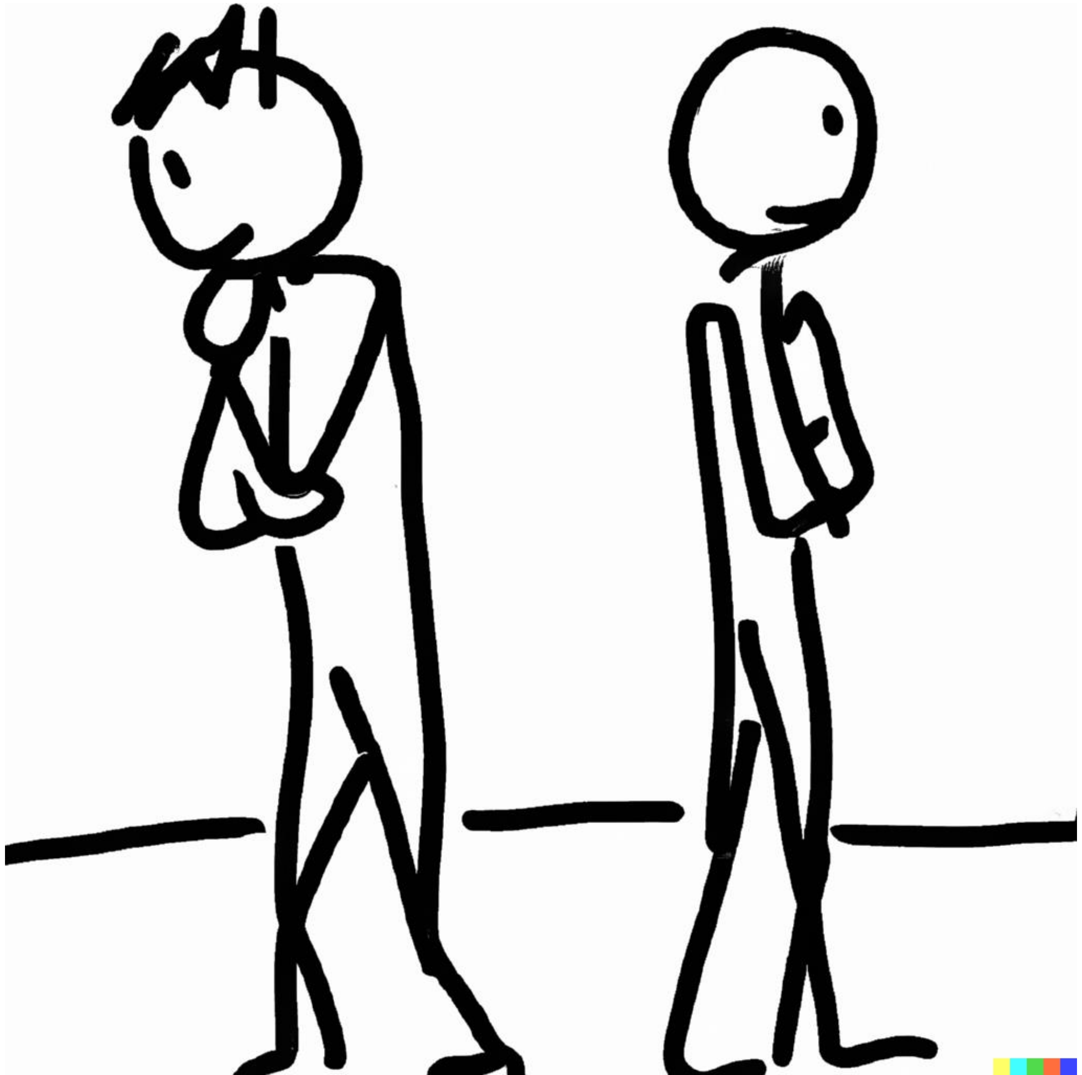
