Problem 1

The following plots represent densities on the real axis.

3 p. 1.1. Mark on the graphs the median of each distribution.

6 p. 1.2. Label the marked points on the y axis with the correct values.

4 p. 1.3. Mark on the graphs the mean of each distribution. (For c, it is sufficient to mark the approximate position w.r.t the ticks)

5 p. 1.4. What is the value of \( P(x > 1) \)?

---

**solution**

The median and mean occur at \( x = 0 \).
The marked point on the y axis is \( \frac{1}{4} \).
\[ P(x > 1) = \frac{1}{4} \]
The median and mean occur at $x = 0$.
The marked point on the y axis is $\frac{1}{2}$.
When $x = 1, f(x) = \frac{1}{4}$. Therefore, $P(x > 1) = \frac{1}{4}(2 - 1)(\frac{1}{4}) = \frac{1}{8}$.

![Graph of a probability distribution]

c. solution

The median occurs at $x = 2$.

The figure can be broken up into six rectangles of width 1 and height $h$. The area under the figure sums to 1. Therefore $6h = 1 \Rightarrow h = \frac{1}{6}$.
The marked points on the y axis are $f(x) = \frac{1}{6}, \frac{1}{3}, \frac{1}{2}$.

The mean is the $x$ of the center of mass of the area under $f(x)$. The area is composed of 3 rectangles with areas 1, 2, 3 and respective centers of mass at $x = 0.5, 1.5, 2.5$. Let $m$ denote the mean. Then

$$1(m - 0.5) + 2(m - 1.5) + 3(m - 2.5) = 0$$

This implies $6m - 11 = 0$ or $m = 6/11$.

Solution using integration: Mean = $E[X] = \int_0^1 \frac{1}{6}x \, dx + \int_1^2 \frac{1}{3}x \, dx + \int_2^3 \frac{1}{2}x \, dx = 1 \frac{5}{6}$

Since there are 5 squares each of area $\frac{1}{6}$ in the region $x > 1$, $P(x > 1) = 5 \times \frac{1}{6} = \frac{5}{6}$.
Problem 2

For this problem, support all your answers and show all your work.

A discrete random variable \(X\) takes values in \(S_X = \{1/3, -1/3\}\). Let

\[ p = P(X = 1/3) \]

3 p. 2.1. Given that \(\text{Var}(X) = E[X^2]\), what is the value of \(p\)?

**solution**

In general \(\text{Var}(X) = E(X^2) - E(X)^2\). Since in our case \(\text{Var}(X) = E(X^2)\),

\[ E(X) = \frac{1}{3}p + \left(-\frac{1}{3}\right)\left(1 - p\right) \]

\[ = 0 \]

\[ \Rightarrow p = \frac{1}{2} \]

3 p. 2.2. Given that \(E[X] = E[X^2]\), what is the value of \(p\)?

**solution**

\[ E(X) = \frac{1}{3}p + \left(-\frac{1}{3}\right)\left(1 - p\right) \]

\[ = \frac{2}{3}p - \frac{1}{3} \]

\[ E(X^2) = \frac{1}{9}p + \frac{1}{9}(1 - p) \]

\[ = \frac{1}{9} \]

\[ \Rightarrow \frac{2}{3}p - \frac{1}{3} = \frac{1}{9} \Rightarrow p = \frac{2}{3} \]
Problem 3

For this problem, support all your answers and show all your work.

Rob is a robot that roams in the basement of Sieg Hall. One of his jobs is to rescue and escort outside the occasional undergraduate student who gets lost in the basement. Rob has noticed that when the Steam Powered Turing Machine (SPTM) is functioning, the rescues tend to take longer. More precisely, let $S$ denote the event ”the SPTM is on”. After years of serving on this job, Rob has learned that $t$ the time (in minutes) he spends on a rescue mission has a discrete exponential distribution

$$P(t|S) = (1 - 2^{-\lambda_1})2^{-\lambda_1(t-1)}, \ t = 1, 2, \ldots$$
$$P(t|\overline{S}) = (1 - 2^{-\lambda_0})2^{-\lambda_0(t-1)}, \ t = 1, 2, \ldots$$

where $\lambda_1 = 1 \quad \lambda_0 = 2$

Let $T = 2$ and denote by $L$ the event “the rescue was long” $\equiv$ “the rescue lasted $\geq T$ minutes”.

