CSE 544, Spring 2020, Probability and Statistics for Data Science

Assignment 2: Random Variables Due: 2/24, in class

(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)

126887696 112674911 113061886 112551470 112684130 112007720

Dhiraj Smit, Adnan Vasanwalla, Akanksha Dara, Anmol Shukla, Gargi Sawhney, Vishal Sarda

1. Transformation of Normal random variables **(Total 5 points)**

(a) If $X \sim Normal(\mu, \sigma^2)$ and $Y = aX + b$, with a > 0, prove that $Y \sim Normal(a\mu + b,\ (a\sigma)^2)$. (3 points)

(b) If $X$and Y are i.i.d. standard Normal RVs, show that $X + Y \sim Normal(0, 2)$. In general, the sum of any two independent Normal RVs is also a Normal. (2 points)

(a) $X \sim Normal(\mu, \sigma^2)$ $\quad Y = aX + b.$

$$f_X(\alpha) = \int_{-\infty}^{\alpha} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$$X = \frac{Y-b}{a}$$

$$dx = \frac{dy}{a}$$

substituting $X = \frac{Y-b}{a}$:

$$f_X(y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{\left(\frac{Y-b}{a} - \mu\right)^2}{\sigma^2 2}} \frac{dy}{a}$$

$$= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}(\sigma a)} e^{-\frac{(y-(b+a\mu))^2}{(\sigma a)^2 2}} dy.$$

Let $\mu_Y = b + \mu a$ and $\sigma_Y = a\sigma$.

$$\therefore f_X(y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}\,\sigma_Y} e^{-\frac{(y-\mu_Y)^2}{\sigma_Y^2 2}} dy$$

PDF for Normal distribution for $Y(\mu_Y, \sigma_Y)$

$$\boxed{\therefore Y \sim Normal(a\mu + b,\ \sigma^2 a^2)}$$ Hence proved.

(b) $X \sim Norm(0,1)$ $\quad Y \sim Norm(0,1)$

MGF of X : $\psi_X(t) = e^{\left(\mu t + \frac{\sigma^2 t^2}{2}\right)}$

$\psi_X(t) = e^{t^2/2} \quad \because \mu = 0, \sigma = 1$

MGF of Y: $\psi_Y(t) = e^{t^2/2}$

MGF of $(Y+X) = \psi_X(t) \cdot \psi_Y(t)$

$= e^{t^2/2} \cdot e^{t^2/2} = e^{t^2}$

Let μ_{x+y} be the mean of this distribution &

σ_{x+y} be the variance.

so, $e^{t^2} = e^{\left(\mu_{x+y} \cdot t + \frac{\sigma^2_{x+y} t^2}{2}\right)}$

$\Rightarrow \boxed{\mu_{x+y} = 0} \quad \boxed{\sigma^2_{x+y} = 2}$

Hence $X+Y \sim Normal(0,2)$

Alternatively, this can be verified as:

$\because$ X & Y are i.i.d std normal $\sim N(0,1)$

~~M~~ $\cdot \mu = 0, \sigma = 1$

$\because E[X] = E[Y] = 0$

$Var[X] = Var[Y] = 1$

Let $Z = X+Y$

Taking expectation

$E[Z] = E[X+Y]$

$\overset{LoE}{=} E[X] + E[Y]$

$= 0 + 0 = 0$

$Z = X+Y$

Taking variance

$Var(Z) = Var(X+Y)$

$= Var(X) + Var(Y)$

$= 1+1 = 2$

$\therefore Z \sim N(0,2)$

2. 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) X – no. of heads in first 2 flips | Y – no. of heads in last 2 flips

X	0	1	2
P(x)	1/4	1/2	1/4

Y	0	1	2
P(y)	1/4	1/2	1/4

$E[X] = \mu_X = \sum x_i p_X(x_i)$

$= 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot 1/4$

$\boxed{\mu_X = 1}$

$E[Y] = \mu_Y = \sum y_i p_Y(y_i)$

$= 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4}$

$\boxed{\mu_Y = 1}$

Ω (sample space)

