Probability and Statistics: MA6.101

Tutorial 4

Topics Covered: Continuous Random Variables, PDF, CDF, Joint random variables

Q1: Let X be a continuous random variable with PDF

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

If $Y = \frac{2}{X} + 3$, find $\text{Var}(Y)$.

A: First, note that

$$\text{Var}(Y) = \text{Var}(\frac{2}{X} + 3) = 4 \cdot \text{Var}(\frac{1}{X}),$$

Thus, it suffices to find $\text{Var}(\frac{1}{X})$, which is given by

$$\text{Var}(\frac{1}{X}) = \mathbb{E}[\frac{1}{X^2}] - (\mathbb{E}[\frac{1}{X}])^2.$$

Using LOTUS (Law of the Unconscious Statistician), we have

$$\mathbb{E}[\frac{1}{X}] = \int_0^1 x\left(2x + \frac{3}{2}\right)\, dx = \frac{17}{12},$$

$$\mathbb{E}[\frac{1}{X^2}] = \int_0^1 \left(2x + \frac{3}{2}\right)\, dx = \frac{5}{2}.$$

Thus,

$$\text{Var}(\frac{1}{X}) = \mathbb{E}[\frac{1}{X^2}] - (\mathbb{E}[\frac{1}{X}])^2 = \frac{71}{144}.$$

So, we obtain

$$\text{Var}(Y) = 4 \cdot \text{Var}(\frac{1}{X}) = \frac{71}{36}.$$

Q2: An absent-minded professor schedules two student appointments for the same time. The appointment durations are independent and exponentially distributed with mean thirty minutes. The first student arrives on time, but the second student arrives five minutes late. What is the expected value of the time between the arrival of the first student and the departure of the second student?

A: The expected value of the time between the arrival of the first student and the departure of the second student is given by:

$E[\text{Time}] = (5 + E[\text{stay of 2nd student}]) \cdot P(\text{1st stays no more than 5 minutes}) +$

$(E[\text{stay of 1st} \mid \text{stay of 1st} \geq 5] + E[\text{stay of 2nd student}]) \cdot P(\text{1st stays more than 5 minutes}).$

We have:

$$E[\text{stay of 2nd student}] = 30,$$

and using the memorylessness property of the exponential distribution:

$$E[\text{stay of 1st} \mid \text{stay of 1st} \geq 5] = 5 + E[\text{stay of 1st}] = 35.$$

Also,

$$P(\text{1st student stays no more than 5 minutes}) = 1 - \exp\left(-\frac{5}{30}\right),$$

$$P(\text{1st student stays more than 5 minutes}) = \exp\left(-\frac{5}{30}\right).$$

Substituting these into the formula, we get:

$$E[\text{Time}] = (5+30) \cdot \left(1 - \exp\left(-\frac{5}{30}\right)\right) + (35+30) \cdot \exp\left(-\frac{5}{30}\right).$$

$$E[\text{Time}] = 35 + 30 \cdot \exp\left(-\frac{5}{30}\right) = 60.394.$$

Q3: Let X be a random variable uniformly distributed in $[0, \frac{\pi}{2}]$. Let $Y = \sin(X)$. Calculate the probability density function (PDF) of Y.

Also, calculate the PDF of Y if X is uniformly distributed in $\left[-\frac{\pi}{2}, \pi\right]$.

A: Let X be a random variable uniformly distributed in $\left[-\frac{\pi}{2}, \pi\right]$, and let $Y = \sin(X)$. We wish to calculate the probability density function (PDF) of Y.

Case 1: X uniformly distributed in $\left[0, \frac{\pi}{2}\right]$

The cumulative distribution function (CDF) of Y can be computed as follows:

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(\sin(X) \leq y).$$

Since X is uniformly distributed in $\left[0, \frac{\pi}{2}\right]$, we have

$$\mathbb{P}(\sin(X) \leq y) = \mathbb{P}(X \leq \arcsin(y)).$$

Thus,

$$F_Y(y) = \frac{\arcsin(y)}{\frac{\pi}{2}} = \frac{2\arcsin(y)}{\pi},$$

for $0 \leq y \leq 1$. The probability density function (PDF) $f_Y(y)$ is the derivative of the CDF:

