

Probability Space :- A probability space models random events and is made up of three parts: Sample Space, Event space & probability function.

Sample Space ( $\Omega$ ): the set of all possible outcomes.

Event Space ( $A$ ): the set of all events.

probability function ( $P$ ): the assignment of probability to the event.

Q1. Find the prob. Space of even no. in a throw of a dice.

$$\text{Sol. } \Omega = \{1, 2, 3, 4, 5, 6\}$$

$$A = \{2, 4, 6\}$$

$$\therefore P(A) = \frac{3}{6} = \frac{1}{2}$$

~~∴~~ The required prob. Space is  $(\Omega, A, P)$

$$\text{where, } \Omega = \{1, 2, 3, 4, 5, 6\}$$

$$A = \{2, 4, 6\}$$

$$P = \frac{1}{2}$$

Q2. If we toss 3 coins simultaneously. Find the probability space of atleast 2 Heads.

# Mutually exclusive Events :- ~~nothing common~~ events

# Mutually exhaustive Events :- ~~something common~~ in events

e.g - ①

$E_1 = \text{Odd no.}$

$E_2 = \text{Even no.}$

$E_1 \cap E_2 = \emptyset$  ] mutually exclusive

$E_1 \cup E_2 = S$

e.g - ②

$E_1 = \{2, 0, 6\}$

$E_2 = \{5, 3\}$

$E_3 = \{1, 2\}$

$E_1 \cap E_2 \cap E_3 = \emptyset$

# Equally likely events :- whose probability is same.

e.g -  $E_1 = \text{Odd no.}$  ] same

$E_2 = \text{Even no.}$  ] probability

# Conditional probability :-

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

already happened.

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Independent Event :-

$$\boxed{P(A \cap B) = P(A) \cdot P(B)}$$

e.g. :-  $P(A) = \frac{1}{3}$  }  $\rightarrow$  problem solved by A.

$P(B) = \frac{2}{3}$  }  $\rightarrow$  problem solved by B.

$$P(\text{problem solved}) = P(A)P(B) + P(\bar{A})P(B)$$
$$+ P(A)P(\bar{B})$$

Multiplicative law

$$P(A \cap B) = P(A|B)P(B)$$

Baye's Theorem :-

If  $E_1, E_2, \dots, E_n$  are mutually exclusive & exhaustive events with  $P(E_i) \neq 0$  ( $i=1, 2, \dots, n$ ) of a random experiment then for any arbitrary event A of the sample space of the above experiment with  $P(A) \geq 0$ , we have

$$P(\bar{E}_i | A) = \frac{P(E_i) P(A|E_i)}{\sum_{i=1}^n P(E_i) P(A|E_i)}$$

proof :- Let S be the sample space of the random experiment.

The events,  $E_1, E_2, \dots, E_n$ , being exhaustive

$$\begin{aligned} S &= E_1 \cup E_2 \cup \dots \cup E_n \\ A &= A \cap S \\ A &= A \cap (E_1 \cup E_2 \cup \dots \cup E_n) \\ &= (A \cap E_1) \cup (A \cap E_2) \cup \dots \cup (A \cap E_n) \quad (\text{Distributive Law}) \end{aligned}$$

$$P(A) = P(E_1) P(A|E_1) + P$$

$$\begin{aligned} P(A) &= P(A \cap E_1) + P(A \cap E_2) + \dots + P(A \cap E_n) \\ &\stackrel{n}{\approx} P(E_1) P(A|E_1) + P(E_2) P(A|E_2) + \dots + P(E_n) P(A|E_n) \\ &\Rightarrow \sum_{i=1}^n P(E_i) P(A|E_i) \quad \boxed{\frac{P(A|E_i)}{P(E_i)}} \end{aligned}$$

$$P(E_i|A) = \frac{P(A \cap E_i)}{P(A)}$$

$$= \frac{P(E_i) P(A|E_i)}{\sum_{i=1}^n P(E_i) P(A|E_i)}$$

$$\sum_{i=1}^n P(E_i) P(A|E_i)$$

- Q1. In a bolt factory, machines A, B, & C manufacture resp. 25%, 35%, & 40% of the total. On their output 5%, 4%, & 2% are defective bolts. A bolt is drawn at random from the product & is found to be defective. What is the probability that it was manufactured by machine B?

Sol-  $P(E_1) = \frac{25}{100}$ ,  $P(E_2) = \frac{35}{100}$ ,  $P(E_3) = \frac{40}{100}$

$P(A|E_1) = \frac{5}{100}$ ,  $P(A|E_2) = \frac{4}{100}$ ,

$P(A|E_3) = \frac{2}{100}$ .

$$P(E_2|A) = \frac{P(E_2) P(A|E_2)}{P(E_1) P(A|E_1) + P(E_2) P(A|E_2) + P(E_3) P(A|E_3)}$$

$$= \frac{\frac{35}{100} \times \frac{4}{100}}{\frac{25}{100} \times \frac{5}{100} + \frac{35}{100} \times \frac{4}{100} + \frac{40}{100} \times \frac{2}{100}}$$

Q1. A university bought 45%, 25%, 30% of computers from HCL, WIPRO and IBM resp, and 2%, 3% & 1% of these were found to be defective. Find the prob. of a comp. Selected at random is found to be defective. What is the prob. that the defective comp. is from HCL?

Sol-  $P(E_1) = \frac{45}{100}$ ,  $P(E_2) = \frac{25}{100}$ ,  $P(E_3) = \frac{30}{100}$

$\underbrace{\qquad\qquad\qquad}_{HCL}$   $\underbrace{\qquad\qquad\qquad}_{WIPRO}$   $\underbrace{\qquad\qquad\qquad}_{IBM}$

$$P(A|E_1) = \frac{2}{100}, \quad P(A|E_2) = \frac{3}{100}, \quad P(A|E_3) = \frac{1}{100}$$

$$P(A) = P(E_1) P(A|E_1) + P(E_2) P(A|E_2) + P(E_3) P(A|E_3)$$

$$\Rightarrow \frac{45}{100} \times \frac{2}{100} + \frac{25}{100} \times \frac{3}{100} + \frac{30}{100} \times \frac{1}{100}$$

$$= \frac{90 + 75 + 30}{10000} = \frac{195}{10000}$$

$$\begin{aligned}
 P(E_1 | A) &= \frac{P(A|E_1) P(E_1)}{P(A)} \\
 &= \frac{\frac{45}{100} \times \frac{2}{100}}{\frac{195}{1000}} = \frac{\frac{45 \times 2}{195}}{\frac{13}{100}} = \frac{6}{13} \\
 &= 0.46.
 \end{aligned}$$

Q2. Each of the identical boxes  $B_1$ ,  $B_2$  &  $B_3$  contain 2 coins  $\Rightarrow B_1$  contain Both gold coin  
 $\Rightarrow B_2$  contain Both silver coin  
 $\Rightarrow B_3 \rightarrow 1$  gold, 1 silver.