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	2/8

$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$

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

$E[XY] = \sum x_i y_i P_{x_i y_i} = 1 \cdot 1 \cdot \frac{1}{8} + 1 \cdot 2 \cdot \frac{1}{8} + 2 \cdot 1 \cdot \frac{1}{8} + 2 \cdot 2 \cdot \frac{1}{8}$

$= 10/8 = 1.25$

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

$= 1.25 - 1 \cdot 1 = 0.25$

$\boxed{Cov(X,Y) = 0.25}$

(b) face value of a fair-sided dice:

X	-5	-2	0	2	5
$P_X(x)$	1/5	1/5	1/5	1/5	1/5

$E[X] = -5.\frac{1}{5} + -2.\frac{1}{5} + 0.\frac{1}{5} + 2.\frac{1}{5} + 5.\frac{1}{5}$

$\boxed{E[X] = 0}$

$Y = X^2$

Y	0	4	25
$P_X(y)$	1/5	2/5	2/5

$E[Y] = 0.\frac{1}{5} + 4.\frac{2}{5} + 25.\frac{2}{5}$

$= 58/5$

$\boxed{E[Y] = 11.6}$

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

$E[XY] = \sum x_i y_i \, P_{XY}(x_i, y_i)$

Y \ X	-5	-2	0	2	5
0	0	0	1/5	0	0
4	0	1/5	0	1/5	~~45~~
25	1/5	0	0	0	1/5

$E[XY] = \left(-2.4.\frac{1}{5}\right) + \left(-5.25.\frac{1}{5}\right) + 0 + \left(2.4.\frac{1}{5}\right) + \left(5.25.\frac{1}{5}\right)$

$\Rightarrow \boxed{E[XY] = 0}$.

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

$= 0 - 0$

$\Rightarrow \boxed{Cov(X,Y) = 0}$

(c) No

eg. Let X & Y be two random variables.

$X = \{-1, 1\}$

X	-1	1
P_X	1/2	1/2

$E[X] = 0$

$Y = \{-1, 1\}$

Y	-1	1
P_Y	1/2	1/2

$E[Y] = 0$.

$XY = \{(-1,1), (-1,-1)\}$

Y \ X	-1	1
-1	1/2	0
1	1/2	0

$E[XY] = -1.1.\frac{1}{2} + 1.1.\frac{1}{2}$

$\boxed{E[XY] = 0}$

$\Rightarrow E[XY] = E[X]E[Y]$

But $P_r(X \cap Y) \neq P_r(X) . P_r(Y) \Rightarrow$ X&Y are not independent

for $X = -1, Y = -1$ $\quad P_r(X \cap Y) = 1/2 \quad P_r(X)P_r(Y) = \frac{1}{2}.\frac{1}{2} = \frac{1}{4}$

5/5

3. 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) $E[X] = \int_{-\infty}^{\infty} x f_x(x) dx = \int_0^{\infty} x f_x(x) dx$ as $x > 0$

Let t be some number b/w 0 and ∞. $(0 < t < \infty)$

then we can write $E[X]$ as:

$$E[X] = \underbrace{\int_0^t x . f(x) dx}_{A} + \underbrace{\int_t^{\infty} x f(x) dx}_{B}$$

for part Ⓐ, we know that $x > 0$ and $f_x(x) \geq 0$ and $t > 0$.

so. $\int_0^t x f_x(x) dx \geq 0$.

$$\therefore E[X] = (\geq 0) + \int_t^{\infty} x f_x(x) dx$$

$$\Rightarrow \boxed{E[X] \geq \int_t^{\infty} x f_x(x) dx}$$ Hence proved.

(b) $$E[X] = \underbrace{\int_0^t x f_x(x) dx}_{\geq 0 \text{ (from A)}} + \int_t^{\infty} x f_x(x) dx$$

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

(using the result from (a) and explaination on the side)

$$E[X] \geq t \int_t^{\infty} f(x) dx$$ (t is const w.r.t. x)

$\because t > 0, x > 0.$

$\Rightarrow x_{min} = t$

$x \geq t$

multiplying by $f(x)$ on both sides: $(f(x) \geq 0)$

$x f(x) \geq t . f(x)$

Integrating both sides:

$$\int_t^{\infty} x f(x) dx \geq \int_t^{\infty} t f(x) dx$$

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

$$E[X] \geq t\, P[X > t].$$

$$\Rightarrow \boxed{P[X > t] \leq \frac{E[X]}{t}}$$

Hence proved.

