Date 24/1/25

RANDOM VARIABLES

Random Variable

If 'S' is a sample space and 'X' is a real valued function defined on 'S' then 'X' is called a ~~sample spa~~ random variable.

$$X : S \longrightarrow R$$

Domain of a random variable is the sample space 'S'.

Range of a random variable is a set of all possible values taken by 'X' which can be either discrete or continuous.

S: sample space

R: random variable

* Types of random variables:

1) Discrete random variable - It is random variable which takes countable number of possible values.

eg, count of customers, no. of items sold, total of roll of two dice.

Q. A balanced coin is tossed four times. decide whether it is discrete or continuous random variables. list the elements of the sample space that are presumed to be equally likely and the corresponding values X of the random variable X which denotes the total no. of heads.

⇒ X = no. of heads in four tosses of a coin.

sample space = {HHHH, HHTH, HHHT, HHTT, HTHH, HTHT, HTTH, HTTT, THHH, THHT, THTH, THTT, TTHH, TTHT, TTTH, TTTT}

values of random variable X associated with each outcome are

X(HHHH) = 4   X(HTHH) = 3

X(HHTH) = 3   X(HTHT) = 2

X(HHHT) = 3   X(HTTH) = 2

X(HHTT) = 2   X(HTTT) = 1

random variable

discrete — continuous

Probability density funⁿ (Pdf)

...ility funⁿ (PMF)

$X(THHH) = 3$ $X(TTHH) = 2$

$X(THHT) = 2$ $X(TTHT) = 1$

$X(THTH) = 2$ $X(TTTH) = 1$

$X(THTT) = 1$ $X(TTTT) = 0$

Continuous random variable:
It is a random variable which takes uncountable no. of possible values.
eg. Temp. on a given day, % obtained in a particular exam, price of stock

Probability distribution of a discrete random Variable:
If X is a discrete random variable, the function given by $f(x) = P(X = x)$ for each x within the range of X is called the probability distribution / Probability mass funⁿ of X.

$f(x) = P(X = x_i) = p_i$

$f(x) \geq 0 \Rightarrow p_i \geq 0$ for all i

$\Sigma f(x) = 1 \Rightarrow \sum_{i=1}^{n} p_i = 1$

A function can serve as the probability distribution of a discrete random Variable X if and only if its value $f(x)$ satisfy the conditions:
① $f(x) \geq 0$ for each value within its domain
② $\Sigma f(x) = 1$, where the summation extends over all the values within its domain.

* Cumulative Distribution function: denoted by F

If "X" is a discrete random variable, the funⁿ given by

$$F(x) = P(X \le x) = \sum_{t \le x} f(t), \quad -\infty < x < \infty$$

where, $f(t)$ is the value of the probability distribution of X at t, is called the distribution funⁿ or the cumulative distribution of X.

The values of $F(x)$ of the distribution function of a discrete random variable X satisfy the condⁿ:

① $F(-\infty) = 0$ and $F(\infty) = 1$

② if $a < b$, then $F(a) \le F(b)$ for any real numbers a and b.

Q. Find the probability distribution of total number of heads obtained in four tosses of a balanced coin. Also find the cumulative distribution function.

→ X are. 0, 1, 2, 3, 4.

$P(X=0) = P(\text{zero heads}) = P(\{TTTT\}) = 1/16$

$P(X=1) = P(\text{one head}) = P(\{HTTT, THTT, TTHT, TTTH\})$
$= 4/16.$

Similarly, $P(X=2) = \frac{6}{16}$, $P(X=3) = \frac{4}{16}$, $P(X=4) = \frac{1}{16}$.

NOTE:

Step I, find sample space

Step II, Apply random variable on every element of sample space so we get out put and denoted by $X = x$.

Step III, find Probability for all $X = x$.

PMF

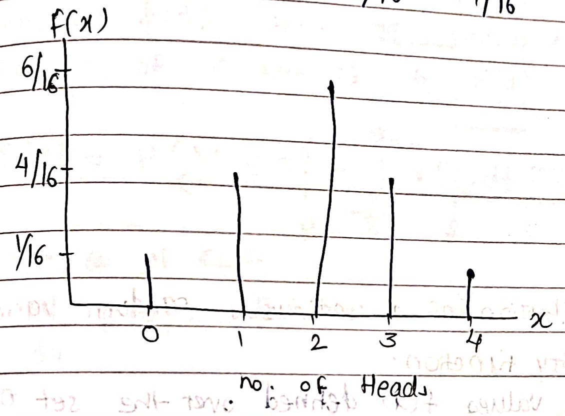
$P(x=2)$

$P(X=x\leq 2)$

$= P(X=0)+P(X=1)+P(X=2)$

Probability Mass function (PMF): denoted by f

X	0	1	2	3	4
$P(X)$	$1/16$	$4/16$	$6/16$	$4/16$	$1/16$



→ Cumulative distribution function is as follows:

