CS577 - Machine Learning – Assignment 1

Alexandros Angelakis csdp1363, angelakis@csd.uoc.gr

Exercise 1 - Probabilities (Theoretical)

a.

Let's define the random variable X = 'Gender of kid'. Since the the probability of being boy or girl is equally likely, X will the following probability mass function (PMF):

$$p_X(x) = \begin{cases} \frac{1}{2} & , x = \text{'boy'} \\ \frac{1}{2} & , x = \text{'girl'} \end{cases}$$

Now, we know that each couple in our population has exactly two kids. The probability of each couple having one daughter and one son can be calculated in two ways:
1. The first child is a boy and the second is a girl
2. The first child is a girl and the second is a boy
Thus, the probability of having one daughter and one son is:

$$P(\text{1Couple1Daugher1Son}) = P(\text{'boy'}) \times P(\text{'girl'}) + P(\text{'girl'}) \times P(\text{'boy'}) = \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2} = 0.5$$

We have found that the probability of one couple having one daughter and one son is $\frac{1}{2}$
For n couples, since the couples are independent, the probability of each couple having one daughter and one son can be calculated by multiplying the probability of one couple having one daughter and one son n times. This gives us:

$$P(\text{nCouples1Daughter1Son}) = \prod_{i=1}^{n} \frac{1}{2} = \frac{1}{2} \times \frac{1}{2} \times \cdots \times \frac{1}{2} = (\frac{1}{2})^n = (0.5)^n, \text{ where } n \in \mathbb{N}$$

b.

Let's define the random variable X = 'Result of a coin flip'. We know that the coin is fair, so X will have the following probability mass function (PMF):

$$p_X(x) = \begin{cases} \frac{1}{2} & , x = \text{'Heads'} \\ \frac{1}{2} & , x = \text{'Tails'} \end{cases}$$

For a single flip, the probability of getting 'Heads' is $P(\text{'Heads'}) = \frac{1}{2}$.
For n flips, since each coin flip is independent, the probability of getting 'Heads' in all n flips can be calculated by multiplying the probability of getting 'Heads' in a single flip n times. This gives us:

$$P(\text{'All Heads'}) = \prod_{i=1}^{n} \frac{1}{2} = \frac{1}{2} \times \frac{1}{2} \times \cdots \times \frac{1}{2} = (\frac{1}{2})^n = (0.5)^n, \text{ where } n \in \mathbb{N}$$

c.

We know that in a box, there are beads of three colours. The probabilities of drawing each colour is given as:

$$P(R) = 0.3, \quad P(G) = 0.5, \quad P(B) = 0.2$$

We are also given that half of the red beads, two-thirds of the blue beads and two-thirds of the green beads are hollow. In terms of probabilities, this means that:

$$P(H|R) = \frac{1}{2}, \quad P(H|G) = \frac{2}{3}, \quad P(H|B) = \frac{2}{3}$$

We want to find the probability of drawing a hollow bead by picking a random bead from the box, ie., the probability $P(H)$. Since there are three possible colours, our sample space is divided into three sets Using the law of total probability, we can calculate $P(H)$ as:

$$P(H) = P(H|R) \times P(R) + P(H|G) \times P(G) + P(H|B) \times P(B) = \frac{1}{2} \times 0.3 + \frac{2}{3} \times 0.5 + \frac{2}{3} \times 0.2 \approx 0.616$$

Thus, the probability of drawing a hollow bead from the box is approximately 0.616.

Exercise 2 - Bayesian Theorem (Programming)

a.

A photon package emmited by a star has a probability of 1×10^{-7} to pass through the Earth's atmoshpere and reach a detector placed on the ground: $P(\text{'received'}) = 1 \times 10^{-7}$
We are also given that:
1. False Positive (FP) rate of 10%, meaning: $P(\text{'detection'}|\text{'not received'}) = 0.1$
2. True Positive (TP) rate of 85%, meaning: $P(\text{'detection'}|\text{'received'}) = 0.85$

We want to find the probability that a photon package was actually received, given that the detector reported a detection, i.e., $P(\text{'received'}|\text{'detection'})$. Using Bayes' Theorem, we can calculate this probability as:

$$P(\text{'received'}|\text{'detection'}) = \frac{P(\text{'detection'}|\text{'received'}) \times P(\text{'received'})}{P(\text{'detection'})}$$

