CSE 544

PROBABILITY AND STATISTICS FOR DATA SCIENCE

ASSIGNMENT 2

Team members:

Name	**SBU ID**
Venkata Satwik Chilukuri	113263262
Akshay Somayaji	113322316
Shardul Churi	114403283

1) Given, $\text{Cov}(X,Y) = E[(X-E[X])(Y-E[Y])] = E[XY] - E[X]E[Y]$
and $\text{Cov}(X,Y) = 0$ if $X \perp Y$

a) now, X - number of heads in first 2 flips
Y - number of heads in last 2 flips
$\therefore$ Ω of both X & Y $= \{0, 1, 2\}$

now, $E[X] = \sum_i x_i p_X(x_i)$, where $p_X(x_i) = Pr(X = x_i)$
→ since X & Y are sampled over 2 flips, possible combinations are $\{HH, HT, TH, TT\}$
$\therefore$ $E[X] = E[Y] = x_0 p_X(x_0) + x_1 p_X(x_1) + x_2 p_X(x_2)$
$= 0 \cdot 1/4 + 1 \cdot \cancel{2/4}\ 1/2 + \cancel{2} \cdot 1/\cancel{4}\ 2$
$= 1/2 + 1/2 = 1$

Now, to find $E[XY]$, let us build a joint dist^n matrix of X & Y, to find the joint pmf
$$p_{XY}(x,y) = Pr(X = x \text{ and } Y = y)$$

This matrix is sampled over all possible combinations of flipping 3 coins $= \Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$

X \ Y	0	1	2
0	1/8	1/8	0
1	1/8	2/8	1/8
2	0	1/8	1/8

Eg calculations -
i) $p(0,0) = Pr(X=0 \cap Y=0)$
$= Pr(\{TTT\}) = 1/8$
ii) $p(0,1) = Pr(\{TTH\}) = 1/8$
iii) $p(0,2)$ = not possible $= 0$
iv) $p(2,0)$, $p(1,0)$ are symmetrical on $p(0,2)$ & $p(0,1)$ respectively
v) $p(2,1) = p(1,2) = p(\{HHT\})$ or $p(\{THH\})$ $= 1/8$
~~vi) $p(1,1) = p(\{HTH, THT\})$~~
$= 2/8$
vii) $p(2,2) = p(\{HHH\})$
$= 1/8$

$\therefore E[XY] = \sum_x \sum_y xy . p_{XY}(x,y)$

$= 1.1.\frac{2}{8} + 1.2.\frac{1}{8} + 2.1.\frac{1}{8} + 2.2.\frac{1}{8} + 0$ (remaining terms)

$\Rightarrow E[XY] = 10/8$

$\therefore Cov(X,Y) = E[XY] - E[X]E[Y]$

$= \frac{10}{8} - 1.1 = \frac{2}{8} = 1/4$

$\therefore \boxed{Cov(X,Y) = 1/4}$

b) $X = \{-5, -2, 0, 2, 5\}$, $Y = X^2$

since X is a fair 5-side die, $p_X(x) = 1/5$

$\Rightarrow E[X] = \sum_i x_i \, p_X(x_i)$

$= 1/5 (-\cancel{5} - \cancel{2} + 0 + \cancel{2} + \cancel{5}) = 0$

$E[XY] = E[X.X^2] = E[X^3]$

now, $X^3 = \{-125, -8, 0, 8, 125\}$

Hence, $E[X^3]$ too will be 0

$\Rightarrow Cov(X,Y) = E[XY] - E[X]E[Y]$

$= E[X^3] - 0.E[Y]$

$= 0 - 0 = 0$

c) In b), the value of Y depended on that of X since $Y = X^2$, that is, X & Y are dependant RVs. Hence, we cannot ~~imply~~ assume that zero covariance always implies that the RVs are independant.

2) Given, X is a non-negative RV, i.e, $X \geq 0$.
It has mean μ & variance σ^2. We have some $t>0$.

a) Now, for continuous RVs, we have

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)\,dx \quad , \text{ where } f_X(x) = \text{pdf}$$
and $f_X(x)$ is always $\geq 0 \;\forall x$

Since $X \geq 0$, we can reduce above as

$$\rightarrow E[X] = \int_0^{\infty} x f_X(x)\,dx \quad (i)$$

Now, t is some real number such that $0 \leq t < \infty$
$\therefore$ We can split integral (i) as

$$\Rightarrow E[X] = \int_0^t x f_X(x)\,dx + \int_t^{\infty} x f_X(x)\,dx \quad (ii)$$

We know that $t > 0$, $x \in X \geq 0$, $f_X(x) \geq 0$
Thus, (ii) becomes

$$\rightarrow E[X] = (\text{+ve value or zero}) + \int_t^{\infty} x f_X(x)\,dx$$

$$\therefore \boxed{E[X] \geq \int_t^{\infty} x f_X(x)\,dx}$$

b) Using a), we have To Prove $Pr(X>t) \leq \frac{E[X]}{t}$
now, in the LHS, we're given $X > t$.
$\rightarrow$ So for some $x \in X$, $x > t$
multiplying both sides by $f(x)$ and integrating wrt x from $t \rightarrow \infty$, since both $f(x) \geq 0$ & $t > 0$, we get

$$\int_t^{\infty} x f(x)\,dx > \int_t^{\infty} t f(x)\,dx \quad (ii)$$

Using (ii) in result of (a), we get

$$E[X] \geq \int_t^{\infty} t f(x)\,dx \quad (iii)$$

In (iii), since t is a constant wrt x, can write as

$$E[X] \geq t\int_t^{\infty} f(x)\,dx \quad \text{(iv)}$$

Now, from the validity test, we know

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

$$\Rightarrow \int_{-\infty}^{t} f(x)\,dx + \int_t^{\infty} f(x)\,dx = 1$$

$$\Rightarrow P_r(X \leq t) + \int_t^{\infty} f(x)\,dx = 1 \quad \left(\text{as CDF} = \int_{-\infty}^{t} f(x)\,dx = F_X(t) = P(X \leq t)\right)$$

$$\Rightarrow \cancel{1} - P_r(X > t) + \int_t^{\infty} f(x)\,dx = \cancel{1}$$

$$\therefore \quad P_r(X > t) = \int_t^{\infty} f(x)\,dx \quad \text{(v)}$$

Replacing (v) in (iv), we get

$$\Rightarrow \quad E[X] \geq t \cdot P_r(X > t)$$

$$\therefore \quad \boxed{P_r(X > t) \leq \frac{E[X]}{t}} \quad \text{(b)}$$

c) Using b), we have To Prove $P_r(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2}$

Now, inequality b) holds good for any RV $X \geq 0$, and $t > 0$.

