Probability & Statistics

Assignment : 3

Ayush Sharma (2019101004)

Q.1 = For RV, 'X' pdf = $f_X(x)$.
& $Y = aX + b$.

(1) finding $f_Y(y)$?

⇒ First let us find out $F_Y(y)$ i.e.

$$P(Y \leq y) = P(aX + b \leq y)$$

⇒

$$F_Y(y) = P\left(X \leq \left(\frac{y-b}{a}\right)\right)$$

Since, X is continuous RV

$\Rightarrow$

$$F_Y(y) = \int_{-\infty}^{(y-b)/a} f_X(x)\,dx$$

$$F_Y(y) = F_X\left(\frac{y-b}{a}\right)$$

Differentiating both side w.r.t y;

$$f_Y(y) = f_X\left(\frac{y-b}{a}\right)\left|\frac{1}{a}\right| .$$

; $Y = aX+b$ & $a \neq 0, b \neq 0$.

(2) $X \sim \text{exponential}(\lambda)$

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} ; & x > 0 \\ 0 ; & \text{otherwise} \end{cases}$$

So, $$f_Y(y) = f_X\left(\frac{y-b}{a}\right) = \begin{cases} \dfrac{\lambda e^{-\lambda\left(\frac{y-b}{a}\right)}}{|a|} ; & y > b \ \& \ a > 0 \\ 0 ; & \text{otherwise} \end{cases}$$

Thus,

RV 'Y' will be exponential only when $b = 0$ & $a > 0$ but a & b are non-zero.

So, Y can never be exponential RV.

PAGE NO.: ______
DATE: / /

(3) $X \sim N(\mu, \sigma^2)$

that is

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad \forall x \in \mathbb{R}$$

and from first part

$$f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right)$$

$$f_Y(y) = \frac{1}{(a\sigma)\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{y-(a\mu+b)}{a\sigma}\right)^2}$$

$$\forall y \in \mathbb{R}$$

$f_Y(y) \sim N(a\mu+b, (a\sigma)^2)$.

(!! always !!)

PAGE NO.:
DATE: / /

Q.2 = $X \sim N(\mu, \sigma^2)$

that means if $Z = N(0,1)$

then, $X = \sigma Z + \mu$.

And already $\phi(-Z) = 1 - \phi(Z)$; $\phi(z) =$ CDF of Z

And $Y = e^X$.

(1) $F_Y(y) = P(Y \le y) = P(e^X \le y) = P(X \le \ln y)$

So, $F_Y(y) = F_X(\ln y)$ —— ①

&

$F_X(x) = P(X \le x) = P(\sigma Z + \mu \le x)$

so, $F_X(x) = F_Z\left(\frac{x-\mu}{\sigma}\right) = \phi\left(\frac{x-\mu}{\sigma}\right)$

—— ⑪

from ① & ⑪

$F_y(y) = F_X(\ln y) = \phi\left(\frac{\ln y - \mu}{\sigma}\right)$

So;

$$F_Y(y) = \begin{cases} \phi\left(\frac{\ln y - \mu}{\sigma}\right) & ; y > 0 \\ 0 & ; \text{otherwise} \end{cases}$$

where $\phi(m_1) = P(Z \le m_1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{m_1} e^{\frac{-m^2}{2}} dm$.

(2) Without method of transformation :-

As we know, $f_Y(y) = F'_Y(y)$

& we from part 1 of this question:

$$F_Y(y) = \phi[(\ln y - \mu)/\sigma] = \phi[k]$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{K} e^{-\frac{m^2}{2}}\, dm$$

where $k = (\ln y - \mu)/\sigma$; $y > 0$

Otherwise 0.

$$f_Y(y) = F'_Y(y)$$

$$= \phi'[k] \cdot \frac{dk}{dy}$$ ~~(using Leibnitz rule)~~

$$= \frac{1}{\sqrt{2\pi}} e^{-k^2/2} \cdot \frac{dk}{dy}$$

And $\therefore$, $dk = \frac{1}{\sigma} \frac{1}{y} dy$

So,
$$f_Y(y) = \begin{cases} \frac{1}{(\sqrt{2\pi})} \cdot \frac{1}{(\sigma y)} \cdot e^{-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2} & ; y > 0 \text{ or } y = e^X. \\ 0 ; & \text{otherwise} \end{cases}$$

With method of transformation :—

$Y = e^X = g(x)$ is continuous & differentiable function hence we can use method of transformation.
[Also, it is strictly increasing].

$$f_Y(y) = \begin{cases} \dfrac{f_X(x_1)}{g'(x_1)} = f_X(x_1) \cdot \dfrac{dx_1}{dy} \; ; \; y = g(x_1) \\ 0 \quad ; \text{if } y = g(x) \text{ does not have solution.} \end{cases}$$

And since $X \sim N(\mu, \sigma^2)$;
hence,
$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left(\frac{x-\mu}{\sigma}\right)^2 \frac{1}{2}} \; ; \forall x$$

And $g'(x) = dy/dx = e^x$

So, $f_Y(y) = \left(\dfrac{1}{\sigma\sqrt{2\pi}} e^{-\left(\frac{x-\mu}{\sigma}\right)^2 \frac{1}{2}} \right) \Big/ e^x$

Putting $x = \ln y$

$$f_Y(y) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} \dfrac{e^{-\left(\frac{\ln y-\mu}{\sigma}\right)^2 \frac{1}{2}}}{(\sigma y)} \; , \quad y = g(x) \text{ or } y > 0. \\ 0 \; ; \quad \text{otherwise} \end{cases}$$

Q.3 = Sol^n ⇒

We can use Normal approximation of binomial distribution because here $n = 12{,}000$ & $p = \frac{1}{6}$

So, $np = 2000 > 5$

& $n(1-p) = 10{,}000 > 5$.

Therefore, also according to central limit theorem

$$Z = \frac{Y - np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0,1)$$

→ mean

→ variance

where Y is RV for no. of 2 in given experiment i.e. rolling a die 12,000 times.

