Statistics for Data Science - 2

Week 5 Graded Assignment solution

Continuous random variable

1. The CDF of a random variable X is

$$F_X(x) = \begin{cases} 1 - e^{-3x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

i) Find $P(X > 4)$.

a) $e^{-3} - e^{-4}$

b) e^{-12}

c) e^{-7}

d) $e^{-3}e^{-4}$

Solution:

$$\begin{aligned} P(X > 4) = 1 - P(X \leq 4) &= 1 - F_X(4) \\ &= 1 - (1 - e^{-3\times 4}) \\ &= e^{-12} \end{aligned}$$

ii) Find the value of $P(-5 < X \leq 6)$.

a) $1 - e^{-18}$

b) $e^{-5} - e^{-18}$

c) e^{-18}

d) e^{-9}

Solution:

$$\begin{aligned} P(-5 < X \leq 6) &= F_X(6) - F_X(-5) \\ &= (1 - e^{-3\times 6}) - 0 \\ &= 1 - e^{-18} \end{aligned}$$

2. Let X be a continuous random variable with the following PDF:

$$f_X(x) = \begin{cases} ke^{-x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

i) Find the value of k.

Solution:

We know that for PDF of the random variable

$$\int_{-\infty}^{\infty} f_X(x) = 1$$
$$\Rightarrow \int_{0}^{\infty} ke^{-x}dx = 1$$
$$\Rightarrow k\left(\frac{e^{-x}}{-1}\right)\Bigg|_0^{\infty} = 1$$

$\Rightarrow k(0+1) = 1$ (As x approaches to ∞, e^{-x} approaches to 0)

$\Rightarrow k = 1$

ii) Find $P(3 < X < 4)$.

a) e^{-1}

b) $e^{-3}e^{-4}$

c) $e^{-3} - e^{-4}$

d) $e^{-4} - e^{-3}$

Hint: Use $\int_a^b e^{-x}dx = e^{-a} - e^{-b}$

Solution:

$$P(3 < X < 4) = \int_3^4 ke^{-x}dx$$
$$= 1 \times \left(\frac{e^{-x}}{-1}\right)\Bigg|_3^4$$
$$= \left(\frac{e^{-4}}{-1} - \frac{e^{-3}}{-1}\right)$$
$$= e^{-3} - e^{-4}$$

3. Let X be a continuous random variable with PDF

$$f_X(x) = \begin{cases} 5x^4 & 0 < x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $P(X \leq \frac{3}{4} \mid X > \frac{1}{4})$.

a) $\dfrac{3}{16}$

b) $\frac{17}{86}$

c) $\frac{22}{93}$

d) $\frac{9}{22}$

Hint: Use $\int_a^b 5x^4 dx = b^5 - a^5$

$$
\begin{aligned}
P(X \leq \frac{3}{4} \mid X > \frac{1}{4}) &= \frac{P(X \leq \frac{3}{4} \text{ and } X > \frac{1}{4})}{P(X > \frac{1}{4})} \\
&= \frac{\int_{1/4}^{3/4} 5x^4 dx}{\int_{1/4}^{1} 5x^4 dx} \\
&= \frac{\left.\frac{5x^5}{5}\right|_{1/4}^{3/4}}{\left.\frac{5x^5}{5}\right|_{1/4}^{1}} \\
\Rightarrow P(X \leq \frac{3}{4} \mid X > \frac{1}{4}) &= \frac{\left.x^5\right|_{1/4}^{3/4}}{\left.x^5\right|_{1/4}^{1}} \\
&= \frac{(\frac{3}{4})^5 - (\frac{1}{4})^5}{1 - (\frac{1}{4})^5} \\
&= \frac{22}{93}
\end{aligned}
$$

4. The lifespan (in hours) of an electronic component used in an electric car has the density function

$$
f_X(x) = \begin{cases} \frac{1}{500} e^{-\frac{x}{500}} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}
$$

Determine the probability that the component lasts more than 200 hours before it needs to be replaced.

a) $e^{-0.4}$

b) e^{200}

c) 0.5

d) $e^{-2.5}$

Solution:
Let X denote the lifespan (in hours) of the electronic component. We have to find the probability that the component lasts more than 200 hours before it needs to be replaced i.e.
$P(X > 200) = 1 - P(X \leq 200)$
Also, we can relate the given density with the exponential distribution with $\lambda = \frac{1}{500}$.

$$\begin{aligned}
\Rightarrow P(X > 200) &= 1 - P(X \leq 200) \\
&= 1 - F_X(200) \\
&= 1 - (1 - e^{-\frac{200}{500}}) \\
&= e^{-0.4}
\end{aligned}$$