So we can replace X with $|X - \mu|$, which will also be a +ve RV in itself

$$\Rightarrow \quad P_r(|X - \mu| > t) \leq \frac{E[|X - \mu|]}{t} \quad \text{(vi)}$$

now, $|X - \mu| > t \Rightarrow (X - \mu)^2 > t^2$

→ Replacing above inequality in (vi), we get

$$P_r((X - \mu)^2 > t^2) \leq \frac{E[(X - \mu)^2]}{t^2} \quad \text{(vii)}$$

(Since $|X - \mu| \rightarrow (X - \mu)^2$ & $t \rightarrow t^2$)

→ now, $E[(X-\mu)^2] = E[X^2 + \mu^2 - 2\mu X]$

Using LOE, $E[(X-\mu)^2] = E[X^2] + E[\mu^2] + E[-2\mu X]$

→ now, $E[c] = c$, where c is a constant

and $E[cX] = c \cdot E[X]$, where c is a constant

$\Rightarrow E[(X-\mu)^2] = E[X^2] + \mu^2 - 2\mu E[X]$

since $E[X] = \mu$, we finally get

$$E[(X-\mu)^2] = E[X^2] + \mu^2 - 2\mu^2$$

$$= E[X^2] - \mu^2 = Var(X) = \sigma^2$$

$\Rightarrow$ Replacing $E[(X-\mu)^2] = \sigma^2$ in (vii),

$$\Rightarrow \Pr((X-\mu)^2 \geq t^2) \leq \frac{\sigma^2}{t^2}$$

$$\Rightarrow \therefore \boxed{\Pr(|X-\mu| \geq t) \leq \frac{\sigma^2}{t^2}}$$

Scanned with CamScanner

3)

(a)

① Given, $f_{X_i}(x) = \lambda_i e^{-\lambda_i x}$, $x \geq 0$ $\forall i \in \{1, 2, \dots k\}$

$$Z = \min(X_1, X_2, \dots, X_k)$$

To find pdf of Z

Now,

$$\text{Cdf of } Z = F_Z(x)$$

$$= Pr(Z \leq x)$$

we know that Z is $\min(X_1, X_2, \dots, X_k)$

$$= Pr(\text{Min}(X_1, X_2, \dots, X_k) \leq x)$$

we know that $Pr(A) = 1 - Pr(\bar{A})$

$$= 1 - Pr(\overline{\text{Min}(X_1, X_2, \dots X_k) \leq x})$$

$$= 1 - Pr(\text{Min}(X_1, X_2, \dots, X_k) > x)$$

Now, if $\text{Min}(X_1, X_2) > \alpha$ then both X_1, X_2 will be greater than α.

$$= 1 - Pr((X_1 > x) \cap (X_2 > x) \dots (X_k > x))$$

Given $x_1, x_2, x_3, \dots x_k$ are independent

So $pr(x_1 \cap x_2 \cap x_3 \dots \cap x_n) = pr(x_1) \cdot pr(x_2) \dots pr(x_k)$

$$= 1 - [pr(x_1 > x) \cdot pr(x_2 > x) \cdot pr(x_3 > x) \dots \dots pr(x_k > x)]$$

Now,

Given $x_i \ \forall i \in \{1,2,3,\dots k\}$ is an exponential Random variable.

So,

$$F_{x_i}(x) = pr(X_i \le x)$$

$$= \int_0^x f(x)\, dx$$ [As $x > 0$ limits start from 0]

$$= \int_0^x \lambda_i e^{-\lambda_i x} dx$$

$$= \lambda_i \int_0^x e^{-\lambda_i x} dx$$

$$= \cancel{\lambda_i} \left[\frac{e^{-\lambda_i x}}{-\cancel{\lambda_i}} \right]_0^x$$ $\left[\int e^{-\lambda x} = \frac{e^{-\lambda x}}{\frac{d}{dx}(-\lambda x)} \right]$

Negating negative symbol and changing limits.

$$= \left[e^{-\lambda_1 x} \right]_x^0$$

$$= 1 - e^{-\lambda_1 x} \quad \text{i.e} \quad P_r(X_1 \le x) = 1 - e^{-\lambda_1 x} \quad \text{——} \; (1)$$

Now,

$$F_Z(x) = 1 - \left[P_r(X_1 > x) \cdot P_r(X_2 > x) \cdots P_r(X_k > x) \right]$$

we know that $P_r(A) = 1 - P_r(\bar{A})$

$$\Rightarrow F_Z(x) = 1 - \left[(1 - P_r(\overline{X_1 > x})) \cdot (1 - P_r(\overline{X_2 > x})) \cdots \cdots (1 - P_r(\overline{X_k > x})) \right]$$

$$= 1 - \left[(1 - P_r(X_1 \le x)) \cdot (1 - P_r(X_2 \le x)) \cdots \cdots (1 - P_r(X_k \le x)) \right]$$

from (1)

$$= 1 - \left[(1 - (1 - e^{-\lambda_1 x})) \cdot (1 - (1 - e^{-\lambda_2 x})) \cdots \cdots (1 - (1 - e^{-\lambda_k x})) \right]$$

$$= 1 - \left[e^{-\lambda_1 x} \cdot e^{-\lambda_2 x} \cdot e^{-\lambda_3 x} \cdots e^{-\lambda_k x} \right]$$

using the formulae $a^m \cdot a^n = a^{m+n}$

$$= 1 - \left[e^{-\lambda_1 x + (-\lambda_2 x) + (-\lambda_3 x) + \cdots + (-\lambda_k x)} \right]$$

Scanned with CamScanner

$$= 1 - e^{(-\lambda_1 - \lambda_2 - \lambda_3 \cdots - \lambda_k)x}$$

$$\therefore F_Z(x) = 1 - e^{-(\lambda_1 + \lambda_2 + \cdots + \lambda_k)x}$$

we know that pdf of $Z = \frac{d}{dx}$ [cdf of Z]

where $Z = \min(X_1, X_2, \ldots, X_k)$

$$\text{pdf of } Z = \frac{d}{dx} [\text{cdf of } Z]$$

$$= \frac{d}{dx}\left[1 - e^{-(\lambda_1 + \lambda_2 + \cdots + \lambda_k)x}\right]$$

$$= \frac{d}{dx}(1) - \frac{d}{dx}\left[e^{-(\lambda_1 + \lambda_2 + \cdots + \lambda_k)x}\right]$$

$$= -e^{-(\lambda_1 + \lambda_2 + \cdots + \lambda_k)x}\left[\frac{d}{dx}\left(-(\lambda_1 + \lambda_2 + \cdots + \lambda_k)x\right)\right]$$

$$= -e^{-(\lambda_1 + \lambda_2 + \cdots + \lambda_k)x}\left[-(\lambda_1 + \lambda_2 + \cdots + \lambda_k)\right]$$

$$= (\lambda_1 + \lambda_2 + \cdots + \lambda_k)\, e^{-(\lambda_1 + \lambda_2 + \cdots + \lambda_k)x}$$

$$\therefore \text{pdf of } Z = (\lambda_1 + \lambda_2 + \cdots + \lambda_k)\, e^{-(\lambda_1 + \lambda_2 + \cdots + \lambda_k)x}$$