Approximation ⇒

$P(1900 < Y < 2150) = ?$ (by continuity correction)

$= P(1900.5 \leq Y \leq 2149.5)$

$$= P\left(\frac{1900.5 - 2000}{\sqrt{2000 \times \frac{5}{6}}} \leq Z \leq \frac{2149.5 - 2000}{\sqrt{2000 \times \frac{5}{6}}}\right)$$

Scanned with CamScanner

$$\Rightarrow P\left(\frac{-99.5}{40.8248} < Z < \frac{149.5}{40.8248}\right)$$

$$\Rightarrow P(-2.43724 < Z < 3.66199)$$

$$\Rightarrow \phi(3.66199) - \phi(-2.43724)$$

$$\Rightarrow \phi(3.66199) + \phi(2.43724) - 1$$

(using Normal Distribution Property)

$$\simeq \phi(3.66) + \phi(2.44) - 1$$

$$\simeq 0.99987 + 0.99266 - 1$$

$$\simeq 0.99253 \simeq 0.992 \quad \text{Ans}$$

Q.4 = Solⁿ

(1) Let B represents boy & G represent girl. Then the possible set for two children is

$U = \{ (B,B), (B,G), (G,B), (G,G) \}$

Given one of them is boy, our sample space reduces to

$S = \{ (B,G), (B,B), (G,B) \}$

P(other is a girl) in the above sample space S is

$P(B,G) + P(G,B)$

$= \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$. Ans

(2) Given that the child, Atsi is boy, now S ~~then~~ ~~be~~ will ~~either~~ be $\{ (B,G), (B,B) \}$

so,

P(other child is girl in S) = P(B,G)

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

(3) Yes, the answer for part a & part b is different.

The reason is :-

when we named the boy i.e. Atsi is a boy, sample space got reduced from

$S \rightarrow S'$ i.e. deleting the option of having (G, B)

due to unique identification of Atsi being a boy.

Hence the probability decreased from $\frac{2}{3}$ to $\frac{1}{2}$.

PAGE NO.:
DATE: / /

(4) Since, we named the child let Atsi be 1^{st}, without loss of generality.

$$S = \{GG, GB, BG, BB\}$$

Let A be an event Atsi is a boy. E be other s.t. other child is a girl given one is boy

$$P(A) = \frac{P(\{BG, GB, BB\})}{P(S)} = \frac{3}{4}$$

$P(E) = 2/3$ from 1^{st} part

Now if other event B is defined as Atsi's sibling is a girl given Atsi is a boy.

then

$$P(B) = P(A \cap E) = P(A)\,P(E)$$

as A & E are independent

$$P(B) = \frac{3}{4} \times \frac{2}{3} = \frac{1}{2}.$$

Ans

Scanned with CamScanner

Q.5 = (1) $Z = X + Y = G(X,Y)$

$X = \text{Exponential}(\lambda)$, $Y = \text{Exponential}(\lambda)$

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} ; & x > 0 \\ 0 ; & x \le 0 \end{cases}$$

$$f_Y(y) = \begin{cases} \lambda e^{-\lambda y} ; & y > 0 \\ 0 ; & y \le 0 \end{cases}$$

We can observe, $0 < Z$ always.

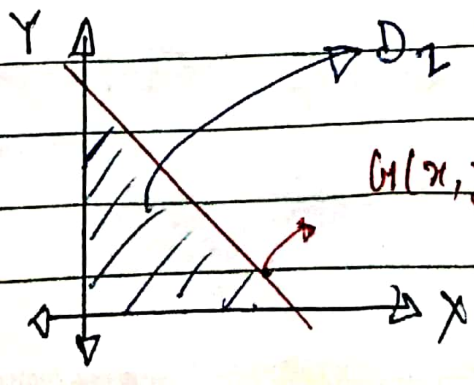
In general,

$$F_Z(z) = P(Z \le z) = P(G(x,y) \le z)$$

$$= P((x,y) \in D_Z)$$

$$= \iint_{x,y \in D_Z} f_{X,Y}(x,y)\, dx\, dy$$

where D_Z is the region in the X-Y plane where $G(X,Y) \le Z$ is satisfied.



(since, $x \not< 0$, $y \not< 0$ & $x+y \not> z$)

$G(x,y) = x + y = z$

Therefore,

$$F_Z(z) = \int_{y=0}^{z} \int_{x=0}^{z-y} f_{XY}(x,y)\, dx\, dy$$

Now, finding $f_{XY}(x,y)$?

Since, X & Y are independent

$$f_{XY}(x,y) = f_X(x) \cdot f_Y(y)$$

$$= \lambda e^{-\lambda x} \cdot \lambda e^{-\lambda y}$$

$$= \lambda^2 e^{-\lambda x} e^{-\lambda y}$$

So, $$F_Z(z) = \int_{y=0}^{z} \lambda^2 e^{-\lambda y} \int_{x=0}^{z-y} e^{-\lambda x}\, dx\, dy$$

$$= \int_{y=0}^{z} \lambda^2 e^{-\lambda y} \left[\frac{1}{\lambda} - \frac{e^{-\lambda(z-y)}}{\lambda} \right] dy$$

$$= \lambda \int_{y=0}^{z} \left(e^{-\lambda y} - e^{-\lambda z} \right) dy$$

$$= \lambda \left[\frac{e^{-\lambda y}}{-\lambda} - (e^{-\lambda z}) y \right]_0^z$$

$$= \lambda \left[\left(\frac{e^{-\lambda z}}{-\lambda} \right) - z(e^{-\lambda z}) + \frac{1}{\lambda} \right]$$

$$= 1 - e^{-\lambda z}(1 + z\lambda)$$

$$; \quad z > 0.$$

So, $f_Z(z) = \frac{d}{dz} F_Z(z)$

$$= \frac{d}{dz} \left(1 - e^{-\lambda z}(1 + z\lambda) \right)$$

$$= \lambda e^{-\lambda z}(1 + \lambda z) - \lambda e^{-\lambda z}$$

$$= \lambda e^{-\lambda z}(1 + \lambda z - 1)$$

$$f_Z(z) = \begin{cases} \lambda^2 z e^{-\lambda z} & ; \quad z > 0 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Scanned with CamScanner

