Statistics for Data Science-2

Week 5 Solve with us

Table of contents

1. The CDF of a random variable X is

$$F_X(x) = \begin{cases} 1 - e^{-7x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Find $P(X \leq 10)$.

a) e^{-70}
b) $1 - e^{-70}$
c) $1 - e^{-7}$
d) $1/2$

Solution:

$$P(X \leq 10) = F_X(10) = 1 - e^{-7 \times 10} = 1 - e^{-70}$$

2. Find the value of $P(-2 < X \leq 4)$.
   a) $1 - e^{-28}$
   b) $e^{14} - e^{-28}$
   c) $e^{-18} - e^{14}$
   d) e^{-28}

Solution:

$$
\begin{aligned}
P(-2 < X \leq 4) &= F_X(4) - F_X(-2) \\
&= (1 - e^{-7\times 4}) - 0 \\
&= 1 - e^{-28}
\end{aligned}
$$

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

$$f_X(x) = \begin{cases} k(x^2+4) & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the value of k.
Hint: $\int_a^b x^n dx = \frac{1}{n+1}(b^{n+1} - a^{n+1})$

a) 3/13
b) 10/13
c) 1/7
d) 6/7

Solution:

We know that for PDF of the random variable

$$\int_{-\infty}^{\infty} f_X(x) = 1$$

$$\Rightarrow \int_{0}^{1} k(x^2+4)dx = 1$$

$$\Rightarrow k\left(\frac{x^3}{3}+4x\right)\Bigg|_0^1 = 1$$

$$\Rightarrow k(\frac{1}{3}+4) = 1$$

$$\Rightarrow k = \frac{3}{13}$$

4. Find the value of $P(0 < X < 1/2)$.

a) 33/104
b) 55/104
c) 49/104
d) 71/104

$$
\begin{aligned}
P(0 < X < 1/2) &= \int_0^{1/2} \frac{3}{13}(x^2+4)dx \\
&= \frac{3}{13}\left(\frac{x^3}{3}+4x\right)\Bigg|_0^{1/2} \\
&= \frac{3}{13}\left(\frac{1}{24}+\frac{4}{2}\right) \\
&= \frac{3}{13}\times\frac{49}{24} \\
&= \frac{49}{104}
\end{aligned}
$$

5. Let X be a continuous random variable with PDF

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

Find $P(X \leq \frac{1}{2} \mid X > \frac{1}{4})$.
Hint: $\int_a^b x^n dx = \frac{1}{n+1}(b^{n+1} - a^{n+1})$

a) $17/63$
b) $46/63$
c) $56/63$
d) $7/63$

Solution:

$$P(X \leq \frac{1}{2} \mid X > \frac{1}{4}) = \frac{P(X \leq \frac{1}{2} \textbf{and} X > \frac{1}{4})}{P(X > \frac{1}{4})}$$

$$= \frac{\int_{1/4}^{1/2} 3x^2 dx}{\int_{1/4}^{1} 3x^2 dx}$$

$$= \frac{\frac{3x^3}{3}\Big|_{1/4}^{1/2}}{\frac{3x^3}{3}\Big|_{1/4}^{1}}$$

$$\Rightarrow P(X \leq \frac{1}{2} \mid X > \frac{1}{4}) = \frac{x^3\Big|_{1/4}^{1/2}}{x^3\Big|_{1/4}^{1}}$$

$$= \frac{\frac{1}{8} - \frac{1}{64}}{1 - \frac{1}{64}}$$

$$= \frac{7}{63}$$

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

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

Define $Y = X^3$. Find the PDF of the random variable Y.
Hint:
Apply the monotonic, differentiable functions theorem and
$\frac{d}{dx}x^3 = 3x^2$

a)

$$f_Y(y) = \begin{cases} \frac{5}{3}y^{2/3} & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

b)

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

c)

$$f_Y(y) = \begin{cases} 4y^3 & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

d)

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

Solution:

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

$f_X(g^{-1}(y)) = f_X(y^{1/3}) = 3y^{2/3}$

Therefore, $f_Y(y) = \dfrac{3y^{2/3}}{3y^{2/3}}$

$\Rightarrow f_Y(y) = 1$

Therefore

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

7. The test scores of the statistics portion in the qualifier exam are normally distributed with a mean score of 46. If the standard deviation of the score is 15, then find approximately how many percent of the scores are between 16 and 61.
Use the following CDF values of standard normal distribution.
$F_Z(-2) = 0.02275, F_Z(-1.5) = 0.06681, F_Z(-1) = 0.15866, F_Z(-0.5) = 0.30854, F_Z(0) = 0.5, F_Z(0.5) = 0.69146,$ and $F_Z(1) = 0.84134$

a) 81.8
b) 18.2
c) 38.39
d) 61.31

Solution:

$\mu = 46$ and $\sigma = 15$

$$\begin{aligned} P(16 < X < 61) &= P\left(\frac{16-46}{15} < \frac{X-46}{15} < \frac{61-46}{15}\right) \\ &= P(-2 < Z < 1) \\ &= F_Z(1) - F_Z(-2) \\ &= 0.84134 - 0.02275 \\ &= 0.81859 \end{aligned}$$

Therefore, approximately 81.8 percent of the scores are between 16 and 61.