(1) Find $E[Z]$ we know that $f_Z(x) = (\sum_{i=1}^{k} \lambda_i) e^{-(\sum_{i=1}^{k}\lambda_i)x}$

Given $x \geq 0$ so limits will be from 0 to ∞ only

$$E[Z] = \int_0^\infty x \cdot f_Z(x)\, dx$$

$$= \int_0^\infty x \cdot \left(\sum_{i=1}^{k} \lambda_i\right) \cdot e^{-(\sum_{i=1}^{k}\lambda_i)x}\, dx$$

$$= \sum_{i=1}^{k} \lambda_i \int_0^\infty x \cdot e^{-(\sum_{i=1}^{k}\lambda_i)x}\, dx$$

using integration by parts.

i.e.

$$\int_b^a u \cdot v\, dx = \left[u \cdot \int v\, dx\right]_b^a - \int_b^a \left[\frac{du}{dx} \cdot \int v\, dx\right] dx$$

$$= \sum_{i=1}^{k} \lambda_i \left[\left[x \cdot \int e^{-(\sum_{i=1}^{k}\lambda_i)x}\, dx \right]_0^\infty - \int_0^\infty \frac{d}{dx} x \cdot \left[\int e^{-(\sum_{i=1}^{k}\lambda_i)x}\, dx\right] dx \right]$$

$$\cancel{= \sum_{i=1}^{k} \lambda_i \left[\left[x \, e^{-(\sum_{i=1}^{k}\lambda_i)x} \left(-\sum_{i=1}^{k}\lambda_i\right) \right]_0^\infty - \int_0^\infty 1 \cdot \left[e^{-(\sum_{i=1}^{k}\lambda_i)x} \left(-\sum_{i=1}^{k}\lambda_i\right)\right] dx \right]}$$

Scanned with CamScanner

we know that $\int e^{-\lambda x} = \frac{e^{-\lambda x}}{-\lambda}$

so,

$$= \sum_{i=1}^{k} \lambda_i \left[\left[x \cdot \frac{e^{-(\sum_{i=1}^{k} \lambda_i) x}}{-(\sum_{i=1}^{k} \lambda_i)} \right]_0^{\infty} - \int_0^{\infty} 1 \cdot \left[\frac{e^{-(\sum_{i=1}^{k} \lambda_i) x}}{-(\sum_{i=1}^{k} \lambda_i)} \right] dx \right.$$

$$= \sum_{i=1}^{k} \lambda_i \left[[0 - 0] - \int_0^{\infty} \left[\frac{1}{-(\sum_{i=1}^{k} \lambda_i)} \times e^{-(\sum_{i=1}^{k} \lambda_i) x} \right] dx \right]$$

$$= \cancel{\sum_{i=1}^{k} \lambda_i} \left[\frac{1}{\cancel{\sum_{i=1}^{k} \lambda_i}} \int_0^{\infty} e^{-(\sum_{i=1}^{k} \lambda_i) x} \, dx \right]$$

$$= \left[\frac{e^{-(\sum_{i=1}^{k} \lambda_i) x}}{-(\sum_{i=1}^{k} \lambda_i)} \right]_0^{\infty}$$

Negating - with changing integral signs

$$= \frac{1}{\sum_{i=1}^{k} \lambda_i} \left[e^{-(\sum_{i=1}^{k} \lambda_i) x} \right]_{\infty}^{0}$$

Scanned with CamScanner

$$= \frac{1}{\sum_{i=1}^{k} \lambda_i} \left[1 - 0 \right]$$

$$= \frac{1}{\sum_{i=1}^{k} \lambda_i}$$

$$\therefore \; E[z] = \frac{1}{\sum_{i=1}^{k} \lambda_i} = \frac{1}{\lambda_1 + \lambda_2 + \cdots + \lambda_k}$$

(iii) To find $Var[z]$

we know that

$$Var[z] = E[z^2] - \left(E[z]\right)^2$$

we already know $E[z] = \frac{1}{\sum_{i=1}^{k} \lambda_i}$

we need to find $E[z^2]$

$$E[z^2] = \int_0^{\infty} x^2 \cdot f_z(x) \, dx \quad \text{as } x \geq 0$$

$$= \int_0^{\infty} x^2 \cdot \left(\sum_{i=1}^{k} \lambda_i \right) \cdot e^{-\left(\sum_{i=1}^{k} \lambda_i\right) x} \, dx$$

for convinience let $\sum_{i=1}^{k} d_i = d$ — ②

$$= \int_0^{\infty} x^2 \cdot d\, e^{-dx}\, dx$$

$$= d \int_0^{\infty} x^2 \cdot e^{-dx}\, dx$$

using integratation by parts i.e

$$\int_b^a u v\, dx = \left[u \cdot \int v\, dx \right]_b^a - \int_b^a \left[\frac{du}{dx} \cdot \int v\, dx \right] dx$$

$$= d \left[\left[x^2 \cdot \int e^{-dx}\, dx \right]_0^{\infty} - \left[\int_0^{\infty} \frac{d}{dx} x^2 \cdot \int e^{-dx}\, dx\, dx \right] \right]$$

$$= d \left[\left[x^2 \cdot \frac{e^{-dx}}{-d} \right]_0^{\infty} - \left[\int_0^{\infty} 2x \cdot \frac{e^{-dx}}{-d}\, dx \right] \right]$$

$$= d \left[[0 - 0] - \left[\frac{-2}{d} \int_0^{\infty} x \cdot e^{-dx}\, dx \right] \right]$$

$$= \frac{2}{d} \left[\int_0^{\infty} x \cdot d\, e^{-dx}\, dx \right]$$

[bringing d inside & $\frac{2}{d}$ outside] now we know that from 3 (a) (i) part

which we solved before

$$E[x] = \int_0^{\infty} x \cdot f_z(x) \, dx$$

$$= \int_0^{\infty} x \cdot \lambda \cdot e^{-\lambda x} \, dx \quad \text{where } \lambda = \sum_{i=1}^{k} \lambda_i$$

So,

$$E[z^2] = \frac{2}{\lambda} \underbrace{\left[\int_0^{\infty} x \cdot \lambda \cdot e^{-\lambda x} \, dx \right]}_{\text{This is } E[x]}$$

$$E[z^2] = \frac{2}{\lambda} \times E[x]$$

we already know that $E[x] = \frac{1}{\lambda}$

where $\lambda = \sum_{i=1}^{k} \lambda_i$

$$= \frac{2}{\lambda^2}$$

$$\therefore E[z^2] = \frac{2}{\lambda^2} \quad \text{where } \lambda = \sum_{i=1}^{k} \lambda_i$$