If a box is chosen at random & a coin is picked at random & the coin is found to be gold, what is the prob. the other coin in the box is also gold.

Sol.  $P(B_1) = P(B_2) = P(B_3) = \frac{1}{3}$   
 $P(A) = \text{prob. of finding gold}$

$$P(A|E_1) = 1, \quad P(A|E_2) = 0, \quad P(A|E_3) = \frac{1}{2}.$$

$$P(E_1 | A) = \frac{\frac{1}{3} \times 1}{\frac{1}{3} \times 1 + \frac{1}{3} \times 0 + \frac{1}{3} \times \frac{1}{2}} = \frac{2}{3}.$$

Ans:

Q3. A man is known to speak truth 3 out of 4 times. He throws a die and reports that it is a 6. Find the prob. that it

~~Sol.~~ is actually six.

~~Sol.~~  $P(S) = \frac{1}{6}, \quad P(S') = \frac{5}{6}$

$$P(E|S) = \frac{3}{4}, \quad P(E|S') = \frac{1}{4}.$$

$$P(S|E) = \frac{P(S) P(E|S)}{P(S) P(E|S) + P(S') P(E|S')}$$

$$= \frac{\frac{1}{6} \times \frac{3}{4}}{\frac{1}{6} \times \frac{3}{4} + \frac{5}{6} \times \frac{1}{4}} = \frac{\frac{3}{24}}{\frac{3+5}{24}} = \frac{3}{8}$$

Ans -

Q4. The prob. of X, Y, & Z becoming managers of a company are  $\frac{4}{9}, \frac{2}{9}, \frac{1}{3}$  resp. The prob. that the bonus scheme will be introduced if X, Y, & Z becomes manager are  $\frac{3}{10}, \frac{1}{2}, \frac{4}{5}$  resp.

(a) find the prob. that bonus scheme will be introduced.

(b) If the bonus scheme has been introduced, find the prob. that manager appointed was X or Y.

$$\text{Soln: } P(X) = \frac{4}{9}, P(Y) = \frac{2}{9}, P(Z) = \frac{3}{9}$$

$$P(B|X) = \frac{3}{10}, P(B|Y) = \frac{1}{2} = \frac{5}{10}, P(B|Z) = \frac{4}{5} = \frac{8}{10}$$

$$(a) P(B) = \frac{4 \times \frac{3}{90}}{90} + \frac{10}{90} + \frac{24}{90} = \frac{46}{90} = \frac{23}{45} = 0.52$$

$$(b) P(X \cup Y|B) = P(X|B) + P(Y|B)$$

$$= \frac{P(X) P(B|X)}{P(B)} + \frac{P(Y) P(B|Y)}{P(B)}$$

$$= \frac{\frac{4}{9} \times \frac{3}{10}}{\frac{46}{90}} + \frac{\frac{2}{9} \times \frac{5}{10}}{\frac{46}{90}}$$

$$= \frac{12+10}{46} = \frac{22}{46} = \frac{11}{23}$$

Q5. An insurance company insured 2000 scooter drivers, 4000 car drivers & 6000

truck drivers. The prob. of any accident involving a scooter, car & truck is 0.01, 0.03, 0.15 resp. One of the insured persons meets with an accident, what is the prob. that he is a scooter driver?

$$\text{Soln? } P(S) = \frac{2000}{12000}, P(C) = \frac{4000}{12000}, P(T) = \frac{6000}{12000}$$

$$P(A|S) = \frac{1}{100}, P(A|C) = \frac{3}{100}, P(A|T) = \frac{15}{100}$$

$$P(S|A) = \frac{P(A|S) P(S)}{P(A|S) P(S) + P(A|C) P(C) + P(A|T) P(T)}$$

$$\rightarrow \frac{\frac{2000}{12000} \times \frac{1}{100}}{\frac{1}{100} \times \frac{2}{12} + \frac{3}{100} \times \frac{4}{12} + \frac{15}{100} \times \frac{6}{12}}$$

$$= \frac{\frac{1}{12} \times \frac{1}{12}}{\frac{1}{12} + \frac{1}{12} + \frac{1}{12}} = \frac{1}{3} = \frac{1}{4} + \frac{9}{4} = \frac{1}{4} + \frac{9}{4} = \frac{10}{4} = \frac{5}{2}$$

$$\therefore P(S|A) = \frac{5}{2} = 2.5$$

$$P(S|A) = \frac{1}{4} + \frac{9}{4} = \frac{10}{4} = \frac{5}{2} = 2.5$$

$$P(S|A) = \frac{1}{4} + \frac{9}{4} = \frac{10}{4} = \frac{5}{2} = 2.5$$

$$(2.5)^2 + (2.5)^2 = (2.5)^2 + (2.5)^2 = (2.5)^2$$

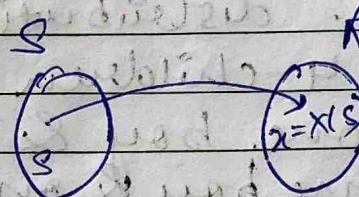
$$2.5^2 + 2.5^2 = 2.5^2 + 2.5^2 = 2.5^2$$

Prob. of an accident involving a scooter, car & truck is 0.01, 0.03, 0.15 resp.

Random variable :- A random variable  $X$  on a sample space  $S$  is a function

$$X: S \rightarrow \mathbb{R}$$

$X(s) = x$  to each sample point.



