MATH1853 Linear Algebra, Probability & Statistics
25/26 Semester 1 – Part II
Assignment 3 (Proposed Solutions)

Remarks: *These are the proposed solutions of the author. They shall by no means be considered as an official solution provided by the department. They may contain errors. You are advised to use with discretion.*

1. At dusk, and on average, one porcupine is spotted on the Peak Circle Walk. State the distribution and parameter(s) you applied to compute the probabilities.

   (a) Find the probability that no porcupine are spotted on a given day.

Solution:
Let X be the number of porcupines spotted on a day, then $\boxed{X \sim \text{Poisson}(\lambda = 1)}$. The PMF is given by

$$p(x) = \frac{e^{-1}(1)^x}{x!}, \quad x = 0, 1, 2, \dots$$

Then,

$$P(X = 0) = p(0) = \frac{e^{-1}(1)^0}{0!} = \boxed{e^{-1}}.$$

   (b) Find the probability that at no porcupines will be spotted on 3 days this week.

Solution:
Assumeing independence, let $Y = Y_1 + Y_2 + \cdots + Y_7$ be the number of days in which no porcupines are spotted, where $Y_i = \begin{cases} 1, & \text{no porcupine spotted on day } i \\ 0, & \text{otherwise} \end{cases}$, $i \in \{1, 2, \dots, 7\}$. Then, $Y_i \sim \text{Bernoulli}(p = e^{-1})$ and $\boxed{Y \sim \text{Binomial}(n = 7, p = e^{-1})}$. For Y, the PMF is given by

$$p(y) = \binom{7}{y}(e^{-1})^y(1 - e^{-1})^{7-y}, \quad y = 0, 1, 2, \dots, 7.$$

Then,

$$P(Y = 3) = p(3) = \binom{7}{3}(e^{-1})^3(1 - e^{-1})^4 = \boxed{35e^{-3}(1 - e^{-1})^4} \approx 0.2782.$$

2. In Major League Baseball, the World Series determines the champion team. The team that wins four games in a series of up to seven games wins the World Series. This year, the defending champion Los Angeles Dodgers faced the Toronto Blue Jays. What was the probability that Shohei Otani leads the Dodgers to back-to-back championships? Assume thae probability that the Dodgers win a game against the Blue Jays is 0.55.

Solution:
Let X count the number of gams until 4 wins by Dodgers. Then $X \sim \text{NegativeBinomial}(n = 4, p = 0.55)$. The PMF is given by

$$p(x) = \binom{x-1}{3}(0.55)^4(0.45)^{x-4}, \quad x = 4, 5, 6, 7.$$

Then,

$$\begin{aligned} P(\text{champion}) &= \sum_{i=4}^{7} p(i) = p(4) + p(5) + p(6) + p(7) \\ &= \binom{3}{3}(0.55)^4(0.45)^0 + \binom{4}{3}(0.55)^4(0.45)^1 + \binom{5}{3}(0.55)^4(0.45)^2 + \binom{6}{3}(0.55)^4(0.45)^3 \\ &\approx 0.09151 + 0.1647 + 0.1853 + 0.1668 = \boxed{0.60831}. \end{aligned}$$

3. The lifespan of a laptop battery in years is exponentially distributed. The expected lifespan of such a battery is 4 years. Given that a laptop battery lasts 4 years, what is the probability that it will last at least 6 years?

Solution:

Let X be the lifespan of a laptop battery in years. Then, $X \sim \text{Exponential}(\lambda)$. Since the expected lifespan is 4 years, we have

$$\mathbb{E}(X) = \frac{1}{\lambda} = 4 \Rightarrow \lambda = \frac{1}{4}.$$

Then, the PDF is given by:

$$f(x) = \begin{cases} \frac{1}{4}e^{-1/4x}, & x \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

By the memoryless property, we have

$$\begin{aligned} P(X \geq 6 | X \geq 4) &= P(X \geq 2) \\ &= \int_2^\infty \frac{1}{4}e^{-\frac{1}{4}x}\,dx \\ &= \lim_{t\to\infty}\left[-e^{-x/4}\right]_2^t \\ &= \boxed{e^{-\frac{1}{2}}}. \end{aligned}$$

4. Assume the weights of Siamese cats are normally distributed with mean 5.5 kg and standard deviation 0.8 kg. Find the probability a randomly selected Siamese cat weights between 5 kg and 6.2 kg.