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}\left(\frac{2\arcsin(y)}{\pi}\right) = \frac{2}{\pi\sqrt{1-y^2}}.$$

Thus, the PDF of Y is:

$$f_Y(y) = \begin{cases} \frac{2}{\pi\sqrt{1-y^2}} & \text{for } 0 \leq y \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Case 2: X uniformly distributed in $\left[-\frac{\pi}{2}, \pi\right]$
The cumulative distribution function (CDF) of Y can be computed as follows:

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(\sin(X) \leq y).$$

Since X is uniformly distributed in $\left[-\frac{\pi}{2}, \pi\right]$, we need to consider two intervals for X when $y \in [-1, 1]$.
Case 2a: $-1 \leq y \leq 0$
In this case, $\sin(X) \leq y$ implies $X \in \left[-\frac{\pi}{2}, \arcsin(y)\right]$. The CDF for this case is:

$$F_Y(y) = \mathbb{P}(X \leq \arcsin(y)) = \frac{\arcsin(y) + \frac{\pi}{2}}{\frac{3\pi}{2}} = \frac{\frac{\pi}{2} + \arcsin(y)}{\frac{3\pi}{2}} = \frac{1}{3} + \frac{2\arcsin(y)}{3\pi}.$$

Case 2b: $0 \leq y \leq 1$
Here, $X \in [0, \arcsin(y)]$ or $X \in [\pi - \arcsin(y), \pi]$.
- The probability for $X \leq \arcsin(y)$ is:

$$\mathbb{P}(X \leq \arcsin(y)) = \frac{\arcsin(y) + \frac{\pi}{2}}{\frac{3\pi}{2}} = \frac{1}{3} + \frac{2\arcsin(y)}{3\pi}.$$

- The probability for $X \geq \pi - \arcsin(y)$ is:

$$\mathbb{P}(X \geq \pi - \arcsin(y)) = \frac{\arcsin(y)}{\frac{3\pi}{2}} = \frac{2\arcsin(y)}{3\pi}.$$

Thus, the total CDF for $0 \leq y \leq 1$ is:

$$F_Y(y) = \mathbb{P}(X \leq \arcsin(y)) + \mathbb{P}(X \geq \pi - \arcsin(y)) = \frac{1}{3} + \frac{4\arcsin(y)}{3\pi}.$$

PDF Calculation
The probability density function (PDF) $f_Y(y)$ is the derivative of the CDF:
For $-1 \leq y \leq 0$:

$$f_Y(y) = \frac{d}{dy}\left(\frac{1}{3} + \frac{2\arcsin(y)}{3\pi}\right) = \frac{2}{3\pi\sqrt{1-y^2}}.$$

For $0 \leq y \leq 1$:

$$f_Y(y) = \frac{d}{dy}\left(\frac{1}{3} + \frac{4\arcsin(y)}{3\pi}\right) = \frac{4}{3\pi\sqrt{1-y^2}}.$$

Thus, the PDF of Y is:

$$f_Y(y) = \begin{cases} \frac{2}{3\pi\sqrt{1-y^2}} & \text{for } -1 \leq y \leq 0, \\ \frac{4}{3\pi\sqrt{1-y^2}} & \text{for } 0 \leq y \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Q4: Let $X_1, X_2, \ldots, X_n$ be n independent exponential random variables with the same parameter λ. Let $Z_{\min} = \min(X_1, X_2, \ldots, X_n)$ and $Z_{\max} = \max(X_1, X_2, \ldots, X_n)$. Calculate the probability density functions of $Z_{\min}$ and $Z_{\max}$.