Given $f(0)=\frac{1}{16}$, $f(1)=\frac{4}{16}$, $f(2)=\frac{6}{16}$, $f(3)=\frac{4}{16}$, $f(4)=\frac{1}{16}$

It follows that,

$F(0)=f(0)=\frac{1}{16}$

$F(1)=f(0)+f(1)=5/16$

$F(2)=f(0)+f(1)+f(2)=11/16$

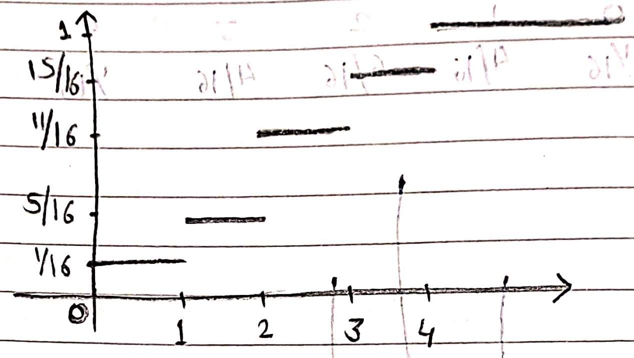
$F(3)=f(0)+f(1)+f(2)+f(3)=15/16$

$F(4)=f(0)+f(1)+f(2)+f(3)+f(4)=1$

⇓

$$F(x)=\begin{cases} 0 & \text{for } x<0 \\ 1/16 & \text{for } 0\leq x<1 \\ 5/16 & \text{for } 1\leq x<2 \\ 11/16 & \text{for } 2\leq x<3 \\ 15/16 & \text{for } 3\leq x<4 \\ 1 & \text{for } x\geq 4 \end{cases}$$

Graph of CDF:



* Probability distribution of a continuous random variable.

→ Probability density function:

A function with values $f(x)$, defined over the set of all real numbers is called a probability density function of the continuous random variable X if and only if

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx$$

for any real constant 'a' and 'b' with $a \leq b$

∴ for a continuous random variable

$$P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$$

→ Properties of Probability density function:

A function can serve as a probability density function of a continuous random variable X if it satisfies the following conditions:

① $f(x) \geq 0$, $-\infty < x < \infty$

② $\int_{-\infty}^{\infty} f(x)\,dx = 1$

* Cumulative distribution function of a continuous random variable:

If X is a continuous random variable with a density function $f(t)$, then the cumulative distribution function (CDF) of X denoted by $F(x)$ is defined as

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)\,dt \quad \text{for } -\infty < x < \infty$$

Properties of CDF:

① $\frac{d}{dx} F(x) = f(x)$ ② $P(a \le X \le b) = F(b) - F(a)$

③ $F(x)$ is a monotonically increasing function

i.e if $a < b$ then $F(a) < F(b)$

④ $\lim_{x \to \infty} F(x) = 1$ ⑤ $\lim_{x \to -\infty} F(x) = 0$

Q. Find PDF and CDF for given function.

If X has a probability density function

$$f(x) = \begin{cases} ke^{-3x} & , x > 0 \\ 0 & , \text{otherwise} \end{cases}$$

find 'k' and $P(0.5 \le x < 1)$

⇒ We know that, $\int_{-\infty}^{\infty} f(x)\,dx = 1$.

$$\Rightarrow \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{\infty} f(x)\,dx = 1$$

$$\Rightarrow \int_{-\infty}^{0} 0\,dx + \int_{0}^{\infty} ke^{-3x}\,dx = 1$$

$$\Rightarrow 0 + k\left[\frac{e^{-3x}}{-3}\right]_0^\infty = 1$$

$$\frac{k}{-3}\left(e^{-\infty} - e^{0}\right) = 1 \Rightarrow \frac{k}{-3}(-1) = 1$$

$$\Rightarrow \underline{\underline{k = 3}}$$

For $p(0.5 \le x < 1)$,

$$= \int_{0.5}^{1} 3e^{-3x}\,dx = 3\left[\frac{e^{-3x}}{-3}\right]_{0.5}^{1} = -1\left(e^{-3x}\right)_{0.5}^{1}$$

$$= -1\left(e^{-3} - e^{-3/2}\right) = 1\left(e^{-3/2} - e^{-3}\right)$$

$$= 1(0.024 - 0.049) = -0.075$$

* Expectation / mean of a random variable:

If X is a discrete random variable taking values x_i, $0 \le i \le 1$ and $f(x_i) = P_i = P(X = x_i)$ is the value of its probability mass function at each x_i then the expected value of X denoted by $E(x)$ or μ is
$$\mu = E(x) = \sum_{i=0}^{n} x_i P_i = \sum_{i=0}^{n} x_i f(x_i)$$