We know all the probabilities except $P(\text{'detection'})$, which can be calculated using the law of total probability. This gives:

$$\begin{aligned} P(\text{'detection'}) &= P(\text{'detection'}|\text{'not received'}) \times P(\text{'not received'}) + P(\text{'detection'}|\text{'received'}) \times P(\text{'received'}) \\ &= 0.1 \times (1 - 1 \times 10^{-7}) + 0.85 \times (1 \times 10^{-7}) \approx 8.5 \times 10^{-8} + 0.1 = 0.100000085 \end{aligned}$$

Substituting the values of the probabilities, we have:

$$P(\text{'received'}|\text{'detection'}) = \frac{0.85 \times 10^{-7}}{0.100000085} = \frac{0.000000085}{0.100000085} \approx 8.5 \times 10^{-7}$$

b.

Assuming we have a photon package composed of 100 photons and each photon has an equal probability of carrying an energy of either 10, 20, 30 or 40 electron-volts (uniform distribution), we define the random variable X as the energy of a photon in a package composed of 100 photons. Since X is uniformly distributed, the probability mass function of X is:

$$p_X(x) = \begin{cases} 0.25 & , x = 10 \\ 0.25 & , x = 20 \\ 0.25 & , x = 30 \\ 0.25 & , x = 40 \end{cases}$$

We can compute the expected value:

$$E[X] = \sum_x x p_X(x) = 0.25 \times 10 + 0.25 \times 20 + 0.25 \times 30 + 0.25 \times 40 = 2.5 + 5 + 7.5 + 10 = 25$$

Next, we calculate the second moment (for variance):

$$E[X^2] = \sum_x x^2 p_X(x) = 0.25 \times 10^2 + 0.25 \times 20^2 + 0.25 \times 30^2 + 0.25 \times 40^2 = 25 + 100 + 225 + 400 = 750$$

Thus, the variance is:

$$Var(X) = E[X^2] - (E[X])^2 = 750 - 625 = 125$$

The total energy of a package (the sum of the energies of the photons composing a package), i.e., random variable Y, will follow a Normal (Gaussian) distribution with $\mu = E[X] \times 100$ and $\sigma^2 = Var(X) \times 100$, according on the Central Limit

Theorem, so $Y \sim \mathcal{N}(2500, 12500)$.
The Central Limit Theorem states that when we have a large number of independent and identically distributed (i.i.d) random variables, the sum of these variables will approximately follow a normal distribution, regardless of the original distribution of the variables.
In our case, the i.i.d random variables are Y_1, Y_2, $\ldots$, Y_n. The sum of these random variables

$$S_n = Y_1 + Y_2 + \cdots + Y_n, \text{ where } n \in \mathbb{N}$$

will follow a Normal Distribution as $n \to \infty$
This can be observed in Figure 1, where, as n increases, the sum of the random variables tends to follow a normal distribution.