A: The cumulative distribution function (CDF) of $Z_{\min}$ is given by:

$$F_{Z_{\min}}(z) = \mathbb{P}(Z_{\min} \leq z) = 1 - \mathbb{P}(Z_{\min} > z).$$

Since $Z_{\min} > z$ if and only if all $X_i > z$, we have:

$$\mathbb{P}(Z_{\min} > z) = \mathbb{P}(X_1 > z) \cdot \mathbb{P}(X_2 > z) \cdots \mathbb{P}(X_n > z).$$

For an exponential random variable with parameter λ,

$$\mathbb{P}(X_i > z) = e^{-\lambda z}.$$

Thus,

$$\mathbb{P}(Z_{\min} > z) = \left(e^{-\lambda z}\right)^n = e^{-n\lambda z}.$$

Therefore,

$$F_{Z_{\min}}(z) = 1 - e^{-n\lambda z}.$$

The probability density function (PDF) is the derivative of the CDF:

$$f_{Z_{\min}}(z) = \frac{d}{dz} F_{Z_{\min}}(z) = n\lambda e^{-n\lambda z}.$$

Thus, the PDF of $Z_{\min}$ is:

$$f_{Z_{\min}}(z) = \begin{cases} n\lambda e^{-n\lambda z} & \text{for } z \geq 0 \\ 0 & \text{otherwise} \end{cases}.$$

The cumulative distribution function (CDF) of $Z_{\max}$ is given by:

$$F_{Z_{\max}}(z) = \mathbb{P}(Z_{\max} \leq z) = \mathbb{P}(X_1 \leq z, X_2 \leq z, \ldots, X_n \leq z).$$

Since the X_i are independent,

$$\mathbb{P}(X_1 \leq z, X_2 \leq z, \ldots, X_n \leq z) = \mathbb{P}(X_1 \leq z) \cdot \mathbb{P}(X_2 \leq z) \cdots \mathbb{P}(X_n \leq z).$$

For an exponential random variable with parameter λ,

$$\mathbb{P}(X_i \leq z) = 1 - e^{-\lambda z}.$$

Thus,

$$\mathbb{P}(Z_{\max} \leq z) = \left(1 - e^{-\lambda z}\right)^n.$$

Therefore,

$$F_{Z_{\max}}(z) = \left(1 - e^{-\lambda z}\right)^n.$$

The probability density function (PDF) is the derivative of the CDF:

$$f_{Z_{\max}}(z) = \frac{d}{dz} F_{Z_{\max}}(z) = n\lambda e^{-\lambda z} \left(1 - e^{-\lambda z}\right)^{n-1}.$$

Thus, the PDF of $Z_{\max}$ is:

$$f_{Z_{\max}}(z) = \begin{cases} n\lambda e^{-\lambda z} \left(1 - e^{-\lambda z}\right)^{n-1} & \text{for } z \geq 0 \\ 0 & \text{otherwise} \end{cases}.$$

Q5: Let X be a non-negative continuous random variable. Show that

$$E[X^2] = \int_{x=0}^{\infty} 2x\mathbb{P}(X > x)\, dx$$

A: Consider a random variable $Y = X^2$. Since Y is a non-negative continuous random variable, we have

$$E[Y] = \int_{y=0}^{\infty} \mathbb{P}(Y > y)dy$$

$$\implies E[X^2] = \int_{x^2=0}^{\infty} \mathbb{P}(X^2 > x^2)dx^2$$

Since X is non-negative and $dx^2 = 2xdx$

$$\implies E[X^2] = \int_{x=0}^{\infty} 2x\mathbb{P}(X > x)dx$$

Q6: Let Y be Geometric(p) where $p = \lambda h$. Define $X = Yh$ where $\lambda, h > 0$. Prove that for any $x \in (0, \infty)$, we have

$$\lim_{h \to 0} F_X(x) = 1 - e^{-\lambda x}$$

A: The CDF of variable Y is

$$F_Y(y) = P(Y \leq y) = 1 - (1-p)^y = 1 - (1 - \lambda h)^y$$

Since $X = Yh$,

$$F_X(x) = P(Yh \leq x) = P\left(Y \leq \frac{x}{h}\right) = F_Y\left(\frac{x}{h}\right) = 1 - (1 - \lambda h)^{\frac{x}{h}}$$

We can now evaluate the limit.

$$\lim_{h \to 0} F_X(x) = \lim_{h \to 0} 1 - (1 - \lambda h)^{\frac{x}{h}}$$

Since $\lim_{m \to \infty}(1 - \frac{\lambda}{m})^{xm} = e^{-\lambda x}$ where $m = \frac{1}{h}$

$$\lim_{h \to 0} F_X(x) = 1 - e^{-\lambda x}$$

We have shown that as $p \to 0$, a geometric random variable becomes exponential.