Correspondingly,

If X is a continuous random variable and $f(x)$ is the value of its probability density at x, the expected value,
$$\mu = E(x) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

Properties of expectation and variance:

If X and Y are any two random variables and 'a' and 'b' are constants then

1. $E(X+Y) = E(X) + E(Y)$
2. $E(a) = a$
3. $E(aX+b) = aE(X) + b$
4. $Var(a) = 0$
5. $Var(aX+b) = a^2 Var(X)$

IMP 6. $Var(X) = E[(X-\mu)^2] = E(X^2) - (E(X))^2$

7. $SD(aX+b) = |a|\, SD(X)$

* Variance and Standard deviation of a random variable

If X is a discrete random variable and $f(x)$ is the value of its probability distribution at 'x', the variance of X, denoted by σ^2 is
$$Var(x) = \sigma^2 = E[(X-\mu)^2] = \sum_{x} (x-\mu)^2 f(x)$$
or $P(x)$

Correspondigly,

If X is a continuous random variable and $f(x)$ is the value of its probability density at x, the variance of X is

$$Var(x) = \sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$$

Standard deviation of a random variable (discrete and continuous) is the positive square root of variance and is denoted by 'σ'.

$$SD = \sigma = +\sqrt{Var(x)}$$

Q. A balanced coin is tossed four times. List the elements of the sample space that are presumed to be equally likely and the corresponding values x of the random variable X which denotes the total number of heads. also find PMF, CDF, mean, variance, SD.

$\Rightarrow$ X = Total no. of heads. $2^4 = 16$

$S = \{$ HHHH, HHHT, HHTH, HHTT, HTHH,
HTHT, HTTH, HTTT, THHH, THHT,
THTH, THTT, TTHH, TTHT, TTTH,
TTTT $\}$

$P(X=0) = 1/16$ ~~i.e X~~
$P(X=1) = 4/16$
$P(X=2) = 6/16$
$P(X=3) = 4/16$
$P(X=4) = 1/16$ i.e $X(HHHH) = 4 \Rightarrow P(X=4) = 1/16$

X	0	1	2	3	4
PMF $\to f(x) = P(x)$	$1/16$	$4/16$	$6/16$	$4/16$	$1/16$
CDF $\to F(x)$	$1/16$	$1/16 + 4/16 = 5/16$	$\frac{1+4+6}{16} = \frac{11}{16}$	$\frac{1+4+6+4}{16} = 15/16$	$\frac{1+4+6+4+1}{16} = 1$

Mean, $\mu = E(x) = \sum_{i=0}^{n} x_i p_i$

$= 0 \times \frac{1}{16} + 1 \times \frac{4}{16} + 2 \times \frac{6}{16} + 3 \times \frac{4}{16} + 4 \times \frac{1}{16}$

$= \frac{4}{16} + \frac{12}{16} + \frac{12}{16} + \frac{4}{16} = \frac{32}{16} = 2$

mean = 2

$var(x) = \sigma^2 = \sum_{x}^{\square} (x-\mu)^2 f(x)$

$= (0-2)^2 \frac{1}{16} + (1-2)^2 \frac{4}{16} + (2-2)^2 \frac{6}{16} + (3-2)^2 \frac{4}{16}$

$+ (4-2)^2 \frac{1}{16}$

$= \frac{4}{16} + \frac{4}{16} + 0 + \frac{4}{16} + \frac{4}{16} = \underline{\underline{1}}$

$S.D = \sigma = \underline{\underline{1}}$

$var(X) = E(X^2) - (E(X))^2$

$= x_i^2 p_i - (x_i p_i)^2$

Q. find PMF, CDF, mean, variance, SD.

$P(x) = kx$, $x = 1, 2, 3, \ldots, 10$

$\Rightarrow$

X	~~0~~	1	2	3	4	5	6	7	8	9	10
P(x)											

$f(x) = P(x) = kx$, $x = 1, 2, 3, \ldots 10$

$P(x=1) = k$ $\qquad$ $P(x=6) = 6k$

$P(x=2) = 2k$ $\qquad$ $P(x=7) = 7k$

$P(x=3) = 3k$ $\qquad$ $P(x=8) = 8k$

$P(x=4) = 4k$ $\qquad$ $P(x=9) = 9k$

$P(x=5) = 5k$ $\qquad$ $P(x=10) = 10k$

X	1	2	3	4	5	6	7	8	9	10
f(x) = P(x)	k	2k	3k	4k	5k	6k	7k	8k	9k	10k.
F(x)	k	3k	6k	10k	15k	21k	28k	36k	45k	55k.

$55k = 1 \Rightarrow k = \frac{1}{55}$.

