Introduction to probability and statistics

Master in Cognitive Science 2025-2026

Lecture 1

October, 2025

Course content

- Basic concepts of probability theory.
- Probability distributions.
- Elementary ***Bayesian inference*** using Rjags.
- Introduction to statistical inference (***frequentist paradigm***): Estimators; hypothesis testing.
- Testing the equality of two means of normal populations.

References

The main references are

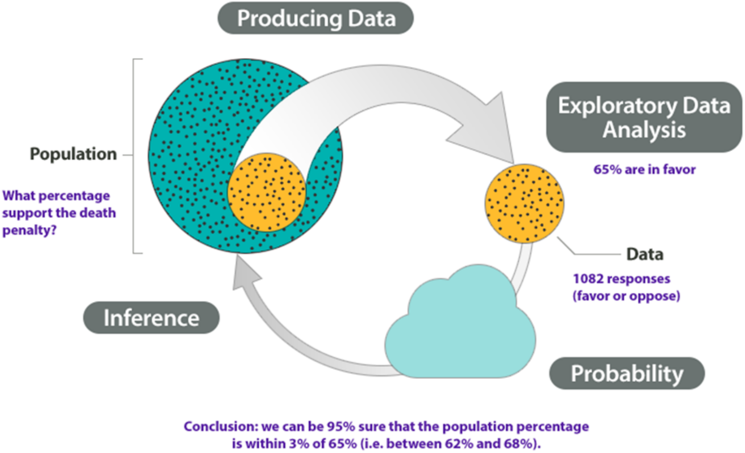
- Kass et al. (2014). Analysis of Neural Data. Springer Series in Statistics.
- Lee et al. (2014). Bayesian cognitive modeling, a practical course. Cambridge university press.
- Bergh et .al. (2021). A tutorial on Bayesian multi-model linear regression with BAS and JASP. Springer.

Guideline

Structure of the course:

- 9h lectures; 12h exercises; 18h practicals on R or Python (Houssam Boukhecham, houssam.boukhecham@univ-amu.fr)
- Evaluation will be partly based on practical reports. Each group at the end of practical sessions will post its report on Ametice. There will be an individual final exam.

Statistics


Producing Data
Population
What percentage support the death penalty?
Exploratory Data Analysis
65% are in favor
Data
1082 responses (favor or oppose)
Inference
Probability
Conclusion: we can be 95% sure that the population percentage is within 3% of 65% (i.e. between 62% and 68%).

On the meanings of probability

- Empirical or frequentist.

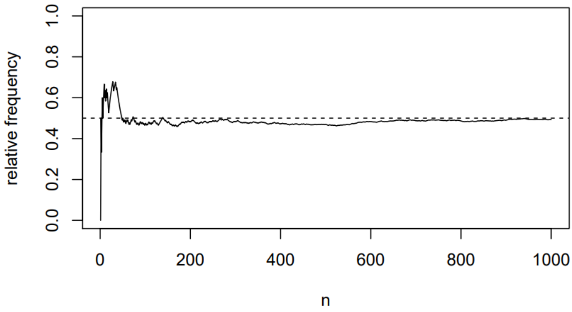


Figure 1: Relative frequency of heads in a fair coin tossing

- Subjective:
  - They are your probabilities, and express **Your relationship** to the event, e.g stakeholders will have different information and different probabilities.
  - All probabilities are conditional on a context H.
  - Is probability Chance or Ignorance? e.g probability that is will rain tomorrow ? What is the height of Nigara Falls.
- ... (For further reading see "Dawid Philip, Probability and Proof".)

Probability space

It starts with a triplet $(\Omega, \mathcal{F}, P)$:

- Ω **sample space**: it is the collection of all possible outcomes of a random experiment. We denote as ω the elementary outcome.
- $\mathcal{F}$ **events space**: a set whose elements $A \in \mathcal{F}$ (called events) are subsets of Ω.
- P **the probability measure**: it is a function from $\mathcal{F}$ into $[0, 1]$ that assigns to any event $A \in \mathcal{F}$, $P(A)$ the probability of the event A.

Examples

1. One trial coin tossing (Head or Tail). $\Omega = \{H, T\}$.
   There are two possible events $A = \{H\}$ and $B = \{T\}$.
2. Tossing a 6-headed dice. $\Omega = \{1, 2, \cdots, 6\}$.
3. Tossing two 6-headed dices. $\Omega = \{(1,1), (1,2), \cdots, (6,6)\}$.
   In this case, the number of possible outcomes is $\#\Omega =??$

Symbol set theory

Consider two events A and B. We can combine these two events using different logical operations. We denote as:

- Ω all possible outcomes of an experiment $\setminus$ certain event.
- $\emptyset$ empty set $\setminus$ impossible event.
- A subset of Ω (write $A \subset \Omega$) $\setminus$ Event; if $\omega \in A$, Event A occurred.
- A^c **complement** of A $\setminus$ No elementary outcome of A occurred.