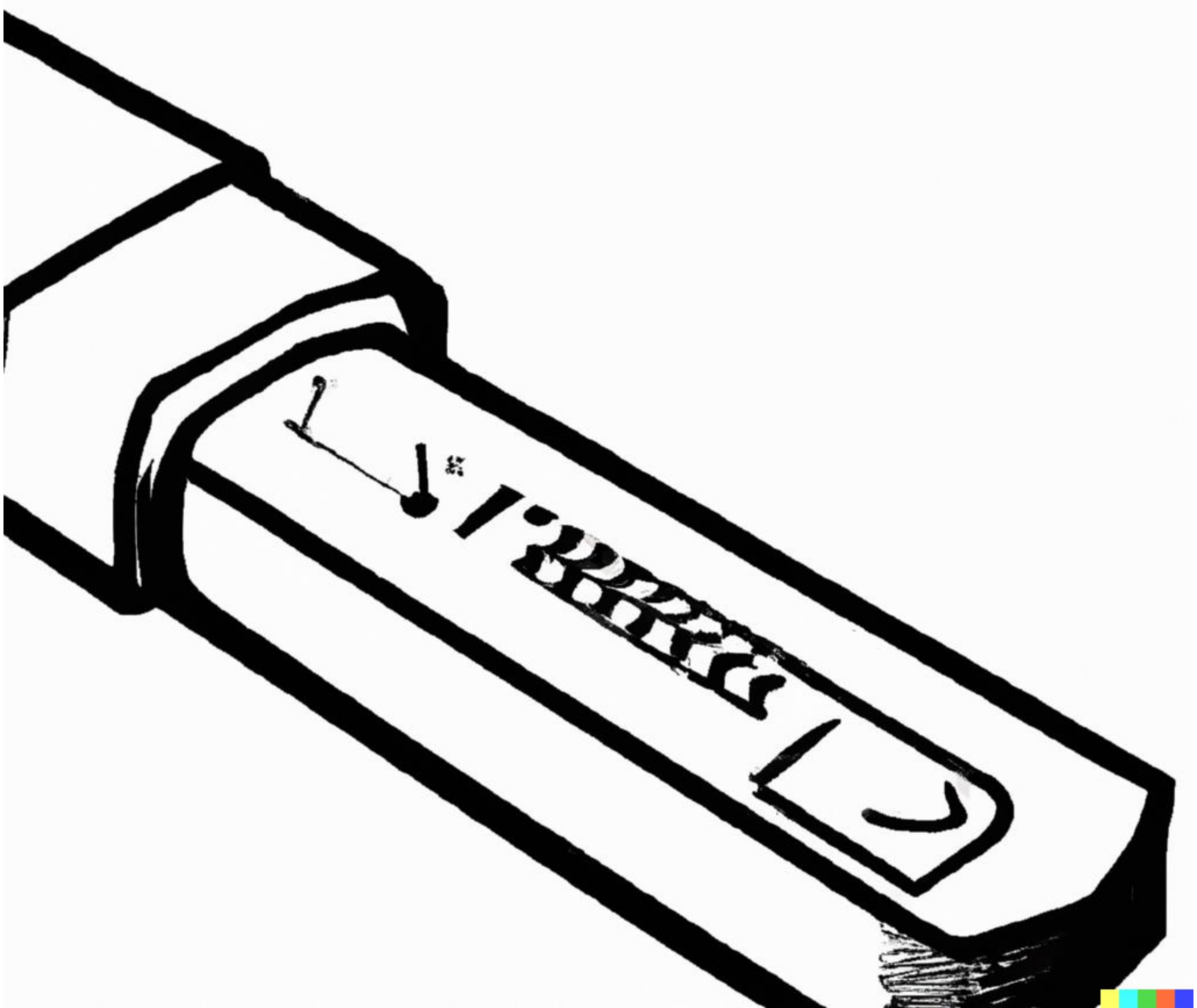
Base-rate neglect