X	1	2	3	4	5	6	7	8	9	10
∴ P(x)	1/55	2/55	3/55	4/55	5/55	6/55	7/55	8/55	9/55	10/55
F(x)	1/55	3/55	6/55	10/55	15/55	21/55	28/55	36/55	45/55	1

Mean, $\mu = E(x) = \sum_{i=0}^{n} x_i P_i$

$= \left(1 \times \frac{1}{55}\right) + \left(2 \times \frac{2}{55}\right) + \left(3 \times \frac{3}{55}\right) + 4\left(\frac{4}{55}\right) + \left(5 \times \frac{5}{55}\right)$

$+ \left(6 \times \frac{6}{55}\right) + \left(7 \times \frac{7}{55}\right) + \left(8 \times \frac{8}{55}\right) + \left(9 \times \frac{9}{55}\right) + \left(10 \times \frac{10}{55}\right)$

$= 7$

$$\text{var}(x) = \sigma^2 = \sum_{x} (x-\mu)^2 f(x)$$

$$= \left[(1-7)^2 \frac{1}{55}\right] + \left[(2-7)^2 \frac{2}{55}\right] + \left[(3-7)^2 \frac{3}{55}\right]$$

$$+ \left[(4-7)^2 \frac{4}{55}\right] + \left[(5-7)^2 \frac{5}{55}\right] + \left[(6-7)^2 \frac{6}{55}\right]$$

$$+ \left[(7-7)^2 \frac{7}{55}\right] + \left[(8-7)^2 \frac{8}{55}\right] + \left[(9-7)^2 \frac{9}{55}\right]$$

$$+ \left[(10-7)^2 \frac{10}{55}\right]$$

$$= \frac{36}{55} + \frac{50}{55} + \frac{48}{55} + \frac{36}{55} + \frac{20}{55} + \frac{6}{55}$$

$$+ 0 + \frac{8}{55} + \frac{36}{55} + \frac{90}{55} = \frac{330}{55} = 6$$

$\text{var}(x) = 6$

$SD = 2.44$

Q. let X be a continuous random variable with probability density function.

$$f(x) = \begin{cases} kx & 0 \le x \le 2 \\ 2k & 2 \le x < 4 \\ -kx + 6k & 4 \le x < 6 \end{cases}$$

find 'k' and the mean.

$\Rightarrow$ $f(x)$ is probability density function.

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

$$\Rightarrow \int_0^2 kx\,dx + \int_2^4 2k\,dx + \int_4^6 (-kx+6k)\,dx = 1$$

$$\Rightarrow \left[\frac{k\cdot x^2}{2}\right]_0^2 + [2kx]_2^4 + \left[\frac{-kx^2}{2}\right]_4^6 + [6kx]_4^6 = 1$$

$$\Rightarrow \frac{k}{2}(4) + 2k(4-2) + \frac{-k}{2}(36-16) + 6k(6-4) = 1$$

$$\Rightarrow 2k + 4k - 10k + 12k = 1.$$

$$8k = 1 \quad \Rightarrow \quad \boxed{k = 1/8}$$

for mean, $E(x) = \int_{-\infty}^{\infty} x f(x)\,dx$ (for continuous random variable).

$$E(x) = \int_0^2 x(kx)\,dx + \int_2^4 2kx\,dx + \int_4^6 (-kx+6k)x\,dx$$

$$= \int_0^2 kx^2\,dx + \int_2^4 2kx\,dx - \int_4^6 kx^2\,dx + \int_4^6 6kx\,dx$$

$$= \frac{k}{3}[x^3]_0^2 + \frac{2k}{2}[x^2]_2^4 - \frac{k}{3}[x^3]_4^6 + \frac{6k}{2}[x^2]_4^6$$

$$= \frac{k}{3}(8) + k(16-4) - \frac{k}{3}(216-64) + 3k(36-16)$$

$$= \frac{8k}{3} + 12k - \frac{152k}{3} + 60k$$

$$= \frac{8}{3}\left(\frac{1}{8}\right) + 12\left(\frac{1}{8}\right) - \frac{152}{3}\left(\frac{1}{8}\right) + 60\left(\frac{1}{8}\right)$$

$$= 3.$$

$\therefore$ $k = \frac{1}{8}$ and Mean $= 3$.

Q. let X be a continuous random variable with the following distribution. $F(x) = \begin{cases} kx , & \text{if } 0 \le x \le 5 \\ 0 , & \text{elsewhere} \end{cases}$

ⓐ find k , ⓑ find $P(2 \le X \le 5)$ ⓒ find cdf
ⓓ Plot the graph of pmf and cdf .