- $A \cup B$: **union** of two events A and B \ elementary outcome lies in A or B (or in both, A and B).
- $A \cap B$: **intersection** of two events A and B \ elementary outcome occurred that lies in A and B.
- $B \backslash A = B \cap A^c$: contains outcome that are in B but not in A.
- A and B are **disjoint** events, if and only if $A \cap B = \emptyset$.
- $A \subset B$. A is subset of B \ if event A occurs, then event B occurs too.

Probability measure

A probability measure P, is a real valued function defined on $\mathcal{F}$, which satisfies the follwing axioms:

- $P(\Omega) = 1$.
- $P(A) \geq 0, \forall A \in \mathcal{F}$.
- If $A_i, \ i = 1, 2, \ldots$ are mutually disjoints then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Proposition

- Complement rule: $P(A^c) = 1 - P(A)$.
- Empty set: $P(\emptyset) = 0$.
- Monotonicity: if $A \subset B$, then $P(A) \leq P(B)$.
- Addition rule: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Computing probability

- For a finite sample space $\Omega = \{\omega_1, \ldots, \omega_n\}$ with outcomes ω_i that are mutually disjoint and with $P(\omega_i) = p_i$, the probability of an event $A \subseteq \Omega$ is

$$P(A) = \sum_{\omega_i \in A} P(\omega_i).$$

- If Ω has n equally likely and mutually exclusive outcomes, and if $A \subseteq \Omega$ contains k such outcomes, then

$$P(A) = \frac{\text{number of favorable outcomes}}{\text{total number of outcomes}} = \frac{|A|}{|\Omega|} = \frac{k}{n}.$$

Example: Two coin tosses

Consider the experiment of tossing a fair coin twice.

- The sample space is

$$\Omega = \{HH, HT, TH, TT\},$$

  where $H =$ heads and $T =$ tails.

- Each outcome has probability

$$P(\omega) = \tfrac{1}{4}, \quad \omega \in \Omega.$$

- Let $A =$ "exactly one head". Then

$$A = \{HT, TH\}, \quad |A| = 2,$$

  so

$$P(A) = \frac{|A|}{|\Omega|} = \frac{2}{4} = \tfrac{1}{2}.$$

Basic rules of counting

Often, all elements of some set X can be produced by making several choices consecutively.

- **Sum rule**: Find how many choices there are for the first step. In each case, how many choices are there for the second step. Continue until done. Finally, add up the choices.
- **Product rule**: If the first step has n choices, and each leads to m choices in the seond step, then there are $n \times m$ possible results.

Examples

- How many four-digit PIN codes can you make of the ten digits ? (**repetition allowed**).
- How many four-digit PIN codes are possible if the repetition of the same number is not allowed ?
- Lotto: There are 49 numbers, $1, 2, \ldots, 49$, and you randomly pick 6 numbers; how many different lotto tickets can you form with and without repetition?

Summary

If you have n different elements and arrange them into k-element sequences, then

- n^k sequences if repetition is allowed.
- $n(n-1)\cdots(n-k+1)$ sequences if repetition is not allowed.
- n distinct elements can be ordered in $n! = n(n-1)\cdots 1$ ways.
- From an n-element set, the number of different k-element subset is

$$\frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1} = \frac{n!}{k!(n-k)!} = \binom{n}{k}.$$

Conditional probability

Given $A, B \subset \Omega$ and $P(B) > 0$, the conditional probability of A **given** B is defined as:

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

Interpret as :"how likely is an event given that another has happened".

Bayes Theorem

How $P(A|B)$ and $P(B|A)$ are related to each other ?

$$\begin{aligned} P(B|A) &= \frac{P(A \cap B)}{P(A)} \\ &= \frac{P(B)}{P(A)} \frac{P(A \cap B)}{P(B)} \\ &= \frac{P(B)}{P(A)} P(A|B). \end{aligned}$$

Example

Let A and B be the following events:

- A: the patient has the Covid.
- B: the test is positive.

Terminology

- The **sensitivity** of the test is $P(B|A)$.
- The **specificity** of the test is $P(B^c|A^c)$.
- The **prevalence** $P(A)$ is the probability of observing the Covid in the population.

We know that the test will be **positive in 85%** of cases when used on **patients with Covid**, while it will be **positive in 5%** of cases when patients **do not have Covid**. The prevalence of Covid in France is **10%**.

Compute the probability of **having Covid** knowing that the **test is positive**.

Further reading

Chapter 3.1 in Kass et al (2009).