Assignment-1 Solution

October 2023

Question 1

Let N be a geometric random variable with parameter p and let $X_i, i \geq 0$ denote i.i.d exponential random variables with parameter λ. Let $Z = X_1+X_2+....+X_N$. Show that Z is an exponential random variable. What is its parameter?

Solution 1

To show that the random variable Z is exponential with parameter λ, we can use the moment-generating function (MGF) method. The MGF of a random variable Z is defined as $M_Z(t) = E[e^{tz}]$. We will first find the MGF for Z and then determine its distribution.

Given that N is a geometric random variable with parameter p, we know that $P(N = k) = (1-p)^{k-1} \cdot p$ for $k = 1, 2, 3, \ldots$.

We also know that X_i are independent and identically distributed exponential random variables with parameter λ. The MGF of each X_i is $M_{X_i}(t) = \frac{\lambda}{\lambda - t}$ for $t < \lambda$.

Now, let's find the MGF for Z:

$$\begin{aligned}
M_Z(t) &= E[e^{tZ}] \\
&= \sum_{k=1}^{\infty} E[e^{t(X_1+X_2+\ldots+X_N)} \mid N = k] \cdot P(N = k) \\
&= \sum_{k=1}^{\infty} E[e^{t(X_1+X_2+\ldots+X_k)} \mid N = k] \cdot P(N = k)
\end{aligned}$$

Now, we can use the fact that X_i are independent and identically distributed:

$$\begin{aligned} M_Z(t) &= \sum_{k=1}^{\infty} E[e^{t(X_1+X_2+\ldots+X_k)} \mid N=k] \cdot P(N=k) \\ &= \sum_{k=1}^{\infty} E[e^{tX_1} \cdot e^{tX_2} \cdot \ldots \cdot e^{tX_k}] \cdot P(N=k) \\ &= \sum_{k=1}^{\infty} E[e^{tX_1}] \cdot E[e^{tX_2}] \cdot \ldots \cdot E[e^{tX_k}] \cdot P(N=k) \\ &= \sum_{k=1}^{\infty} (M_{X_1}(t) \cdot M_{X_2}(t) \cdot \ldots \cdot M_{X_k}(t)) \cdot P(N=k) \end{aligned}$$

Now, let's substitute the expressions for the MGF of X_i and the probability mass function of N:

$$\begin{aligned} M_Z(t) &= \sum_{k=1}^{\infty} \left(\frac{\lambda}{\lambda-t}\right)^k \cdot (1-p)^{k-1} \cdot p \\ &= \frac{p}{1-p} \sum_{k=1}^{\infty} \left(\frac{\lambda \cdot (1-p)}{\lambda - t}\right)^k \end{aligned}$$

To evaluate the sum, we can use the formula for the sum of a geometric series:

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r}$$

In this case, $r = \frac{\lambda}{\lambda-t} \cdot (1-p)$. So, we have:

$$\begin{aligned} M_Z(t) &= \frac{p}{1-p} \cdot \frac{\lambda(1-p)}{\lambda-t} \sum_{k=0}^{\infty} \left(\frac{\lambda(1-p)}{\lambda-t}\right)^k \\ &= \frac{p\lambda}{\lambda-t} \cdot \frac{1}{1-\frac{\lambda(1-p)}{\lambda-t}} \\ &= \frac{p\lambda}{p\lambda - t} \end{aligned}$$

Now, we have the MGF of Z, which is $M_Z(t) = \frac{p\lambda}{p\lambda-t}$.

The MGF of an exponential random variable with parameter λ is $M_Y(t) = \frac{\lambda}{\lambda-t}$. Comparing $M_Z(t)$ to the MGF of an exponential random variable, we see that Z follows an exponential distribution with parameter λp.

Therefore, Z is an exponential random variable with parameter λp.