Random variable

Discrete  
fixed

Continuous

X	0	1	2
$P(X)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

→ Discrete prob. Distribution

Properties

$$1) f(x) \geq 0$$

Binomial

$$2) \sum f(x) = 1$$

Poisson Distn.

$$3) P(X=x) = f(x)$$

where  $f(x)$  is known as probability Mass function (pmf)

→ Continuous prob. Distribution  
Properties

$$E(x) = \int x f(x)$$

$$1) f(x) \geq 0$$

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

Exponential  
Normal  
Uniform distn.

$$3) P(a < x < b) = \int_a^b f(x) dx$$

where,  $f(x)$  is known as prob. density func. (pdf)

- Q1. Find the prob. distribution & prob. that in a family of 4 children there will be (a) at least one boy & (b) at least one boy & one girl, assuming that the prob. of male birth is  $\frac{1}{2}$ .

Sol:-

$X$	0	1	2	3	4
$P(X)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

$$(a) P(\text{at least 1B}) = 1 - P(\text{No Boy}) = 1 - \frac{1}{16} = \frac{15}{16}$$

$$(b) P(\text{at least 1B & 1G}) = P(1) + P(2) + P(3) \\ \Rightarrow \frac{14}{16}$$

- Q2. Out of 24 eggs, 6 are rotten. If 2 eggs are drawn at random, find the prob. distribution of the no. of rotten eggs that can be drawn.

Sol:-

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

$$P(X=0) = \frac{18C_2}{24C_2} \Rightarrow \frac{\frac{18 \times 17 \times 16!}{16!}}{\frac{24 \times 23 \times 22!}{22!}} = \frac{18 \times 17}{24 \times 23}$$

$$P(X=1) = \frac{18C_1 \times 6C_1}{24C_2} \Rightarrow \frac{\frac{18 \times 17!}{17!} \times \frac{6 \times 5!}{5!}}{24 \times 23 \times 22!}$$

$$P(X=2) = \frac{6C_2}{24C_2} \Rightarrow \frac{\frac{6 \times 5 \times 4!}{4!}}{24 \times 23 \times 22!} = \frac{6 \times 5}{92} = \frac{5}{92}$$

Answers of 23. New Ex. Selection methods & their probabilities

$$\begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|} \hline & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ \hline \text{Ans.} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & (x) \\ \hline \end{array}$$

(23x2x2), (2x2x2), (2x2x2) standing (i)

23x2x2 = 10 ways combining with hand (ii)

$$L < (x > x) \dots$$

$$1 = (n+3)^{10}$$

$$1 = 1 + 3 + 3^2 + 3^3 + 3^4 + 3^5 + 3^6 + 3^7 + 3^8 + 3^9 + 3^{10}$$

$$1 = 1 + 3 + 3^2 + 3^3 + 3^4 + 3^5 + 3^6 + 3^7 + 3^8 + 3^9 + 3^{10}$$

$$1 = 1 + 3 + 3^2 + 3^3 + 3^4 + 3^5 + 3^6 + 3^7 + 3^8 + 3^9 + 3^{10}$$

$$1 = 1 + 3 + 3^2 + 3^3 + 3^4 + 3^5 + 3^6 + 3^7 + 3^8 + 3^9 + 3^{10}$$

$$1 = 1 + 3 + 3^2 + 3^3 + 3^4 + 3^5 + 3^6 + 3^7 + 3^8 + 3^9 + 3^{10}$$

# # Probability Distribution

pmf  
condition  
i)  $p_i \geq 0$

- Discrete probability Distribution
  - i)  $p(x_i) = f(x_i) \geq 0$
  - ii)  $\sum p(x_i) = 1$

- Continuous probability Distribution.

$$F(x) = p(x \leq x) = \int_{-\infty}^x f(x) dx.$$

- i)  $0 \leq f(x) \leq 1$
- ii)  $F(x) = f(x)$
- iii)  $\int_{-\infty}^{\infty} f(x) dx = 1$

pdf  
Condition  
1)  $f(x) \geq 0, -\infty < x < \infty$   
2)  $\int_{-\infty}^{\infty} f(x) dx = 1$

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

Ques: A random variable  $x$  has the following probability funcn.

$x$	0	1	2	3	4	5	6	7
$p(x)$	0	$R$	$2R$	$2R$	$3R$	$R^2$	$2R^2$	$7R^2 + R$

- i) find  $R$
- ii) Evaluate  $p(x < 6)$ ,  $p(x \geq 6)$ ,  $p(3 < x \leq 6)$
- iii) find the minimum value of  $x$  so that  $p(x \leq x) > \frac{1}{2}$

Sol<sup>n</sup>:  $\sum p(x) = 1$

$$R + 2R + 2R + 3R + R^2 + 2R^2 + 7R^2 + R = 1$$

$$9R + 10R^2 = 1$$

$$\cancel{9R + 10R^2 = 1} \quad 9R + 10R^2 - 1 = 0$$

$$\cancel{-9R - 10R^2 = 1} \quad 10R^2 + 9R - 1 = 0$$

$$10R^2 + 10R - R - 1 = 0$$

$$10R(R+1) - 1(R+1) = 0$$

$$(R+1)(10R-1) = 0$$

$$R = \frac{1}{10} ; -1 \text{ min.} \quad [R \neq -1]$$

so,  $\boxed{R = \frac{1}{10}}$

ii)  $P(X < 6) = P(x) = P \rightarrow 1 - P(X \geq 6)$

$$P(X \geq 6) = P(6) + P(7) = 0.02 + 0.17 = 0.19.$$

$$P(3 < X \leq 6) = 0.37.$$

$$P(X \leq 1) = P(X=0) + P(X=1) = 0 + \frac{1}{10} = \frac{1}{10} < \frac{1}{2}$$

$$P(X \leq 2) = P(X=0) + P(X=1) + P(X=2) = 0 + \frac{1}{2} + \frac{2}{10} = \frac{3}{5} < \frac{1}{2}$$

$$P(X \leq 3) = P(X=0) + P(X=1) + P(X=2) + P(X=3) \\ = 0 + \frac{1}{10} + \frac{2}{10} + \frac{2}{10} = \frac{8}{10} = \frac{4}{5}$$

minimum value of  $x$  is 4.