Q.5 ⇒

(ii) $Z = X + Y = G(X,Y)$

$X = \text{Uniform}(0,1)$, $Y = \text{Uniform}(0,1)$

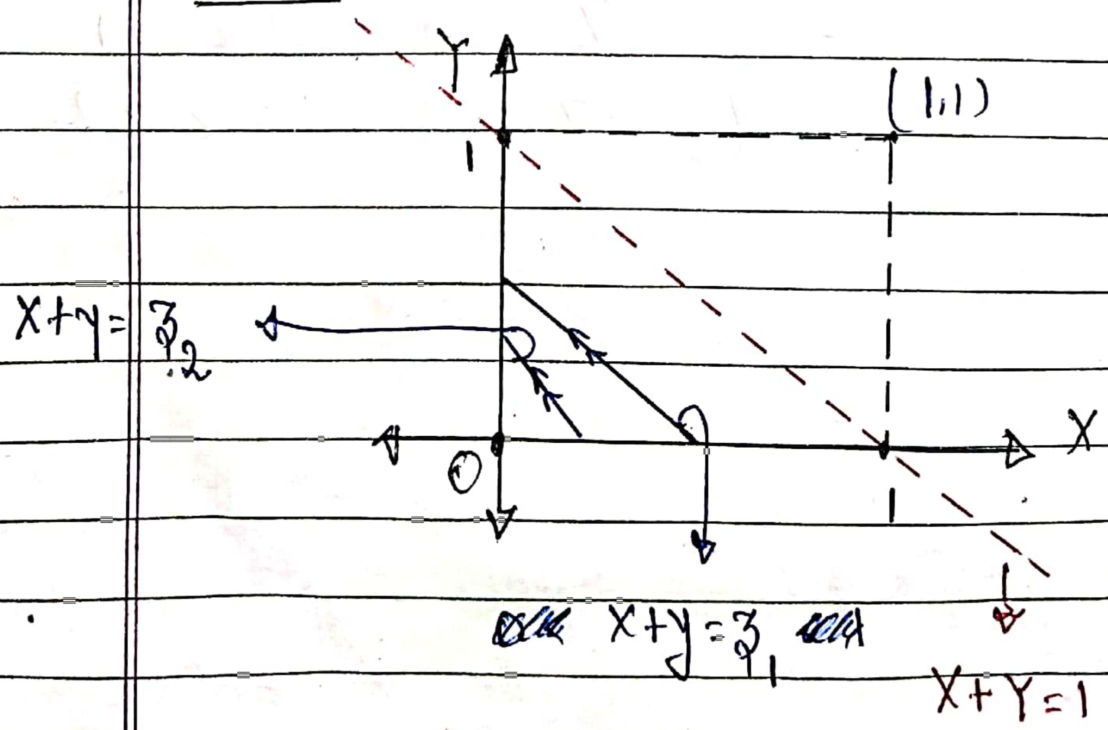
$$f_Y(y) = \begin{cases} 1 ; & y > 0 \ \& \ y < 1 \\ 0 ; & \text{otherwise} \end{cases}$$

$$f_X(x) = \begin{cases} 1 ; & 0 \leq x \leq 1 \\ 0 ; & \text{otherwise} \end{cases}$$

We can directly say $0 \leq Z \leq 2$.

Lets break Z in two range ① $0 \leq Z \leq 1$
② $1 \leq Z \leq 2$

<u>Case : 1</u> :- $0 \leq Z \leq 1$



DATE: / /

In general,

$$F_Z(z) = P(Z \le z) = P(G(X,Y) \le z)$$

$$= P((X,Y) \in D_z)$$

$$= \iint_{x,y \in D_Z} f_{XY}(x,y)\, dx\, dy$$

where D_Z in the X,Y plane represents region such that $G(x,y) \le z$ is satisfied.

Now observing the graph:
we can directly say $1 > z_1 > z_2$.

Therefore our D_Z will be area formed by following lines:-

$(X=0)$, $(Y=0)$ & $(x+y=z \;; 0<z<1.)$

$$\text{So, } F_Z(z) = \int_{y=0}^{z} \int_{x=0}^{z-y} f_{XY}(x,y)\, dx\, dy \quad \text{--- (1)}$$

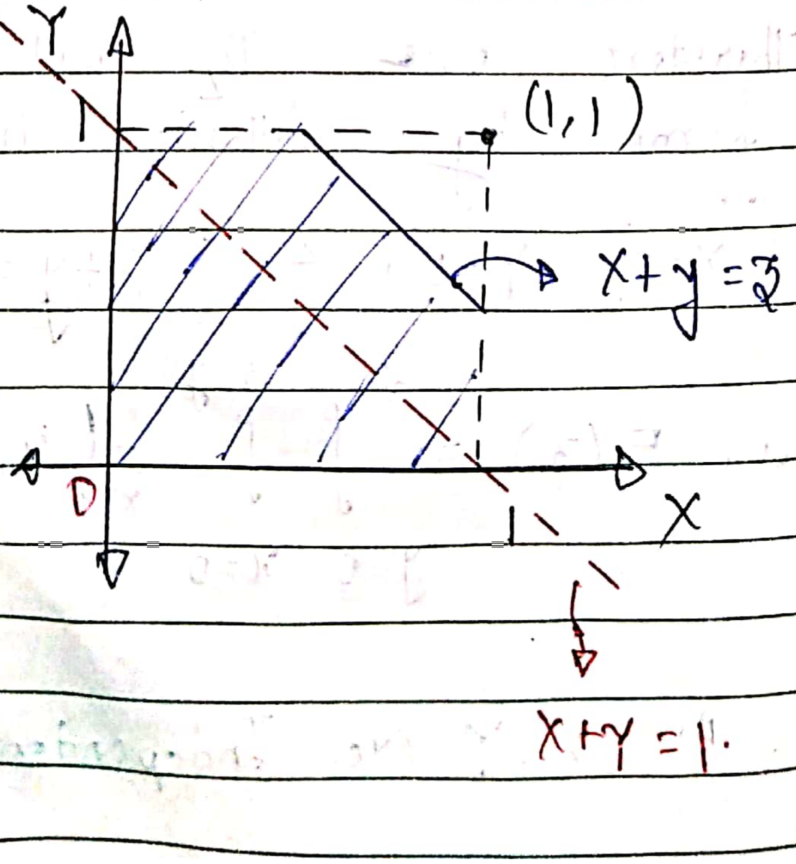
As X, Y are Independent;