→ ⓐ X is a continuous random variable then the total probability must be '1'.

$$\int_0^5 kx \, dx = 1 \Rightarrow \frac{k}{2}\left[x^2\right]_0^5 \Rightarrow \frac{25}{2} k = 1.$$

$$\boxed{k = \frac{2}{25}}$$

ⓑ $$\int_2^5 kx \, dx = \int_2^5 \frac{2}{25} x \, dx = \frac{2}{25}\left[\frac{x^2}{2}\right]_2^5$$

$$= \frac{1}{25}\left[x^2\right]_2^5 = \frac{1}{25}(25-4) = \frac{21}{25}$$

ⓒ $$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t) \, dt$$

$F(x) = 0$, $x < 0$

$F(x) = kx$, $0 \le x \le 5$

$$\int_{-\infty}^{0} f(t)\,dt + \int_{0}^{x} f(t)\,dt + \int_{x}$$

* Bi Variate / Joint Probability distribution:

If X and Y are discrete random variables taking values $x_1, x_2, \dots x_m$ and $y_1, y_2, \dots y_n$ respectively then the fun^n given by $f(x_i, y_j) = P_{ij} = P(X = x_i, Y = y_j)$

for each pair of values (x_i, y_j) within the range of X and Y is called the joint probability distribution of X and Y where

1) $f(x_i, y_j) \geq 0 \quad \forall \quad 0 \leq i \leq m, \quad 0 \leq j \leq n$

2) $\sum_{j=0}^{n} \sum_{i=0}^{m} f(x_i, y_j) = 1$

Joint Probability Table

X \ Y	y_1	y_2	...	y_n	Marginal prob. of x_i (sume of entire row)
x_1	$f(x_1, y_1)$	$f(x_1, y_2)$	...	$f(x_1, y_n)$	$f_1(x_1)$
x_2	$f(x_2, y_1)$	$f(x_2, y_2)$		$f(x_2, y_n)$	$f_1(x_2)$
:	:	:			:
x_m	$f(x_m, y_1)$	$f(x_m, y_2)$		$f(x_m, y_n)$	$f_1(x_m)$
	$f_2(y_1)$	$f_2(y_2)$	...	$f_2(y_n)$	1

Marginal prob. of y_i (sume of entire column)

always 1

* Marginal Probabilities:

The probability that $X = x_i$, $0 \leq i \leq m$ is obtained by adding all entries in the row corresponding to x_i and is given by

$$f_1(x_i) = P(X = x_i) = \sum_{j=0}^{n} f(x_i, y_j)$$

This is called as marginal probability of $X = x_i$, $0 \leq i \leq m$

similarly, the probability that $Y = y_j$, $0 \leq j \leq n$ is obtained by adding all entries in the column corresponding to y_j and is given by

$$f_2(y_j) = P(Y = y_j) = \sum_{i=0}^{m} f(x_i, y_j)$$

This is called the marginal probability of $Y = y_j$, $0 \le j \le$

The joint distribution function of X and Y is defined

$$F(x_i, y_j) = P(X \le x_i, Y \le y_j) = \sum_{u \le x_i} \sum_{v \le y_j} f(u, v)$$

* Independent random Variables:

Two discrete random variables X and Y are said to be independent if

$$P(X=x, Y=y) = P(X=x) \cdot P(Y=y) \quad \text{, for all}$$

Equivalently,

If X and Y are continuous random variables then they are said to be indepent if

$$P(X \le x, Y \le y) = P(X \le x) \cdot P(Y \le y) \quad \text{, for all } x$$

Q.1) The joint probability function of two discrete random variables X and Y is given by $f(x,y) = c(2x+y)$ where x and y can assume all integers such that $0 \le x \le 2$, $0 \le y \le 3$ and $f(x,y) = 0$ otherwise.

(a) find the value of the constant c.

(b) find $P(X=2, Y=1)$

(c) find $P(X \ge 1, Y \le 2)$

(d) find the marginal probabilities of X and Y.

(e) Check if the variables X and Y are independent.

$$f(x,y) = c(2x+y), \quad 0 \le x \le 2, \; 0 \le y \le 3$$
$$f(x,y) = 0 \quad \text{otherwise.}$$

ⓐ

$x \backslash y$	0 (y_1)	1 (y_2)	2 (y_3)	3 (y_4)	
$x_1 \to 0$	0	0	2c	3c	6c
x_2 1	2c	3c	4c	5c	14c
x_3 2	4c	5c	6c	7c	22c
	6c	9c	12c	15c	42c

$\to$ CMF for x

CMF for y. ↳ always equal to 1.

$42c = 1. \Rightarrow \boxed{c = 1/42}$

ⓑ

$x \backslash y$	0	1	2	3	
0	0	1/42	2/42	3/42	6/42
1	2/42	3/42	4/42	5/42	
2	4/42	5/42	6/42	7/42	
	6/42				

$\to$ Marginal prob. of (X) ~~or CMF for 'x'~~ at $x=0$

similarly for all $X = x$.