Solution:

Let X be the weight. Then $X \sim \text{Normal}(\mu = 5.5, \sigma = 0.8^2)$. Let $Z = \frac{X-5.5}{0.8}$, then $Z \sim \text{Normal}(0, 1)$.

$$\begin{aligned} P(5 < X < 6.2) &= P\left(\frac{5-5.5}{0.8} < Z < \frac{6.2-5.5}{0.8}\right) \\ &= P(-0.625 < Z < 0.875) \\ &= P(Z < 0.875) - P(Z < -0.625) \\ &= P(Z < 0.875) - [1 - P(Z < 0.625)] \\ &= \Phi(0.875) - 1 + \Phi(0.625) \\ &\approx 0.8078 - 1 + 0.7324 \\ &= \boxed{0.5402} \end{aligned}$$

5. The processing time X (in minutes) for a single student's course registration on the university portal is a random variable with mean $\mu = 4.5$ minutes and standard deviation $\sigma = 1.8$ minutes. The distribution of X is unknown. What is the approximate probability that the average processing time for 50 students is more than 5 minutes?

Solution:

Let $\overline{X}$ be the mean of a random sample of size $n = 50$. As $n = 50 \geq 30$, by the Central Limit Theorem, $\overline{X} \overset{\text{approx}}{\sim} \text{Normal}(4.5, \frac{1.8^2}{50})$. Let $Z = \frac{\overline{X}-4.5}{1.8/\sqrt{50}} \overset{\text{approx}}{\sim} \text{Normal}(0, 1)$.

Then,

$$\begin{aligned} P(\overline{X} > 5) &\approx P\left(Z > 1.964\right) = 1 - P\left(Z \leq 1.964\right) = 1 - \Phi(1.964) \\ &\approx 1 - 0.9750 = \boxed{0.0250}. \end{aligned}$$

6. A university reports that 30% of its students use the library's online study resources on a daily basis. If 200 students are randomly selected, what is the probability that more than 70 of them use the library's online resources daily? State the underlying distribution and parameters, and use a normal distribution to approximate the desired probability.

Solution:

Let $R \sim \text{Binomial}(200, 0.3)$ count the number of students using the online resources daily.

Check: $np = 200 \times 0.3 = 60 \geq 5$ and $n(1-p) = 200 \times 0.7 = 140 \geq 5$. So we can use normal approximation.

Let $\overline{R} = \frac{R}{200}$. Then, by Central Limit Theorem, $\overline{R} \overset{\text{approx}}{\sim} \text{Normal}(60, 42)$. Let $Z = \frac{\overline{R}-60}{\sqrt{42}} \overset{\text{approx}}{\sim} \text{Normal}(0,1)$.

Then,

$$\begin{aligned} P(R \geq 71) \approx P(\overline{R} \geq 70.5) &= P\left(Z \geq \frac{70.5-60}{\sqrt{42}}\right) \\ &\approx P(Z \geq 1.62) = 1 - P(Z \leq 1.62) \\ &\approx 1 - 0.9474 = \boxed{0.0526}. \end{aligned}$$

7. A study on student wellness during exam season is conducted. In a random sample of 64 university students, the sample mean sleep duration in the last 24 hours is found to be 5.2 hours with a sample standard deviation of 1.8 hours. Construct a 97% confidence interval for the true mean sleep duration of all university students during exam season. An influencer claims, "During exams, university students get less than 5 hours of sleep on average." Does your confidence interval support this claim?

Solution:

$$\alpha = 1 - 0.97 = 0.03 \Rightarrow z_{\alpha/2} = z_{0.015}$$

Then,

$$\begin{aligned} P(Z < z_{0.015}) - P(Z < -z_{0.015}) &= 0.97 \\ -1 + 2P(Z < z_{0.015}) &= 0.97 \\ \Phi(z_{0.015}) &= 0.985 \\ z_{0.015} &= 2.17 \end{aligned}$$

The 97% confidence interval is given by

$$\mu = 5.2 \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 5.2 \pm 2.17\frac{1.8}{\sqrt{64}} = 5.2 \pm 0.48825 = \boxed{(4.712, 5.688)}$$

Since 5 is contained in the confidence interval, we do not have enough evidence to support the claim that university students get less than 5 hours of sleep on average during exam season. Therefore, $\boxed{\text{not}}$ supported.

8. Suppose a random variable X has the moment generating function given by:

$$M_X(t) = \frac{1}{(1-5t)^3}, \quad t < \frac{1}{5}.$$

(a) Compute $\mathbb{E}(X)$.

Solution:
We have:

$$M'_X(t) = \frac{15}{(1-5t)^4} \implies M'_X(0) = 15.$$

Therefore, $\mathbb{E}(X) = \boxed{15}$.

(b) Compute $\text{Var}(X)$.

Solution:
We have:

$$M''_X(t) = \frac{300}{(1-5t)^5} \implies M''_X(0) = 300.$$

Therefore,

$$\text{Var}(X) = M''_X(0) - (M'_X(0))^2 = 300 - 15^2 = 75.$$

Thus, $\text{Var}(X) = \boxed{75}$.