PRATIK NAGELIA - 114122014
AMEYA SANKHE - 113219562
SHUBHAM AGRAWAL - 113166701
RANJAN KUMAR - 113262786

CSE 544, Spring 2021, Probability and Statistics for Data Science

Assignment 2: Random Variables Due: 3/09, 1:15pm, via Blackboard
(7 questions, 70 points total)

I/We understand and agree to the following:
(a) Academic dishonesty will result in an 'F' grade and referral to the Academic Judiciary.
(b) Late submission, beyond the 'due' date/time, will result in a score of 0 on this assignment.

(write down the name of all collaborating students on the line below)
PRATIK NAGELIA - 114122014, AMEYA SANKHE, SHUBHAM AGRAWAL, RANJAN KUMAR

1. Introduction to Covariance (Total 5 points)

The covariance of two RVs X and Y is defined as: Cov(X,Y) = E[(X - E[X]) (Y - E[Y])] = E[XY] - E[X] E[Y].
Covariance of independent RVs is always zero.

(a) In an experiment, an unbiased/fair coin is flipped 3 times. Let X be the number of heads in the first two flips and Y be the number of heads in the last two flips. Calculate Cov(X,Y). (2 points)
(b) Let X be a fair 5-sided dice with face values {-5, -2, 0, 2, 5}. Let Y = X^2. Calculate Cov(X,Y). (2 points)
(c) Does a zero covariance imply that the RVs are independent? Justify your answer. (1 point)

(a) Given:
$$Cov(X,Y) = E[(X-E[X])(Y-E[Y])] = E[XY] - E[X]E[Y]$$
Let the flipping of coins 3 times be F_1, F_2 & F_3.
Let the number of heads in first two flips be denoted by X
Let the number of heads in last two flips be denoted by Y

	F_1	F_2	F_3	X	Y	XY
1.	T	T	T	0	0	0
2.	T	T	H	0	1	0
3.	T	H	T	1	1	1
4.	T	H	H	1	2	2
5	H	T	T	1	0	0
6	H	T	H	1	1	1
7	H	H	T	2	1	2
8	H	H	H	2	2	4
Σ				8	8	10

From, the table above:

$$\Sigma X = 8$$

$$E[X] = \frac{\Sigma X}{n} = \frac{8}{8} = 1$$

Similarly, $E[Y] = \frac{8}{8} = 1$

$$\Sigma[XY] = 10$$

$$E[XY] = \frac{\Sigma[XY]}{n} = \frac{10}{8} = \frac{5}{4}$$

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

$$= \frac{5}{4} - 1 \times 1 = \frac{5}{4} - 1 = \frac{1}{4} \text{ Ans.}$$

(b) Given:

$$X = \{-5, -2, 0, 2, 5\}$$

$$Y = X^2 = \{25, 4, 0, 4, 25\}$$

$$XY = \{-125, -8, 0, 8, 125\}$$

$$E[X] = \frac{\Sigma X}{n} = 0, \quad E[Y] = \frac{\Sigma Y}{n} = \frac{58}{5}$$

$$E[XY] = \frac{\Sigma[XY]}{n} = 0$$

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

$$= 0 - 0 \times \frac{58}{5} = 0 \text{ Ans.}$$

(c) Zero Covariance does not imply that the RVs are independent.

In part (b) of this question, $Cov(X, Y) = 0$.
This does not mean that X and Y are independent.
There can be two theories to justify the same:

(i) To show that X and Y are not dependent, we can show that product rules fails, i.e.

$P(\cancel{X}Y) \neq P(X) \cdot P(Y)$

In the above question, $P(X=-2, Y=0) = 0$

$P(X = -2) = \frac{1}{5}$

$P(Y=0) = \frac{1}{5}$

$P(X=-2) \cdot P(Y=0) = \frac{1}{5} \times \frac{1}{5} = \frac{1}{25}$

Hence $P(XY) \neq P(X) \cdot P(Y)$ proved!

(ii) In the above question, we are finding covariance of X and Y where $Y = X^2$. X and X^2 are dependent on each other because, we can know the value of X^2 if we know the value of X.
Hence, we can say that X and X^2 are dependent and Cov (X, Y) is still 0.

2. Inequalities **(Total 10 points)**

Let X be a non-negative RV with mean μ and variance σ^2, and let t > 0 be some real number.

(a) Prove that $E[X] \geq \int_t^{\infty} xf(x)dx$. (3 points)

(b) Using part (a), prove that $\Pr(X > t) \leq \frac{E[X]}{t}$ (3 points)

(c) Using part (b), prove that $\Pr(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2}$ (4 points)

(a) $E[x] = \int_{-\infty}^{\infty} x f(x)dx$

Since, x is a non-negative RV

$E[x] = \int_0^{\infty} xf(x)dx$

$= \int_0^t xf(x)dx + \int_t^{\infty} xf(x)dx$ [Since $t > 0$ is a real number between 0 and ∞]

$\Rightarrow E[x] - \int_0^t xf(x)dx = \int_t^{\infty} xf(x)dx$

$\Rightarrow E[x] \geq \int_t^{\infty} xf(x)dx$ proved! [since, $\int_0^t xf(x)dx \geq 0$]

(b) Using part (a)

$E[x] \geq \int_t^{\infty} xf(x)dx \geq t \int_t^{\infty} f(x)dx$

$\therefore E[x] \geq t \Pr(x > t)$

$\therefore \Pr(x > t) \leq \frac{E[x]}{t}$ proved!