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

$$= \frac{2}{\lambda^2} - \left[\frac{1}{\lambda}\right]^2$$

$$= \frac{2}{\lambda^2} - \frac{1}{\lambda^2}$$

$$= \frac{1}{\lambda^2} \quad \text{where } \lambda = \sum_{i=1}^{k} \lambda_i$$

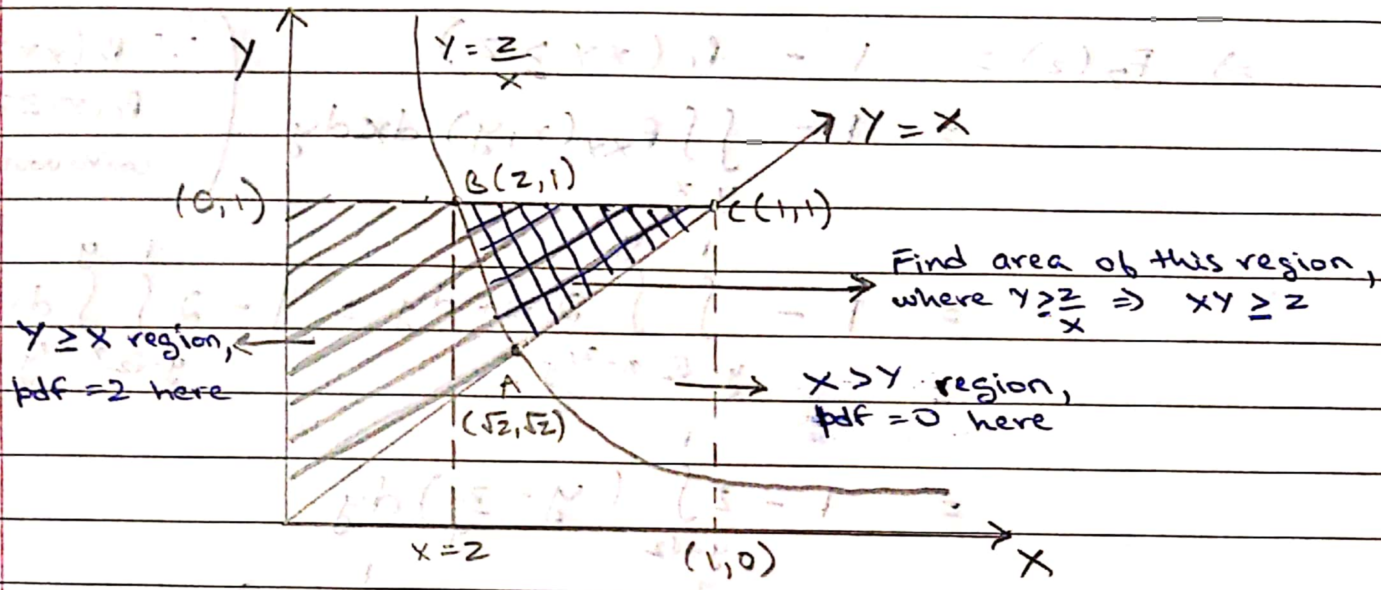
$$\therefore \text{var}[z] = \frac{1}{(\lambda_1 + \lambda_2 + \dots + \lambda_k)^2}$$

3b) X and Y are RVs with joint pdf given by

$$f_{XY}(x,y) = \begin{cases} 2, & 0 \le x \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

To find - $f_Z(XY)$, that is, pdf of $Z = XY$

→ We can graphically represent the joint pdf as below



→ Now, $Z = XY \Rightarrow Y = Z/X$ represents a family of hyperbolas. The joint pdf holds a value of 2 only in this range of $0 \le x \le y \le 1$, that is, the area above the line $x = y$ upto $y = 1$, as shown in graph.

→ Now, CDF of Z can be written as

$$F_Z(z) = Pr(Z \le z) = Pr(XY \le z)$$

$$= 1 - Pr(XY > z)$$

$$F_Z(z) = 1 - \text{area bound by region ABC} \quad (i)$$

Now, we have $B = (z,1)$ & $C = (1,1)$

A is the point of intersection of $Y = X$ & $Y = z/x$

$\Rightarrow Y^2 = X^2 = z \Rightarrow x = y = \sqrt{z}$

$\therefore A = (\sqrt{z}, \sqrt{z})$

Now, to find area of region ABC, let us compute the limits.

- We can see limit of $y \rightarrow \sqrt{z}$ to 1
- For x, lower limit can be taken from $xy > z$, i.e, $x > z/y$

upper limit can be taken from $x \le y$

$\therefore$ limit of $x \rightarrow z/y$ to y

$$\Rightarrow F_Z(z) = 1 - Pr(xy \geq z) \qquad \left(\because Pr(xy > z) = Pr(xy \geq z) \text{ for continuous RVs} \right)$$

$$= 1 - \int_y \int_x f_{xy}(x,y)\, dx\, dy$$

$$= 1 - \int_{y=\sqrt{z}}^{1} \int_{x=z/y}^{y} 2\, dx\, dy = 1 - 2 \int_{y=\sqrt{z}}^{1} \int_{x=z/y}^{y} dx\, dy$$

$$= 1 - 2 \int_{y=\sqrt{z}}^{1} \left(y - \frac{z}{y} \right) dy$$

$$= 1 - 2 \left(\int_{y=\sqrt{z}}^{1} y\, dy - z \int_{y=\sqrt{z}}^{1} \frac{dy}{y} \right)$$

$$= 1 - 2 \left(\left. \frac{y^2}{2} \right|_{\sqrt{z}}^{1} - z \left(\ln(y) \Big|_{\sqrt{z}}^{1} \right) \right)$$

$$= 1 - 2 \left(\frac{1-z}{2} - z \left(\ln(1) - \ln(\sqrt{z}) \right) \right)$$

now, $\ln(1) = 0$, & $\ln(\sqrt{z}) = \ln(z^{1/2}) = \ln(z)/2$

$$\Rightarrow F_Z(z) = 1 - \cancel{2} \left(\frac{1-z}{\cancel{2}} + \frac{z\ln(z)}{\cancel{2}} \right)$$

$$\therefore F_Z(z) = \cancel{1} - \cancel{1} + z - z\ln(z)$$

$$\Rightarrow F_Z(z) = z - z\ln(z).$$

Now, pdf of $z = f_Z(z) = \frac{d(F_Z(z))}{dz}$

$$= \frac{d(z - z\ln(z))}{dz} = 1 - \frac{d}{dz}(z\ln(z))$$

$$= 1 - \left(\cancel{z} \cdot \frac{1}{\cancel{z}} + \ln(z) \cdot 1 \right)$$ (differentiation ~~derivation~~ by parts)

$$= \cancel{1} - \cancel{1} - \ln(z)$$

$$\boxed{\therefore \quad f_2(z) = -\ln(z)}$$

Question 4 : Daenerys returns to King's Landing, almost.

a) To return to Kings Landing, Daenerys is to leave Meereen.