↳ Marginal prob. of y' or CMF for 'y'. for $Y=0$ similarly for all $Y=y$

$$P(X=2, Y=1) = 5c = \frac{5}{42}$$

ⓒ $P(X \ge 1, Y \le 2) = \sum_{u \ge 1} \sum_{v \le 2} f(u,v)$

$= P(X=1, X=2, Y=0, Y=1, Y=2)$

$= (2c + 3c + 4c) + (4c + 5c + 6c)$

$= 24c = \frac{24}{42} = \frac{4}{7}$

ⓓ Marginal probability function for 'x' is given by $P(X=x) = f_1(x)$ and can be obtained from the margin totals in the right hand column.

$$f_1(x) = \begin{cases} 6C = 6/42 = 1/7 & , \text{at } x = 0 \\ 14C = 14 \times \frac{1}{42} = 1/3 & , \text{at } x = 1 \\ 22C = 22 \times 1/42 = 11/21 & , \text{at } x = 2 \end{cases}$$

$$f_2(y) = \begin{cases} 6C = 6/42 = 1/7 & , \text{at } y = 0 \\ 8C = 8/42 = 4/21 & , \text{at } y = 1 \\ 12C = 12/42 = 6/21 & , \text{at } y = 2 \\ 15C = 15/42 = 5/14 & , \text{at } y = 3 \end{cases}$$

Q. let 'X' and 'Y' be independent random variables with the following distributions: find the joint distribution of X and Y.

X	1	2
$p(X=x_i)$	0.6	0.4

Y	5	10	15
$p(Y=y_j)$	0.2	0.5	0.3

⇒

X \ Y	5	10	15	
1	0.12	0.30	0.18	0.6
2	0.08	0.20	0.12	0.4
	0.2	0.5	0.3	

Covariance

Covariance of two random variables X and Y denoted by $Cov(X,Y)$ or σ_{xy} is defined as

$$\sigma_{xy} = Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$$

where, μ_x and μ_y are means of X and Y respectively.

NOTE: The simplified expression for the above formula is

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

↳ joint distribution mean.

NOTE: ① The covariance of two random variables describes their behaviour, i.e. it measures to what extent the random variables vary together.

② If $Cov(X,Y) = 0$, then X and Y are uncorrelated.

③ Positive value of $Cov(X,Y)$ indicates that Y tends to increase as X does, whereas a negative value indicates that Y tends to decrease as X increases

$E[XY] \rightarrow$ Mean of joint distribution.

* Properties of Covariance

① $Cov(X,Y) = Cov(Y,X)$

② $Cov(X,X) = Var(X)$

③ $Cov(aX,Y) = a\,Cov(X,Y)$

④ $Cov(X+Z,Y) = Cov(X,Y) + Cov(Z,Y)$

⑤ $Cov(aX,bY) = ab\,Cov(X,Y)$

⑥ $Cov(X+a, Y+b) = Cov(X,Y)$

Q. The joint and marginal probabilities of X and Y, the numbers of aspirin and sedative caplets among the two caplets drawn at random from a bottle containing three aspirin, two sedative and four laxative caplets are recorded as follows:

		$x=0$	$x=1$	$x=2$	
y	0	1/6	1/3	1/12	7/12
	1	2/9	1/6		7/18
	2	1/36			1/36
		5/12	1/2	1/12	

find mean, standard deviatation and covariance of x & y

$\Rightarrow$ Covariance $(XY) = E(XY) - E(X)E(Y)$

$$E(X) = \left(0 \times \frac{5}{12}\right) + \left(1 \times \frac{1}{2}\right) + \left(2 \times \frac{1}{12}\right)$$

$$= \frac{1}{2} + \frac{1}{6} = \frac{4}{6} = \frac{2}{3}.$$

$$E(Y) = \left(0 \times \frac{7}{12}\right) + \left(1 \times \frac{7}{18}\right) + \left(2 \times \frac{1}{36}\right)$$

$$= \frac{7}{18} + \frac{1}{18} = \frac{8}{18} = \frac{4}{9}$$

$$E(XY) = \left(0 \times 0 \times \frac{1}{6}\right) + \left(0 \times 1 \times \frac{1}{3}\right) + \left(0 \times 2 \times \frac{1}{12}\right) +$$

$$\left(1 \times 0 \times \frac{2}{9}\right) + \left(1 \times 1 \times \frac{1}{6}\right) + \left(2 \times 0 \times \frac{1}{36}\right)$$

$$= 0 + 0 + 0 + 0 + \frac{1}{6} + 0 = \frac{1}{6}$$

$$\text{covariance}(XY) = \frac{1}{6} - \left(\frac{2}{3}\right)\left(\frac{4}{9}\right) = \frac{1}{6} - \frac{8}{27} = \frac{-7}{54}$$

The -ve result suggests that the more aspirin tablets we get fewer sedative tablets we will get and vice versa and this of course make sense.