(c) Let $\mu = E[X]$ and $\sigma^2 = Var(X)$.
from Markov's inequality in (b), we know that

$$P(X > t) \leq \frac{E[X]}{t} \qquad (t > 0)$$

So,

$$P(|X - \mu| \geq t) \leq \frac{E[X - \mu]}{t}$$

Squaring both sides:

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

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

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

Hence proved.

$t > 0$
$|X - \mu| > 0$.
from not possible $|X - \mu| \leq -t$
or
$|X - \mu| \geq t$
$\therefore$ we only consider $P(|X - \mu| \geq t)$

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

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

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:

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

(a) Let $X_1, X_2 \ldots X_k$ be k independent RV's.

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

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

$F_{X_i}(\alpha) = \int_0^\alpha \lambda_i e^{-\lambda_i x} dx$

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

CDF $\boxed{F_{X_i}(\alpha) = 1 - e^{-\lambda_i \alpha}} = P(X \leq \alpha)$

$\Rightarrow P(X > \alpha) = 1 - P(X \leq \alpha)$

$\boxed{P(X > \alpha) = e^{-\lambda_i \alpha}}$

$P_r(Z > \alpha) = P_r(\min(X_1, X_2 \ldots X_k) > \alpha)$

if $X_1 > \alpha, X_2 > \alpha \ldots X_k > \alpha \Rightarrow$ then $\min(X_1, X_2 \ldots X_k) > \alpha$

then we can write $P_r(Z > \alpha)$ as:

$P_r(Z > \alpha) = P_r((X_1 > \alpha) \cap (X_2 > \alpha) \ldots \cap (X_k > \alpha))$

As $X_1, X_2 \ldots X_k$ are independent random variables:

$P_r(Z > \alpha) = P_r(X_1 > \alpha) \cdot P_r(X_2 > \alpha) \cdot P_r(X_3 > \alpha) \ldots P_r(X_k > \alpha)$

$= e^{-\lambda_1 \alpha} \cdot e^{-\lambda_2 \alpha} \cdot e^{-\lambda_3 \alpha} \ldots e^{-\lambda_k \alpha}$

$= \prod_{i=1}^{k} e^{-\lambda_i \alpha} = e^{-\sum_{i=1}^{k} \lambda_i \alpha}$.

CDF. $P_r(Z \leq \alpha) = 1 - P_r(Z > \alpha)$

$\boxed{F_Z(\alpha) = P(Z \leq \alpha) = 1 - e^{-\sum_{i=1}^{k} \lambda_i \alpha}}$

PDF. $f_Z(\alpha) = \frac{d F_Z(\alpha)}{d\alpha} = 0 - \left(-\sum_{i=1}^{k} \lambda_i \alpha\right) e^{-\sum_{i=1}^{k} \lambda_i \alpha}$

$$\therefore \boxed{PDF = f_Z(\alpha) = \left(\sum_{i=1}^{k} \lambda_i\right) e^{-\sum_1^k \lambda_i \alpha}} \quad (i)$$

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

then $f_X(x)$ is an exponential distribution

$$E[X] = \int_{-\infty}^{\infty} x \underbrace{\lambda e^{-\lambda x} dx}_{dg}.$$

let $g = -e^{-\lambda x}$, $dg = \lambda e^{-\lambda x} dx$

Integration by parts:

$$\int f dg = fg - \int g df$$

$$E[X] = x \cdot (-e^{-\lambda x}) \Big|_0^\infty + \int_0^\infty e^{-\lambda x} dx$$

$$= 0 + \frac{e^{-\lambda x}}{-\lambda} \Big|_0^\infty.$$

$$\boxed{E[X] = 1/\lambda}.$$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 \lambda e^{-\lambda x} dx$$

$$= \int_0^\infty x^2 dg.$$

$$= x^2 (-e^{-\lambda x}) \Big|_0^\infty + \int_0^\infty e^{-\lambda x}$$

$$= 0 + \frac{2}{\lambda^2}$$

$$= 2/\lambda^2$$

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

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

(iii) $\boxed{Var(X) = 1/\lambda^2}$.

