CSE3504 Homework 1

Isaac Piegat

October 10, 2024

1 Problem 1:

Let X be a continuous random variable with the probability density function (PDF):

$$f_X(x) = \begin{cases} 6x(1-x), & 0 \le x \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find the CDF of X, $F_X(x)$.

The cumulative distribution function (CDF) is obtained by integrating the PDF:

$$F_X(x) = \int_0^x 6t(1-t)\, dt.$$

First, expand and integrate:

$$F_X(x) = 6\int_0^x (t-t^2)\, dt = 6\left[\frac{t^2}{2} - \frac{t^3}{3}\right]_0^x = 6\left(\frac{x^2}{2} - \frac{x^3}{3}\right).$$

Thus, the CDF is:

$$F_X(x) = 3x^2 - 2x^3.$$

(b) Find $P(X < 1/4)$ and compute the probability.

The probability $P(X < 1/4)$ is given by $F_X(1/4)$:

$$F_X(1/4) = 3\left(\frac{1}{4}\right)^2 - 2\left(\frac{1}{4}\right)^3 = 3\left(\frac{1}{16}\right) - 2\left(\frac{1}{64}\right) = \frac{3}{16} - \frac{2}{64} = \frac{3}{16} - \frac{1}{32} = \frac{6}{32} - \frac{1}{32} = \frac{5}{32}.$$

Thus, $P(X < 1/4) = \frac{5}{32}$.

(c) Find $P(X > 1/2)$ and compute the probability.

The probability $P(X > 1/2)$ is given by:

$$P(X > 1/2) = 1 - P(X \leq 1/2) = 1 - F_X(1/2).$$

First, compute $F_X(1/2)$:

$$F_X(1/2) = 3\left(\frac{1}{2}\right)^2 - 2\left(\frac{1}{2}\right)^3 = 3\left(\frac{1}{4}\right) - 2\left(\frac{1}{8}\right) = \frac{3}{4} - \frac{1}{4} = \frac{2}{4} = \frac{1}{2}.$$

Thus,

$$P(X > 1/2) = 1 - \frac{1}{2} = \frac{1}{2}.$$

(d) Find the median of X.

The median is the value m such that $P(X \leq m) = 0.5$, or equivalently $F_X(m) = 0.5$. From part (c), we found that $F_X(1/2) = 0.5$, so the median of X is:

$$m = \frac{1}{2}.$$

Problem 2

A large lot of marbles have diameters which are approximately normally distributed with a mean of 1 cm. One third of the marbles have diameters greater than 1.1 cm.

(a) Find the standard deviation of the distribution.

We are given that one third of the marbles have diameters greater than 1.1 cm. This implies:

$$P(X > 1.1) = \frac{1}{3}.$$

Using the standard normal distribution, we need to find the corresponding z-score for which $P(Z > z) = \frac{1}{3}$. From the z-table, this corresponds to:

$$z = -0.43.$$

Now, we use the z-score formula to find the standard deviation σ:

$$z = \frac{1.1 - 1}{\sigma}.$$

Substituting the values:

$$-0.43 = \frac{0.1}{\sigma},$$

$$\sigma = \frac{0.1}{0.43} \approx 0.2326.$$

Thus, the standard deviation is approximately:

$$\sigma \approx 0.2326\,\text{cm}.$$

(b) Proportion whose diameters are within 0.2 cm of the mean.

We are asked to find $P(0.8 \leq X \leq 1.2)$. First, convert the values to z-scores:

$$P\left(\frac{0.8-1}{\sigma} \leq Z \leq \frac{1.2-1}{\sigma}\right) = P\left(\frac{-0.2}{0.2326} \leq Z \leq \frac{0.2}{0.2326}\right) = P(-0.86 \leq Z \leq 0.86).$$

Using the z-table or R function to find the cumulative probabilities:

$$P(-0.86 \leq Z \leq 0.86) = P(Z \leq 0.86) - P(Z \leq -0.86) = 0.8051 - 0.1949 = 0.6102.$$

Thus, the proportion of marbles with diameters within 0.2 cm of the mean is approximately:

$$P(0.8 \leq X \leq 1.2) \approx 0.6102.$$

To compute this in R:

Compute the proportion within 0.2 cm of the mean pnorm(0.2 / 0.2326) - pnorm(-0.2 / 0.2326)

(c) Diameter that is exceeded by 75% of the marbles.

