

Submission for Deep Learning Exercise 1

Team: dl2024-sigma_learning

Students: Uralp Ergin (5975013), İsmail Karabaş (7654321), Ahmet Yaşar Ayfer (5986167)

October 18, 2024

1 Bertrand's Box Paradox Solution

1.1 Question 1

Probabilities of drawing a black or white side

Let's define our sample space and events:

- B : Event of drawing a black side
- W : Event of drawing a white side

There are 6 sides in total: 3 black and 3 white.

Probability of drawing a black side: $P(B) = \boxed{\frac{1}{2}}$

Probability of drawing a white side: $P(W) = \boxed{\frac{1}{2}}$

We can also calculate the probability by considering each card type:

- Black-Black card: $P(B) = \frac{1}{3} \cdot 1 = \frac{1}{3}$
- White-White card: $P(B) = \frac{1}{3} \cdot 0 = 0$
- Black-White card: $P(B) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$

Total probability: $P(B) = \frac{1}{3} + 0 + \frac{1}{6} = \frac{1}{2}$

This confirms our initial result: $P(B) = P(W) = \frac{1}{2}$

1.2 Question 2

Probability that the other side is black, given that we see a black side

Let's define the events:

- B_1 : Event of seeing a black side
- B_2 : Event that the other side is also black

We need to calculate $P(B_2|B_1)$

Using Bayes' theorem:

$$P(B_2|B_1) = \frac{P(B_1|B_2)P(B_2)}{P(B_1)}$$

- $P(B_1|B_2) = 1$ (if both sides are black, we'll definitely see a black side)
- $P(B_2) = \frac{1}{3}$ (probability of choosing the all-black card)
- $P(B_1) = \frac{1}{2}$ (from question 1)

$$P(B_2|B_1) = \frac{1 \cdot \frac{1}{3}}{\frac{1}{2}} = \boxed{\frac{2}{3}}$$

We can also calculate this probability by considering each possible black side:

Let's label the sides of the cards as follows:

- B1-B2 (all-black card)
- B3-W1 (mixed card)
- W2-W3 (all-white card)

When we choose a black side, there are three equally likely possibilities:

- B1: $P(\text{other side black} \mid B1) = 1$
- B2: $P(\text{other side black} \mid B2) = 1$
- B3: $P(\text{other side black} \mid B3) = 0$

The total probability is thus:

$$P(\text{other side black} \mid \text{black side}) = \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 0 = \frac{2}{3}$$

This confirms our result from Bayes' theorem.

1.3 Question 3

Find the probability that the other side of the card is black if the card shows a white side.

Let's define the events:

- W_1 : Event of seeing a white side
- B_2 : Event that the other side is black

We need to calculate $P(B_2|W_1)$

Using Bayes' theorem:

$$P(B_2|W_1) = \frac{P(W_1|B_2)P(B_2)}{P(W_1)}$$

- $P(W_1|B_2) = 1$ (if one side is black and we see white, the other side must be black)
- $P(B_2) = \frac{1}{6}$ (probability of choosing the black-white card and having the black side down)
- $P(W_1) = \frac{1}{2}$ (from question 1)

Substituting:

$$P(B_2|W_1) = \frac{1 \cdot \frac{1}{6}}{\frac{1}{2}} = \boxed{\frac{1}{3}}$$

We can also calculate this probability by considering each possible black side:

Let's label the sides of the cards as follows:

- B1-B2 (all-black card)
- B3-W1 (mixed card)
- W2-W3 (all-white card)

When we choose a black side, there are three equally likely possibilities:

- W1: $P(\text{other side black} \mid \text{W1}) = 1$
- W2: $P(\text{other side black} \mid \text{W2}) = 0$
- W3: $P(\text{other side black} \mid \text{W3}) = 0$

The total probability is thus:

$$P(\text{other side black} \mid \text{White side}) = \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 0 + \frac{1}{3} \cdot 0 = \frac{1}{3}$$

This confirms our result from Bayes' theorem.

2 Question 7.2

Bayesian Linear Regression (BLR), How do the prior and posterior distributions differ? Why?

Differences Between the Prior and the Posterior Distributions:

Prior Distribution represents our initial estimations about the parameters before observing any data. It is defined by a mean (μ_{pre}) and covariance (Σ_{pre}), often with large variance reflecting uncertainty.

Posterior Distribution represents our updated estimations after observing data. The posterior mean (μ_{post}) moves towards values that better fit the data, while the posterior covariance (Σ_{post}) becomes smaller, indicating reduced uncertainty.

The key difference is the role of data. The prior reflects uncertainty before seeing data, while the posterior refines that estimation using observed data, reducing uncertainty and providing more accurate estimates. Bayesian inference updates our understanding by balancing prior knowledge with new information.