4 p. 3.1. Compute the conditional probabilities $P(L|S)$ and $P(L|\overline{S})$.

solution

$$P(t = 1|S) = (1 - 2^{-1}) \cdot 2^{-1} = \frac{1}{2}$$
$$P(L|S) = 1 - P(t = 1|S) = 1 - \frac{1}{2} = \frac{1}{2}$$

$$P(t = 1|\overline{S}) = (1 - 2^{-2}) \cdot 2^{-2} = \frac{2}{8}$$
$$P(L|\overline{S}) = 1 - P(t = 1|\overline{S}) = 1 - \frac{2}{8} = \frac{3}{4}$$

6 p. 3.2 Rob has to refuel occasionally; call $R$ the event “Rob needs refueling”. Refueling happens only after a rescue is completed, with the following probabilities

$$P(R|L) = \frac{2}{3} \quad P(R|\overline{L}) = \frac{1}{3}$$

Where $L$ is the event “the previous rescue was long”. Assume that $R$ is conditionally independent of anything else given that we know whether $L$ is true or false.

Compute the probabilities $P(R|S)$ and $P(R|\overline{S})$ the conditional probabilities that Rob has to refuel after the current rescue, given the state of the Steam Powered Turing Machine.

solution
\[ P(R|S) = P(R|L, S) P(L|S) + P(R|\bar{L}, S) P(\bar{L}|S) \]
\[ = P(R|L) P(L|S) + P(R|\bar{L}) P(\bar{L}|S) \]
\[ = \frac{2}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2} \]
\[ = \frac{3 + 6}{6} \]
\[ = \frac{1}{2} \]

\[ P(R|\bar{S}) = P(R|L, \bar{S}) P(L|\bar{S}) + P(R|\bar{L}, \bar{S}) P(\bar{L}|\bar{S}) \]
\[ = P(R|L) P(L|\bar{S}) + P(R|\bar{L}) P(\bar{L}|\bar{S}) \]
\[ = \frac{2}{3} \cdot \frac{1}{4} + \frac{1}{3} \cdot \frac{3}{4} \]
\[ = \frac{6 + 4}{12} \]
\[ = \frac{5}{12} \]

3 p. 3.3 Now compute the conditional probability \( P(S|R) \) representing the probability that the SPTM is on, given that Rob had to refuel after the current rescue. The probability that the SPTM is on is

\[ P(S) = \frac{1}{2} \]

**solution**

\[ P(S|R) = \frac{P(R|S) P(S)}{P(R)} \]
\[ = \frac{P(R|S) P(S)}{P(R|S) P(S) + P(R|\bar{S}) P(\bar{S})} \]
\[ = \frac{\frac{1}{2} \cdot \frac{1}{2}}{\frac{1}{2} \cdot \frac{1}{2} + \frac{5}{12} \cdot \frac{1}{2}} \]
\[ = \frac{\frac{1}{4}}{\frac{1}{4} + \frac{5}{24}} \]
\[ = \frac{\frac{1}{4}}{\frac{11}{24}} \]
\[ = \frac{6}{11} \]
Problem 4

For this problem, support all your answers and show all your work.

In an experiment, the observed data are real numbers in \((-\infty, \infty)\). The data are generated by two sources with known distributions. Source \(S_1\) generates data that is normal with mean 0 and variance \(\sigma^2\); source \(S_2\) generates data that is normal with mean 0 and variance \(2\sigma^2\). When a data point \(x\) is observed your task is to decide which of the sources \(S_1, S_2\) has generated \(x\) by the Maximum Likelihood (ML) method.