We are asked to find the diameter d such that $P(X > d) = 0.75$. This is equivalent to finding the z-score for which $P(Z > z) = 0.75$, or equivalently $P(Z \leq z) = 0.25$.

Using the z-table, we find that the z-score corresponding to $P(Z \leq z) = 0.25$ is:

$$z = -0.6745.$$

Now, we use the z-score formula to relate z, d, the mean, and the standard deviation:

$$z = \frac{d-1}{\sigma}.$$

Substituting the known values:

$$-0.6745 = \frac{d-1}{0.2326}.$$

Solving for d:

$$d - 1 = -0.6745 \times 0.2326 \approx -0.157,$$

$$d = 1 - 0.157 = 0.843.$$

Thus, the diameter that is exceeded by 75% of the marbles is approximately:

$$d \approx 0.843\,\text{cm}.$$

Problem 3

Transistors produced by one machine have a lifetime which is exponentially distributed with rate $\lambda = 0.01\,\text{hour}^{-1}$. Transistors produced by a second machine have an exponentially distributed lifetime with rate $\lambda = 0.005\,\text{hour}^{-1}$. A package of 12 transistors contains 4 produced by the first machine and 8 produced by the second machine.

(a) Probability that the lifetime of a randomly chosen transistor exceeds 100 hours.

The probability that the lifetime T of an exponentially distributed random variable exceeds t hours is given by:

$$P(T > t) = e^{-\lambda t}.$$

We need to calculate the overall probability for the randomly chosen transistor, which could come from either machine. The probability that a randomly chosen transistor comes from the first machine is $\frac{4}{12} = \frac{1}{3}$, and from the second machine is $\frac{8}{12} = \frac{2}{3}$.

For the first machine ($\lambda_1 = 0.01$):

$$P(T_1 > 100) = e^{-0.01\times 100} = e^{-1} \approx 0.3679.$$

For the second machine ($\lambda_2 = 0.005$):

$$P(T_2 > 100) = e^{-0.005\times 100} = e^{-0.5} \approx 0.6065.$$

Now, we calculate the weighted average:

$$P(T > 100) = \frac{1}{3}P(T_1 > 100) + \frac{2}{3}P(T_2 > 100) = \frac{1}{3}\times 0.3679 + \frac{2}{3}\times 0.6065 = 0.4470.$$

Thus, the probability that the lifetime of a randomly chosen transistor exceeds 100 hours is approximately:

$$P(T > 100) \approx 0.4470.$$

To compute this in R:

Calculate the probabilities using the exponential distribution $P_{T1} < -exp(-0.01 * 100) P_{T2} < -exp(-0.005 * 100)$

Compute the weighted probability $P_1 00 < -(1/3) * P_T 1_1 00 + (2/3) * P_T 2_1 00$

(b) Probability that the lifetime exceeds 200 hours, given it has already exceeded 150 hours.

For an exponential distribution, the memoryless property holds, which means that:

$$P(T > t + s \mid T > s) = P(T > t).$$

In this case, we want to find $P(T > 200 \mid T > 150)$, which simplifies to $P(T > 200 - 150) = P(T > 50)$.