Ques. If the func" defined as follows a density func.

$$f(x) = \begin{cases} e^{-x}, & x \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

If so, find  $P(1 \leq x \leq 2)$

sol.  $\int_{-\infty}^{\infty} f(x) dx = \int_0^{\infty} e^{-x} dx + \int_0^{\infty} e^{-x} dx = 1$

$$\int_0^{\infty} e^{-x} dx = -[e^{-x}]_0^{\infty} = -[e^{-2} - e^0] = e^2 - 1$$

$$\text{Gamma Function} \quad \int_0^{\infty} x^{n-1} e^{-ax} dx = \frac{\Gamma(n)}{a^n}$$

DATE \_\_\_\_\_  
PAGE NO. \_\_\_\_\_

Ques. 'X' is a r.v. with pdf given.

$$f(x) = \begin{cases} Rx, & \text{if } 0 \leq x < 2 \\ 2R, & \text{if } 2 \leq x < 4 \\ -Rx + 6R, & \text{if } 4 \leq x < 6 \end{cases}$$

Find R & mean value of X.

Sol'n

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

$$\int_{-\infty}^{\infty} [0 \cdot dx + \int_0^2 Rx \cdot dx + \int_2^4 2R \cdot dx + \int_4^6 (-Rx + 6R) \cdot dx] = 1$$

$$\int_0^2 0 \cdot dx = 1$$

$$0 + \left[ \frac{x^2}{2} \right]_0^2 + (2R[x]^4)_2^4 + \int_2^4 Rx \cdot dx + \int_4^6 6R \cdot dx = 1$$

$$R[2] + 4R - R\left[\frac{x^2}{2}\right]_2^4 + 6R[x]_4^6 = 1$$

$$2R + 4R - R[18 - 8] + 6R[x]_4^6 = 1$$

$$2R + 4R - 10R + 12R = 1$$

$$K = 1$$

$E(\phi(x)) \rightarrow$  Expected value  
 D.R.V      C.R.V      mean

$$\sum_x \phi(x) P(x) \quad \int_{-\infty}^{\infty} \phi(x) f(x) dx$$

$$\text{If } \phi(x) = x \quad E(x) = \sum_x x P(x)$$

$$\text{mean} = E(x)$$

$$E(\bar{x}) = \bar{x}$$

$$\text{variance} = E(x - \bar{x})^2$$

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

$$E(x) = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

## # Mean & variance of Random variable Discrete prob. Distribution:

$X$	$x_1$	$x_2$	...	$x_n$
$P(X)$	$p(x_1)$	$p(x_2)$		$p(x_n)$

$$\text{mean } (\mu) = \frac{\sum p_i x_i}{\sum p_i} = \sum p_i x_i$$

[Expected value  $E(x)$  / Average]

$$\text{variance} = V(X) = \sigma^2 = \sum p_i (x_i - \mu)^2$$

if  $\mu$  is not a whole number then  $\sigma^2$

$$= \sum p_i x_i^2 - \mu^2$$

standard Deviation,  $\sigma = \sqrt{V(X)}$ .

moments about the mean ( $M_x$ ).

$$x^k \text{ moment about the mean } (M_x)$$

$$M_x = \sum (x_i - \mu)^k F(x_i) \text{ where } F(x_i) = p_i$$

$$\text{For } k=0, M_0 = \sum F(x_i) = \sum p_i = 1$$

$$\text{For } k=1, M_1 = \sum (x_i - \mu) F(x_i) = \sum (x_i - \mu) p_i$$

$$= \sum x_i p_i - \mu \sum p_i = \mu - \mu = 0$$

$$\text{For } k=2, M_2 = \sum (x_i - \mu)^2 F(x_i) = \sigma^2$$

$$\text{For } k=3, M_3 = \sum (x_i - \mu)^3 F(x_i) = \sum (x_i - \mu)^3 p_i$$

$$\text{For } k=4 = M_4 = \sum (x_i - \mu)^4 F(x_i) = \sum (x_i - \mu)^4 p_i$$

$$\text{Mean Deviation about mean} = \sum |x_i - \mu| F(x_i)$$

## # Continuous prob. Distribution:

$$E(x) = \int_{-\infty}^{\infty} x F(x) dx$$

$$E(x^2) = \int_{-\infty}^{\infty} x^2 F(x) dx$$

$$\text{Var}(x) = \int_{-\infty}^{\infty} (x - \mu)^2 F(x) dx$$

$$= E(x^2) - [E(x)]^2$$

Mean Deviation about mean =  $\int |x - \mu| F(x) dx$

$$M_d = \int_{-\infty}^{\infty} |(x - \mu)|^r F(x) dx$$

If  $\mu = 0$  and if we put it in  $M_d$  then we will get  $E(|x|)$

(Expected value is the moment about origin)

### # Moment Generating funct.

$$M_a(t) = \sum P_i e^{t(x_i - a)} \quad M_a(t) = (Pe^t + q)^n$$

$$M_d = \left[ \frac{d^r}{dt^r} M_a(t) \right]_{t=0} \quad r^{\text{th}} \text{ moment about } a$$

$$M_a(t) = \sum P_i e^{t(x_i - a)} = e^{-at} \sum P_i e^{tx_i} = e^{-at} M_a(0)$$

$$M_a(t) \approx \int_{-\infty}^{\infty} e^{t(x-a)} f(x) dx$$

$$1) E(ax+b) = aE(x) + b$$

$$2) \text{Var}(ax+b) = a^2 \text{Var}(x)$$

Relation b/w Central moments ( $U_{i,j}$ ) & moments about any arbitrary origin 'a'.