(c) Using part (b)

$$\Pr(X > t) \leq \frac{E[X]}{t}$$

$$\Rightarrow \Pr(|X - \mu| > t) \leq \frac{E[|X - \mu|]}{t}$$ [as eq. by question]

$\Rightarrow$ Squaring both sides, we get

$$\Pr((X - \mu)^2 \geq t^2) \leq \frac{E[(X - \mu)^2]}{t^2}$$

$\Rightarrow$ we know that : $E[(X - \mu)^2] = \sigma^2$

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

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

3. Functions of RVs **(Total 10 points)**

(a) Let $X_1, X_2, \dots, X_k$ be k independent exponential random variables with pdfs given by $f_{X_i}(x) = \lambda_i e^{-\lambda_i x},\ x \ge 0,\ \forall\, i \in \{1, 2, \dots, k\}$. Let $Z = \min(X_1, X_2, \dots, X_k)$.

i. Find the pdf of Z. (3 points)

ii. Find E[Z]. (1 point)

iii. Find Var(Z). (2 points)

(b) Let X and Y be two random variables with joint density function:

$f_{XY}(x,y) = \begin{cases} 2, & 0 \le x \le y \le 1 \\ 0, & otherwise \end{cases}$. Find the pdf of $Z = XY$. (4 points)

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

Let CDF of Z be $F_Z(\alpha)$

$$F_Z(\alpha) = P(Z \le \alpha)$$
$$= P(\min(X_1, X_2, \dots\dots, X_K) \le \alpha)$$
$$= 1 - P(\min(X_1, X_2, \dots, X_K) > \alpha)$$
$$= 1 - P((X_1 > \alpha) \cap (X_2 > \alpha) \cap \dots\dots \cap (X_K > \alpha))$$

$\Rightarrow$ as $X_1, X_2 \dots\dots X_K$ are independent

$\therefore P(AB) = P(A) \cdot P(B)$

$$F_Z(\alpha) = 1 - \left(P(X_1 > \alpha) * P(X_2 > \alpha) * \dots\dots * P(X_K > \alpha)\right)$$

We know, cdf X_K, $F_{X_K}(\alpha) = P(X_K \le \alpha)$ for $K = 1, 2, \dots\dots, k$

$$F_{X_K}(\alpha) = 1 - P(X_K > \alpha)$$

$$\therefore P(X_K > \alpha) = 1 - F_{X_K}(\alpha) \quad \text{for } K = 1, 2, \dots K$$

$$F_Z(\alpha) = 1 - \left((1 - F_{X_1}(\alpha)) * (1 - F_{X_2}(\alpha)) \dots\dots (1 - F_{X_K}(\alpha))\right) \text{———— } ①$$

Now $F_{X_K}(\alpha) = P(X_K \le \alpha)$

$$F_{X_K}(\alpha) = \int_0^{\alpha} \lambda_K e^{-\lambda_K x} dx$$

$$= \lambda_K \left[\frac{e^{-\lambda_K x}}{-\lambda_K} \right]_0^{\alpha}$$

$$= \left[e^{-\lambda_K x} \right]_{\alpha}^{0}$$

as it has negative term, to remove it we reverse the limits

$$= e^{-\lambda_K \times 0} - e^{-\lambda_K \alpha}$$

$$= 1 - e^{-\lambda_K \alpha}$$

for $K = 1, 2, 3 \dots K$

$$\therefore 1 - F_{X_K}(\alpha) = e^{-\lambda_K \alpha} \qquad \text{(a)}$$

for $K = 1, 2, 3 \dots K$

$$\therefore F_Z(\alpha) = 1 - \left((e^{-\lambda_1 \alpha}) * (e^{-\lambda_2 \alpha}) \dots (e^{-\lambda_K \alpha}) \right)$$

by subsituting (a) in (1)

$$= 1 - e^{-\alpha(\lambda_1 + \lambda_2 + \dots + \lambda_K)}$$

$\dots e^a * e^b = e^{a+b}$

$$F_Z(\alpha) = 1 - e^{-\alpha(\sum_{i=1}^{K} \lambda_i)}$$

$\therefore$ Cdf of Z ($F_Z(\alpha)$) is also exponential

$$\therefore F_Z(x) = 1 - e^{-x(\sum_{i=1}^{K} \lambda_i)}$$

To get Pdf we differentiate Cdf by x

$$\therefore \frac{\partial F_Z(x)}{\partial x} = -\left(-\sum_{i=1}^{K} \lambda_i\right) e^{-x(\sum_{i=1}^{K} \lambda_i)}$$