Marking Scheme :-

- 1 mark for writing MGF of Z as summation for all k
- 1 mark for expanding formula using MGF of X_i and using P(N=k)
- 2 marks for deriving final MGF of Z to prove it exponential
- 1 mark for writing parameter for Z

Question 2

Let Z = X + Y where X and Y are independent exponential random variables with parameters λ_1 and λ_2 respectively. Obtain the pdf, cdf and MGF of Z.

Solution 2

For the PDF of Z, we have

$$\begin{aligned}
f_Z(z) &= \int_{-\infty}^{\infty} f_X(x) \cdot f_Y(z-x)\, \mathrm{d}x \\
&= \int_0^z \lambda_1 e^{-\lambda_1 x} \cdot \lambda_2 e^{-\lambda_2 (z-x)}\, \mathrm{d}x \text{ when } x > 0 \text{ , } z - x > 0 \\
&= \int_0^z \lambda_1 \lambda_2 e^{-\lambda_2 z} e^{-\lambda_1 x + \lambda_2 x}\, \mathrm{d}x \\
&= \lambda_1 \lambda_2 e^{-\lambda_2 z} \int_0^z e^{(\lambda_2 - \lambda_1) x}\, \mathrm{d}x \\
&= \lambda_1 \lambda_2 e^{-\lambda_2 z} \frac{\left[e^{(\lambda_2 - \lambda_1)x}\right]_0^z}{\lambda_2 - \lambda_1}
\end{aligned}$$

Therefore,

$$f_Z(z) = \begin{cases} \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} \left[e^{-\lambda_1 z} - e^{-\lambda_2 z}\right] & z > 0 \\ 0 & \text{otherwise} \end{cases}$$

For the CDF of Z, we have

$$\begin{aligned}
F_Z(z) &= \int_{-\infty}^{z} f_Z(z)\, \mathrm{d}z \\
&= \int_0^z \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} \left[e^{-\lambda_1 z} - e^{-\lambda_2 z}\right] \\
&= \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} \left[\frac{e^{-\lambda_2 z}}{\lambda_2} - \frac{e^{-\lambda_1 z}}{\lambda_1}\right]_0^z \\
&= \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} \left(\frac{e^{-\lambda_2 z}}{\lambda_2} - \frac{e^{-\lambda_1 z}}{\lambda_1}\right) + 1
\end{aligned}$$

For the MGF of Z, we have

$$\begin{aligned}
\mathrm{M}_Z(t) &= \mathbb{E}(e^{tZ}) \\
&= \int_{-\infty}^{\infty} e^{tz} f_Z(z)\, \mathrm{d}z \\
&= \int_{0}^{\infty} \frac{\lambda_1\lambda_2}{\lambda_2-\lambda_1}\left[e^{(t-\lambda_1)z} - e^{(t-\lambda_2)z}\right] \mathrm{d}z \\
&= \frac{\lambda_1\lambda_2}{\lambda_2-\lambda_1}\left[\frac{e^{(t-\lambda_1)z}}{t-\lambda_1} - \frac{e^{(t-\lambda_2)z}}{t-\lambda_2}\right]_0^{\infty} \\
&= \frac{\lambda_1\lambda_2}{\lambda_2-\lambda_1}\left[\frac{1}{t-\lambda_1} - \frac{1}{t-\lambda_2}\right] \text{ when } \lambda_1, \lambda_2 > t \\
&= \frac{\lambda_1\lambda_2}{(t-\lambda_1)(t-\lambda_2)}
\end{aligned}$$

Possible Alternative Solutions

- ∗ MGF of Z could also be calculated by calculating MGFs of X and Y separately and then using properties of MGF.
- ∗ CDF (and PDF) of Z can also be calculated via the joint probability formulation.

Marking Scheme :-

- 2 marks for finding MGF
- 1.5 marks for finding pdf
- 1.5 marks for finding cdf

Question 3

Let $Z = X - Y$ where X and Y are independent and geometric(p), Find pmf of Z.