$$u_i = u - a$$

$$U_2 = (u'_2) - (u'_1)^2$$

$$U_3 = u'_3 - 3u'_2 u'_1 + 2(u'_1)^3$$

$$U_4 = u'_4 - 4u'_3 u'_1 + 6u'_2 (u'_1)^2 - 3(u'_1)^4$$

14/09/2023

Q. The density function of a random variable  $x$  is given by  $f(x) = Kx(2-x)$ ,  $0 \leq x \leq 2$ .

Find

$$(i) K$$

$$(ii) \text{ Variance}$$

$$(iii) \text{ Mean Deviation about mean.}$$

$$\textcircled{1} \quad \int_{-\infty}^{\infty} f(x) dx = 1 \quad \int_0^2 Kx(2-x) dx = 1$$

$$K = 3/4$$

$$\textcircled{2} \quad \text{Variance} = E(x^2) - [E(x)]^2$$

$$E(x) = \int_0^2 x f(x) dx = \frac{3}{4} \int_0^2 x^2 (2-x) dx = 1$$

$$E(x^2) = \int_0^2 x^2 f(x) dx = \frac{3}{4} \int_0^2 x^2 (2-x) dx = \frac{6}{5}$$

$$\text{Variance} = \frac{6}{5} - 1^2 = \frac{1}{5}$$

$$\textcircled{3} \quad \text{Mean deviation about mean}$$

$$= \int_0^2 |x - \bar{x}| f(x) dx$$

$$= \int_0^2 |x - 1| f(x) dx$$

$$\Rightarrow \int_0^1 (1-x) f(x) dx - \int_1^2 (x-1) f(x) dx$$

= 38

8. Find the MGF of the exponential distribution  $f(x) = \frac{1}{c} e^{-x/c}$ ,  $0 \leq x \leq \infty$

$c > 0$ , hence find its mean of SD.

Sol<sup>n</sup> The MGF about origin is  $M_0(t)$ .

$$\begin{aligned} M_0(t) &= \int_0^\infty e^{tx} f(x) dx = \int_0^\infty e^{tx} \cdot \frac{1}{c} e^{-x/c} dx \\ &= \frac{1}{c} \int_0^\infty e^{(t-\frac{1}{c})x} dx = \frac{1}{c} \left[ \frac{e^{(t-\frac{1}{c})x}}{t-\frac{1}{c}} \right]_0^\infty \end{aligned}$$

$$\begin{aligned} &\Rightarrow (1-ct)^{-1} \\ &= 1 + ct + c^2 t^2 + c^3 t^3 + \dots \end{aligned}$$

$$\mu'_1 = \left[ \frac{d}{dt} M_0(t) \right]_{t=0} = [c + 2c^2 t + 3c^3 t^2 + \dots]_{t=0} = c$$

$$\text{Mean} = 0 + \mu'_1$$

$$\begin{aligned} \mu'_2 &= \left[ \frac{d^2}{dt^2} M_0(t) \right]_{t=0} = (2c^2 + 6c^3 t + \dots)_{t=0} = 2c^2 \\ &\Rightarrow E(X^2) - [E(X)]^2 \end{aligned}$$

$$\text{Variance} \sigma^2 = \mu'_2 - (\mu'_1)^2$$

$$\Rightarrow 2c^2 - c^2 = c^2$$

$$SD = \sqrt{\frac{\sigma^2}{c}}$$

Binomial Distribution  $\rightarrow B(n, p)$ Trials =  $n$ Probability of success =  $p$ Probability of failure =  $q$ 

$$\boxed{p+q = 1}$$

$$P(X=x) = {}^n C_x p^x q^{n-x}, \quad x = 0, 1, 2, \dots, n$$

probability mass func

Binomial Expansion / Distribution  $(p+q)^n$ Recurrence or Revision Formula for  $B(n/p)$ 

$$P(r) = {}^n C_r q^{n-r} p^r = \frac{n!}{r!(n-r)!} q^{n-r} p^r$$

$$\begin{aligned} P(r+1) &= {}^n C_{r+1} q^{n-r-1} p^{r+1} \\ &= \frac{n!}{(n-r-1)!(r+1)!} q^{n-r-1} p^{r+1} \end{aligned}$$

$$\frac{P(r+1)}{P(r)} = \frac{n-r}{r+1} \cdot \frac{p}{q}$$

$$P(r+1) = \frac{n-r}{r+1} \cdot \frac{p}{q} P(r)$$

# Binomial Distribution  $B(n, p)$ 

$$P(X=x) = {}^n C_x P^x q^{n-x}$$

$$\text{mean} = np$$

$$\text{variance} = npq$$

# Recurrence relation of Binomial Distribution.

$$P(x_{t+1}) = \frac{n-x}{x+1} \cdot \frac{p}{q} P(x),$$

Q1 If six dice are thrown 729 times. How many times do you expect at least 3 dice show a 5 or 6?

Sol<sup>n</sup>

$$n=6$$

$$p = \frac{2}{3} = \frac{1}{3},$$

$$q = 1-p = 1-\frac{1}{3} = \frac{2}{3}$$

$$P(\text{at least three success}) = P(x=3) + P(x=4) + \\ P(x=5) + P(x=6) \quad (\text{or}) \\ \Rightarrow 1 - P(x=0) - P(x=1) - P(x=2)$$

$$\therefore \frac{233}{729} = 0.3196$$

Q2. Determine the Binomial Distribution whose mean is 9 & S.D is  $\frac{3}{2}$ .

Sol<sup>n</sup>

$$\mu = 9 \Rightarrow np = 9$$

$$\sigma = S.D = \frac{3}{2} = \sqrt{npq}$$

$$\sigma^2 = \text{var} = \frac{9}{4}$$

$$\cancel{n=12} \quad npq = \frac{9}{4}, \quad n=12$$

$$q = \frac{1}{4} \Rightarrow p = \frac{3}{4}.$$

Q3. If the sum of the mean & var. of a binomial dist^n for 5 trials is 1.8, find the distribution.

Sol^n.  $\mu = np, \sigma^2 = npq, n=5$

ATQ

$$\begin{aligned}\mu + \sigma^2 &= 1.8 \\ np + npq &= 1.8 \\ np(1+q) &= 1.8 \\ np(1+(1-p)) &= 1.8 \\ np(2-p) &= 1.8 \\ sp(2-p) &= 1.8\end{aligned}$$

$$p = 0.1.8 \text{ or } 0.2$$

$p \neq 1.8$  as  $0 \leq p \leq 1$

$$p = 0.2 = \frac{1}{5}$$

$$q = 0.8 = \frac{4}{5}$$

$$\left(\frac{1}{5} + \frac{4}{5}\right)^5$$

Q4. calculate  $P(x)$  for  $x=1, 2, 3, 4 \& 5$  taking  $n=5, p=\frac{1}{5}$  with the help of the recurrence formula of the binomial distribution.