$$= \left(\sum_{i=1}^{K} \lambda_i\right) e^{-x(\sum_{i=1}^{K} \lambda_i)}$$

$$\therefore P(z) = \left(\sum_{i=1}^{K} \lambda_i\right) e^{-(\sum_{i=1}^{K} \lambda_i) \cancel{x} z}$$

$$\therefore \text{Pdf of } Z = \min(X_1, X_2 \dots, X_K) \Rightarrow P(z) = \left(\sum_{i=1}^{K} \lambda_i\right) e^{-(\sum_{i=1}^{K} \lambda_i) \cancel{x} (z)}$$

$\therefore$ Pdf of Z has ~~Geometric~~ Exponential Distribution

We know

When $X = \lambda e^{-\lambda X}$

$$E(X) = \frac{1}{\lambda}$$

$$Var(X) = \frac{1}{\lambda^2}$$

$\therefore$

ii when $Z = \sum_{i=1}^{K} \lambda_i e^{-\left(\sum_{i=1}^{K}\lambda_i\right)(Z)}$

$$\boxed{E(Z) = \frac{1}{\sum_{i=1}^{K}\lambda_i}}$$

iii) $$\boxed{Var(Z) = \frac{1}{\left(\sum_{i=1}^{K}\lambda_i\right)^2}}$$

b) Given

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

Let $Z = XY$

Cdf Z i.e $F_Z(z) = P(Z \le z)$
$= P(XY \le z)$

Now we know $0 \le x \le y \le 1$ and $xy \le z$

If

y limits are $0 \rightarrow 1$

Then x limits are $0 \rightarrow y$ as $x \leq y$

We know maximum value of $y = 1$

putting $y = 1$ in $xy \leq z$

then $x \leq z$

Now x limits are $0 \rightarrow z$

$$P(XY \leq z) = \int_{x=0}^{z} \int_{y=0}^{1} f_{XY}(x,y)\, dx\, dy$$

$$= \int_{x=0}^{z} \int_{y=0}^{1} 2\, dx\, dy$$

$$= 2 \int_{x=0}^{z} dx \left[y\right]_0^1$$

$\int dy = y + c$

$$= 2 \int_{x=0}^{z} dx [1-0]$$

$$= 2 \left[x\right]_0^z$$

$$P(XY \leq z) = 2z$$

$\int dx = x + c$

$\therefore$ Cdf of Z i.e $P(XY \leq z) = 2z$

To get pdf we differentiate Cdf of Z with z

$$\therefore \frac{\partial P(XY \leq z)}{\partial z} = P(z) = 2$$

$$\boxed{\therefore \text{Pdf } P(Z) = 2}$$

4. Daenerys returns to King's Landing, almost. **(Total 10 points)**

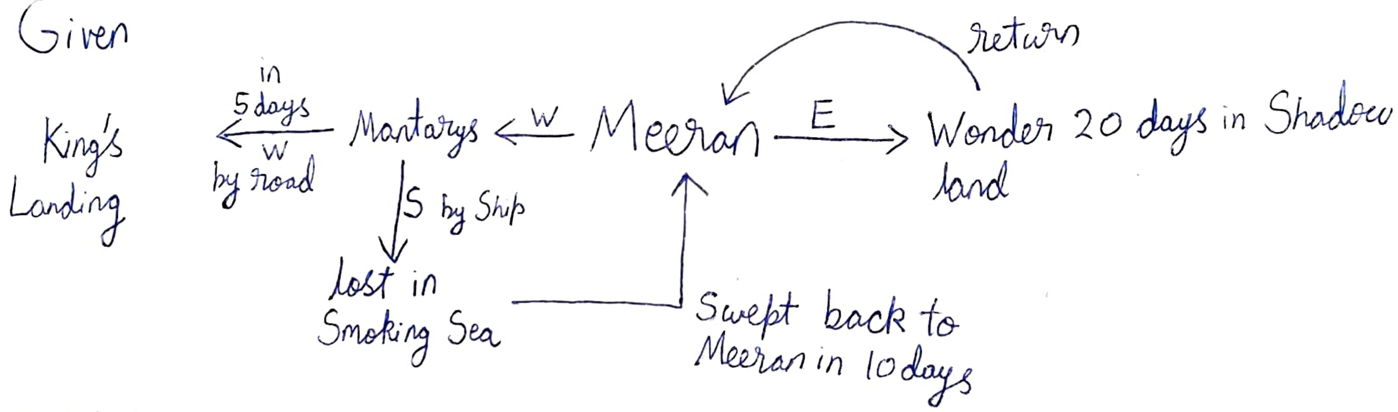
In an alternate universe of Game of Thrones (or A Song of Ice and Fire, for fans of the books), Daenerys Targaryen is finally ready to leave Meereen and return to King's Landing. However, she does not know the way. From Meereen, if she goes East, she will wander around for 20 days in the Shadow Lands and return back to Meereen. If she goes West from Meereen, she will immediately arrive at the city of Mantarys. From Mantarys , she can go West by road or South via ship. If she goes South, her ship will get lost in the Smoking Sea and will be swept back to Meereen after 10 days. However, if she goes West from Mantarys, she will eventually reach King's Landing in 5 days. Let X denote the time spent by Daenerys before she reaches King's Landing. Assume that she is equally likely to take either of two paths whenever presented with a choice and has no memory of prior choices.