1 p. 4.1. Make a neat drawing of the two distributions.

**solution** (ignore the numbers on the axes)

![Diagram of two normal distributions](image)

1 p. 4.2. What is the likelihood of \(S_1\) for a given \(x\)?

**solution**

\[
\text{likelihood} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}
\]

1 p. 4.3. Which source has a greater likelihood of generating \(x = 0\)?

**solution**

From the figure, we can see that \(S_1\) has the larger likelihood.

3 p. 4.4. Denote by \(f(x|S_1), f(x|S_2)\) the likelihoods of \(x\) under models \(S_1, S_2\)
respectively. What are the values of \( x \) for which

\[
f(x|S_1) = f(x|S_2)
\]

**solution**

Equating the likelihoods of \( S_1 \) and \( S_2 \), we get

\[
\frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \\
\Rightarrow \sqrt{2} e^{-\frac{x^2}{2\sigma^2}} = e^{-\frac{x^2}{2}} \\
\Rightarrow \ln \sqrt{2} - \frac{x^2}{2\sigma^2} = -\frac{x^2}{4\sigma^2} \\
\Rightarrow \frac{1}{2} \ln 2 = \frac{x^2}{4\sigma^2} \\
\Rightarrow 2\sigma^2 \ln 2 = x^2 \\
\Rightarrow x = \pm \sigma \sqrt{2 \ln 2} = \pm 2\sigma \sqrt{\ln 2}
\]

1 p. 4.5. Which source has a larger likelihood of generating \( x = 2\sigma \)? (Hint: \( \sqrt{2} \approx 1.41, e \approx 2.71 \))

**solution**

From the previous answer, we know that the two likelihoods are equal at \( x_0 = \pm 2\sigma \sqrt{\ln 2} \).

Since \( \sqrt{2} < e \), we have that \( \ln 2 < 1 \Rightarrow \sigma > x_0 = 2\sigma \sqrt{\ln 2} \).
From the drawing, we know that \( x = 2\sigma \) lies to the right of the positive intersection point, and can see that \( S_2 \) has the larger likelihood.

Another solution:

\[
\frac{f(2\sigma|S_1)}{f(2\sigma|S_2)} = \sqrt{2} e^{-\frac{4\sigma^2}{2\sigma^2} + \frac{4\sigma^2}{4\sigma^2}} = \frac{\sqrt{2}}{e} < 1
\]

It follows that source 2 has higher likelihood.

3 p. 4.6. Assume that a fair coin is flipped to decide which source will generate the next data point. You cannot observe the outcome of the coin flip but you observe that \( x = 0 \). What is the probability that \( x = 0 \) was generated by the first source?

**solution**

\[
P(S_1|x = 0) = \frac{f(x = 0|S_1)P(S_1)}{f(x = 0|S_1)P(S_1) + f(x = 0|S_2)P(S_2)}
\]
\[
= \frac{\frac{1}{2} \cdot \frac{1}{\sigma \sqrt{2\pi}}}{\frac{1}{2} \cdot \frac{1}{\sigma \sqrt{2\pi}} + \frac{1}{2} \cdot \frac{1}{\sqrt{2} \sigma \sqrt{2\pi}}} \\
= \frac{1}{1 + \frac{1}{\sqrt{2}}} \\
= 2 - \sqrt{2}
\]

Problem 5

Each of the following plots represent the joint density \( f_{XY} \) of two continuous random variables \( X, Y \). The density is uniform in the shaded areas and 0 elsewhere.

Are \( X, Y \) independent?

\[\begin{array}{cc}
\text{a} & \text{b} \\
\text{2 p. Yes} & \text{2 p. No} \\
\text{c} & \text{d} \\
\text{2 p. Yes} & \text{2 p. No}
\end{array}\]