5. The number of days in advance by which airline tickets are purchased by travelers is exponentially distributed with an average of 28 days. If there is an 80% chance that a traveler will purchase tickets fewer than d days in advance, then what is the value of d? Write your answer to the nearest integer.
Solution:
Let X be the number of days in advance by which airline tickets are purchased by travelers.
We need to find the value of d.
Given that average is 28 days, so $\lambda = \frac{1}{28}$ and there is 80% chance that a traveler will purchase tickets fewer than d days in advance.

$$\begin{aligned}
\Rightarrow P(X < d) &= 0.80 \\
\Rightarrow 1 - e^{-\frac{d}{28}} &= 0.80 \\
\Rightarrow e^{-\frac{d}{28}} &= 0.20 \\
\Rightarrow \frac{d}{28} &= -ln(0.20) \\
\Rightarrow d &= 28 \times (1.609) \\
\Rightarrow d &= 45.052
\end{aligned}$$

Rounding off to the nearest integer, $d = 45$ days.

6. A firm produces machines with a lifespan, whose distribution has a mean of 200 months and standard deviation of 50 months. The firm wishes to introduce a warranty scheme in which it would like to replace all the dysfunctional machines with new ones within warranty period. But they do not wish to do so for more than 11.9% of the machines they produce. If the lifespan of the machine is assumed to follow a normal distribution,

how long a guarantee period should be offered? (Answer is expected in months)
Hint: Use $P(Z < -1.18) = 0.119$, where Z represents the standard normal distribution.
Solution:
Let X denote the lifespan of the machines in months. Given that $\mu = 200$ and $\sigma = 50$.
The firm did not wish to replace more than 11.9% of the machines they produce.
If m be the guarantee period (in months), then

$$\begin{aligned} P(X \leq m) &= 0.119 \\ \Rightarrow P\left(\frac{X-200}{50} \leq \frac{m-200}{50}\right) &= 0.119 \end{aligned}$$

Comparing this equation with the given value of standard normal distribution we will get

$$\begin{aligned} \frac{m-200}{50} &= -1.18 \\ \Rightarrow m &= 141 \end{aligned}$$

7. Let X be a continuous random variable with the following PDF:

$$f_X(x) = \begin{cases} 3(1-x)^2 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Define $Y = (1-X)^3$. Find the PDF of the random variable Y.

a)
$$f_Y(y) = \begin{cases} 1 & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

b)
$$f_Y(y) = \begin{cases} (1-y)^3 & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

c)
$$f_Y(y) = \begin{cases} y^3 & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

d)
$$f_Y(y) = \begin{cases} 3y^{2/3} & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Hint:
Apply the monotonic, differentiable function theorem and $\frac{d}{dx}(1-x)^3 = -3(1-x)^2$

Solution:
We know that in the range (0, 1), $(1-x)^3$ is monotonic (decreasing function).
Therefore, we can use the formula, $f_Y(y) = \dfrac{1}{|g'(g^{-1}(y))|} f_X(g^{-1}(y))$
Given $Y = (1-X)^3 = g(X)(\text{let})$
$\Rightarrow y^{1/3} = 1-x, \Rightarrow x = 1-y^{1/3} = g^{-1}(y)$
Therefore $g^{-1}(y) = 1-y^{1/3}$
$g(x) = (1-x)^3 \Rightarrow g'(x) = -3(1-x)^2$, since $\dfrac{d}{dx}(1-x)^3 = -3(1-x)^2$
And
$g'(g^{-1}(y)) = g'(1-y^{1/3}) = -3(1-(1-y^{1/3}))^2 = -3y^{2/3}$
$|g'(g^{-1}(y))| = 3y^{2/3}$, since $y^{2/3}$ is positive in the range $(0,1)$.
$f_X(g^{-1}(y)) = f_X(1-y^{1/3}) = 3(1-(1-y^{1/3}))^2 = 3y^{2/3}$
Therefore, $f_Y(y) = \dfrac{3y^{2/3}}{3y^{2/3}}$
$\Rightarrow f_Y(y) = 1$
Therefore

$$f_Y(y) = \begin{cases} 1 & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

8. Let X be a continuous random variable with the following PDF:

$$f_X(x) = \begin{cases} x^2/81 & -6 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

Define $Y = \frac{1}{3}(12 - X)$. Find the PDF of the random variable Y.

a)
$$f_Y(y) = \begin{cases} (12-3y)^2/27 & -6 < y < 3 \\ 0 & \text{otherwise} \end{cases}$$

b)
$$f_Y(y) = \begin{cases} (12-3y)^2/27 & 3 < y < 6 \\ 0 & \text{otherwise} \end{cases}$$

c)
$$f_Y(y) = \begin{cases} (12-3y)/27 & -6 < y < 3 \\ 0 & \text{otherwise} \end{cases}$$

d)
$$f_Y(y) = \begin{cases} (12-3y)/27 & 3 < y < 6 \\ 0 & \text{otherwise} \end{cases}$$