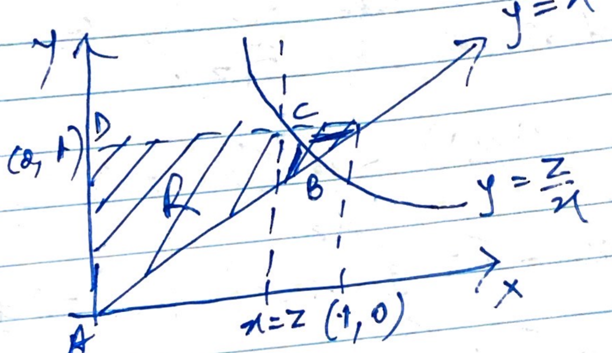
A4.b) $f_{XY}(x, y) = \begin{cases} 2 & 0 \le x < y \le 1 \\ 0 & \text{otherwise} \end{cases}$

$Z = XY$

CDF of $Z = F(z) = F(Z \le z)$

$= \iint_R f_X(x, y)\, dy\, dx$

To find the region under the curve, let us consider the digram below:-



R = region bounded by points A, B, C and D

To find co-ordinates of B :-

$y = \frac{z}{x} = x \Rightarrow \therefore x = \sqrt{z}$

$y = \sqrt{z}$

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

Splitting the region R at ~~point~~ $x = z$ line, we get,

$F(z) = \int_{x=0}^{z} \int_{y=x}^{1} 2\, dy\, dx + \int_{x=z}^{\sqrt{z}} \int_{y=x}^{z/x} 2\, dy\, dx$

$= 2\left[\int_{x}^{1} (1 - x)\, dx \right] + 2\left[\int_{z}^{\sqrt{z}} \left(\frac{z}{x} - x \right) dx \right]$

$$= 2\left[x - \frac{x^2}{2}\right]_0^z + 2\left[z\ln(x) - \frac{x^2}{2}\right]_z^{\sqrt{z}}$$

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

$$= 2z - \cancel{z^2} + 2z\ln(\sqrt{z}) - 2z\ln(z) - z + \cancel{z^2}$$

$$= z + 2z\ln(\sqrt{z}) - 2z\ln(z)$$

$$= z + z\ln(z) - 2z\ln(z)$$

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

Now to find the pdf of z, we take the derivative of $F(z)$.

$$\therefore f(z) = \frac{d}{dz}(F(z)) = \frac{d}{dz}[z - z\ln(z)]$$

$$= \frac{dz}{dz} - \left[z\frac{d}{dz}\ln(z) + \ln(z)\frac{dz}{dz}\right]$$

$$= 1 - \left[\cancel{z} \times \frac{1}{\cancel{z}} + \ln(z)\right]$$

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

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

5. **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.)

(a) By law of total expectation,

$E[X] = E[X \mid \text{she goes East from Mereen}] \cdot Pr(\text{she goes east from Mereen}) + E[X \mid \text{she goes west from Mereen}] \cdot Pr(\text{she goes west from Mereen})$ → (1)

Now,

$E[X \mid \text{she goes East from Mereen}] = E[X] + 20$

$E[X \mid \text{she goes west from Mereen}] = E[X \mid \text{she goes west from Mantarys}] \times Pr(\text{she goes west from Mantarys}) + E[X \mid \text{she goes south from mantarys}] \times Pr(\text{she goes south from Mantarys})$

$E[X \mid \text{she goes south from Mantarys}] = E[X] + 10$

$E[X \mid \text{she goes west from Mantarys}] = 5$

Since all the probabilities are equally likely of choosing any of the directions,

$Pr(\text{she goes east from Mereen}) = Pr(\text{she goes west from Mereen}) = \frac{1}{2}$

$Pr(\text{she goes south from mantarys}) = Pr(\text{she goes west from Mantarys}) = \frac{1}{2}$