$$f_{XY}(x,y) = f_X(x) \cdot f_Y(y)$$
$$= 1.1$$
$$= 1.$$

Therefore,

$$F_Z(z) = \int_{y=0}^{z} \int_{x=0}^{z-y} 1.\, dx\, dy$$
$$= \int_{y=0}^{z} (z-y)\, dy$$
$$= z^2/2.$$

<u>Case : 2</u> :- $1 \leq z \leq 2.$



In this case;
for finding $F_Z(z)$, our D_Z will be the area in the above X-Y plane where (x,y) lie inside square formed by points $(0,0), (0,1), (1,0), (1,1)$ And $x+y < z$.
That is shaded region.
Calculating that way is very complex.
Otherway we can say,

$$F_Z(z) = P(Z < z) = 1 - P(Z > z)$$

& $$P(Z > z) = P((x,y) \in D'_Z)$$

$$= \iint_{x,y \in D'_Z} f_{XY}(x,y)\, dx\, dy$$

where D'_Z is the X-Y plane represents region s.t $G(X,Y) > Z$ & $0 < X < 1$ & $0 < Y < 1$.

i.e. unshaded part in the formed square.

So, $$P(Z > z) = \int_{y=z-1}^{1} \int_{x=z-y}^{1} f_{XY}(x,y)\, dx\, dy$$ from case (I)

$$= \int_{y=z-1}^{1} \int_{x=z-y}^{1} 1 \cdot dx \cdot dy \quad \left\{ f_{XY}(x,y) = 1. \right.$$

Scanned with CamScanner

So, $P(Z > z) = \int_{y=z-1}^{1} (1 - z + y)\, dy$

$$= \frac{(2-z)^2}{2} \; ; \quad 1 \le z < 2.$$

So, $F_Z(z) = 1 - [P(Z > z)]$

$$= 1 - (2-z)^2/2 \; ; \quad 1 \le z < 2.$$

Since,

$$f_Z(z) = \frac{d}{dz}\left[F_Z(z)\right]$$

$$= \begin{cases} z & ; \; 0 \le z \le 1 \\ 2 - z & ; \; 1 \le z \le 2 \\ 0 & ; \; \text{otherwise} \end{cases}$$

An

Q. 6 =

(1) $T = T_1 - T_2 = G(T_1, T_2)$

$T_1 = \text{Exponential}(\lambda)$, $T_2 = \text{Exponential}(\lambda)$

$$T_1 = \begin{cases} \lambda e^{-\lambda t_1} ; & t_1 > 0 \\ 0 ; & \text{otherwise} \end{cases}$$

$$T_2 = \begin{cases} \lambda e^{-\lambda t_2} ; & t_2 > 0 \\ 0 ; & \text{otherwise} \end{cases}$$

We can observe $T \in \mathbb{R}$.
Finding $F_T(t) \ \forall \ t \geq 0$. ?

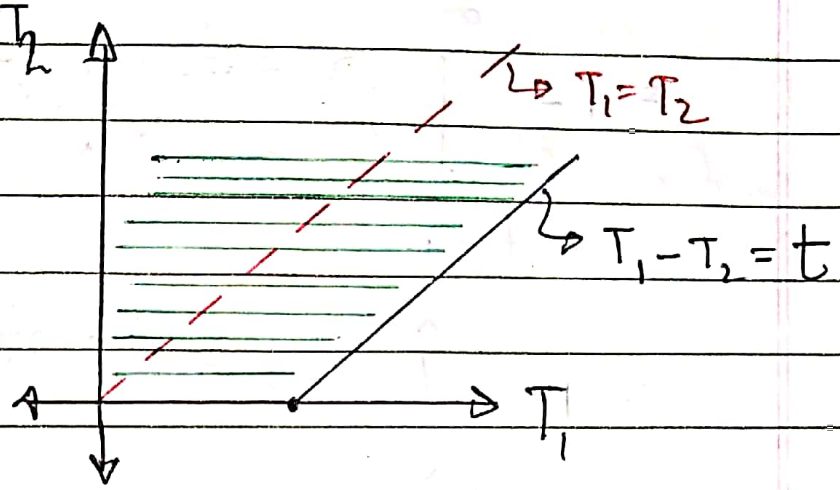
$$F_T(t) = P(T \leq t) = P(\,G(T_1, T_2) \leq t)$$

$$= P((t_1, t_2) \in D_z)$$

$$= \iint_{t_1, t_2 \in D_z} f_{T_1, T_2}(t_1, t_2)\, dT_1\, dT_2$$

where D_z is the regions in the $T_1 - T_2$ plane where $G(t_1, t_2)$ $\leq T$ is satisfied.

D_2 = ~~the~~ shaded portion of graph below.