2 paths from Meereen, East and West

East leads to Shadow Lands and back to Meereen in 20 DAYS

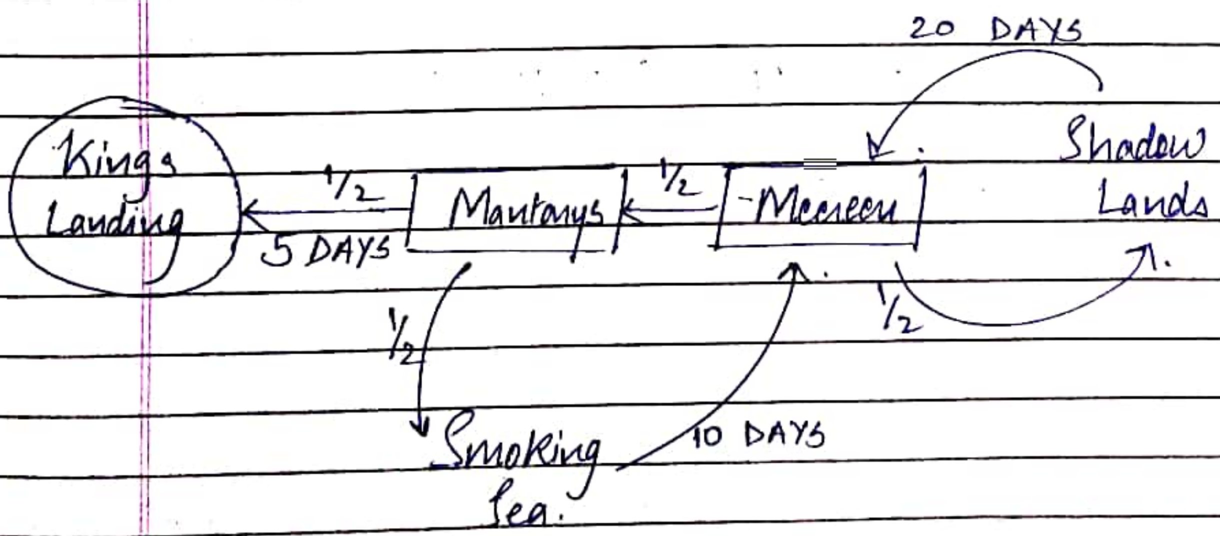
West leads to the city of MANTARYS

Mantarys again has 2 paths, West and South

South leads to Smoking Sea and back to Meereen in 10 DAYS

West leads to Kings Landing in 5 DAYS

To Consider : 1. Probabilities of both choices are equal.
2. Past choices aren't remembered.



From Meereen,

$P(\text{Shadow Lands}) = P(\text{Mantarys}) = 0.5$

From Mantarys

$P(\text{Sinking Seas}) = P(\text{King's Landing}) = 0.5$

To find, Expected no. of days in which Daenerys will return to King's Landing.

$$E(X) = P(\text{Shadow Lands}) \times E(X/\text{Shadow Lands}) + P(\text{Mantarys}) \times E(X/\text{Mantarys})$$

$$= P(\text{Shadow Lands}) \times E(X/\text{Shadow Lands}) + \left(P(\text{Sinking Seas}) \times E(X/\text{Sinking Seas}) + P(\text{King's Landing}) \times E(X/\text{Kings Landing}) \right)$$

$$\therefore E[X] = \left(0.5(E[X]+20)\right) + \left(0.5\left(0.5(E(X)+10) + 0.5(5)\right)\right)$$

$$= 0.5(E[X]) + 10 + 0.25(E[X]) + 2.5 + \cancel{2.5}\ 1.25$$

$$\therefore E[X] = 0.75(E[X]) + 13.75$$

$$\therefore E[X] = 55$$

(b) To Find, $V(X)$

$$V(X) = E[X^2] - (E[X])^2$$

$$E[X^2] = P(\text{Shadow Lands}) \times E(X^2|\text{Shadow Lands}) + P(\text{Mantarys}) \times E(X^2|\text{Mantarys})$$

$$= P(\text{Shadow Lands}) \times E(X^2|\text{Shadow Lands}) + \left(P(\text{Sinking Seas}) \times E(X^2|\text{Sinking Seas}) + P(\text{Kings Landings}) \times E(X^2|\text{Kings Landing})\right)$$

$$\therefore E[X^2] = \frac{1}{2}\left(E[X+20]^2\right) + \frac{1}{2}\left(\frac{1}{2}\left(E[X+10]^2\right) + \frac{1}{2}(5^2)\right)$$

$$= \frac{1}{2}\left(E[X^2] + 40E[X] + 400\right) + \frac{1}{2}\left(\frac{1}{2}\left(E[X^2] + 20E[X] + 100\right) + \frac{1}{2} \times 25\right)$$

$$= \frac{2E[X^2] + 80E[X] + 800}{4} + \frac{E[X^2] + 20E[X] + 100 + 25}{4}$$

$$\therefore 4E[X^2] = 3E[X^2] + 100E[X] + 925$$

Substituting value of $E[X] = 55$ from PART a

$$\therefore E[X^2] = 55(100) + 925$$
$$= 5500 + 925$$
$$= 6425$$

$$E[X]^2 = 55^2 = 3025$$

$$\therefore V[X] = E[X^2] - (E[X])^2 = 6425 - 3025$$
$$= \underline{3400}$$

5) We have 2 weather states, $C \rightarrow$ Clear & $S \rightarrow$ Snowy

Now, the transition probabilities are read as -

Pr[Weather tomorrow is x_{i+1}, given that weather today is x_i and weather yesterday was x_{i-1}] $\rightarrow Pr[x_{i+1} | x_i x_{i-1}]$,

where $X \rightarrow \{C, S\}$. We're also given,

$\rightarrow Pr(C|CC) = 0.9 \Rightarrow Pr(S|CC) = 1 - 0.9 = 0.1$

$\rightarrow Pr(C|CS) = 0.8 \Rightarrow Pr(S|CS) = 1 - 0.8 = 0.2$

$\rightarrow Pr(C|SC) = 0.5 \Rightarrow Pr(S|SC) = 1 - 0.5 = 0.5$

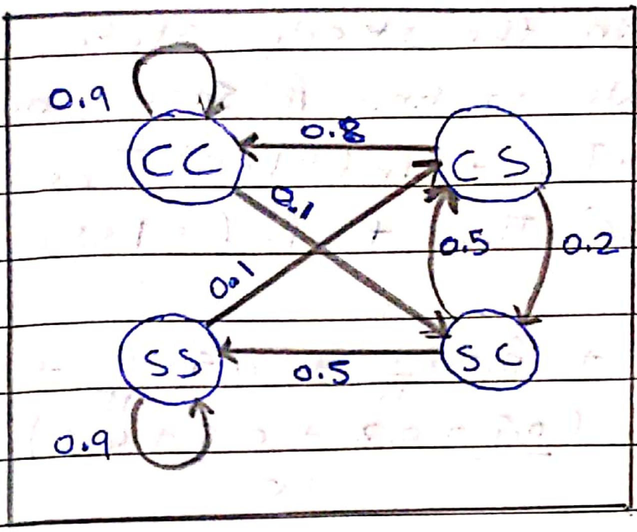
$\rightarrow Pr(C|SS) = 0.1 \Rightarrow Pr(S|SS) = 1 - 0.1 = 0.9$

a) We need to find steady state probabilities, π_{CC}, π_{CS}, π_{SC}, π_{SS}. There are thus 4 states, $\{CC, CS, SC, SS\}$, and we represent the transitions as

