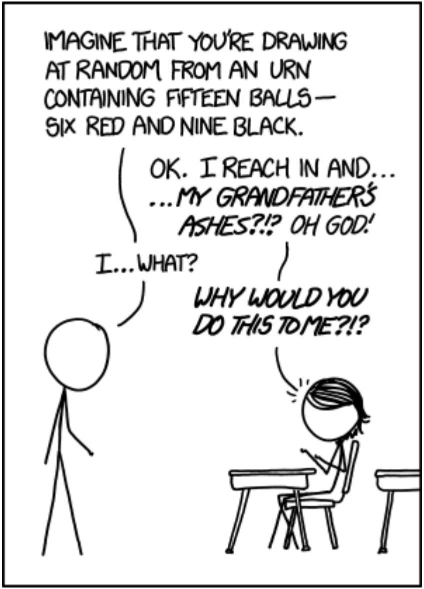
Berner Fachhochschule
Haute école spécialisée bernoise
Bern University of Applied Sciences

Applied Statistics

Conditional Probability & Independence

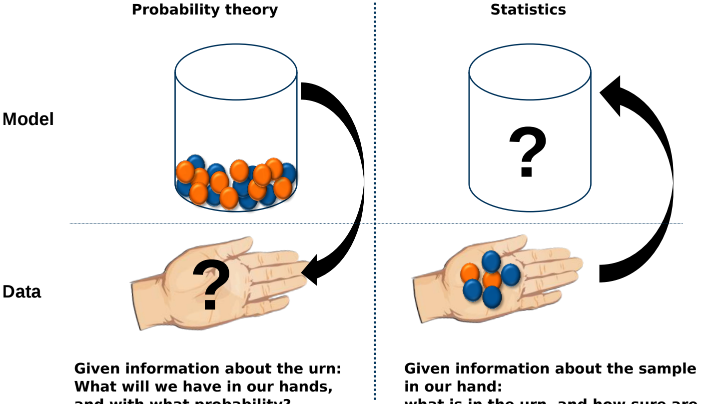
Release FS24

Recap



(xkcd.com)

Recap



**Given information about the urn:
What will we have in our hands,
and with what probability?**

**Given information about the sample
in our hand:
what is in the urn, and how sure are
we about that?**

Birthday problem – Problem 2.13

Probability that at least 2 out of 23 people have the same birthday:

$$1 - \frac{365 \cdot 364 \ldots 343}{365^{23}} \approx 0.51$$

What about you? We have exactly 23 students in this module.

Learning objectives

- Know the mathematical concept of independence and conditional probability
- Draw and read probability trees
- Calculate conditional probabilities using Bayes' theorem
- Apply Bayes' theorem to implement a simple spam filter

Independence

Definition (Independence)

Two events A and B are called **independent** if $\mathrm{P}(A \cap B) = \mathrm{P}(A) \cdot \mathrm{P}(B)$.

Careful: *Independent* events and *disjoint* events are not the same!

- A, B disjoint: $\mathrm{P}(A \cap B) = 0$, i.e. if either event occurs, then the other cannot occur
- A, B independent: $\mathrm{P}(A \cap B) = \mathrm{P}(A) \cdot \mathrm{P}(B)$, i.e. if either event occurs, this gives no information about the other

Examples

- Sex / Handedness: 10% of all people are left-handed; 12% of all men are left-handed. *What do you conclude?*
  Sex and handedness are **not** independent.
- Bonnie & Clyde: Bonnie attends class with probability 0.6; Clyde with probability 0.5. Both of them attend class with probability 0.4. *What do you conclude?*
  Attendance of Bonnie and Clyde are **not** independent. If they were, they would both attend class with probability $0.6 \cdot 0.5 = 0.3$.