We can write :-

~~$F_T(t) = 1 - P((T_1, T_2) \in D_2')$~~

~~where D_2' is region where $G(T_1, T_2) > T$.~~

$$\text{So, } F_T(t) = \int_{t_2=0}^{\infty} \int_{t_1=0}^{t+t_2} f_{T_1T_2}(t_1, t_2)\, dT_1\, dT_2$$

Since, T_1 & T_2 are independent hence

$$f_{T_1T_2}(t_1, t_2) = f_{T_1}(t_1) \cdot f_{T_2}(t_2)$$

$$= \lambda^2 e^{-\lambda t_1} e^{-\lambda t_2}$$

So,

$$F_T(t) = \int_{t_2=0}^{\infty} \lambda^2 e^{-\lambda t_2} \left[\int_{t_1=0}^{t+t_2} e^{-\lambda t_1} dt_1 \right] dt_2$$

$$F_T(t) = \int_{t_2=0}^{\infty} \frac{\lambda^2 e^{-\lambda t_2}}{\lambda} \left[1 - e^{-\lambda(t+t_2)} \right] dt_2$$

$$F_T(t) = \lambda \int_{t_2=0}^{\infty} \left[e^{-\lambda t_2} - e^{-\lambda(t+2t_2)} \right] dt_2$$

$$F_T(t) = \lambda \left[\frac{e^{-\lambda t_2}}{(-\lambda)} - \frac{e^{-\lambda(t+2t_2)}}{(-2\lambda)} \right]_0^{\infty}$$

$$F_T(t) = \lambda \left[(0 - 0) - \left(\frac{1}{(-\lambda)} - \frac{e^{-\lambda t}}{-2\lambda} \right) \right]$$

$$F_T(t) = 1 - \frac{e^{-\lambda t}}{2} \quad ; \quad t \geq 0.$$

Ans

(2) $T < 0$

$T = T_1 - T_2$

& assume $T' = T_2 - T_1$

then

$$P(T > t) = P(T' > t)$$

due to independence of T_1, T_2 and also due to symmetry.

Now, $P(T \leq t) = P(T' \geq -t)$

$= P(T \geq -t)$

$= 1 - P(T \leq (-t))$

$$P(T \leq t) + P(T \leq (-t)) = 1.$$

So,

$$F_T(t) + F_T(-t) = 1$$

Thus,

$$F_T(-t) = 1 - F_T(t)$$

as for $t \geq 0$, $F_T(t) = 1 - \frac{e^{-\lambda t}}{2}$.

So, $F_T(-t) = \frac{e^{-\lambda t}}{2}$; $t \geq 0$

or

$$F_T(q) = \frac{e^{\lambda q}}{2} \; ; \; q \leq 0.$$

Ans

(3) $$F_T(t) = \begin{cases} 1 - \dfrac{e^{\lambda t}}{2} & ; \; t \geq 0 \\ \dfrac{e^{\lambda t}}{2} & ; \; t < 0 \end{cases}$$

$$f_T(t) = \frac{d}{dt}\left(F_T(t)\right)$$

$$= \begin{cases} \dfrac{\lambda e^{-\lambda t}}{2} & ; \; t \geq 0 \\ \dfrac{\lambda e^{\lambda t}}{2} & ; \; t < 0 \end{cases}$$

So,

$$f_T(t) = \left(\frac{\lambda}{2}\right) e^{-\lambda|t|}$$

$$\forall \; t.$$

(4) It is Laplace distribution pdf where $\mu = 0$ & $b = \frac{1}{\lambda}$.

Laplace pdf:

$$f(x \mid \mu, b) = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}}.$$

Scanned with CamScanner

Q.7 = Solⁿ ⇒

(1) Let us consider the waiting time of Atsi for the next train to be T minutes where $T \in [0, 10)$

Now as the distribution of arrival of Atsi is uniform, values taken by T are also uniformly distributed between 0 to 10.

So, T is a RV in range $[0, 10)$.

Thus,

mean waiting time $= E[T]$

so, $E[T] = \int_0^{10} t \cdot \left(\frac{1}{10}\right) dt$

$$= \frac{1}{20}\left[t^2\right]_0^{10}$$

$$= \frac{100}{20}$$

$$= 5 \text{ min}$$ Ans

(2) Answer = $P(T \geq 6+3 \mid T \geq 6)$

$$= \frac{P((T \geq 9) \cap (T \geq 6))}{P(T \geq 6)}$$

$$= \frac{P(T \geq 9)}{P(T \geq 6)}$$

$$= \frac{\frac{10-9}{10-0}}{\frac{10-6}{10-0}}$$

$$= \frac{1/10}{4/10}$$

$$= 1/4 \text{ Ans}$$

(3) Let the time after which the train arrives after Atsi reach the station be T.

Now it's given that T' is exponential & $E[T'] = 10$ min ~~and~~ & where T' is time of arrival of second train after 1st train left.

Since, exponential distribution is memoryless. Atsi's waiting time i.e. T will also be exponential with same mean $E[T] = 10$ min.

Let λ be parameter of distribution.

$$E[T] = \int_0^\infty t \cdot \lambda e^{-\lambda t} \, dt = 10$$

$$\Rightarrow \left[-t e^{-\lambda t}\right]_0^\infty + \int_0^\infty e^{-\lambda t} \, dt = 10$$

$$\Rightarrow \left[\frac{e^{-\lambda t}}{-\lambda}\right]_0^\infty = 10$$

$$\Rightarrow \frac{1}{\lambda} = 10 \quad \Rightarrow \lambda = \frac{1}{10} = 0.1.$$

Hence, Atsi's waiting time ~ Exponential (0.1)

Q.8 = Sol^n

(1) Claim : For a poisson process with parameter λ, the time for a new arrival is exponentially distributed with same parameter λ.

Proof :-

Let for a poisson process (λ) an event occur at time t_1. Assume exactly after t unit time from now another event occur. i.e at '$t+t_2$' time.

Required to show : $f_T(t) = \lambda' e^{-\lambda' t}$.

where $\lambda' = \lambda$.

where T is RV for waiting time.

Proof :-

$P(t > a) = P($no. of arrival at time t = no. of arrival at time '$t+a$')

$= P[$no. of arrival in time interval 'a' $= 0)$

$= [(\lambda a)^0 e^{-\lambda a}]/(0)!$

So, $P(t > a) = e^{-\lambda a}$

So,

$$P(T \leq t) = 1 - P(T > t)$$

$$F_T(t) = 1 - e^{-\lambda t}$$

Differentiating both side w.r.t t;

$$f_T(t) = \begin{cases} \lambda e^{-\lambda t} & ; t > 0 \end{cases} \sim \text{exponential}(\lambda).$$

Hence, proved that waiting time for a poisson distribution $P(\lambda)$ is Exponentially distributed (λ')

& $\lambda' = \lambda$.