Sol^n  $P(x+1) = \frac{n-x}{x+1} \cdot \frac{p}{q} \cdot P(x)$

$$n=5, p=\frac{1}{6}, q=\frac{5}{6}$$

$$P(0) = {}^5C_0 \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^5 = \left(\frac{5}{6}\right)^5$$

$$P(0+1) = P(1) = \frac{5-0}{0+1} \times \frac{1}{5} P(0) = P(0) = \left(\frac{5}{6}\right)^5$$

$$P(2) = P(1+1) = (5-1) \times P(1)$$

Q. Fit a binomial distribution to the following data:-

$x$	0	1	2	3	4
$f$	30	62	46	10	2

Soln

$x$	$f$	$f \cdot x$	
0	30	0	$\mu = \sum f x = 192$
1	62	62	$\sum f = 150$
2	46	92	$= 1.28$
3	10	30	$np = 1.28$
4	2	8	$4p = 1.28$

$p = 0.32$ ,  
 $q = 0.68$ ,  
 $N = 150$ .

$$\begin{aligned} \text{Binomial distribution} &= N(p+q)^n \\ &= 150 \left(0.32 + 0.68\right)^4 \end{aligned}$$

## # Poisson Distribution

A limiting case of binomial distribution.

$$n \rightarrow \infty$$

$$P \rightarrow 0$$

$$\lambda = np \text{ (fixed)}$$

$$\boxed{\sigma^2 = \mu = \lambda}$$

$$P(X=x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$

Recurrence formula.

$$P(x) = \frac{e^{-\lambda} \lambda^x}{x!}$$

$$P(x+1) = \frac{e^{-\lambda} \lambda^{x+1}}{(x+1)!}$$

$$\boxed{\frac{P(x+1)}{P(x)} = \frac{\lambda}{x+1} \cdot P(x)}$$

$$P(x=x) = \frac{n!}{x!} q^{n-x} p^x$$

$$= \frac{n(n-1)(n-2)\cdots(n-x+1)(n-x)}{x!(n-x)!} \cdot x(1-p)^{n-x} p^x$$

$$\approx \frac{n(n-1)(n-2)\cdots(n-x+1)x}{x!} \left(1 - \frac{\lambda}{n}\right)^{n-x} \times \left(\frac{\lambda}{n}\right)^x$$

$$\therefore \lambda = np$$

$$p = \frac{\lambda}{n}$$

$$\Rightarrow \frac{x^x}{x!} \left[ \left( \frac{n}{n} \right) \left( 1 - \frac{1}{n} \right) \left( 1 - \frac{2}{n} \right) \cdots \left( 1 - \frac{x-1}{n} \right) \right] \cdot \left( 1 - \frac{x}{n} \right)^n$$

$$\frac{x^x}{x!} \left[ \left( 1 - \frac{1}{n} \right) \left( 1 - \frac{2}{n} \right) \cdots \left( 1 - \frac{x-1}{n} \right) \right] \left[ \left( 1 - \frac{x}{n} \right)^{\frac{n}{n}} \right]^x$$

As  $n \rightarrow \infty$ , each  $(x-1)$  factors  $\left( 1 - \frac{1}{n} \right), \left( 1 - \frac{2}{n} \right), \dots, \left( 1 - \frac{x-1}{n} \right)$  tends to 1.

Let  $x \rightarrow \pm \infty$   $\left( 1 + \frac{1}{x} \right)^x = e$ , then

$$\left[ \left( 1 - \frac{x}{n} \right)^{-\frac{n}{n}} \right]^{-x} \rightarrow e^{-x} \text{ as } n \rightarrow \infty.$$

$$\Rightarrow P(X=x) = \frac{x^x e^{-x}}{x!}$$

can be x

Q. Assume that the prob. of an individual coal miner being killed in a mine accident during a year is  $\frac{1}{2400}$ . Find the prob.

that in a mine employing 200 miners there will be at least one fatal accident a year.

Sol:

$$n = 200$$

$$x = n \times p$$

$$p = \frac{1}{2400}$$

$$= \frac{1}{12}$$

$$= 0.083$$

$$P(X \geq 1) = 1 - P(X=0)$$

D 1 Y  
=

Q2. Six coins are tossed 6400 times. Using the Poisson distribution determine the approx. prob. of getting six heads  $x$  times.

Sol

$$n = 6400$$

$$p = \left(\frac{1}{2}\right)^6 \rightarrow \text{six coins.}$$

$\rightarrow$  probability of getting head.

Q3. Fit a Poisson distribution & also find theoretical frequency.

$\mu$	0	1	2	3	4	Total	$N$
$f$	109	65	22	3	1	200	

$$\underline{\text{Sol}} \quad \mu = \frac{\sum xf}{\sum f} = \frac{122}{200} = 0.61$$

$$\lambda = \mu = 0.61$$

$$N = 200$$

$$P(x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad \Rightarrow \quad N \cdot \frac{e^{-\lambda} \lambda^x}{x!}$$

$x$	$N \cdot P(x)$	Theoretical frequency
0	108.7	109
1	66.3	66
2	20.2	20
3	4.1	4
4	0.7	1
		200

## # Normal Distribution

Limiting case of binomial distribution  
when  $n \rightarrow \infty$ ,  $P \rightarrow \frac{1}{2}$

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2} = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} z^2}$$

$$-\infty < x < \infty$$

$$\text{mean} = \mu$$

$$\text{SD} = \sigma$$

$$N(\mu, \sigma^2)$$

$$-\infty < \mu < \infty$$

$$\sigma > 0$$

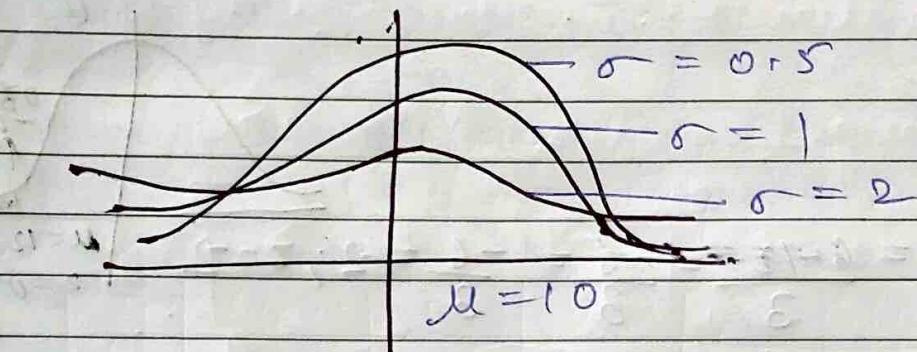
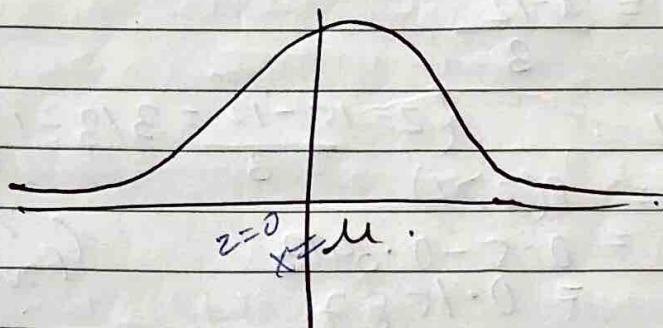
## # NOTES