$\sigma_{xy} = cov(x,y) = E(XY) - E(X)E(Y)$

$\sigma_x = cov(x) = E(x^2) - [E(x)]^2$

$\sigma_y = Cov(Y) = E(y^2) - [E(Y)]^2$

* Correlation Coefficient

→ The strength of the linear relationship between X and Y is indicated by the correlation between X and Y.

→ It provides a measure of dependence between two random variables.

→ It is a dimensionless quantity obtained by dividing the covarience by the product of the S.D of X and Y.

$$\therefore \rho = Corr(X,Y) = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

Where,

σ_{xy} is the covariance betn X and Y.

σ_x is the SD of X.

σ_y is the SD of Y.

* Properties of Correlation coefficient.

1) $-1 \leq \rho \leq 1$

2) $\rho = -1$ indicates perfect negative linear relationship betn X and Y.

3) $\rho = 1$ indicates perfect +ve linear relationship betn X and Y.

4) $\rho = 0$ indicates that the variables X and Y are uncorrelated

5) If X and Y are independent then they are uncorrelated i.e $\rho = 0$ however the converse is not true.

Q.

y \ x	0	1	2	
0	1/6	1/3	1/12	7/12
1	2/9	1/6		7/18
2	1/36			1/36
	5/12	1/2	1/12	

find the covarience of X and Y.

$\Rightarrow$ $E(X) = \left(0 \times \frac{5}{12}\right) + \left(1 \times \frac{1}{2}\right) + \left(2 \times \frac{1}{12}\right) = \frac{1}{2} + \frac{1}{6} = \frac{4}{6} = \frac{2}{3}.$

$E(Y) = \left(0 \times \frac{7}{12}\right) + \left(1 \times \frac{7}{18}\right) + \left(2 \times \frac{1}{36}\right) = \frac{7}{18} + \frac{1}{18} = \frac{8}{18} = \frac{4}{9}$

$E(XY) = \left(0 \times 0 \times \frac{1}{6}\right) + \left(0 \times 1 \times \frac{1}{3}\right) + \left(2 \times 0 \times \frac{1}{12}\right) + \left(1 \times 0 \times \frac{2}{9}\right)$

$+ \left(1 \times 1 \times \frac{1}{6}\right) + \left(2 \times 0 \times \frac{1}{36}\right) = 0 + 0 + 0 + 0 + \frac{1}{6} + \cancel{\frac{1}{18}}$

$\cancel{= \frac{4}{18}} = \cancel{\frac{2}{9}} \quad = \frac{1}{6}$

$cov(X,Y) = \sigma_{xy} = E(XY) - E(X)E(Y)$

$= \frac{1}{6} - \left(\frac{2}{3}\right)\left(\frac{4}{9}\right) = \frac{1}{6} - \frac{8}{27} = \frac{-7}{54}.$

$\sigma_X = cov(X) = E(X^2) - [E(X)]^2$

$= \left[\left(0^2 \times \frac{5}{12}\right) + \left(1^2 \times \frac{1}{2}\right) + \left(2^2 \times \frac{1}{12}\right)\right] -$

$\left[\left(0 \times \frac{5}{12}\right)^2 + \left(1 \times \frac{1}{2}\right)^2 + \left(2 \times \frac{1}{12}\right)^2\right]$

$= \left[\frac{1}{2} + \frac{1}{3}\right] - \left[\cancel{\frac{25}{144}} + \frac{1}{4} + \left(\frac{1}{36}\right)\right]$

$= \frac{5}{6} - \left[\cancel{\frac{25}{144}} + \frac{1}{4} + \frac{1}{36}\right] = \frac{120 - 25 - 36 - 4}{144}$

$= \frac{55}{144} = \underline{\underline{0.55}}$

$\sigma_y = \sqrt{0.55} = 0.74$

$$\sigma_y = cov(y) = E(y^2) - [E(y)]^2$$

$$= \left[\left(0^2 \times \frac{7}{12}\right) + \left(1^2 \times \frac{7}{18}\right) + \left(2^2 \times \frac{1}{36}\right)\right] - \left[\left(0 \times \frac{7}{12}\right)^2 + \left(1 \times \frac{7}{18}\right)^2 + \left(2 \times \frac{1}{36}\right)^2\right]$$

$$= \left(\frac{7}{18} + \frac{1}{9}\right) - \left[\cancel{\frac{49}{144}} + \frac{49}{324} + \frac{1}{324}\right]$$

$$= \frac{1}{2} - \left(\frac{49}{324} + \frac{1}{324}\right)^2 = \frac{1}{2} - \frac{50}{324} = 0.34 \Rightarrow \sigma_y = \sqrt{0.34} = 0.58$$