Scanned with CamScanner

(2) <u>Claim</u> : the waiting time for a Bernoulli process is Geometric.

<u>Proof</u> :-

For a bernoulli process ; [RV 'X']

$$P_X(x) = \begin{cases} p ; & x=1 \text{ i.e. arrival} \\ 1-p ; & x=0 \text{ i.e. not arrival} \\ 0 ; & \text{otherwise} \end{cases}$$

Let 'Y' be RV which shows the waiting time.

$Y=y$ implies that no arrival occurred till $(y-1)^{th}$ time unit & arrived at y^{th} unit time.

So,

$$P(Y=y) = \begin{cases} (1-p)^{(y-1)} (p) ; & y=1,2,3 \ldots \\ 0 ; & \text{otherwise.} \end{cases}$$

Thus, we can see $f_Y(y) = P(Y=y)$ is a Geometric distribution.

Hence, our claim is correct.

(3) Claim :- An exponential distribution can be obtained as a limit of a Geometric Distribution.

Proof :-

As we know, Geometric Distribution is

$$f_X(x) = p(1-p)^x \quad ; \quad x \geq 0$$

and can be interpreted as number of failures in a sequence of Bernoulli trials until the 1st success.

Now, say p becomes very small & let $n \times p = \lambda$; n is total trials & $n \to \infty$.

So, $P(X=x) = f_X(x) = \left(\frac{\lambda}{n}\right)\left(1-\frac{\lambda}{n}\right)^x$

Now when limit $n \to \infty$, $\sum_{x=0}^{\infty} f_X(x) = ?$

$$\lim_{n\to\infty} \sum_{x=0}^{\infty} \left(\frac{\lambda}{n}\right)\left(1-\frac{\lambda}{n}\right)^{x \cdot \frac{1}{n} \cdot n \cdot \frac{1}{\lambda} \cdot \lambda}$$

$$= \lim_{n\to\infty} \sum_{x=0}^{\infty} (\lambda)\left(1-\frac{\lambda}{n}\right)^{\left(\frac{n}{\lambda}\right)\left(\frac{x}{n}\right)\cdot\lambda} \left(\frac{1}{n}\right)$$

$$= \lim_{n\to\infty} \sum_{y=0}^{\infty} (\lambda)\left(1-\frac{\lambda}{n}\right)^{\frac{n}{\lambda}\cdot y \cdot \lambda} \left(\frac{1}{n}\right)$$

(Assumed $y = \frac{x}{n}$)

Therefore,
we can write

$$\lim_{n\to\infty} \sum_{x=0}^{\infty} f_X(x) = \lim_{n\to\infty} \sum_{y=0}^{\infty} (\lambda)\left(1-\frac{\lambda}{n}\right)^{\left(\frac{n}{\lambda}\right)\cdot y\cdot\lambda} \left(\frac{1}{n}\right)$$

where, $y = \frac{x}{n}$.

$\Rightarrow$

$$\lim_{n\to\infty} \sum_{x=0}^{\infty} f_X(x) =$$

$$\int_0^{\infty} \lambda \lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^{\frac{n}{\lambda}\cdot y\cdot\lambda} dy$$

$$= \int_0^{\infty} \lambda \cdot e^{\lim_{n\to\infty} \frac{n}{\lambda}\cdot y\cdot\lambda\cdot\left(-\frac{\lambda}{n}\right)} dy$$

$$= \int_0^{\infty} \lambda \cdot e^{-y\lambda} dy$$

Therefore
for $n\to\infty$ & $p \ll 1$,

$$F_X(\alpha) = \int_0^{\alpha} \lambda\cdot e^{-y\lambda} dy = 1 - e^{-\alpha\lambda}$$

So, $f_X(\alpha) = \lambda e^{-\alpha\lambda}$ i.e. exponential distribution.

Q.9 = Soln (1)

(a) $R(x;\sigma) = \frac{x}{\sigma^2} e^{\frac{-x^2}{2\sigma^2}} \; ; x \geq 0$

$$\text{mean} = \int_{-\infty}^{\infty} x . R(x;\sigma)\, dx = E$$

$$E = \int_{0}^{\infty} \left(\frac{x^2}{\sigma^2}\right) e^{\left(\frac{-x^2}{2\sigma^2}\right)} dx$$

Say $x = \sigma y$ then $dx = \sigma\, dy$

So, $E = \sigma \int_{0}^{\infty} y^2 e^{-y^2/2} dy = \sigma \int_{0}^{\infty} y^2 e^{-y^2/2} dy$

Say, $y^2 = z \Rightarrow 2y\,dy = dz$

$\Rightarrow dy = \frac{dz}{2\sqrt{z}}$

So, $E = \sigma \int_{0}^{\infty} z\, e^{-z/2} \frac{dz}{2\sqrt{z}}$

$$E = \frac{\sigma}{2} \int_{0}^{\infty} \sqrt{z}\, e^{-z/2} dz$$

$$E = \sigma\sqrt{2}\int_{0}^{\infty}\sqrt{(z/2)}\; e^{-(z/2)}\left(\frac{dz}{2}\right)$$

$$E = \sigma\sqrt{2}\;\underbrace{\Gamma(3/2)}_{\text{Gamma function.}}$$

$$E = \sigma\sqrt{2}\;\frac{1}{2}\;\Gamma(1/2) \qquad \left\{ \Gamma(\alpha+1) = \alpha\,\Gamma(\alpha). \right.$$

$$E[x] = \sigma\sqrt{2}\;\frac{1}{2}\;\sqrt{\pi}$$

$$E[x] = E = \sigma\sqrt{\frac{\pi}{2}} \quad \text{Ans}$$

(b) Variance $= ? = E[x^2] - (E[x])^2$

Finding $E[x^2]$? $\quad (x \geq 0)$

$$E[x^2] = \int_{-\infty}^{\infty} \frac{x^3}{\sigma^2}\, e^{-\frac{x^2}{2\sigma^2}}\, dx$$

$$E[x^2] = \int_{0}^{\infty} \frac{x^3}{\sigma^2}\, e^{-\frac{x^2}{2\sigma^2}}\, dx$$

Let $\frac{x^2}{2\sigma^2} = y$.

