

Introduction

Statistics:

- a **mathematical** science
- to **describe** what has happened and
- to assess what **may** happen in **the future**
- relies on the **observation** of natural phenomena in order to propose an interpretation, often through **probabilistic models**

Statistics:

- a **mathematical** science
- to **describe** what has happened and
- to assess what **may** happen in **the future**
- relies on the **observation** of natural phenomena in order to propose an interpretation, often through **probabilistic models**

Frequentist statistics:

- Neyman & Pearson
- **deterministic** view of the parameters
- **Maximum Likelihood Estimation**
- statistical **test theory** & **confidence interval**



Bayes' theorem

Reverend Thomas Bayes posthumous article in 1763



$$\Pr(A|E) = \frac{\Pr(E|A) \Pr(A)}{\Pr(E|A) \Pr(A) + \Pr(E|\bar{A}) \Pr(\bar{A})} = \frac{\Pr(E|A) \Pr(A)}{\Pr(E)}$$

(conditional probability formula: $\Pr(A|E) = \frac{\Pr(A \cap E)}{\Pr(E)}$)

Bayes' theorem

Reverend Thomas Bayes posthumous article in 1763



$$\Pr(A|E) = \frac{\Pr(E|A) \Pr(A)}{\Pr(E|A) \Pr(A) + \Pr(E|\bar{A}) \Pr(\bar{A})} = \frac{\Pr(E|A) \Pr(A)}{\Pr(E)}$$

(conditional probability formula: $\Pr(A|E) = \frac{\Pr(A \cap E)}{\Pr(E)}$)

In practice:

Last time you visited the doctor, you got **tested for a rare disease**. Unluckily, the result was positive. . .

Given the test result, what is the probability that I actually have this disease?

(Medical tests are, after all, not perfectly accurate.)

→ *Seeing Theory*, Brown University

Bayes theorem: exercise

In June 2022, about 0.33% of the French population was estimated to have COVID-19.

Rapid tests have the following statistical properties:

- if someone has COVID-19, its test will come out positive 71% of the time
- if someone does not have the disease, its test will come out negative 98% of the time

Given that someone got a positive result, what is his/her probability to truly have COVID-19 ?

Bayes theorem: exercise

In June 2022, about 0.33% of the French population was estimated to have COVID-19.

Rapid tests have the following statistical properties:

- if someone has COVID-19, its test will come out positive 71% of the time
- if someone does not have the disease, its test will come out negative 98% of the time

Given that someone got a positive result, what is his/her probability to truly have COVID-19 ?

$$\Pr(D=+) = 0.0033$$

$$\Pr(T=+|D=+) = 0.71$$

$$\Pr(T=-|D=-) = 0.98$$

Bayes theorem: exercise

In June 2022, about 0.33% of the French population was estimated to have COVID-19.

Rapid tests have the following statistical properties:

- if someone has COVID-19, its test will come out positive 71% of the time
- if someone does not have the disease, its test will come out negative 98% of the time

Given that someone got a positive result, what is his/her probability to truly have COVID-19 ?

$$\Pr(D=+) = 0.0033$$

$$\Pr(T=+|D=+) = 0.71$$

$$\Pr(T=-|D=-) = 0.98$$

$$\Pr(D=+|T=+) = ?$$