- $C|CC \xrightarrow{0.9} CC$, $S|CC \xrightarrow{0.1} SC$
- $C|CS \xrightarrow{0.8} CC$, $S|CS \xrightarrow{0.2} SC$
- $C|SC \xrightarrow{0.5} CS$, $S|SC \xrightarrow{0.5} SS$
- $C|SS \xrightarrow{0.1} CS$, $S|SS \xrightarrow{0.9} SS$

Scanned with CamScanner

We represent the transitions in the Markov chain below -



Transition Matrix P

$$\begin{array}{c|cccc} & CC & CS & SC & SS \\ \hline CC & 0.9 & 0.0 & 0.1 & 0.0 \\ CS & 0.8 & 0.0 & 0.2 & 0.0 \\ SC & 0.0 & 0.5 & 0.0 & 0.5 \\ SS & 0.0 & 0.1 & 0.0 & 0.9 \end{array}$$

Using Global balance to solve for the steady-state,

as $n \to \infty$, $\quad P^{n-1} \cdot P = P^n$

$$\Rightarrow \begin{bmatrix} \pi_{cc} & \pi_{cs} & \pi_{sc} & \pi_{ss} \\ \pi_{cc} & \pi_{cs} & \pi_{sc} & \pi_{ss} \\ \pi_{cc} & \pi_{cs} & \pi_{sc} & \pi_{ss} \\ \pi_{cc} & \pi_{cs} & \pi_{sc} & \pi_{ss} \end{bmatrix} \times \begin{bmatrix} 0.9 & 0 & 0.1 & 0 \\ 0.8 & 0 & 0.2 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0 & 0.1 & 0 & 0.9 \end{bmatrix} = \begin{bmatrix} \pi_{cc} & \pi_{cs} & \pi_{sc} & \pi_{ss} \\ \pi_{cc} & \pi_{cs} & \pi_{sc} & \pi_{ss} \\ \pi_{cc} & \pi_{cs} & \pi_{sc} & \pi_{ss} \\ \pi_{cc} & \pi_{cs} & \pi_{sc} & \pi_{ss} \end{bmatrix}$$

$\rightarrow$ Solving for above matrix mult$^{\underline{n}}$, we get

$\rightsquigarrow$ $0.9\pi_{cc} + 0.8\pi_{cs} = \pi_{cc} \Rightarrow \pi_{cc} = 8\pi_{cs}$ ①

$\rightsquigarrow$ $0.5\pi_{sc} + 0.1\pi_{ss} = \pi_{cs}$ ②

$\rightsquigarrow$ $0.1\pi_{cc} + 0.2\pi_{cs} = \pi_{sc}$ ③

$\rightsquigarrow$ $0.5\pi_{sc} + 0.9\pi_{ss} = \pi_{ss} \Rightarrow \pi_{ss} = 5\pi_{sc}$ ④

$\rightarrow$ Replacing ④ in ②, we get

$0.5\pi_{sc} + 0.1(5\pi_{sc}) = \pi_{cs} \Rightarrow \pi_{sc} = \pi_{cs}$ ⑤

$\rightarrow$ Now, we also have $\pi_{cc} + \pi_{cs} + \pi_{sc} + \pi_{ss} = 1$

Substituting ① - ⑤ in above, we get

$8\pi_{cs} + \pi_{cs} + \pi_{sc} + 5\pi_{sc} = 1$

$\Rightarrow 15\pi_{cs} = 1 \Rightarrow$

$$\boxed{\begin{aligned} \pi_{cs} &= \pi_{sc} = 1/15 \\ \pi_{cc} &= 8/15 \\ \pi_{ss} &= 5/15 \end{aligned}}$$

b) We have to find the prob^y that it is snowing 3 days from now, in the steady state.

We can use TP rule to find Pr(snow 3 days from now)

$$\Rightarrow Pr_{xxS} = Pr(S|CC)\cdot\pi_{CC} + Pr(S|CS)\cdot\pi_{CS} + Pr(S|SC)\cdot\pi_{SC} + Pr(S|SS)\cdot\pi_{SS}$$

3rd day = S

(x = clear or snowy first 2 days)

$$= 0.1\times\frac{8}{15} + 0.2\times\frac{1}{15} + 0.5\times\frac{1}{15} + 0.9\times\frac{5}{15}$$

$$= \frac{(0.8 + 0.2 + 0.5 + 4.5)}{15}$$

$$\boxed{Pr_{xxS} = 6/15 = 0.4}$$

c) Steady State: Power iteration >> [0.53333337, 0.06666666, 0.06666667, 0.3333333]

Since we're in steady state, and we have a 2-day steady state probability, the probability of snow on any day will be $\pi_{sc} + \pi_{ss}$, that is, steady state probability of snow today and snow or clear yesterday.

$$\therefore Pr_{xxs} = \pi_{sc} + \pi_{ss}$$
$$= 0.06666667 + 0.3333333$$
$$= 0.40$$

6)

(a)

Given,

$X = (X_1, X_2, \ldots, X_k)$ is a Multivariate Normal

we know that for X to be a multivariate Normal

we need $t_1 X_1 + t_2 X_2 + t_3 X_3 + \cdots + t_k X_k$ to be a

normal distribution for any real values of $t_1, t_2, .. t_k$

let us Assume $t = (t_1, t_2, t_3, \ldots t_k)$

Now let us take the case where

$t_2, t_3, \ldots t_k = 0$ & $t_1 = 1$

So, $t = (1, 0, 0, \ldots, 0)$

Now,

$$\text{linear Combination} = \sum_{i=1}^{k} t_i X_i$$

$$= t_1 X_1 + t_2 X_2 + \cdots + t_k X_k$$

$$= 1 \cdot X_1 + 0 + 0 + \cdots + 0$$

$$= X_1$$

Given $X = \{X_1, X_2, \dots, X_K\}$ is a multivariate Normal

Here, $t_1 X_1 + t_2 X_2 + \dots + t_k X_K = X_1$. So, X_1 has to be a normal distribution

So, if you change "t" which is $(t_1, t_2, \dots, t_k)$ accordingly we can say that Any X_j where $j = 1, 2, 3, \dots k$ is Normal distribution

(b)

Given $X \sim Norm(0,1)$; $Y = SX$

To prove, (X, Y) is not a Multivariate Normal.

$$Y = SX$$

Given

$$S = \begin{cases} 1 & \text{with probability } \frac{1}{2} \\ -1 & \text{with probability } \frac{1}{2} \end{cases}$$

To prove that (X, Y) is not a multivariate Normal we need to prove that atleast one combination of $t_1 X + t_2 Y$ to not be a normal distribution for any Real values of t_1, t_2

with probability $y = -x$ i.e

$$t_1 x + t_2 y$$

$$= t_1 x - t_2 x$$

$$= (t_1 - t_2) x$$

when ~~t~~ $t_1 = t_2$ the linear combination

$t_1 x + t_2 y \nsim$ Normal distribution.