Notes on Behavioural Economics





The cab problem

A cab was involved in a hit and run accident at night. Two cab companies, the Green and the Blue, operate in the city. Participants are given the following data:

1. 85% of the cabs in the city are Green, 15% are Blue.
2. A witness identified the cab as Blue. The court tested the reliability of the witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colours 80% of the time.

What is the probability that the cab involved in the accident was Blue rather than Green?

The cab problem

$$\underbrace{P(claim\ blue \mid is\ blue)}_{80\%} \neq \underbrace{P(is\ blue \mid claim\ blue)}_{\text{Requires Bayes' rule}}$$

The cab problem

$$P(\text{blue}|\text{claim blue}) = \frac{P(\text{claim blue}|\ \text{blue})P(\text{blue})}{P(\text{claim blue})}$$

The cab problem

$$P(\text{blue}|\text{claim blue}) = \frac{P(\text{claim blue}|\text{ blue})P(\text{blue})}{P(\text{claim blue})}$$

$$= \frac{P(\text{claim blue}|\text{ blue})P(\text{blue})}{P(\text{claim blue}|\text{blue})P(\text{blue}) + P(\text{claim blue}|\neg\text{ blue})P(\neg\text{ blue})}$$

The cab problem

$$P(\text{blue}|\text{claim blue}) = \frac{P(\text{claim blue}|\text{ blue})P(\text{blue})}{P(\text{claim blue})}$$

$$= \frac{P(\text{claim blue}|\text{ blue})P(\text{blue})}{P(\text{claim blue}|\text{blue})P(\text{blue}) + P(\text{claim blue}|\neg\text{ blue})P(\neg\text{ blue})}$$

$$= \frac{0.8 \times 0.15}{0.8 \times 0.15 + 0.2 \times 0.85}$$

The cab problem

$$P(\text{blue}|\text{claim blue}) = \frac{P(\text{claim blue}|\text{ blue})P(\text{blue})}{P(\text{claim blue})}$$

$$= \frac{P(\text{claim blue}|\text{ blue})P(\text{blue})}{P(\text{claim blue}|\text{blue})P(\text{blue}) + P(\text{claim blue}|\neg\text{ blue})P(\neg\text{ blue})}$$

$$= \frac{0.8\ \times\ 0.15}{0.8\ \times\ 0.15 + 0.2\ \times\ 0.85}$$

$$= 0.41$$

Medical diagnosis

You test yourself for COVID-19. The following information is known:

- The probability that a person has COVID-19 is 1% (the prevalence).
- If a person has COVID-19, the probability that they test positive is 90% (the sensitivity).
- If a person does not have COVID-19, the probability that they nevertheless test positive is 9% (the false positive rate).

You test positive. What is the chance that you have COVID-19?

Medical diagnosis

$$P(\mathrm{COVID}|{+\mathrm{ve}}) \neq P({+\mathrm{ve}}|\mathrm{COVID})$$

Medical diagnosis

$$P(\mathrm{COVID}|\text{+ve}) = \frac{P(\text{+ve}|\mathrm{COVID})P(\mathrm{COVID})}{P(\text{+ve})}$$

Medical diagnosis

$$P(\mathrm{COVID}|\text{+ve}) = \frac{P(\text{+ve}|\mathrm{COVID})P(\mathrm{COVID})}{P(\text{+ve})}$$

$$= \frac{P(\text{+ve}|\mathrm{COVID})P(\mathrm{COVID})}{P(\text{+ve}|\mathrm{COVID})P(\mathrm{COVID}) + P(\text{+ve}|\neg\,\mathrm{COVID})P(\neg\mathrm{COVID})}$$

Medical diagnosis

$$P(\mathrm{COVID}|\text{+ve}) = \frac{P(\text{+ve}\,|\mathrm{COVID})P(\mathrm{COVID})}{P(\text{+ve})}$$

$$= \frac{P(\text{+ve}\,|\mathrm{COVID})P(\mathrm{COVID})}{P(\text{+ve}\,|\mathrm{COVID})P(\mathrm{COVID}) + P(\text{+ve}\,|\neg\,\mathrm{COVID})P(\neg\mathrm{COVID})}$$

$$= \frac{0.9\ \times\ 0.01}{0.9\ \times\ 0.01 + 0.09\ \times\ 0.99}$$

Medical diagnosis

$$P(\mathrm{COVID}|\text{+ve}) = \frac{P(\text{+ve}\,|\mathrm{COVID})P(\mathrm{COVID})}{P(\text{+ve})}$$

$$= \frac{P(\text{+ve}\,|\mathrm{COVID})P(\mathrm{COVID})}{P(\text{+ve}\,|\mathrm{COVID})P(\mathrm{COVID}) + P(\text{+ve}\,|\neg\,\mathrm{COVID})P(\neg\mathrm{COVID})}$$

$$= \frac{0.9\ \times\ 0.01}{0.9\ \times\ 0.01 + 0.09\ \times\ 0.99}$$

$$= 0.092$$

Natural frequencies

You test yourself for COVID-19. The following information is known:

- Ten in every 1000 people have COVID-19 (the prevalence).
- Of these 10 people with COVID-19, nine will test positive (the sensitivity).
- Of the 990 people without COVID-19, about 89 nevertheless test positive (the false positive rate).