$\Rightarrow$

$$x\,dx = \sigma^2 dy$$

So, $$E[x^2] = \int_0^\infty 2y\, e^{-y}\, \sigma^2 dy$$

$$E[x^2] = 2\sigma^2 \int_0^\infty y e^{-y} dy$$

$$E[x^2] = 2\sigma^2\, \Gamma(2) = 2\sigma^2$$

because $\Gamma(2) = 1! = 1$.

So, Variance $= 2\sigma^2 - \left(\sigma\sqrt{\frac{\pi}{2}}\right)^2$

$$= \sigma^2\left(2 - \frac{\pi}{2}\right). \quad \text{Ans}$$

(c) Mode of distribution = ?

Mode is the value of x at which pdf is max.

$$R(x;\sigma) = \frac{x}{\sigma^2} e^{-x^2/2\sigma^2} \quad ; (x \geq 0).$$

for max/min : $\frac{d}{dx} R(x;\sigma) = 0.$

$$\Rightarrow \frac{d}{dx}\left(\frac{x}{\sigma^2} e^{-x^2/2\sigma^2}\right) = 0$$

$$\Rightarrow \frac{e^{-x^2/2\sigma^2}}{\sigma^2} + \frac{x}{\sigma^2} e^{-x^2/2\sigma^2}\left(\frac{-x}{\sigma^2}\right) = 0$$

$$\Rightarrow \frac{e^{-x^2/2\sigma^2}}{\sigma^2}\left[1 - \frac{x^2}{\sigma^2}\right] = 0.$$

$$\Rightarrow \frac{e^{-x^2/2\sigma^2}}{\sigma^2} = 0 \quad \text{or} \quad x^2 = \sigma^2$$

$\downarrow$ at $x \to \infty$ $\qquad\qquad$ $\downarrow$ $x = \sigma$ $(x \geq 0)$. $(\sigma \geq 0)$

at $x \to \infty$

$\frac{d^2}{dx^2} R(x;\sigma) \to 0.$

at $x = \sigma$

$\frac{d^2}{dx^2} R(x;\sigma) < 0$ i.e. maxima at $x = \sigma$

Hence, mode $= \sigma$. Ans

Q.9 = (2) Sol^n =>

$$C(x;\gamma) = \frac{1}{\pi\gamma\left[1+\left(\frac{x-x_0}{\gamma}\right)^2\right]}$$

(a) mean $= E[x] = \int_{-\infty}^{+\infty} \frac{x\,dx}{\pi\gamma\left[1+\left(\frac{x-x_0}{\gamma}\right)^2\right]}$

Let $\frac{x-x_0}{\gamma} = z$

$dx = \gamma\, dz$

So, $E[x] = \int_{-\infty}^{\infty} \frac{\cancel{\gamma}(\gamma z + x_0)\,dz}{\pi\cancel{\gamma}[1+z^2]}$

$$\pi E[x] = \int_{-\infty}^{\infty} \frac{\gamma z\,dz}{1+z^2} + \int_{-\infty}^{\infty} \frac{x_0\,dz}{[1+z^2]}$$

$$\pi E[x] = \frac{\gamma}{2}\left[\ln(1+z^2)\right]_{-\infty}^{\infty} +$$

$$x_0\left[\tan^{-1} z\right]_{-\infty}^{\infty}$$

Therefore, $E[x] = \frac{\gamma}{2\pi}\left[\ln[1+z^2]\right]_{-\infty}^{+\infty}$

$+ \quad x_0 \pi$

Since, $\left[\ln(1+z^2)\right]_{-\infty}^{\infty} = \ln(\infty) - \ln(\infty)$

$= \infty - \infty$

(undefined)

Hence, mean for Cauchy's distribution is undefined and does not exist.

(b) Characteristic function of

$$C(x;\gamma) = \frac{1}{\pi\gamma\left[1+\left(\frac{x-x_0}{\gamma}\right)^2\right]}$$

is $\phi_x(t; x_0, x) = E[e^{ixt}]$

$$= \int_{-\infty}^{\infty} e^{ixt}\, C(x;\gamma)$$

$$\phi_x(t; x_0, x) = e^{ix_0 t - \gamma|t|}$$

(i) $t \to 0^+$ ($t = 0+h$, $h>0$ & $h \ll 1$)
$|t| = h$.

$$\frac{d}{dt}\phi_x(t; x_0, x) = \frac{d}{dh}\left(e^{ix_0 h - \gamma h}\right)\Big|_{h=0}$$

$$= (ix_0-\gamma)\, e^{(ix_0-\gamma)h}\Big|_{h=0}$$

$$= ix_0 - \gamma$$

(ii) $t = 0^-$ $\quad (t = 0 - h,\ h > 0,\ h << 1,\ |t| = h)$

$dt = -dh$

$$\frac{d}{dt}\left(\phi_X(t; x_0, x)\right) = \frac{d}{dt}\left(e^{ix_0(-h) - \gamma h}\right)\Bigg|_{h=0}$$

$$= \frac{-d}{dh}\left(e^{ix_0(-h) - \gamma h}\right)\Bigg|_{h=0}$$

$$= -\left(-ix_0 - \gamma\right)$$

$$= ix_0 + \gamma.$$

PAGE NO.:
DATE: / /

Hence, derivative does not exist.

* As we know mean is the 1st derivative of characteristic function at $t=0$, this shows that mean doesnot exist.

(C) The cauchy distribution doesnot have finite moments of any order. Higher even powered raw moments evaluate to ∞ while odd powered raw moments, are undefined because their value are essentially equivalent to $\infty - \infty$, since the two half of the integral diverge & have opposite sign. We can also say that central moments & standardized moments are undefined, since they are all based on mean & mean is undefined. The variance which is the second central moment, is likewise non-existent. As variance is the expected square of the deviation from mean, it is not defined, mean is not defined.
And

(c) remaining :-

The consequence undefined variance is that central limit theorem doesn't apply & certain parameter estimation techniques will therefore breakdown

PAGE NO.:
DATE: / /

Q.10 = Sol^n

Say the required rectangle have points $P_1(x_0, y_0)$, $P_2(x_1, y_1)$, $A(x_0, y_1)$ & $B(x_1, y_0)$.