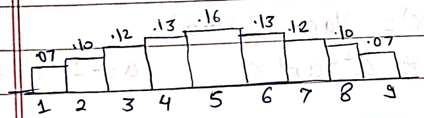
$\cancel{Cov(X,Y) = E[XY] - E[X]E[Y]}$

$$\rho = \frac{Cov(X,Y)}{\sigma_x \sigma_y} = \frac{-0.129}{0.74 \times 0.58} = \frac{-0.129}{0.4292} = -0.3 \quad \rho = -ve$$

means $Y \downarrow$ as $x \uparrow$

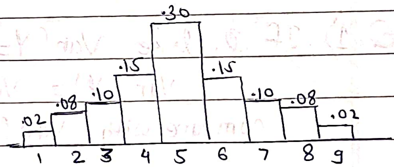
* **Interpretation of Variance:**

Variance represents the spread or dispersion of the distribution of a random variable.



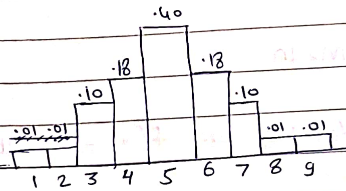
$\mu = 5$, $\sigma^2 = 5.26$

(i)



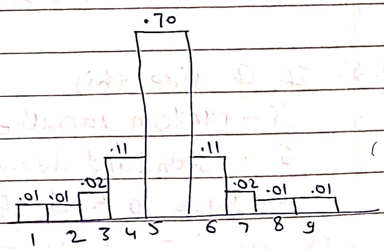
$\mu = 5$ & $\sigma^2 = 3.18$

(ii)



$\mu = 5$, $\sigma^2 = 1.68$

(iii)



$\mu = 5$ & $\sigma^2 = 0.88$

(iv)

i) This distribution is the most spread out.
Values are evenly distributed across the range.
The probability of different values is more balanced.

ii) The peak is slightly higher, more values are concentrated near the mean.

iii) The spread is much smaller than the previous two graphs, more values are clustered near the center, reducing dispersion.

iv) This distribution has the smallest variance, it is most tightly packed around the mean.
The peak is highest, most values are very close to the mean.

* Higher variance → More spread out values, flatter distribution.
Lower variance → More concentrated values, taller peak.
Same mean ($\mu = 5$) → center remains unchanged, only spread changes.

Q.1) If Q. like $Var(Y = 6-2X)$ then use

$$Var(Y) = Var(6-2X) = Var(-2X+6)$$

compare with $Var(aX+b) \Rightarrow a^2 Var(X)$. $a=-2,\ b=6$.

for calculate, $P(X \geq Y)$ & $Y = 6-2X$.

$$P(X \geq 6-2X) = P(3X \geq 6) = P(X \geq 2)$$

Q.2) If Q like this.
$X \rightarrow$ random variable, $\mu \rightarrow$ mean.
$\sigma \rightarrow$ standard deviation.
We have to find $E(X)$ & $Var(X)$ of $f(x) = (x-\mu)/\sigma$.
so, let $Y = (x-\mu)/\sigma$

$$E(Y) = E\left(\frac{x}{\sigma} - \frac{\mu}{\sigma}\right) \text{ as } E(ax+b)$$

$$= \frac{1}{\sigma}E(X) - \frac{\mu}{\sigma} = \frac{1}{\sigma}\mu - \frac{\mu}{\sigma} = 0$$

for $Var(X)$ same. $Var(Y) = var\left(\frac{x}{\sigma} - \frac{\mu}{\sigma}\right) = \frac{1}{\sigma^2} Var(x) = \frac{1}{\sigma^2}\sigma^2 = 1$

Q.3) ⇒ If Q like this, given PDF $f(x) = ke^{-|x|}, -\infty < x < \infty$

then this function is symmetric about $x=0$

then split it from $-\infty$ to 0 & 0 to ∞.

$$f(x) = \begin{cases} ke^{-(-x)}, & -\infty < x < 0 \Rightarrow ke^{x}, +\infty < x < 0 \\ ke^{-(x)}, & 0 < x < \infty \Rightarrow ke^{-x}, 0 < x < \infty \end{cases}$$

and solve.

for finding 'k'. we know that in continuous random variable PDF is sum of all funⁿ integration equal to 1.

$$\int_{-\infty}^{0} ke^{x}\,dx + \int_{0}^{\infty} ke^{-x}\,dx = 1$$

for finding mean, $\mu = E(x) = \int_a^b x\, f(x)\, dx$.

for finding variance.

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

$$= \int_a^b x^2 f(x)\,dx - \left[\int_a^b x\, f(x)\,dx\right]^2.$$