⇒ Substituting all the values in (1):

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

$$= \frac{1}{2}E[X] + 10 + \frac{5}{4} + \frac{1}{4}E[X] + \frac{10}{4} = \frac{3}{4}E[X] + 10 + \frac{15}{4}$$

$$\frac{1}{4}E[X] = \frac{55}{4} \quad \Rightarrow E[X] = 55$$

(b) E[X] value from previous question (5a)

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

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

↳ Linearity of expectation can be used to obtain

$$var(X) = E[var(X|Y)] + var(E[X|Y]) = E[X^2] - (E[X])^2$$

Now,

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

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

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

↳ By linearity of expectation

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

$$\frac{E[X^2]}{4} = 225 + \frac{25}{4} + 25E[X]$$

$$E[X^2] = 900 + 25 + 100E[X]$$

$$E[X^2] = 900 + 25 + 100 \times 55$$

$$E[X^2] = 6425$$

$$\therefore var(X) = E[X^2] - (E[X])^2$$

$$= 6425 - (55)^2$$

$$= 6425 - 3025$$

$$\boxed{var(X) = 3400}$$ Ans

y Away From Stocks! (Total 15 points)

always been your dream to own a CMW car, which costs \$Y, but unfortunately you only have \$X, . X < Y both being positive integers. To overcome this shortage, you decide to bet on the stock rket and buy shares of a stock for \$X. The stock value is known to change as a simple random walk, ., its value either increases or decreases by \$1 every day with probability 0.5. Assume that (i) you will iquidate your stock (convert stock into cash) once your stock reaches the target value of \$Y, and (ii) if your stock value decreases to \$0, then you can't recover (the two stopping criteria). Model the scenario as a discrete time Markov chain to answer the questions below.

(a) What is the probability that your stock value reaches the high of $\$Y$? (4 points)

(b) What is the probability that your stock value reaches low of $\$0$? (3 points)

(c) What is the expected value of your stock at the end? (2 points)

(d) Solve parts (a), (b), and (c) above via simulation (in python). Simulate the stock value as a random walk. To simulate a random walk, start with an initial value as the current state. Generate a uniform random variable $u = U[0, 1]$, if $u < p$ (p is the probability of increase in stock value), then increase the current state by 1, else decrease it by one. Keep repeating this process until you meet either of the stopping criteria. To calculate the above-mentioned probability of events and the expected values, we will go with the frequentist interpretation of probability based on large number of repeated trials. $P(A) = \frac{N_A}{N}$ and $E[X] = \frac{\sum X_i}{N}$ where N_A is the number of favorable events and N is the number of trials and X_i is the outcome of i^{th} trial. So essentially you will be repeating the random walk N times ($N \gg 1$) and calculate the quantities asked for in 6(a) -- 6(c).

Submit your code via the google form link https://forms.gle/uFF4U4Th7YYAhxRn6. For this question, your code submission should include a python file, a2_6.py. The script should have a function `[a, b, c]← rand_walk(init_val, final_high, final_low, prob, N)` where the returned values `a, b, c,` are the answers for 6(a) -- 6(c), respectively, and the function arguments `init_val, final_high, final_low, prob` and `N` are initial stock value, final stock high, low values (stopping criteria), probability of upward movement, and the number of trials, respectively. For 6(d), also mention the final answers in your hardcopy submission for the following test cases in the specified format. (6 points)

Test cases

Case #	Init_val	Final_high	Final_low	Prob	N
1	100	150	0	0.50	10000
2	100	200	0	0.52	10000
3	200	250	0	0.54	10000

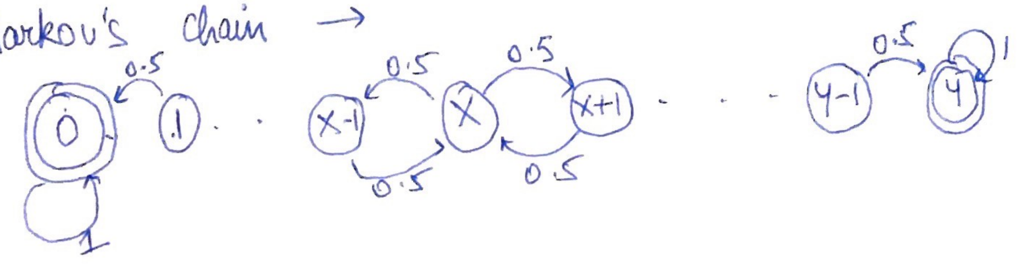
Output format

TEST CASE 1 >> (answer for 6(a)) (answer for 6(b)) (answer for 6(c))