Let F be an event in which rectangle fits completely in the circle. Due to symmetry, the probability of F taking first point from 1st quadrant will be equal to probability of F taking 1st point (P_1) from any quadrant.

Let us suppose, P_1 falls on the quarter circle in 1st quadrant. We know,

$$y_0 = \sqrt{1 - x_0^2} \quad \text{—} \ (1)$$

For rectangle to be fully inside the circle :—

$$x_0^2 + y_1^2 \leq 1 \quad \& \quad x_1^2 + y_0^2 \leq 1$$

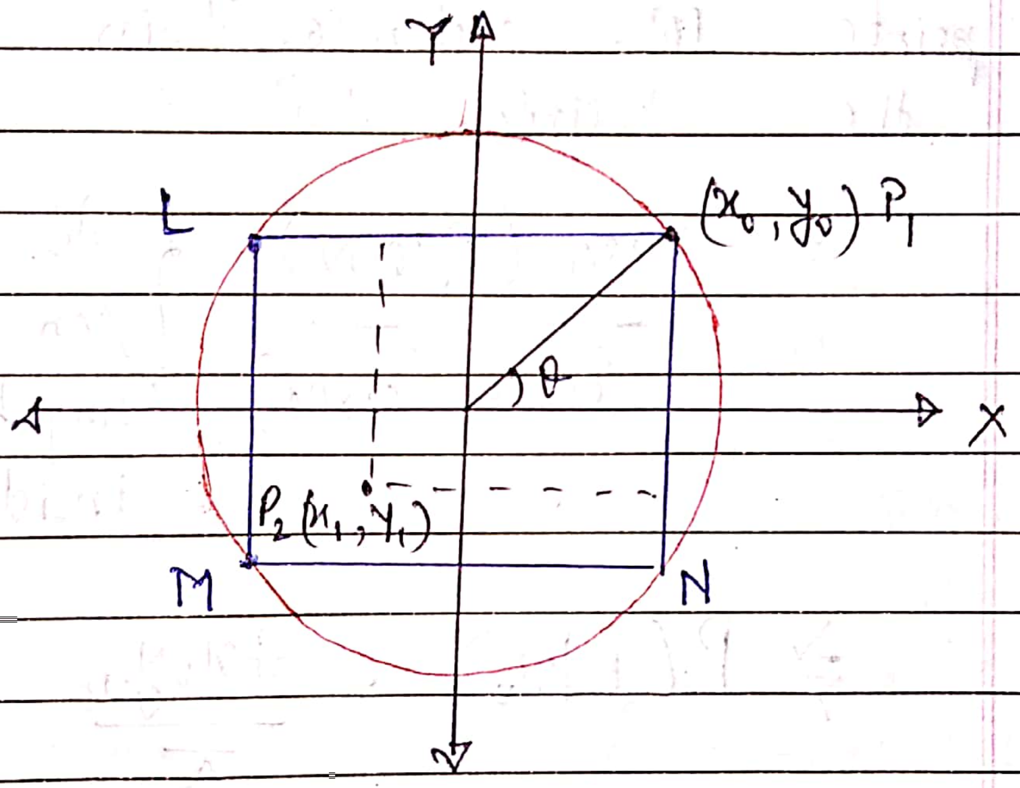
because pt. A & B should be inside/on the circle

$\Rightarrow$ ~~x_1~~

$(1-y_0^2) + y_1^2 \leq 1$ & $x_1^2 + (1-x_0^2) \leq 1$

$\Rightarrow$ $y_1^2 \leq y_0^2$ & $x_1^2 \leq x_0^2$

$\Rightarrow$ $|y_1| \leq |y_0|$ — (II) & $|x_1| \leq |x_0|$ — (III)



Area (LP_1NM) = Area formed by $|y_1| \leq |y_0|$ & $|x_1| \leq |x_0|$

Hence, for the required rectangle point P_2 should lie inside/on rectangle P_1LMN

$P(F|x_0)$ is the probability that rectangle ~~P_1~~ ~~AB~~ (which is asked in Question) falls inside the circle 'C' given the value of x_0.

Therefore, the probability that no points ~~of~~ rectangle lies outside the circle is

$$= \frac{\text{Area}(P_1LMN)}{\text{Area (circle)}}$$ As $P_2(x_1, y_1)$ can be chosen anywhere inside P_1LMN.

$$\Rightarrow P(F|x_0) = \frac{4x_0y_0}{\pi}$$

taking, $x_0 = r\cos\theta$, $y_0 = r\sin\theta$
$= \cos\theta$ $= \sin\theta$

$$P(F|\theta) = \frac{4\cos\theta\sin\theta}{\pi}$$

We also know that,

$P[\, F/x_0 \mid 1^{st} \text{ quadrant} \,] = P[\, F/x_0 \mid \text{whole circle} \,]$

$$\Rightarrow P(F) = \frac{\int_0^{\pi/2} P(F/\theta)\, d\theta}{\pi/2}$$

$$\Rightarrow P(F) = \frac{2}{\pi} \int_0^{\pi/2} \frac{4\cos\theta \sin\theta}{\pi}\, d\theta$$

$$\Rightarrow P(F) = \frac{4}{\pi^2} \int_0^{\pi/2} \sin 2\theta \, d\theta$$

$$\Rightarrow P(F) = \frac{4}{\pi^2} \left[-\frac{\cos 2\theta}{2} \right]_0^{\pi/2}$$

$$= \frac{4}{\pi^2} \left[\left(\frac{-(-1)}{2} \right) - \left(\frac{-1}{2} \right) \right]$$

$$\Rightarrow P(F) = \frac{4}{\pi^2}$$ Ans