- Normal Curve is a Symmetrical bell shaped curve.
- Normal Curve is symmetrical about mean  $\mu$ ;
- Two tails of the normal curve tends to

$-\infty \& +\infty$ 

- The curve is unimodal, which is nothing but its mean.
- The median of curve is also coincide with mean.
- Total area is 1.
- Standard normal variate,

$$Z = \frac{x - \mu}{\sigma}$$



The mean deviation about mean of ND is  $\frac{4}{3}$  of its SD

$$\int_{-\infty}^{\infty} |x - \mu| f(x) dx$$

$$\# P(z_1 \leq Z \leq z_2)$$

$$P(z_1 < Z \leq z_2)$$

$$P(z_1 \leq Z < z_2)$$

$$P(z_1 \leq Z \leq z_2)$$

All are same :

Q A sample of 100 dry battery cells tested to find the length of life produced the following result  $\bar{x} = 12$  hours,  $s = 3$  hours.

Assuming the data to be normal distributed, what percentage of battery cells are expected to have life:

(i) More than 15 hours

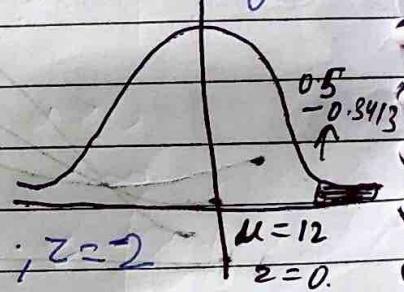
(i.) Less than 6 hours  $P(x < 6)$

(ii) Between 10 & 14 hours  $P(10 < x < 14)$

Sol  $z = \frac{x - \bar{x}}{s} = \frac{x - 12}{3}$

(i) For  $x = 15$ ,  $z = 1$   $\left( z = \frac{15 - 12}{3} = 1 \right)$  Normal distribution  
 $P(x > 15) = P(z > 1)$   
 $= 0.5 - 0.3413$   
 $= 0.1587$  (From table)  
Google

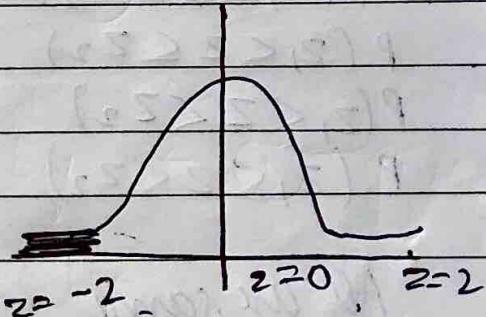
$\boxed{\% = 15.87\%}$



(ii) For  $x = 6$ ,  $z = \frac{6 - 12}{3} = \frac{-6}{3} = -2$ ;  $z = 2$

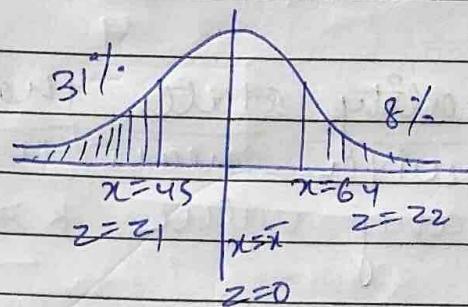
$$\begin{aligned} P(z < 6) &= P(z < -2) \\ P(x < 6) &= P(z > 2) \\ &= 0.5 - P(z = 2) \\ &= 0.5 - 0.9772 \\ &= 0.0228 \end{aligned}$$

$\boxed{\% = 2.28\%}$



Q1. In a normal distribution 31% of the items are under 45 & 8% are over 64. Find the mean & S.D of the Distribution.

Sol<sup>n</sup> Let  $\bar{x}$  &  $\sigma$  be the mean & S.D resp.



when  $x=45$ , let  $z=z_1$   
 $P(z_1 < z < 0) = 0.5 - 0.31 = 0.19$

From the table the value of  $z$  corresponding to the area is 0.5.

$$z_1 = -0.5 \quad (\because z_1 < 0)$$

when  $x=64$  let  $z=z_2$

$$P(0 < z < z_2) = 0.5 - 0.08 = 0.42$$

From table

$$z_2 = 1.4$$

Since,  $z = \frac{x - \bar{x}}{\sigma}$

$$-0.5 = \frac{45 - \bar{x}}{\sigma} \quad \& \quad 1.4 = \frac{64 - \bar{x}}{\sigma}$$

$$45 - \bar{x} = -0.5\sigma \quad \text{--- (1)}$$

$$64 - \bar{x} = 1.4\sigma \quad \text{--- (2)}$$

Solving (1) & (2), we get  
 $\bar{x} = 50, \& \sigma = 10.$

# Chebychev's Inequality $\sigma \rightarrow \infty$ ↓ limited case  
of Markov's inequality→ Prob. of getting values farther away from  $\mu$ .The result is helpful when actual distribution of  $X$  is not known.