TEST CASE 2 >> (answer for 6(a)) (answer for 6(b)) (answer for 6(c))

TEST CASE 3 >> (answer for 6(a)) (answer for 6(b)) (answer for 6(c))

Markov's Chain →



(a) If the person ~~reaches~~ starts from state Y, probability of reaching state Y

If the person ~~reaches~~ starts from state 0, probability of reaching state Y =

To find, the probability of reaching Y starting from X. Consi

to be P(X).

We can write the recursion as,

$$P(X) = 0.5P(X+1) + 0.5P(X-1)$$

At $X=1$, $P(1) = 0.5P(2) + 0.5\underbrace{P(0)}_{0}$

$$\Rightarrow P(1) = 0.5P(2)$$

$$\boxed{P(2) = 2P(1)}$$

At $X=2$, $P(2) = 0.5P(1) + 0.5P(3)$

$$\Rightarrow 2P(1) = 0.5P(1) + 0.5P(3)$$

$$\Rightarrow \boxed{P(3) = 3P(1)}$$

At $X=3$, $\underset{3P(1)}{P(3)} = 0.5P(2) + 0.5P(4)$

$$= 0.5 \times 2P(1) + 0.5P(4)$$

$$\Rightarrow \boxed{P(4) = 4P(1)}$$

This forms an Arithmetic Progression with first term $\boxed{P(1)} = a$

and difference $= P(2) - P(1)$

$$= 2P(1) - P(1)$$

$$= \boxed{P(1)} = d$$

Now, $P(X) = a + (n-1)d$

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

$$= P(1) + XP(1) - P(1)$$

$$\boxed{P(X) = XP(1)} \qquad (1)$$

Similarly, $P(Y) = a + (n-1)d$

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

$$P(Y) = YP(1) = 1$$

$$\boxed{P(1) = \frac{1}{Y}} \qquad (2)$$

Substituting value of $P(1)$ from equation (2) in (1), we get

$$\boxed{P(X) = \frac{X}{Y}}$$

Now, we need to find the probability of reaching low of \$0 and . Since there are only 2 states possible, one way to look is, has to reach either 0 or Y from X.

Hence, probability of reaching 0 + probability of reaching Y = 1

$\Rightarrow$ Probability of reaching $0 + \frac{X}{Y} = 1$

$\Rightarrow$ $\boxed{\text{Probability of reaching } 0 = 1 - \frac{X}{Y}}$

Also, mathematically,

Now, probability of reaching state 0 when person ~~has reached~~ starts from state Y, $P(Y) = 0$.

probability of reaching state 0 when person ~~has reached~~ starts from state 0, $P(0) = 1$

Let probability of reaching state 0 when person ~~has read~~ starts from state X be $P(X)$

We can define recursion as,

$$P(x) = 0.5\,P(X+1) + 0.5\,P(X-1)$$

At $X = Y-1$,

$$P(Y-1) = 0.5\underbrace{P(Y)}_{0} + 0.5\,P(Y-2)$$

$$\Rightarrow P(Y-1) = 0.5\,P(Y-2)$$

$$\Rightarrow \boxed{P(Y-2) = 2P(Y-1)}$$

At $X = Y-2$,

$$P(Y-2) = 0.5\,P(Y-1) + 0.5\,P(Y-3)$$

$$2P(Y-1) = 0.5\,P(Y-1) + 0.5\,P(Y-3)$$

$$\boxed{P(Y-3) = 3P(Y-1)}$$

At $X = Y-3$,

$$P(Y-3) = 0.5\,P(Y-2) + 0.5\,P(Y-4)$$

$$\boxed{P(Y-4) = 4P(Y-1)}$$

This forms an arithmetic progression, ~~with~~

Similarly $P(Y-Z) = Z\,P(Y-1)$

Consider $Z = \boxed{Y-X}$

$$P(Y-(Y-X)) = (Y-X)\,P(Y-1)$$

$$\boxed{P(X) = (Y-X)\,P(Y-1)} \quad \text{——— (3)}$$

Proof: $P(y-y) = y\,P(y-1)$

$P(0) = yP(y-1) = 1$

$\boxed{P(y-1) = 1/y}$ ——(4)

