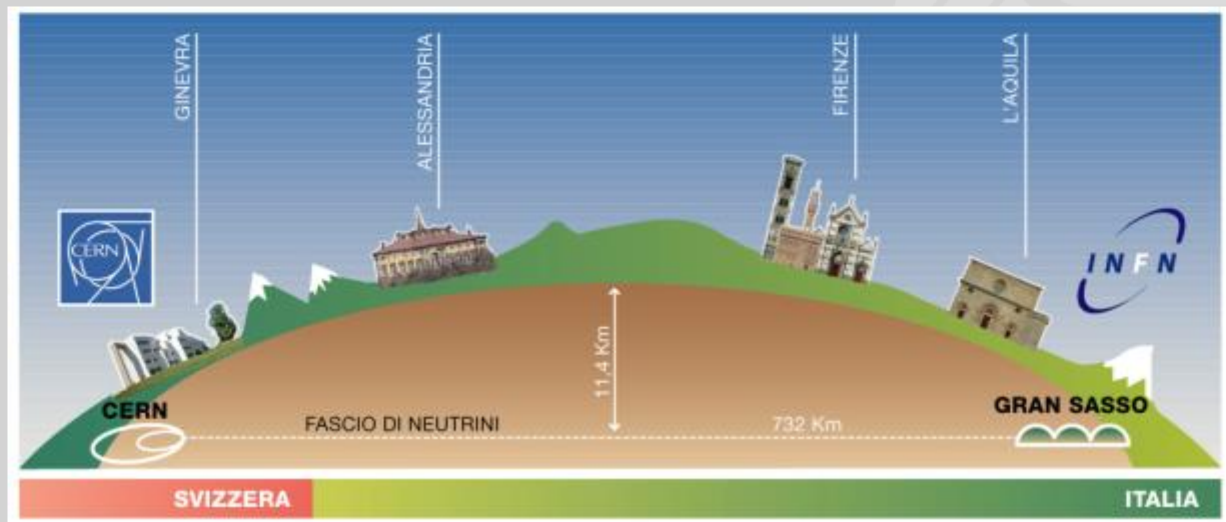
Particles Found to Travel Faster Than Speed of Light

Neutrino results challenge a cornerstone of Albert Einstein's special theory of relativity, which itself forms the foundation of modern physics

By Geoff Brumfiel, Nature magazine on September 22, 2011

Neutrinos wurden am CERN produziert und 730km entfernt registriert. Scheinbar erreichten sie den Detektor 60ns „zu früh“. Ihre Geschwindigkeit betrug $v = 299.800\,\frac{km}{s}$

$$c_{Licht} = 299.792\frac{km}{s}$$



Diese Messung widersprach der etablierten speziellen Relativitätstheorie (SR)!

$$\frac{P(v>c|D)}{P(\mathrm{SR}|D)} = \frac{P(v>c)}{P(\mathrm{SR})} \cdot \frac{P(D|v>c)}{P(D|\mathrm{SR})}$$

BERGISCHE UNIVERSITÄT WUPPERTAL

$$\frac{P(v > c|D)}{P(\mathrm{SR}|D)} = \frac{P(v > c)}{P(\mathrm{SR})} \cdot \frac{P(D|v > c)}{P(D|\mathrm{SR})} = 1:5.000.000$$

$$P(SR) \approx 0{,}999999$$
$$P(v > c) \approx 0{,}000001$$

$$\Rightarrow \frac{P(v > c)}{P(\mathrm{SR})} \approx \frac{1}{1.000.000}$$

UPDATE 8 June 2012

Neutrinos sent from CERN to Gran Sasso respect the cosmic speed limit

Zusammenfassung

- Die Frage „Was ist Wahrscheinlichkeit“ hat keine eindeutige Antwort:

  - Laplace Interpretation. **Anwendung**: Glücksspiele.
  - Frequentistische Interpretation („relative Häufigkeit“). **Anwendung**: Beliebige Zufallsversuche.
  - Subjektivistische Interpretation („Bayes‘sche Statistik“). **Anwendung**: Hypothesen und andere „einmalige“ Ereignisse.
- **Alle** diese Interpretationen erfüllen jedoch die Axiome von Kolmogorov!

- Der Satz von Bayes (zusammen mit der subjektivistischen Interpretation) beschreibt wichtige Aspekte des wissenschaftlichen Vorgehens („Konsensbildung trotz Unsicherheit“).

BERGISCHE UNIVERSITÄT WUPPERTAL

Vielen Dank für ihre Aufmerksamkeit!