Conditional probabilities

Definition (Conditional probability)

Let A and B be events (with $\mathrm{P}(B) > 0$). The **conditional probability of A given B** is defined as

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

Consequence: If A and B are independent, then

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

Example

Bonnie & Clyde: Bonnie attends class with probability 0.6; Clyde with probability 0.5. Both of them attend class with probability 0.4.

$$\begin{aligned}
\mathrm{P}(\text{Bonnie attends}|\text{Clyde attends}) &= \frac{\mathrm{P}(\text{Bonnie attends} \cap \text{Clyde attends})}{\text{Clyde attends}} \\
&= \frac{0.4}{0.5} \\
&= 0.8
\end{aligned}$$

Example

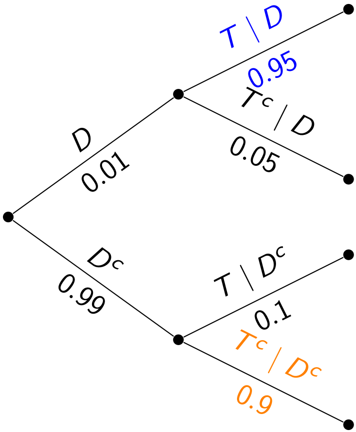
Bonnie & Clyde: Bonnie attends class with probability 0.6; Clyde with probability 0.5. Both of them attend class with probability **0.3**.

$$\begin{aligned}
\mathrm{P}(\text{Bonnie attends}|\text{Clyde attends}) &= \frac{\mathrm{P}(\text{Bonnie attends} \cap \text{Clyde attends})}{\text{Clyde attends}} \\
&= \frac{0.3}{0.5} \\
&= 0.6 \\
&= \mathrm{P}(\text{Bonnie attends})
\end{aligned}$$

Example: PCR test

- Consider PCR-based test to assess whether someone is suffering from COVID-19. This medical test is quite accurate: it detects COVID-19 with 95% probability (**sensitivity** of the test), and indicates the *absence* of COVID-19 with 90% probability (**specificity** of the test).
- Notation: event D: person has COVID-19; event T: test is positive (i.e., indicates infection)
- Assume COVID-19 has incidence of 1%: $\mathrm{P}(D) = 0.01$.
- What is the probability that a random person gets a positive test result?

Easy visualization with a probability tree



$\mathrm{P}(T \cap D) = 0.01 \cdot 0.95 = 0.0095$

$\mathrm{P}(T^c \cap D) = 0.01 \cdot 0.05 = 5e - 04$

$\mathrm{P}(T \cap D^c) = 0.99 \cdot 0.1 = 0.099$

$\mathrm{P}(T^c \cap D^c) = 0.99 \cdot 0.9 = 0.891$

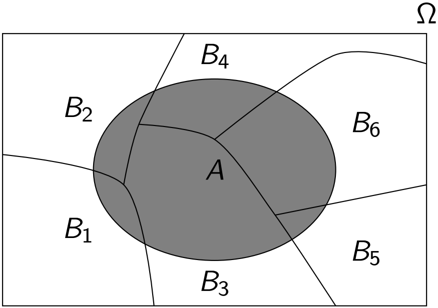
$\mathrm{P}(T) = \mathrm{P}(T \mid D)\mathrm{P}(D) + \mathrm{P}(T \mid D^c)\mathrm{P}(D^c) = 0.0095 + 0.099 =$
0.1085

Law of total probability

Proposition (Law of total probability)

Assume $B_1, B_2, \ldots, B_k$ are disjoint events with $B_1 \cup B_2 \cup \ldots \cup B_k = \Omega$. Then the probability of any event A is

$$\mathrm{P}(A) = \sum_{i=1}^{k} \mathrm{P}(A \cap B_i) = \sum_{i=1}^{k} \mathrm{P}(A|B_i)\mathrm{P}(B_i) \ .$$



Bayes' theorem

Theorem (Bayes' theorem)

Let A and B be events with $\mathrm{P}(A) > 0$ and $\mathrm{P}(B) > 0$. Then we have:

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

In the setting of the law of total probability, we have

$$\mathrm{P}(B_i|A) = \frac{\mathrm{P}(A \cap B_i)}{\mathrm{P}(A)} = \frac{\mathrm{P}(A|B_i)\mathrm{P}(B_i)}{\sum_{j=1}^{k} \mathrm{P}(A|B_j)\mathrm{P}(B_j)} \ .$$

Example: PCR test

The PCR test from the previous example gives a positive result (i.e. indicates an infection).

What is the probability that you actually suffer from COVID-19?

$$\mathrm{P}(D|T) = \frac{\mathrm{P}(T|D) \cdot \mathrm{P}(D)}{\mathrm{P}(T)} = \frac{0.95 \cdot 0.01}{0.1085} = 0.0876$$