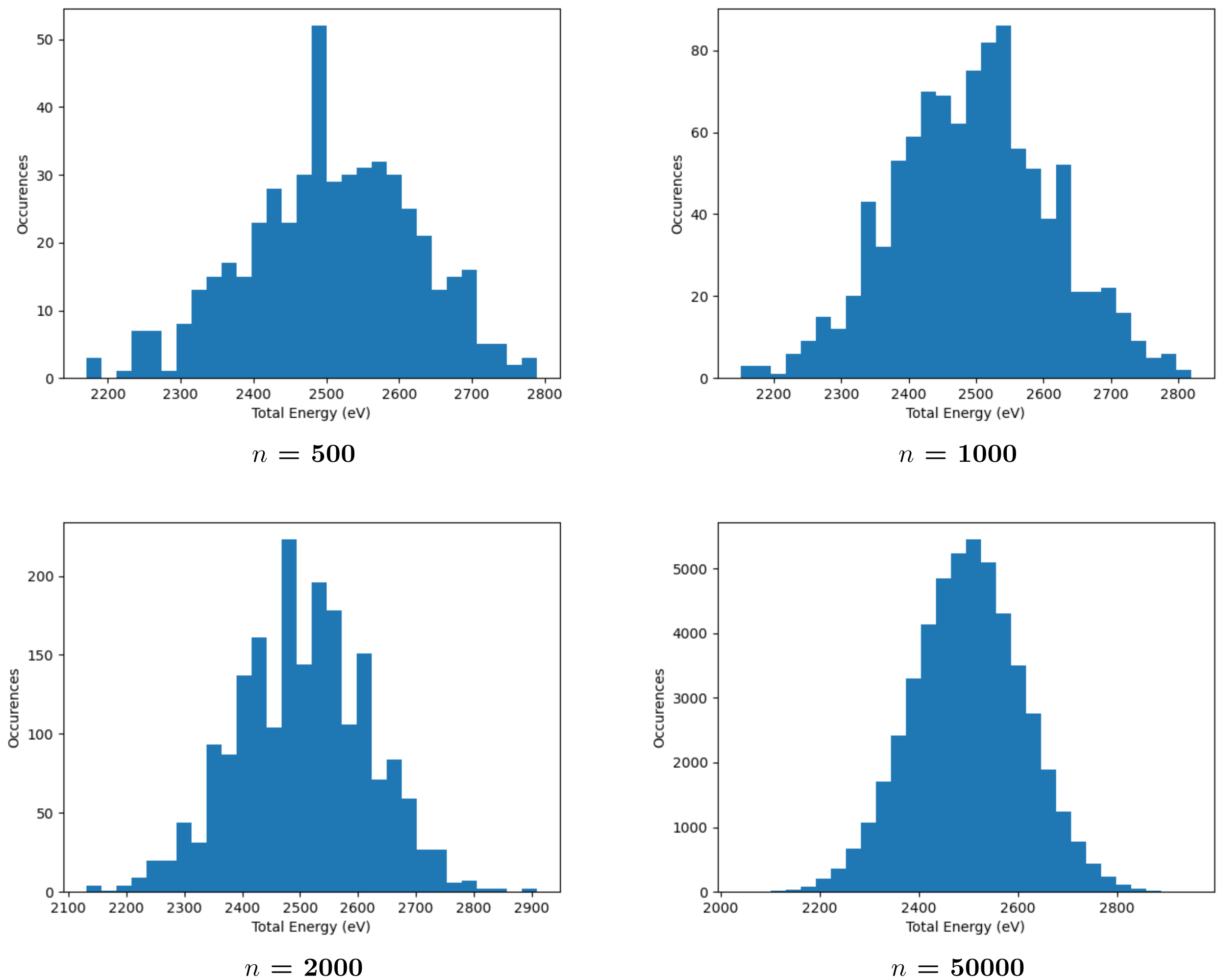

Figure 1: The histograms of the total photon packages energies for increasing n

c.

We now assume that the probability of a photon package reaching the ground follows a normal distribution $\mathcal{N}(\mu, \sigma)$, centered at $\mu = 1e-7$ and $\sigma = 9e-8$. As a result, the posterior probability is not a single value but a distriubtion, as shown in Figure 2

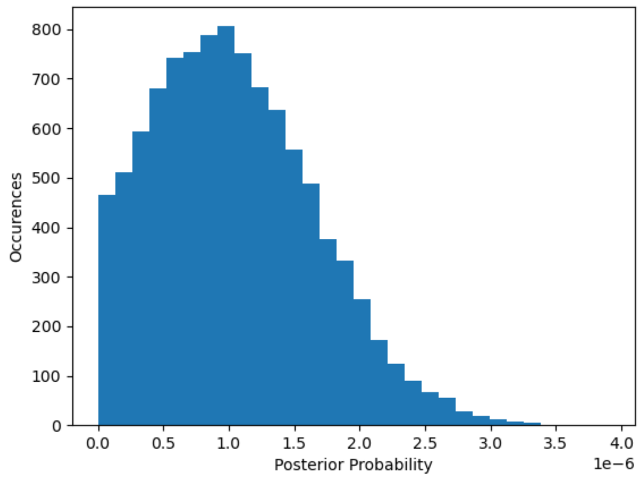

Figure 2: The histogram of the posterior probabilities with priors drawn from a $\mathcal{N}(1e-7, (9e-8)^2)$

The plot makes sense because samples drawn from a Normal distribution are also normally distributed. Since our initial μ is very close to zero, many of samples that are negative. We reject these negative values because probabilities cannot be negative, which is why the left tail of the normal distribution is trimmed.