Solution 3

The geometric distribution with parameter p is defined as:

$$P(X = k) = P(Y = k) = (1-p)^{(k-1)}p$$

where k = 1,2,3 and so on.
Now, we want to find $P(Z = z)$ where $Z = X - Y$, so:

$$P(Z = z) = \sum P(X = z + y, Y = y)$$

Since X and Y are independent, we can separate the joint probability into the product of their individual probabilities:

$$P(Z = z) = \sum P(X = z + y)P(Y = y)$$

Case 1: z is positive
Now, substitute the geometric distribution probabilities:

$$P(Z = z) = \sum_{y=1,2,3..} (1 - p)^{(z+y-1)} p(1 - p)^{(y-1)} p$$

The summation is over all non-negative integers y:

$$P(Z = z) = \sum_{y=1,2,3..} (1 - p)^{(z+2y-2)} p^2$$

$$P(Z = z) = p^2 \sum_{y=1,2,3..} (1 - p)^{(z+2y-2)}$$

The summation represents a geometric series.

$$P(Z = z) = p^2 (1 - p)^{(z-2)} \sum_{y=1,2,3..} (1 - p)^{2y}$$

$$P(Z = z) = p^2 (1 - p)^{(z-2)} \frac{(1 - p)^2}{1 - (1 - p)^2}$$

$$P(Z = z) = \frac{p^2 (1 - p)^z}{2p - p^2}$$

$$P(Z = z) = \frac{p(1 - p)^z}{2 - p}$$

Case 2: z is negative
When $Z \geq 0$, the above formula holds, but when Z is negative Y should start at a value that assures X is positive. That is, y takes value starting from $|z| + 1$.

$$P(Z = z) = p^2 \sum_{y=|z|+1,|z|+2..} (1 - p)^{(z+2y-2)}$$

$$P(Z = z) = p^2 \sum_{k=1,2,3..} (1 - p)^{(z+2|z|+2k-2)}$$

$$P(Z = z) = p^2 (1 - p)^{(|z|-2)} \sum_{k=1,2,3..} (1 - p)^{2k}$$

$$P(Z = z) = p^2 (1 - p)^{(-z-2)} \frac{(1 - p)^2}{1 - (1 - p)^2}$$

$$P(Z=z) = \frac{p^2(1-p)^{-z}}{2p-p^2}$$
$$P(Z=z) = \frac{p(1-p)^{-z}}{2-p}$$

To summarize,

$$P(Z=z) = \frac{p(1-p)^{|z|}}{2-p}, \, k \in Z$$

and 0 otherwise.

Marking Scheme :-

- 1 mark for geometric distribution formula.
- 1 mark for solving the equation in terms of either x or y.
- 1 mark mentioning the two cases.
- 2 marks for rest of the calculation (and independence property).

Question 4

Break a stick of length L at a point which is chosen uniformly over its length. Now take the piece that has the right end point and repeat this once again (that is break it again at a uniformly chosen point and retain the piece with the right end). What is the expected length of the remaining stick?

Solution 4

We start with a stick of length L and break it at a point, which we denote as x. Now, if we were to choose the "left" end of the stick then the distribution of the length of stick will be $U[0, L]$, similarly if we were choose the "right" end of the stick then also the distribution of the length of the stick will be $U[0, L]$ since the choice of origin is arbitrary and we can exploit this symmetry and keep it at the "right" end of the stick.
This tells us that the expected value of the length after the second cut will remain the same regardless of the choice of the "left" or the "right" end as long as we are consistent with both the cuts, the reason for this is that the distribution of the length of the second cut only depends on the distribution of the length of stick after the first cut.
Let X be the random variable for the length of the stick after the first cut and let Y be the random variable for the length of the stick after the second cut. We are given that $X \sim U[0, L]$ and $Y \sim U[0, X]$. The pdf of X and $Y \mid X$ will then become:

$$f_X(x) = \begin{cases} \frac{1}{L} & \text{for } x \in [0, l] \\ 0 & \text{otherwise} \end{cases}$$

$$f_{Y|X=x}(y) = \begin{cases} \frac{1}{x} & \text{for } y \in [0, x] \\ 0 & \text{otherwise} \end{cases}$$