(a) What is E[X]? (3 points)

(b) What is Var[X]? (7 points)

(Hint: Be careful with Var[X]. You want to use conditioning.)

Given



Let X be the time spent by Daenerys before reaching King's Landing from Meeran

∴ Probability of Going East $P(E) = 1/2$

∴ Probability of Going West and then South ⇒ $P(WS) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

Probability of Going West and then West ⇒ $P(WW) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

When She goes East, She wonders for 20 days and return back to Meeran

∴ $\cancel{E(E)} = E(X+20)$ ...... Since we add 20 days to the days to reach King's landing from Meeran as in 20 days we reach back to Meeran

$E(X/E)$

Similarly when she goes West and South, she reaches back to Meeran in 10 days

$\therefore$ Expectation of days to reach King's landing if she goes West and South

$\cancel{E(WS)} = E(X+10)$

$E(X/WS)$

...... Since we add 10 days to the days to reach King's Landing from Meeran as in 10 days we again return to Meeran

If She goes West and West,

Expectation of days to reach King's Landing

$\cancel{E(WW)} = 5$

$E(X/WW)$

$\therefore$ Now,

$\cancel{E(X) = P(E) \times E(E) + P}$

$$E(X) = E(X/_E) \times P(E) + E(X/_{WS}) \times P(WS) + E(X/_{WW}) \times P(WW)$$

$$= \frac{1}{2}(E(x+20)) + \frac{1}{4}(E(x+10)) + \frac{1}{4} \times 5$$

$$= \frac{1}{2}(E(x)+20) + \frac{1}{4}(E(x)+10) + \frac{5}{4}$$

by linearity of Expectation

$E(x+Y) = E(x) + E(Y)$

$$\therefore E(X) = \frac{E(x)}{2} + 10 + \frac{E(X)}{4} + \frac{10}{4} + \frac{5}{4}$$

$$\therefore E(X) = \frac{3*E(X)}{4} + \frac{55}{4}$$

$$\therefore \frac{E(X)}{4} = \frac{55}{4}$$

$$\boxed{\therefore E(X) = 55 \text{ days}}$$

b) $Var(X) = E(X^2) - (E(X))^2$

Now,

if she goes East

$P(E) = \frac{1}{2}$, $E(X^2/E) = E[(X+20)^2]$

if she goes West and South

$P(W,S) = \frac{1}{4}$, $E(X^2/W,S) = E[(X+10)^2]$

if she goes West and West

$P(W,W) = \frac{1}{4}$, $E(X^2/W,W) = E(5^2) = 25$

$$\therefore E(X^2) = E(X^2/E) \times P(E) + E(X^2/W,S) \times P(W,S) + E(X^2/W,W) \times P(W,W)$$

$$= \frac{1}{2} E[(X+20)^2] + \frac{1}{4} \times E[(X+10)^2] + \frac{25}{4}$$

$$= \frac{E(X^2+400+40X)}{2} + \frac{E[(X^2+100+20X)]}{4} + \frac{25}{4}$$

$$= \frac{E(X^2)+400+40E(X)}{2} + \frac{E(X^2)+100+20E(X)}{4} + \frac{25}{4}$$

by linearity of Expectation

$E(X+Y) = E(X) + E(Y)$

$E(\text{Constant}) = \text{Constant}$

$E(\text{Constant} \times X) = \text{Constant} \times E(X)$

$$E(X^2) = \frac{E(X^2)}{2} + 200 + 20E(X) + \frac{E(X^2)}{4} + 25 + 5E(X) + \frac{25}{4}$$

$$E(x^2) = \frac{3}{4}E(x^2) + \left(200 + 25 + \frac{25}{4}\right) + 25 \times E(X)$$

$$\frac{1}{4} \times E(X^2) = \left(225 + \frac{25}{4}\right) + 25 \times 55$$

$$E(x^2) = 4 \times 225 + 25 + 4 \times 25 \times 55$$

$$= 900 + 25 + 5500$$

$$E(x^2) = 6425$$

Now, $Var(X) = E(X^2) - (E(X))^2$

$$= 6425 - 3025$$

$$= 3,400$$

$$\boxed{\therefore Var(X) = 3,400 \text{ days}}$$

5. **Dependence on past 2 states** **(Total 15 points)**

Consider the Clear-Snowy problem from class. However, this time, assume that the weather tomorrow depends on the weather today AND the weather yesterday. While this does not seem to follow the Markovian property, you can modify the state space to work around this issue. Use the following notation and transition probability values:

Pr[Weather tomorrow is X_{i+1}, given that weather today is X_i and weather yesterday was X_{i-1}]
= Pr[X_{i+1} | X_i X_{i-1}] (note that each X is either *c* or *s*).
Pr[c | c c] = 0.9; Pr[c | c s] = 0.8; Pr[c | s c] = 0.5; Pr[c | s s] = 0.1.