So, we can say that (x, y) is not a multivariate normals

(c)

Given,

$z \sim$ Normal $(0,1)$ $w \sim$ Normal $(0,1)$

z, w are independent and identical normal distributions. [iid]

So,

$$E[z] = E[w]$$

$$Var[z] = Var[w]$$

(i) To prove (z, w) are Multivariate Normal

z, w are both independent

Now,

Using the weighted sum of independent Normals

if, $N_1 \sim Norm(\mu_1, \sigma_1^2) \perp N_2 \sim Norm(\mu_2, \sigma_2^2)$

then,

$$a_1 N_1 + a_2 N_2 \sim Norm(a_1\mu_1 + a_2\mu_2, a_1^2\sigma_1^2 + a_2^2\sigma_2^2) \quad \text{— (1)}$$

So,

for (z, w) to be Multivariate Normal, we need to say that any linear combination of z, w has a Normal distribution [$t_1 z + t_2 w \sim$ Normal distribution]

from (1) we can say that any $t_1 z + t_2 w$ is a Normal distribution.

So,

(z, w) is a Multivariate Normal distribution.

(ii) To prove that $(z+2w, 3z+5w)$ are Multivariate Normal

z, w are both independent

Using the weighted sum of independent Normals

if

$$N_1 \sim Norm(\mu_1, \sigma_1^2) \perp N_2 \sim Norm(\mu_2, \sigma_2^2)$$

then

$$a_1 N_1 + a_2 N_2 \sim Norm\left((a_1 \mu_1 + a_2 \mu_2), (a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2)\right)$$

so,

for $Z + 2W$

$$Z \sim Norm(0,1) \qquad W \sim Norm(0,1)$$

$$Z+2W \sim Norm\left(((1\times 0)+(2\times 0)), ((1\times 1) + ((2)^2 \times 1))\right)$$

i.e,

$$Z+2W \sim Normal(0, 5)$$

for $3Z + 5W$

$$Z \sim Norm(0,1) \qquad W \sim Norm(0,1)$$

$$3Z+5W \sim Norm\left((3\times 0 + 5 \times 0), ((3^2 \times 1) + (5^2 \times 1))\right)$$

i.e,

$$3Z+5W \sim Normal(0, 34)$$

both $z+2w$, $3z+5w$ are Normal distribution

for $(z+2w, 3z+5w)$ to be multivariate normal every combination of $t_1(z+2w) + t_2(3z+5w)$ should be Normal distribution for any real values of $t_1, t_2, \ldots, t_k$

$$= t_1(z+2w) + t_2(3z+5w)$$

$$= t_1 z + 2t_1 w + 3t_2 z + 5t_2 w$$

$$= (t_1+3t_2) z + (2t_1+5t_2) w$$

Now,

let $t_1 = 3$, $t_2 = -1$

$= (3-3) z + (1) w$

$= w$

$\downarrow$

Normal distribution

let $t_1 = 5$, $t_2 = -2$

$= -z + (10-10) w$

$= -z \rightarrow$ Normal distribution.

Using the weighted sum of independent Normals, we can say that for any values of t_1, t_2 the linear combination $t_1(z+2w) + t_2(3z+5w)$ will be Normal distribution. So, we can say that

$(z+2w, 3z+5w)$ is multivariate Normal distribution

(d)

Given,

$x = \{x_1, x_2, \ldots, x_n\}$, $y = \{y_1, y_2, \ldots y_m\}$ are multivariate normal with x and y are independent

let function $N(x)$ be any linear combination of x

let function $M(y)$ be any linear combination of y

i.e,

$$N(x) = t_1 x_1 + t_2 x_2 + \cdots + t_n x_n$$

$$M(y) = a_1 y_1 + a_2 y_2 + \cdots + a_m y_m$$

Since x, y are multivariate normals then any $N(x)$, any $M(y)$ will be Normal distribution

Since x, y are independent; $N(x)$, $M(y)$ is also independent

<u>proof:</u>

let A, B subset of Real nos 'R'

$$pr(N(x) \in A \cap M(y) \in B) = pr(x \in N^{-1}(A) \cap y \in M^{-1}(B))$$

Since x, y are independent

$$P(x\, y) = P(x) \cdot pr(y)$$

$= \Pr(x \in N^{-1}(A)) \cdot \Pr(y \in M^{-1}(B))$

$= \Pr(N[x] \in A) \cdot \Pr(M[y] \in B)$

i.e,

$N[x]$, $M[y]$ are independent.

i.e,

$t_1x_1 + t_2x_2 + \dots + t_nx_n$, $a_1y_1 + a_2y_2 + \dots + a_my_m$

are independent and both are normal distributions

for any value of $(t_1, t_2, \dots t_n)$ & $(a_1, a_2, \dots a_m)$

so According to weighted sum of independent Normals

i.e

if $N_1 \sim Norm(\mu_1, \sigma_1^2)$ $\perp$ $N_2 \sim Norm(\mu_2, \sigma_2^2)$

then $N_1 + N_2 \sim Norm(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

So,

$t_1x_1 + t_2x_2 + \dots + t_nx_n + a_1y_1 + a_2y_2 + \dots + a_my_m$ is also Normal

for any real values of

$t_1, t_2, \dots t_n, a_1, a_2, \dots a_m$

So, The Concatenated vector $w = (x_1, x_2, \dots x_n, y_1, y_2, \dots y_m)$ is a multivariate Normal.

(e)

Given,

<u>fact 1:</u> if x is a multivariate Normal that can be written as $x = \{x_1, x_2\}$ and every component of x_1 is uncorrelated with every component of x_2 then x_1, x_2 are independent.