Substituting the value of $P(y-1)$ from equation (4) to equation get

$P(X) = (y-x)\,P(y-1)$

$= (y-x)\left(\frac{1}{y}\right) = \boxed{1 - \frac{x}{y}}$

(c) Expected value of stock at end $= 0 \times P(0) + y \times P(y)$

$= 0 \times \left(1 - \frac{x}{y}\right) + y \times \left(\frac{x}{y}\right)$

$\boxed{= x}$

(d) Output format

	6(a)	6(b)	6(c)
TEST CASE 1 >>	0.66980	0.33020	100.4700
TEST CASE 2 >>	0.99970	0.00030	99.94000
TEST CASE 3 >>	1.0000	0.0000	250.0000

→ the probabilities of Pr[cc].

7. 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 via the google form link https://forms.gle/uFF4U4Th7YYAhxRn6. For this question, your code submission should include a python file, a2_7.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) C → clear, S → snowy

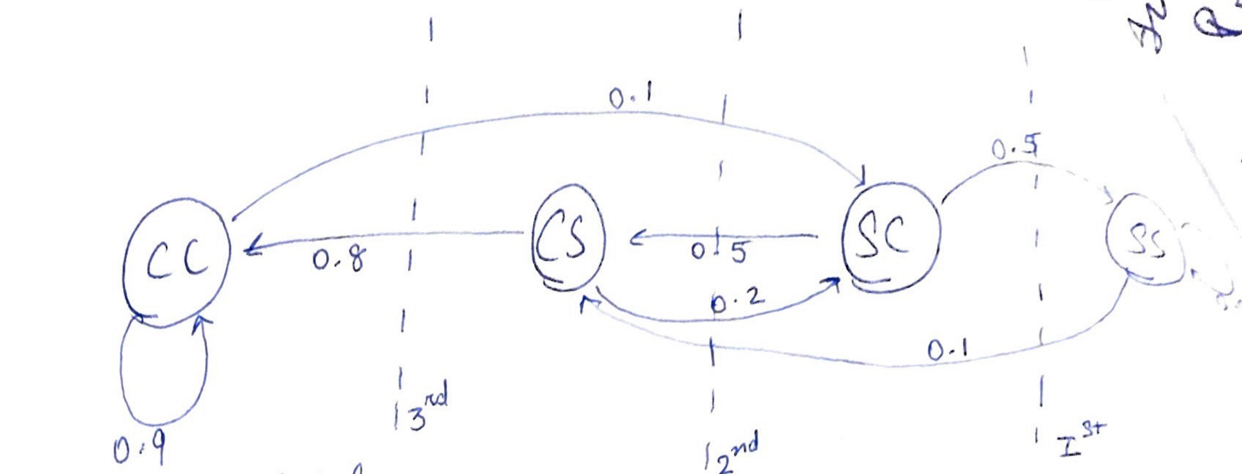
$Pr(X_{i+1} | X_i X_{i-1}) = Pr$ (weather tomorrow is X_{i+1}, given weather today is X_i and weather yesterday was X_{i-1})

$Pr(c|cc) = 0.9$, $Pr[c|cs] = 0.8$, $Pr[c|sc] = 0.5$, $Pr[c|ss] = 0.1$

$Pr(s|cc) = 1 - 0.9 = 0.1$, $Pr(s|cs) = 0.2$, $Pr(s|sc) = 0.5$, $Pr(s|ss) = 0.9$

$c|cc \xrightarrow{0.9} cc$ $\quad c|cs \xrightarrow{0.8} cc$ $\quad c|sc \xrightarrow{0.5} cs$ $\quad c|ss \xrightarrow{0.1} cs$