(a) Find the eventual (steady-state) Pr[c c], Pr[c s], Pr[s c], and Pr[s s]. Show your Markov chain and the transition probabilities. (7 points)

(b) In steady-state, what is the probability that it will be snowy 3 days from today. (3 points)

(c) Solve the problem of finding the steady state probability via simulation (in python). You need to find the steady state by raising the transition matrix to a high power ($\pi = P^k; k \gg 1$) and then take any row of the exponentiated matrix ($\pi[i,:]$) as the steady state. For taking power of matrix in python, you can use **np.linalg.matrix_power(matrix, power)**. After you obtain the steady state distribution, solve part (b) numerically. (5 points)

Submit your code along with your solution as part of the zip/tar file on BB. Name your python file a2_5.py. The script should have a function **a ← `steady_state_power` (transition matrix)**, where **`steady_state_power`** **()** should have the implementation of Power method and the return value **a** is the final steady state. Also, in the hardcopy submission, you should mention the final steady state you obtained in the following format:

`Steady_State: Power iteration >> [xx, xx, xx, xx]`

(a) Pr [cc] = Probability of clear day given previous day is clear.

Pr [cs] = Probability of snowy day given ~~probab~~ previous day is clear.

Pr [sc] = Probability of clear day given previous day is snowy

Pr [ss] = Probability of snowy day given previous day is snowy

Showing transitions on a tabular form for better understanding.

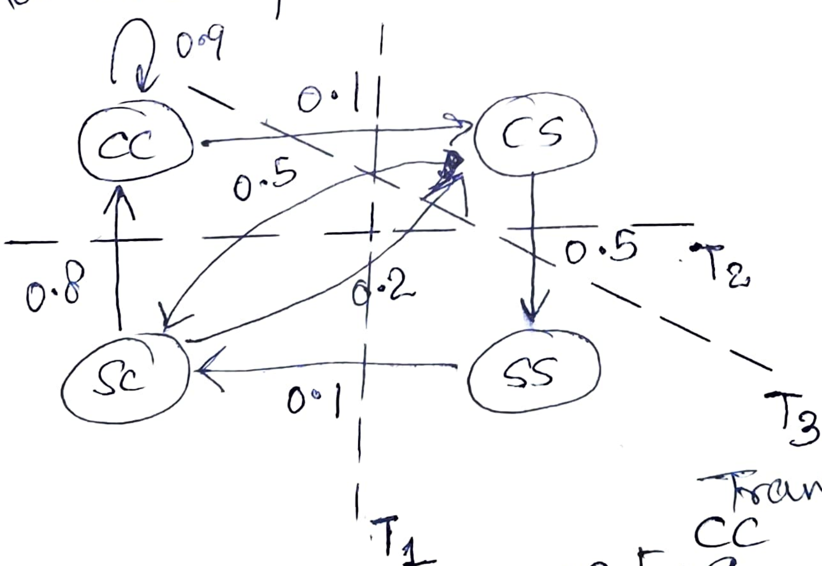
Given:

i−1	i	i+1	Probability
c	c	c	0.9
s	c	c	0.8
c	s	s	0.5
s	s	c	0.1

(1) To find transitions for Pr[CC], Pr[CS], Pr[SC], Pr[SS]

i−1	i	i+1	Probability	
C	C	C	0.9	Given
C	C	S	0.1	(1−0.9)
C	S	C	0.5	Given
C	S	S	0.5	(1−0.5)
S	S	C	0.1	Given
S	S	S	0.9	(1−0.1)
S	C	C	0.8	Given
S	C	S	0.2	(1−0.2)

Markov chain from the above table:



Transition Probabilities

$$
\begin{array}{c|cccc}
 & CC & CS & SS & SC \\
\hline
\begin{array}{c} CC \\ CS \\ SS \\ SC \end{array} &
\multicolumn{4}{c}{\begin{bmatrix} 0.9 & 0.1 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.9 & 0.1 \\ 0.8 & 0.2 & 0 & 0 \end{bmatrix}}
\end{array}
$$

From T_1; taking Local Balance

$$\pi_{cc} \times 0.1 + \pi_{sc} \times 0.2 = \pi_{cs} \times 0.5 + \pi_{ss} \times 0.1 \quad \text{--- (1)}$$