You test positive. What is the chance that you have COVID-19?

Natural frequencies

$$\hat{P}(\text{ COVID}|\text{+ve}) = \frac{n(\text{+ve} \cap \text{COVID})}{n(\text{+ve})}$$

Natural frequencies

$$\hat{P}(\,\mathrm{COVID}|\text{+ve}) = \frac{n(\text{+ve} \cap \mathrm{COVID})}{n(+\mathrm{ve})}$$

$$= \frac{9}{9 + 89}$$

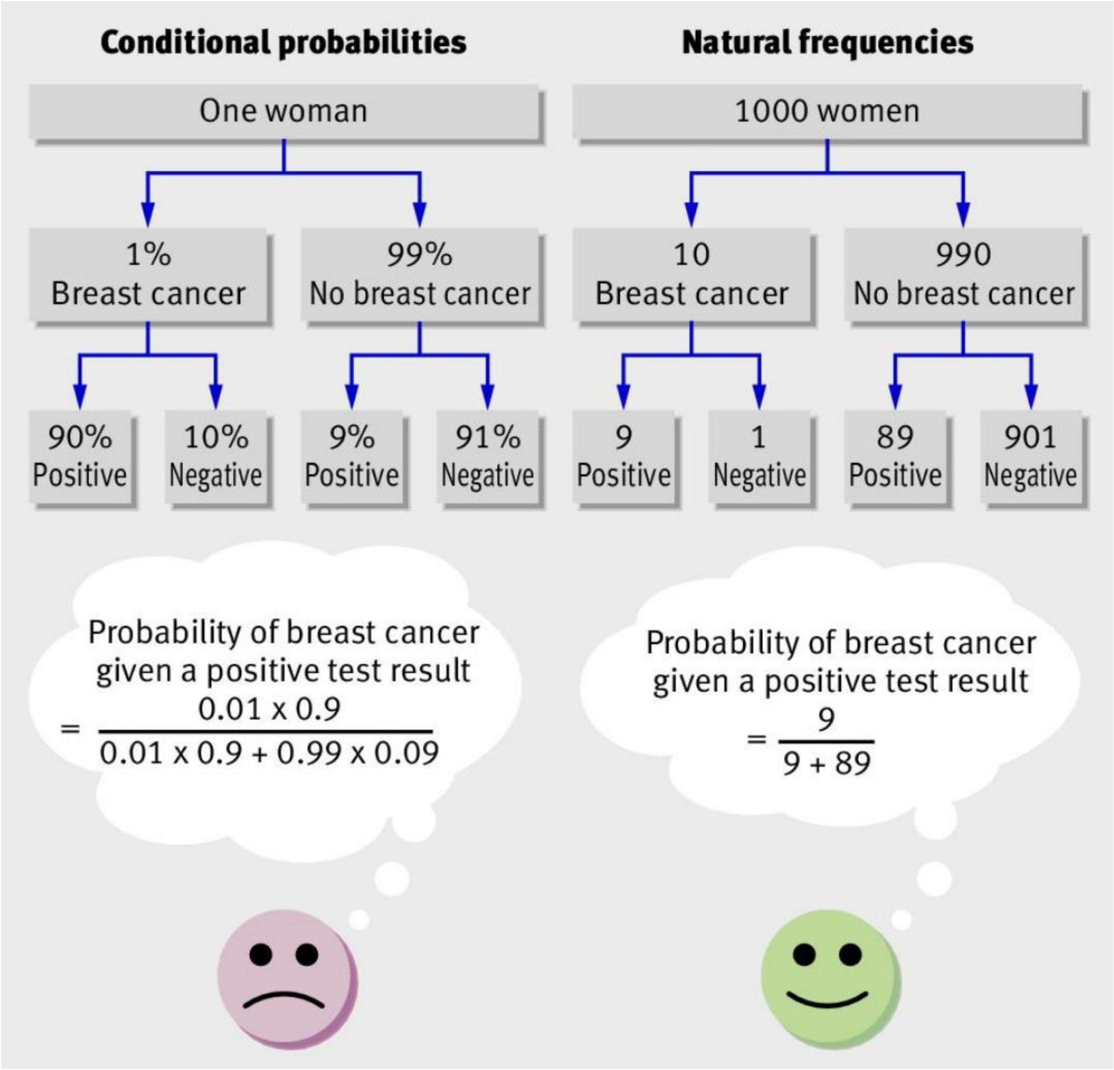
Natural frequencies

$$\hat{P}(\text{ COVID}|\text{+ve}) = \frac{n(\text{+ve} \cap \text{COVID})}{n(\text{+ve})}$$

$$= \frac{9}{9 + 89}$$

$$= 0.092$$

Natural frequencies



Identifying a finch

You know the following about the finches in your area:

- 99% of the finches are Wallace finches. The remaining 1% are Darwin finches.
- If you spot a Darwin finch, you will correctly identify that it is a Darwin finch 95% of the time. The other 5% of the time you identify it as a Wallace finch.
- If you spot a Wallace finch, you will correctly identify that it is a Wallace finch 95% of the time. The other 5% of the time you identify it as a Darwin finch.

You spot a finch and identify it as a Darwin finch.

Identifying a finch

$$P(\mathrm{D}|\mathrm{I}) = \frac{P(\mathrm{I}|\mathrm{D})P(\mathrm{D})}{P(\mathrm{I})}$$

Identifying a finch

$$P(\mathrm{D}|\mathrm{I}) = \frac{P(\mathrm{I}|\mathrm{D})P(\mathrm{D})}{P(\mathrm{I})}$$