$s|cc \xrightarrow{0.1} sc$ $\quad s|cs \xrightarrow{0.2} sc$ $\quad s|sc \xrightarrow{0.5} ss$ $\quad s|ss \xrightarrow{0.9} ss$



Using Local Balance,

$$\text{rate}_{l\to r} = \text{rate}_{r\to l}$$

From 1st cut $\Rightarrow$ $\pi_{SC} \times 0.5 = \pi_{SS} \times 0.1$ $\Rightarrow$ $\boxed{\pi_{SS} = 5\pi_{SC}}$ —(1)

From 2nd Cut $\Rightarrow$ $\pi_{CS} \times 0.2 + \pi_{CC} \times 0.1 = \pi_{SC} \times 0.5 + \pi_{SS} \times 0.1$

$\Rightarrow$ $\pi_{CS} \times 0.2 + \pi_{CC} \times 0.1 = \pi_{SC} \times 0.5 + (5\pi_{SC}) \times 0.1$

$\Rightarrow$ $\pi_{CS} \times 0.2 + \pi_{CC} \times 0.1 = \pi_{SC}$

From 3rd Cut $\Rightarrow$ $\pi_{CC} \times 0.1 = \pi_{CS} \times 0.8$

$\Rightarrow$ $\boxed{\pi_{CC} = 8\pi_{CS}}$ —(2)

Putting value of π_{CC} from (2) in 2nd Cut equation,

$\Rightarrow$ $\pi_{CS} \times 0.2 + (8\pi_{CS}) \times 0.1 = \pi_{SC}$

$\Rightarrow$ $\boxed{\pi_{CS} = \pi_{SC}}$ —(3)

Now, we know that the probability of all the events sum upto 1,

$\Rightarrow$ $\pi_{CC} + \pi_{CS} + \pi_{SC} + \pi_{SS} = 1$

$\Rightarrow$ $8\pi_{CS} + \pi_{CS} + \pi_{SC} + 5\pi_{SC} = 1$ [from equ. 1, 2 & 3]

$\Rightarrow$ $8\pi_{SC} + \pi_{SC} + \pi_{SC} + 5\pi_{SC} = 1$

$\Rightarrow$ $\boxed{\pi_{SC} = 1/15}$

$\boxed{\pi_{CS} = 1/15}$ [from equ (3)]

$\boxed{\pi_{CC} = 8/15}$ [from equ (2)]

$\boxed{\pi_{SS} = 5/15}$ [from equ (1)]

So, the steady state matrix looks like

$$\begin{array}{cccc} CC & CS & SC & SS \\ \end{array}$$
$$\begin{bmatrix} 8/15 & 1/15 & 1/15 & 5/15 \\ 8/15 & 1/15 & 1/15 & 5/15 \\ 8/15 & 1/15 & 1/15 & 5/15 \\ 8/15 & 1/15 & 1/15 & 5/15 \end{bmatrix}$$

In steady state, we got the probabilities of $Pr[cc]$, $Pr[cs]$, $Pr[sc]$ and $Pr[ss]$.

Using these probabilities, we got

$$P_{sss} = P_{s|cc} \cdot \pi_{cc} + P_{s|cs} \cdot \pi_{cs} + P_{s|sc} \cdot \pi_{sc} + P_{s|ss} \cdot \pi_{ss}$$

$$= (0.1) \cdot \left(\frac{8}{15}\right) + (0.2)\left(\frac{1}{15}\right) + (0.5)\left(\frac{1}{15}\right) + (0.9) \cdot \left(\frac{5}{15}\right)$$

$$= \frac{0.8 + 0.2 + 0.5 + \cancel{0.45}\ 4.5}{15}$$

$$= \frac{6}{15} = \boxed{0.4}$$

So, probability of snow for 3 days in a row $= \boxed{0.4}$.

(c) Steady state: Power Iteration >> [0.53333 0.06667 0.06667 0.33333]

Solving part (b) numerically using the above result;

Probability it will be snowy 3 days from today $= 0.4$