From T_2; taking Local Balance

$$\pi_{sc} \times 0.8 + \pi_{sc} \times 0.2 = \pi_{cs} \times 0.5 + \pi_{cs} \times 0.5 \quad \text{--- (2)}$$

$$\Rightarrow \quad \pi_{sc} = \pi_{cs} \quad \text{--- (3)}$$

Putting the value of (3) in (1)

$$\pi_{cc} \times 0.1 + \pi_{sc} \times 0.2 = \pi_{sc} \times 0.5 + \pi_{ss} \times 0.1$$

$$\Rightarrow \quad 0.1\,\pi_{cc} - 0.1\,\pi_{ss} = 0.3\,\pi_{sc} \quad \text{--- (4)}$$

we know, from sum of probabilities:

$$\pi_{ss} + \pi_{cs} + \pi_{sc} + \pi_{cc} = 1 \quad \text{--- (5)}$$

Putting the value of (3) in (5)

$$\Rightarrow \quad \pi_{ss} + 2\pi_{sc} + \pi_{cc} = 1 \quad \text{--- (6)}$$

From T_3; taking Local Balance

$$0.1 \times \pi_{cc} + 0.2 \times \pi_{sc} = 0.5\,\pi_{cs} + 0.5\,\pi_{cs}$$

$$\Rightarrow \quad 0.1\,\pi_{cc} + 0.2\,\pi_{sc} = \pi_{cs} \quad \text{--- (7)}$$

Putting the value of (3) in (7)

$$\Rightarrow \quad \pi_{sc} = 0.1\,\pi_{cc} + 0.2\,\pi_{sc}$$

$$\Rightarrow \quad 0.8\,\pi_{sc} = 0.1\,\pi_{cc}$$

$$\Rightarrow \quad \pi_{cc} = 8\,\pi_{sc} \quad \text{--- (8)}$$

Using 8 in 6

$\Pi_{ss} + 2\,\Pi_{sc} + 8\,\Pi_{sc} = 1$

$\Rightarrow \Pi_{ss} + 10\,\Pi_{sc} = 1$ ——— 9

Using 8 in 4

$\Pi_{cc} - \Pi_{ss} = 3\,\Pi_{sc}$

$\Rightarrow 8\,\Pi_{sc} - \Pi_{ss} = 3\,\Pi_{sc}$

$\Rightarrow 5\,\Pi_{sc} - \Pi_{ss} = 0$ ——— 10

Solving 10 and 9

$$
\begin{aligned}
10\,\Pi_{sc} + \cancel{\Pi_{ss}} &= 1 \\
5\,\Pi_{sc} - \cancel{\Pi_{ss}} &= 0 \\
\hline
15\,\Pi_{sc} &= 1 \\
\Pi_{sc} &= 1/15
\end{aligned}
$$

Using 3 $\Pi_{sc} = \Pi_{cs}$

$\therefore \Pi_{cs} = \frac{1}{15}$

Using 8 $\Pi_{cc} = 8\,\Pi_{sc} = 8 \times \frac{1}{15} = \frac{8}{15}$

Using 9 $\Pi_{ss} = 1 - 10\,\Pi_{sc}$

$= 1 - 10 \times \frac{1}{15} = \frac{15-10}{15} = \frac{5}{15}$

$$\boxed{\Pi_{sc} = \frac{1}{15} \quad \Pi_{cs} = \frac{1}{15} \quad \Pi_{cc} = \frac{8}{15} \quad \Pi_{ss} = \frac{5}{15}}$$

b) Probability that it will be snowy in 3 days from today in steady state

$$P(\text{Snow in 3 days}) = \Pi_{SC} \times P_{S|CS} + \Pi_{CS} \times P_{S|SC} + \Pi_{CC} \times P_{S|CC} + \Pi_{SS} \times P_{S|SS}$$

$P_{S|CS} = 0.2$, $P_{S|SC} = 0.5$, $P_{S|CC} = 0.1$, $P_{S|SS} = 0.9$

...... from ①

$$P(\text{Snow in 3 days}) = \frac{0.2}{15} + \frac{0.5}{15} + \frac{8 \times 0.1}{15} + \frac{5 \times 0.9}{15}$$

$$= \frac{6}{15}$$

$$= \frac{2}{5}$$

$$\boxed{P(\text{Snow in 3 days in steady state}) = \frac{2}{5}}$$

(c) Final steady state obtained from the python code:

Steady_State: Power iteration >> [0.53, 0.067, 0.33, 0.067]

Explaination:
index represents respective states.
Default power value provided in the code is taken as 150.

6. Multivariate Normal **(Total 10 points)**

A random vector $\mathbf{X} = (X_1, \ldots, X_k)$ is said to have a Multivariate Normal distribution if every linear combination of Xj has a Normal distribution. That is, we require $t_1X_1 + t_2X_2 + \ldots + t_kX_K$ to have a Normal distribution for any real values of $t_1, \ldots, t_k$. As a special case, we consider $t_1X_1 + t_2X_2 + \ldots + t_kX_K$ to be a degenerate normal distribution with variance 0 if $t_1X_1 + t_2X_2 + \ldots + t_kX_K$ is a constant (such as when all t_j's are 0).