$$= \frac{P(\mathrm{I}|\mathrm{D})P(\mathrm{D})}{P(\mathrm{I}|\mathrm{D})P(\mathrm{D}) + P(\mathrm{I}|\neg\,\mathrm{D})P(\neg\mathrm{D})}$$

Identifying a finch

$$P(\mathrm{D}|\mathrm{I}) = \frac{P(\mathrm{I}|\mathrm{D})P(\mathrm{D})}{P(\mathrm{I})}$$

$$= \frac{P(\mathrm{I}|\mathrm{D})P(\mathrm{D})}{P(\mathrm{I}|\mathrm{D})P(\mathrm{D}) + P(\mathrm{I}|\neg\mathrm{D})P(\neg\mathrm{D})}$$

$$= \frac{0.95 \times 0.01}{0.95 \times 0.01 + 0.05 \times 0.99}$$

Identifying a finch

$$P(\mathrm{D}|\mathrm{I}\ ) = \frac{P(\mathrm{I}|\mathrm{D})P(\mathrm{D})}{P(\mathrm{I})}$$

$$= \frac{P(\mathrm{I}|\mathrm{D})P(\mathrm{D})}{P(\mathrm{I}|\mathrm{D})P(\mathrm{D}) + P(\mathrm{I}|\neg\ \mathrm{D})P(\neg\mathrm{D})}$$

$$= \frac{0.95\ \times\ 0.01}{0.95\ \times\ 0.01 + 0.05\ \times\ 0.99}$$

$$= 0.16$$

Identifying a finch

- Suppose there are 10,000 finches.
- That would mean there are 100 Darwin finches and 9,900 Wallace finches.
- If I spotted these 100 Darwin finches, I would identify 95 as Darwin finches.
- If I spotted a Wallace Finch, I would identify $0.05 \times 9900 = 495$ as Darwin Finches.
- That means 95 of the $95 + 495 = 590$ birds I identify as Darwin finches would be Darwin finches.

Identifying a finch

$$P(\mathrm{D}|\mathrm{I}) = \frac{95}{590}$$

Identifying a finch

$$P(\mathrm{D}|\mathrm{I}) = \frac{95}{590}$$

$$= 0.16$$