By the law of total expectation, we know that $\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y \mid X]]$, so we will first find $\mathbb{E}[Y \mid X]$:

$$\begin{aligned} \mathbb{E}[Y \mid X = x] &= \int_0^x y f_{Y|X}(y)\, dy \\ &= \frac{1}{x} \int_0^x y\, dy \\ &= \frac{x}{2} \end{aligned}$$

Now, $E[Y]$ will become:

$$\begin{aligned} \mathbb{E}[Y] &= \mathbb{E}[\mathbb{E}[Y \mid X = x]] \\ &= \int_0^L f_X(x) \mathbb{E}[Y \mid X = x]\, dx \\ &= \frac{1}{L} \int_0^L \frac{x}{2}\, dx \\ &= \frac{L}{4} \end{aligned}$$

Marking Scheme :-

- 1 mark for writing the pdf of $Y \mid X$
- 1 mark for calculating $\mathbb{E}[Y|X]$
- 3 marks for using law of total expectation and calculating $\mathbb{E}[Y]$

Question 5

1. Find the MGF of gamma distribution X with parameters α and β.
2. Let Y be gamma distribution with parameters λ and β. Define Z = X + Y, find the MGF of Z, and show that it is also a gamma distribution. Report the parameters of that gamma distribution.

Solution 5

1. We know MGF = $E[e^{tx}]$ and gamma distribution is defined as -

$$Gamma(\alpha, \beta) = \begin{cases} \dfrac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta x} x^{\alpha-1}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

   Calculating the MGF is straightforward. Just note that the domain of integration is $[0, \infty]$, as $x < 0$ pdf is 0 anyway. For the below calculation assume $\beta > t$. We will later see what happens otherwise.

$$\begin{aligned} E[e^{tx}] &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{tx} e^{-\beta x} x^{\alpha-1} dx \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{-x(\beta-t)} x^{\alpha-1} dx \\ &\text{Let } x' = x(\beta - t) \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{-x'} \frac{x'^{\alpha-1}}{(\beta-t)^{\alpha-1}} \frac{dx'}{\beta - t} \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)(\beta-t)^\alpha} \int_0^\infty e^{-x'} x'^{\alpha-1} dx' \\ &= \left(\frac{\beta}{\beta - t}\right)^\alpha \end{aligned}$$

   If we take $\beta = t$, it is just integral of $x^{\alpha-1}$ which is x^α, on limits 0 to ∞ it blows up. Now if $\beta < t$, it is $\int_0^\infty e^{kx} x^{k'} dx$, where k is positive. As exponential dominates polynomials, the product blows up to infinity again. Hence, the region of convergence is $t < \beta$.

2. Let Z = X + Y. MGF of Z = $E[e^{tz}]$, which is nothing but $E[e^{tx}e^{ty}]$. As X and Y are independent, we can write it as $E[e^{tx}]E[e^{ty}]$. Which is the product of there respective MGFs. Substituting the formula for MGF of gamma distribution we get

$$\begin{aligned} MGF_Z(t) &= \left(\frac{\beta}{\beta - t}\right)^\alpha \left(\frac{\beta}{\beta - t}\right)^\lambda \\ &= \left(\frac{\beta}{\beta - t}\right)^{\alpha+\lambda} \end{aligned}$$

   Clearly if Z $\sim$ $Gamma(\alpha + \lambda, \beta)$ then it would have above MGF. By uniqueness property, Z is the aforementioned distribution itself.

Marking Scheme :-

1. Find MGF of X - 2.5 marks
2. Find ROC of MGF of X - 2.5 marks
3. Find MGF of Z - 2.5 marks
4. Show Z is Gamma distribution and write its parameters - 2.5 marks