(a) If $\mathbf{X} = (X_1, \ldots, X_k)$ is a Multivariate Normal, show that the distribution of any X_j is Normal. (1 point)

(b) It is possible to have normally distributed random variables $X_1, \ldots, X_k$ such that $(X_1, \ldots, X_k)$ is not Multivariate Normal: Let X = Normal(0, 1) and S = 1 with probability ½ and -1 with probability ½. Then Y = SX is normal due to the symmetry of the Standard Normal. In this case show that (X, Y) is not a Multivariate Normal. (2 points)

(c) Let Z, W be i.i.d Normal(0, 1) random variables. Show that (Z, W) and (Z + 2W, 3Z + 5W) are Multivariate Normals. (2 points)

(d) If $\mathbf{X} = (X_1, \ldots, X_n)$ and $\mathbf{Y} = (Y_1, \ldots, Y_m)$ are Multivariate Normal vectors with X independent of Y, then show that the concatenated vector $W = (X_1, \ldots, X_n, Y_1, \ldots, Y_m)$ is also a Multivariate Normal. (2 points)

(e) <u>Fact 1 (Uncorrelated implies independence):</u> If X is a Multivariate Normal that can be written as $\mathbf{X} = (\mathbf{X_1}, \mathbf{X_2})$, where $\mathbf{X_1}$ and $\mathbf{X_2}$ are subvectors, and every component of $\mathbf{X_1}$ is uncorrelated with every component of $\mathbf{X_2}$, then $\mathbf{X_1}$ and $\mathbf{X_2}$ are independent.

<u>Fact 2 (Property of Covariance):</u> For any random variables X, Y, W, and V, we have:
Cov(aX + bY, cW + dV) = ac Cov(X, W) + ad Cov(X, V) + bc Cov(Y, W) + bd Cov(Y, V).

Let X, Y be i.i.d. standard Normals. Use Fact 1 and Fact 2 to show that (X+Y, X-Y) is a Multivariate Normal. (3 points)

(a) If $X = (X_1 \ldots X_K)$ is a Multivariate Normal, then By definition, $t_1X_1 + t_2X_2 + \ldots t_KX_K$ is a Normal Distribution.

By Weighted Sum of Normals Rule, if any X_j is Normal, then their linear combination $(t_1 X_1 + t_2 X_2 \ldots t_K X_K)$ is a Normal Distribution.

Hence Proved, Any X_j is Normal.

5(d) Given $X=(X_1 \dots X_n)$ and $Y=(Y_1 \dots Y_m)$ are multivariate Normal vectors with X independent of Y.

$\therefore$ By definition,

$S_1X_1 + S_2X_2 + \dots S_nX_n$ have a Normal Distribution for an real values of $S_1 \dots S_n$.

$\therefore$ Similary $t_1Y_1 + \dots t_mY_m$ have a normal distribution. for all real values of $t_1 \dots t_n$

Since they are idependent,

$\therefore$ There sum is also a Normal Distribution by.

Weighted Sum of Normals Rule

$\therefore S_1X_1 + S_2X_2 \dots S_nX_n + t_1Y_1 + \dots t_mY_m \sim$ Normal Distribution

$\therefore W = (X_1, \dots X_n, Y_1 \dots Y_m)$ is also a Multivariate Normal.

6-(b) $X \sim$ Normal $(0,1)$

$$S = \begin{cases} 1 & \text{with probality } 1/2 \\ -1 & \text{with probality } 1/2. \end{cases}$$

$Y = SX$ is normal due to Symmetry of Standard Normal.

$\therefore$ Let $Z = X + Y$

$$Z = \begin{cases} 2X & \text{with probablity } 1/2 \\ 0 & \text{with probablity } 1/2 \end{cases}$$

$\therefore X+Y$ is clearly not a Normal Distribution.

$\therefore (X, Y)$ is not a Multivariate Normal.

Q6.(c) Given, $Z, W \overset{iid}{\sim} Normal(0,1)$

By the property, Weighted Sum of Independent Normals,

If $N_i \sim Normal(\mu_i, \sigma_i^2)$, for $i = 1 \dots n$, and all N_i are $\perp$

then $\sum_{i=1}^{n} a_i N_i \sim Normal\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$

Hence, any linear combination of Z & W is a normal

$\therefore$ $t_1 Z + t_2 W \sim$ Normal

$\therefore$ Now since the above is normal,

$\therefore$ (Z, W) is a Multivariate Normal.

(As per definition in the question!)