Solution:
We know that in the range (-6, 3), $\frac{1}{3}(12 - x)$ is monotonic (decreasing function).
Therefore, we can use the formula, $f_Y(y) = \dfrac{1}{|g'(g^{-1}(y))|} f_X(g^{-1}(y))$
Given $Y = \frac{1}{3}(12 - X) = g(X)$(let)
$\Rightarrow 3y = 12 - x, \Rightarrow x = 12 - 3y = g^{-1}(y)$
Therefore $g^{-1}(y) = 12 - 3y$
$g(x) = \frac{1}{3}(12 - x) \Rightarrow g'(x) = -\frac{1}{3}$
And
$g'(g^{-1}(y)) = g'(12 - 3y) = -\frac{1}{3}$
$|g'(g^{-1}(y))| = \frac{1}{3}$
$f_X(g^{-1}(y)) = f_X(12 - 3y) = \dfrac{(12 - 3y)^2}{81}$
Therefore, $f_Y(y) = \dfrac{\frac{(12 - 3y)^2}{81}}{\frac{1}{3}}$
$\Rightarrow f_Y(y) = \dfrac{(12 - 3y)^2}{27}$
When $x = -6, y = 6$ and $x = 3, y = 3$.
Therefore

$$f_Y(y) = \begin{cases} \dfrac{(12 - 3y)^2}{27} & 3 < y < 6 \\ 0 & \text{otherwise} \end{cases}$$

9. Let X be a continuous random variable with the following PDF:

$$f_X(x) = \begin{cases} x^3(6x^2 + 5x - 4) & 0 < x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the value of $E[X]$.

a) $\dfrac{523}{210}$
b) $\dfrac{23}{210}$
c) $\dfrac{173}{210}$
d) $\dfrac{187}{210}$

Hint: Use $\int_a^b x^n dx = \frac{1}{n+1}(b^{n+1} - a^{n+1})$

Solution:

$$
\begin{aligned}
E[X] &= \int_{-\infty}^{\infty} x f_X(x) dx \\
&= \int_0^1 x \times x^3(6x^2+5x-4)dx \\
&= \int_0^1 (6x^6+5x^5-4x^4)dx \\
&= \frac{6x^7}{7}\Big|_0^1 + \frac{5x^6}{6}\Big|_0^1 - \frac{4x^5}{5}\Big|_0^1 \\
&= \frac{6}{7}+\frac{5}{6}-\frac{4}{5} \\
&= \frac{187}{210}
\end{aligned}
$$

10. Let X be a continuous random variable with the following PDF:

$$
f_X(x) = \begin{cases} x & 0 \le x \le 1 \\ 2-x & 1 < x \le 2 \\ 0 & \text{otherwise} \end{cases}
$$

Define $Y = 6X + 5$. Find the variance of Y.

6

Use $\int_a^b x^n dx = \frac{1}{n+1}(b^{n+1} - a^{n+1})$

Also, $\int_a^b x^n dx = \int_a^c x^n dx + \int_c^b x^n dx$ where $a < c < b$.

Solution:

$\text{Var}(Y) = \text{Var}(6X+5) = 36\text{Var}(X)$

And $\text{Var}(X) = E[X^2] - (E[X])^2$

$$\begin{aligned}
E[X] &= \int_{-\infty}^{\infty} x f_X(x)dx \\
&= \int_0^2 x f_X(x)dx \\
&= \int_0^1 x f_X(x)dx + \int_1^2 x f_X(x)dx \\
&= \int_0^1 x.xdx + \int_1^2 x(2-x)dx \\
&= \frac{x^3}{3}\bigg|_0^1 + \frac{2x^2}{2}\bigg|_1^2 - \frac{x^3}{3}\bigg|_1^2 \\
&= \frac{1}{3} + (2^2 - 1^2) - \frac{(2^3 - 1^3)}{3} \\
&= \frac{1}{3} + 3 - \frac{7}{3} \\
&= 1
\end{aligned}$$

$$\begin{aligned}
E[X^2] &= \int_{-\infty}^{\infty} x^2 f_X(x)dx \\
&= \int_0^2 x^2 f_X(x)dx \\
&= \int_0^1 x^2 f_X(x)dx + \int_1^2 x^2 f_X(x)dx \\
&= \int_0^1 x^2.xdx + \int_1^2 x^2(2-x)dx \\
&= \frac{x^4}{4}\bigg|_0^1 + \frac{2x^3}{3}\bigg|_1^2 - \frac{x^4}{4}\bigg|_1^2 \\
&= \frac{1}{4} + \frac{2}{3}(2^3 - 1^3) - \frac{1}{4}(2^4 - 1^4) \\
&= \frac{1}{4} + \frac{14}{3} - \frac{15}{4} \\
&= \frac{7}{6}
\end{aligned}$$

Therefore,
$\text{Var}(X) = \frac{7}{6} - 1 = \frac{1}{6}$
$\Rightarrow \text{Var}(Y) = 36 \times \frac{1}{6} = 6$