<u>fact 2:</u> for any RV x, y, w & v we have

$$\text{Cov}(ax+by, cw+dv) = ac\,\text{Cov}(x,w) + ad\,\text{Cov}(x,v) + bc\,\text{Cov}(y,w) + bd\,\text{Cov}(y,v)$$

Given,

$x \sim \text{Norm}(0,1)$ $\quad y \sim \text{Norm}(0,1)$; x, y are iid

So,

$$E[x] = E[y]$$

$$\text{var}[x] = \text{var}[y] \quad \text{i.e} \quad \text{Cov}[x\,x] = \text{Cov}[y\,y] \quad \text{— (1)}$$

By weighted sum of independent normals,

if $N_1 \sim \text{Norm}(\mu_1, \sigma_1^2) \perp N_2 \sim \text{Norm}(\mu_2, \sigma_2^2)$

then $a_1 N_1 + a_2 N_2 \sim \text{Norm}\left((a_1\mu_1 + a_2\mu_2), (a_1^2\sigma_1^2 + a_2^2\sigma_2^2)\right)$

so using the above formulae,

$$X+Y \sim \text{Norm}\left((0\times1 + 0\times1) , ((1^2 \times 1^2) + (1^2 \times 1^2)) \right)$$

$$X+Y \sim \text{Norm}(0, 2)$$

using the weighted sum of independent normals,

$$X-Y = X + (-1) Y$$

here $a_1 = 1$, $a_2 = -1$

$$X-Y \sim \text{Norm} \left((1\times0 + (-1\times0)) , (((1)^2 \times 1) + ((-1)^2 \times 1)) \right)$$

$$X-Y \sim \text{Norm}(0, 2)$$

$X+Y$, $X-Y$ are identically distributed.

<u>using fact-2</u>

$$\text{cov}(X+Y, X-Y) = \left[(1\times1)\ \text{cov}(X, X)\right] + \left[(1)(-1)\ \text{cov}(X, Y)\right]$$

$$+ \left[(1\times1)\ \text{cov}(Y, X)\right] + \left[(1)(-1)\ \text{cov}(Y, Y)\right]$$

$$= \text{cov}(X, X) - \text{cov}(X, Y) + \text{cov}(Y, X) + \text{cov}(Y, Y)$$

we know that

$$\text{cov}(X, Y) = \text{cov}(Y, X)$$

$= cov(x,x) - cov(x,y) + cov(x,y) - cov(y,y)$

$= cov(x,x) - cov(y,y)$

from ① i.e x, y are independent so $cov(x,x) = cov(y,y)$

$= 0$

So,

$$cov(x+y, x-y) = cov(x-y, x+y) = 0$$

There is no corrdelation between $x+y$, $x-y$ So they are independent. [According to Assignment-2 Q1]

<u>using fact-1:</u>

Now,

$(x+y, x-y)$ both are normally distributed and both are having no corrdelation so According to fact-1 we can say that

$(x+y, x-y)$ is a multivariate normal

7) Given, there are n distinct types of Pokemon, with equal probability of catching each of the n types. Using X, which is the discrete RV that represents the number of days needed to catch all n types, we need to find $E[X]$ and $Var(X)$

a) Now, we start with 0 Pokemon. The first one we catch will always be unique. There will then be a certain number of days elapsing before we catch the second distinct, and then another interval of days between the 2^{nd} and 3^{rd} distinct Pokemon, and so on until we catch all n after X days from the start. If we use X_i to denote the number of days between catching the $(i-1)^{th}$ and i^{th} pokemon for $i = 1, 2 \dots n$, we can denote X as

$$X = X_1 + X_2 + \dots X_n = \sum_{i=1}^{n} X_i,$$

where $X_1 = 1$ is the base case

$$\therefore \quad E[X] = E\left[\sum_{i=1}^{n} X_i\right] \quad (i)$$

Using Linearity of Expectation, (i) simplifies to

$$E[X] = \sum_{i=1}^{n} E[X_i] = E[X_1] + E[X_2] + \dots E[X_n]$$

(ii)

→ Now, we have $X_1 = 1$. Expectation of a constant is the constant, i.e, $E[c] = c \Rightarrow E[X_1] = E[1] = 1$

→ Now, there are $n-1$ distinct pokemon left to be caught. The prob^y of catching one of them will be $n-1/n$. Similary, if i distinct Pokemon have already been caught, the probability of catching the remaining distinct ones will be $n-i/n$

Also, between catching $(i-1)^{th}$ pokemon to the i^{th} one, there are X_i number of days where we catch already caught pokemon. This follows a Geometric Dist^n, since there are X_i number of failure days before we catch a distinct new Pokemon i.

If C_i denotes the event of catching the i^{th} Pokemon, given we've already caught $i-1$,

$$Pr(C_i) = \frac{n-(i-1)}{n} \quad \text{(iii)}$$

Now, we know that the Expectation of a Geometric RV is 1/success probability.

$$\Rightarrow E[X_i] = 1/Pr(C_i) = \frac{1}{(n-(i-1)/n)} = \frac{n}{n-(i-1)}$$

Substituting above in (ii), we get

$$E[X] = \frac{n}{n-(1-1)} + \frac{n}{n-(2-1)} + \frac{n}{n-(3-1)} + \dots \frac{n}{n-(n-1)}$$

$$= \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \dots \frac{n}{1}$$

$$= n\left(\frac{1}{n} + \frac{1}{n-1} + \frac{1}{n-2} + \dots 1\right)$$

reversing the order of summation $\stackrel{\sim}{=} n\left(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n-1} + \frac{1}{n}\right)$

$$\boxed{E[X] = n\sum_{i=1}^{n} 1/i} \quad (1)$$

Page

b) From (iii), $Pr(C_i)$ is the success probability.
Using the variance definition for Geometric RVs,
we get $Var(X_i) = \frac{1 - Pr(C_i)}{(Pr(C_i))^2}$

$$\Rightarrow Var(X_i) = \frac{1 - \frac{n-(i-1)}{n}}{\left(\frac{n-(i-1)}{n}\right)^2} = \frac{\frac{\cancel{n} - \cancel{n} + i - 1}{\cancel{n}}}{\frac{(n-i+1)^2}{n^{\cancel{2}}}}$$

$$\Rightarrow Var(X_i) = \frac{n(i-1)}{(n-i+1)^2} \quad (iv)$$

$\longrightarrow$ Using Linearity of variance, we get
$$Var(X) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$$
(Since we know $X_1 \perp X_2 \dots \perp X_n$)

$$\Rightarrow Var(X) = \sum_{i=1}^{n} \frac{n(i-1)}{(n-i+1)^2} = n \sum_{i=1}^{n} \frac{i-1}{(n-i+1)^2}$$

$$= n\left(\frac{1-1}{(n-1+1)^2} + \frac{2-1}{(n-2+1)^2} + \dots + \frac{n-1-1}{(n-(n-1)+1)^2} + \frac{n-1}{(n-n+1)^2}\right)$$

reversing the summation
$$\overset{\sim}{=} n\left(\frac{n-1}{(1)^2} + \frac{n-2}{(2)^2} + \dots + \frac{n-(n-1)}{(n-1)^2} + \frac{n-n}{n^2}\right)$$

$$= n \sum_{i=1}^{n} \frac{n-i}{i^2}$$

$$= n^2 \sum_{i=1}^{n} 1/i^2 - n \sum_{i=1}^{n} 1/i$$

$$\therefore \boxed{Var[X] = n^2 \sum_{i=1}^{n} \frac{1}{i^2} - E[X]} \quad \text{(from (1))}$$

Scanned with CamScanner