Along the same lines,

In order to $(Z+2W, 3Z+5W)$ to be Multivariate Normal,

we need to ~~prove~~ establish that

① ---- $t_1(Z+2W) + t_2(3Z+5W)$ has to be a Normal distⁿ

This can be written as

$= t_1 Z + 2W t_1 + 3t_2 Z + 5t_2 W$

$= (t_1 + 3t_2) Z + (2t_1 + 5t_2) W$ ---- ②

With Weighted ~~som~~ Sum of Independent Normals property,

② is a Normal Distribution.

$\therefore$ $(Z+2W, 3Z+5W)$ is a Multivariate Normal.

Q6. (e) $Cov(aX + bY, cW + dV) = ac\,Cov(X, W) + ad\,Cov(X, V)$
$+ bc\,Cov(Y, W) + bd\,Cov(Y, V)$

Let $a = 1$, $b = 1$, $c = 1$, $d = -1$

$\therefore Cov(X+Y, X-Y)$
$= Cov(X, X) - Cov(X, Y) + Cov(Y, X) - Cov(Y, Y)$

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

[Since $Cov(X, Y) = Cov(Y, X)$]

$= Var(X) - Var(Y)$ [Since $cov(X, X) = Var(X)$

$= 0$ [Since X, Y are iid Standard Normal, $Var(X) = Var(Y$

$\therefore$ ~~X is not correlated with Y~~ $\therefore$ $X+Y$ is not correlated with $X-Y$ as $COV(X+Y, X-Y) = 0$

Now from Fact 1,
since $X+Y$ is not correlated with $X-Y$,
therefore $(X+Y)$ is independent to $(X-Y)$

Also since, X, Y are iid standard Normal, then $X+Y$ & $X-Y$ are normals [by Weighted Sum of Independent Normal property)

Hence, By definition of the Weighted Sum of Independent Normal, $(X+Y, Y-Y)$ is a Multivariate Normal.

7. Pokémon Go fanatic **(Total 10 points)**

Let us assume there are only n distinct types of Pokémons to capture in the entire Pokémon world, though there is an infinite supply of each type. Every day, you capture exactly one Pokémon. The Pokémon that you capture could be any one of the n types of Pokémons with equal probability. Your goal is to capture at least one Pokémon of all n distinct types. Let X denote the number of days needed to complete your goal.

(a) What is E[X]? (5 points)

(b) What is Var[X]? (5 points)

We do not need closed-forms here for parts (a) and (b).

(a) Ans. Let X be the no of days required to collect all n distinct types of Pokemon.

Let x_i be the random variable which denotes the time taken to collect i^{th} pokemon after $(i-1)$ pokemon had been collected.

So we have $X = x_1 + x_2 + x_3 + \cdots + x_n$ — ①

Let p_i be the probability of collecting new coupon.

So $p_i = \dfrac{\text{No of new coupons left}}{\text{Total no of coupons}} = \dfrac{n-(i-1)}{n} = \dfrac{n-i+1}{n}$

since i^{th} pokemon is collected only after $(i-1)^{th}$ pokemon had been collected. As there is an infinite supply of each pokemon of type i, hence for pokemon of type i, we need to make sure that pokemon of type i has not been collected before it. Hence for first success of pokemon of type i, there needs to be $(i-1)$ failure before the first success of pokemon of type i. So x_i has a geometric progression and Mean for any Random variable x_i of geometric progression is given $E(x_i) = \dfrac{1}{p_i} = \dfrac{n}{n-i+1}$ — ②

So From Equation 1, we get :-

$E(X) = E[x_1 + x_2 + x_3 + \cdots + x_n]$

From Linearity of Expectation, we get :-

or, $E(X) = E[x_1] + E[x_2] + \cdots + E[x_n]$

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

$= n\left[\frac{1}{n} + \frac{1}{n-1} + \frac{1}{n-2} + \cdots + \frac{1}{1}\right]$

(b) Ans. For any geometric random variable x_i,

$var(x_i) = \frac{q_i}{p_i^2}$ — ③ where $q_i = (1 - p_i)$

From Equation ①, we get :-

$Var(X) = Var(x_1 + x_2 + x_3 + \cdots + x_n)$

or, $Var(X) = Var(x_1) + var(x_2) + \cdots + var(x_n)$ — ④

{using Linearity of variance. since x_i are independent}

From Equation ③, we get :-

$var(x_i) = \frac{q_i}{p_i^2} = \frac{i}{n} \times \frac{n^2}{(n-i)^2} = \frac{ni}{(n-i)^2}$

Since x_i are independent,

So $Var(X) = \sum_{i=0}^{n-1} \frac{ni}{(n-i)^2} = \sum_{i=1}^{n} \frac{n(n-i)}{i^2}$

or, $Var(X) = \sum_{i=1}^{n} \frac{n^2 - ni}{i^2} = \sum_{i=1}^{n} \frac{n^2}{i^2} - n\sum \frac{i}{i^2}$

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

or, $Var(X) = n^2\left[\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots + \frac{1}{n^2}\right] - n\left[\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots \frac{1}{n}\right]$