Digital Skills

Christos Dimitrakakis

November 17, 2022

Outline

Bayesian analysis

- Recall the definition of Conditional probability:

$$P(A|B) = P(A \cap B)/P(B)$$

i.e. the probability of A happening if B happens.

Bayesian analysis

- Recall the definition of Conditional probability:

$$P(A|B) = P(A \cap B)/P(B)$$

  i.e. the probability of A happening if B happens.
- It is also true that:

$$P(B|A) = P(A \cap B)/P(A)$$

Bayesian analysis

- Recall the definition of Conditional probability:

$$P(A|B) = P(A \cap B)/P(B)$$

  i.e. the probability of A happening if B happens.

- It is also true that:

$$P(B|A) = P(A \cap B)/P(A)$$

- Combining the two equations, reverse the conditioning:

$$P(A|B) = P(B|A)P(A)/P(B)$$

Bayesian analysis

- Recall the definition of Conditional probability:

$$P(A|B) = P(A \cap B)/P(B)$$

  i.e. the probability of A happening if B happens.
- It is also true that:

$$P(B|A) = P(A \cap B)/P(A)$$

- Combining the two equations, reverse the conditioning:

$$P(A|B) = P(B|A)P(A)/P(B)$$

- So we can reverse the order of conditioning, i.e. relate to the probability of A given B to that of B given A.

The murder problem

- Somebody saw somebody matching their description and he was found in the neighbourghood. There is no other evidence.

Prior elicitation

- All those that think the accused is guilty, raise their hand.
- Divide by the number of people in class
- Let us call this $P(H_1)$.
- This is a purely subjective measure!

The murder problem

- Somebody saw somebody matching their description and he was found in the neighbourghood. There is no other evidence.
- There are two possibilities:

What is your belief that they have committed the crime?

Prior elicitation

- All those that think the accused is guilty, raise their hand.
- Divide by the number of people in class
- Let us call this $P(H_1)$.
- This is a purely subjective measure!

The murder problem

- Somebody saw somebody matching their description and he was found in the neighbourghood. There is no other evidence.
- There are two possibilities:
  - H_0: They are innocent.

What is your belief that they have committed the crime?

Prior elicitation

- All those that think the accused is guilty, raise their hand.
- Divide by the number of people in class
- Let us call this $P(H_1)$.
- This is a purely subjective measure!

The murder problem

- Somebody saw somebody matching their description and he was found in the neighbourghood. There is no other evidence.
- There are two possibilities:
  - H_0: They are innocent.
  - H_1: They are guilty.

  What is your belief that they have committed the crime?

Prior elicitation

- All those that think the accused is guilty, raise their hand.
- Divide by the number of people in class
- Let us call this $P(H_1)$.
- This is a purely subjective measure!

DNA test

- Let us now do a DNA test on the suspect

DNA test properties

- D: Test is positive

Run the test

DNA test

- Let us now do a DNA test on the suspect

DNA test properties

- D: Test is positive
- $P(D|H_0) = 1\%$: False positive rate

Run the test

DNA test

- Let us now do a DNA test on the suspect

DNA test properties

- D: Test is positive
- $P(D|H_0) = 1\%$: False positive rate
- $P(D|H_1) = 100\%$: True positive rate

Run the test

DNA test

- Let us now do a DNA test on the suspect

DNA test properties

- D: Test is positive
- $P(D|H_0) = 1\%$: False positive rate
- $P(D|H_1) = 100\%$: True positive rate

Run the test

- The result is either positive or negative ($\neg D$).

DNA test

- Let us now do a DNA test on the suspect

DNA test properties

- D: Test is positive
- $P(D|H_0) = 1\%$: False positive rate
- $P(D|H_1) = 100\%$: True positive rate

Run the test

- The result is either positive or negative ($\neg D$).
- What is your belief now that the suspect is guilty?

Everybody is a suspect

- Run a DNA test on everybody.

Everybody is a suspect

- Run a DNA test on everybody.
- What is different from before?

Everybody is a suspect

- Run a DNA test on everybody.
- What is different from before?
- Who has a positive test?

Everybody is a suspect

- Run a DNA test on everybody.
- What is different from before?
- Who has a positive test?
- What is your belief that the people with the positive test are guilty?

Explanation

- Probability of innocence:
  $P(H_0|D) = P(D \cap H_0)/P(D) = P(D|H_0)P(H_0)/P(D)$

Explanation

- Probability of innocence:
  $P(H_0|D) = P(D \cap H_0)/P(D) = P(D|H_0)P(H_0)/P(D)$
- Marginal probability:
  $P(D) = P(D|H_0)P(H_0) + P(D|H_1)P(H_1)$

Explanation

- Probability of innocence:
  $P(H_0|D) = P(D \cap H_0)/P(D) = P(D|H_0)P(H_0)/P(D)$
- Marginal probability:
  $P(D) = P(D|H_0)P(H_0) + P(D|H_1)P(H_1)$
- Bayes's theorem: $P(H_0|D) = \frac{P(D|H_0)P(H_0)}{P(D|H_0)P(H_0)+P(D|H_1)P(H_1)}$

Explanation

- Probability of innocence: $P(H_0|D) = P(D \cap H_0)/P(D) = P(D|H_0)P(H_0)/P(D)$
- Marginal probability: $P(D) = P(D|H_0)P(H_0) + P(D|H_1)P(H_1)$
- Bayes's theorem: $P(H_0|D) = \frac{P(D|H_0)P(H_0)}{P(D|H_0)P(H_0)+P(D|H_1)P(H_1)}$
- Assuming $P(D|H_1) = 1$, and setting $P(H_0) = q$, this gives

$$P(H_0|D) = \frac{0.1q}{0.1q + 1 - q} = \frac{q}{10 - 9q}$$

Python example

Let us assume $P(D|H_0) = 10^{-6}$.

```
import numpy as np
base_rate = 0.000001
population_size = 10000000
n_positive_tests = np.random.binomial(population_size, base_
return n_positive_tests
```

The covid test problem

10% of the class has covid, i.e. P(covid) = 0.1. Each one of you performs a covid test. If you have covid, the test is correct 80% of the time, i.e. P(positive | covid) = 0.8. Conversely, if you do not have covid, there is still a 10% chance of a positive test, with P(positive | not-covid) = 0.1

How likely is it that you have covid if your test is positive or negative, i.e. P(covid | positive), vs. P(covid | negative)?

First of all, each one of you should independently generate a uniform random number between 1 and 10. For that, you can each throw a die, and record the outcome.

Then you throw a second die, and record that as well.

I will now pass over the tables and tell each one of you if they have a positive test.

Now, everybody with a positive test raises their hand. I expect it to be slightly more than 10% (but it depends).

The cards problem

1. Print out a number of cards, with either [A|A], [A|B] or [B|B] on their sides.
2. Get a card (say with face A), and ask what is the probability the other side is the same.
3. Have the students perform the experiment with:
   3.1 Draw a random card.
   3.2 Count the number of people with A.
   3.3 What is the probability that somebody with an A on one side will have an A on the other?
   3.4 Half of the people should have an A?

The prior probabilities

A	A	2/6
A	B	1/6
B	A	1/6
B	B	2/6

The posterior probabilities if you have an A