Now, we calculate $P(T > 50)$ for each machine:

For the first machine with $\lambda_1 = 0.01$:

$$P(T_1 > 50) = e^{-0.01 \times 50} = e^{-0.5} \approx 0.6065.$$

For the second machine with $\lambda_2 = 0.005$:

$$P(T_2 > 50) = e^{-0.005 \times 50} = e^{-0.25} \approx 0.7788.$$

Since 1/3 of the transistors come from the first machine and 2/3 come from the second machine, the weighted probability is:

$$P(T > 50) = \frac{1}{3} P(T_1 > 50) + \frac{2}{3} P(T_2 > 50) = \frac{1}{3} \times 0.6065 + \frac{2}{3} \times 0.7788.$$

Carrying out the calculation:

$$P(T > 50) = 0.7214.$$

Thus, the probability that the lifetime exceeds 200 hours, given that it has already exceeded 150 hours, is approximately:

$$P(T > 200 \mid T > 150) \approx 0.7214.$$

Problem 5

Three students take equivalent standardized tests.

- On test 1, student 1 scores 144 with a mean of 128 and a standard deviation of 34.
- On test 2, student 2 scores 90 on a test with a mean of 86 and a standard deviation of 18.
- On test 3, student 3 scores 18 with a mean of 15 and a standard deviation of 5.

All the test scores are normally distributed. We are asked to determine which student's score is the most impressive.

To compare the test scores, we can compute the z-scores for each student. The z-score is given by the formula:

$$z = \frac{x - \mu}{\sigma},$$

where x is the observed score, μ is the mean of the test, and σ is the standard deviation.

Student 1

For student 1, we have $x = 144$, $\mu = 128$, and $\sigma = 34$:

$$z_1 = \frac{144 - 128}{34} = \frac{16}{34} \approx 0.47.$$

Student 2

For student 2, we have $x = 90$, $\mu = 86$, and $\sigma = 18$:

$$z_2 = \frac{90 - 86}{18} = \frac{4}{18} \approx 0.22.$$

Student 3

For student 3, we have $x = 18$, $\mu = 15$, and $\sigma = 5$:

$$z_3 = \frac{18 - 15}{5} = \frac{3}{5} = 0.6.$$

Conclusion

Comparing the z-scores of the three students:

$$z_1 \approx 0.47, \quad z_2 \approx 0.22, \quad z_3 = 0.6.$$

Since student 3 has the highest z-score ($z_3 = 0.6$), student 3's score is the most impressive relative to the other students' scores.

Problem 6

The time taken by a computer technician to fix a laptop is uniformly distributed between 15 minutes and 1 hour 15 minutes.

The range of the uniform distribution is $a = 15$ minutes and $b = 75$ minutes (since 1 hour 15 minutes is 75 minutes).

(a) Find the probability that it takes the technician less than 30 minutes to fix the laptop.

For a uniform distribution, the probability that a random variable X falls between two values is given by:

$$P(a \leq X \leq b) = \frac{X - a}{b - a}.$$

We are asked to find $P(X < 30)$. Applying the formula:

$$P(X < 30) = \frac{30 - 15}{75 - 15} = \frac{15}{60} = 0.25.$$

Thus, the probability that it takes the technician less than 30 minutes is:

$$P(X < 30) = 0.25.$$

The corresponding R code to compute this probability using the ‘punif‘ function is:

"‘R Probability of fixing the laptop in less than 30 minutes punif(30, min = 15, max = 75)

(b) Find the probability that it takes the technician between 45 minutes and one hour to fix the laptop.

We are asked to find $P(45 \leq X \leq 60)$, where 60 minutes corresponds to 1 hour. For a uniform distribution, the probability that a random variable X falls between two values is given by:

$$P(a \leq X \leq b) = \frac{X_2 - X_1}{b - a}.$$

Here, $X_1 = 45$ minutes, $X_2 = 60$ minutes, $a = 15$ minutes, and $b = 75$ minutes. Using this, we get:

$$P(45 \leq X \leq 60) = \frac{60 - 45}{75 - 15} = \frac{15}{60} = 0.25.$$

Thus, the probability that it takes the technician between 45 minutes and 1 hour to fix the laptop is:

$$P(45 \leq X \leq 60) = 0.25.$$

The corresponding R code to compute this probability using the ‘punif‘ function is:

Probability of fixing the laptop between 45 minutes and 60 minutes punif(60, min = 15, max = 75) - punif(45, min = 15, max = 75)