Chebychev's inequality states that:-

If  $X$  is a r.v. with mean  $\mu$  & variance  $\sigma^2$  then for any value  $k > 0$ .

$$P\{|x - \mu| \geq k\} \leq \frac{\sigma^2}{k^2}$$

Proof:- we prove it for the case when  $X$  is continuous with density func  $f(x)$ for a non-negative r.v.  $X$  &  $a > 0$ .

$$E(X) = \int_0^\infty x f(x) dx + \int_a^\infty x f(x) dx \geq \int_a^\infty x f(x) dx \\ \geq a \int_a^\infty f(x) dx$$

$$= a P\{X \geq a\}$$

$$\text{Thus, } P\{X \geq a\} \leq \frac{E(X)}{a} \quad \text{--- (1)}$$

→ Markov's inequality

Replacing  $x$  by  $(x - \mu)^2$  &  $a$  by  $k^2$ , (1) gives

$$P\{(x - \mu)^2 \geq k^2\} \leq \frac{E((x - \mu)^2)}{k^2}$$

$$P\{|x-\mu| \geq k\} \leq \frac{\sigma^2}{k^2} \quad \text{H.P}$$

1) By replacing  $k$  by  $k\sigma$  in chebyshew's inequality, can be expressed as.

$$P(|x-\mu| \geq k\sigma) \leq \frac{1}{k^2} \rightarrow \text{Surf}$$

2) Sometimes result is applicable in the form  $P\{|x-\mu| \geq k\sigma\} \geq 1 - \frac{1}{k^2}$

3) The inequality  $P(x \geq a) \leq \frac{E(x)}{a}$  is known as markov's inequality.  $a$  &  $E(x)$  is of practical importance.

Q1. The no. of items cleared by an assembly line during a week is a r.v. with mean 50 & variance 25.

(a) what is the prob. that this week items cleared cleared will exceed 75?

(b) what can be said about the prob. that this week's clearance will be b/w 40 to 60.

Sol: Let  $x$  be the r.v. denoting the no. of items cleared in a week.

$$E(x) = 50 \quad \& \quad \text{Var}(x) = 25$$

(a). As  $P\{x \geq a\} \leq \frac{E(x)}{a}$ ,  $a > 0$ .

$$P(x > 75) \leq \frac{50}{75} = \frac{2}{3}$$

(b)

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

$$K = 10 \longrightarrow \begin{array}{l} 50 - 10 = 40 \\ 50 + 10 = 60 \end{array}$$

$$P\{|X-50| \geq 10\} \leq \frac{25}{100} = \frac{1}{4}$$

$$P\{|X-50| < 10\} \geq \frac{3}{4}$$

$$P\{40 < X < 60\} \geq \frac{3}{4} = 0.75$$

# Exponential Distribution  $\rightarrow$  Depends on only 1 parameter

A continuous random variable  $X$  is said to have exponential (negative exponential) distribution with parameter ' $a$ '. If it assumes only real positive values and its p.d.f is given by,

$$f(x) = \begin{cases} ae^{-ax}, & \text{if } x > 0 \\ 0, & \text{otherwise.} \end{cases}$$

where,  $a > 0$  is the prescribed parameter.

$$\text{mean} = 1/a \quad \& \quad \text{variance} = \frac{1}{a^2}$$

Q) The time b/w breakdown of a machine follows an exponential distribution with a mean of 14 days. Find the prob. that a machine breaks down in a 15 days period?

Soln

$$\mu = 17$$

$$\frac{1}{a} = 17$$

$$a = \frac{1}{17}$$

$$\text{pdf} = ac^{-ax} = \frac{1}{17} e^{-t/17}, \text{ for } t \geq 0$$

$$\begin{aligned} P(0 \leq t \leq 15) &= \int_0^{15} f(t) dt \\ &= \int_0^{15} \frac{1}{17} e^{-t/17} dt \\ &= \left[ -e^{-t/17} \right]_0^{15} = 1 - e^{-15/17} \end{aligned}$$

$$\Rightarrow 0.5862$$

∴ There is 58.62% chances.

### # Gamma Distribution.

A cts. r.v.  $X$  defined over the range  $x \geq 0$ .  
s.t.  $X$  Gamma distributed if the p.d.f. is of  
the form

$$f(x) = \frac{\lambda^x x^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x}, \quad x \geq 0$$

where  $\alpha > 0$  &  $\lambda > 0$  &  $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$ .

$\Gamma(\alpha) = (\alpha-1)$  If  $\alpha$  has an integer value.

## # CORRELATION →

↳ coeff. of correlation ( $r$ )  $[-1 \leq r \leq 1]$

↳ types of correlation

- |                     |              |
|---------------------|--------------|
| (1) +ve             | $0 > r > 1$  |
| (2) Strictly +ve    | $r = +1$     |
| (3) -ve             | $-1 < r < 0$ |
| (4) Strictly -ve    | $r = -1$     |
| (5) No correlation. | $r = 0$      |

## # Karl Pearson's coeff. of correlation.

$$\rho = \text{covariance } (x, y) = \frac{\text{cov}(x, y)}{\sqrt{\text{var } x} \sqrt{\text{var } y}} = \frac{\sum xy}{\sigma_x \sigma_y} = \frac{(\sum x^2)(\sum y^2)}{N}$$

$$x = x - \bar{x}$$

$$y = y - \bar{y}$$

$$\text{cov} = \frac{\sum xy}{N}$$

Total no. of terms.

For assumed mean.

$$x' = x - a$$

$$y' = y - b$$

$$\rho = \frac{N \sum x'y' - \sum x' \sum y'}{\sqrt{N(\sum (x')^2) - (\sum x')^2} \sqrt{N(\sum (y')^2) - (\sum y')^2}}$$

Q1. Calculate the correlation coeff. b/w the following data:-

x	5	9	13	17	21
y	12	20	25	33	35

Sol<sup>n</sup> .  $\bar{x} = 13$  ,  $\bar{y} = 25$

$$X = x - \bar{x} , Y = y - \bar{y}$$

$x$	$y$	$X$	$Y$	$x^2$	$y^2$	$XY$
5	12	-8	-13	64	169	104
9	20	-4	-5	16	25	20
13	25	0	0	0	0	0
17	33	4	8	16	64	32
21	35	8	10	64	100	80
		160	358		236	

$$r = \frac{\sum XY}{\sqrt{\sum X^2 \sum Y^2}} = \frac{236}{\sqrt{160 \times 358}} = \frac{236}{239.33} = 0.986.$$

Q2.

$x$	21	23	30	54	57	58	72	78	87	90
$y$	60	71	72	83	110	84	100	92	113	135

$$\bar{x} = 57 , \bar{y} = 92$$

$$x = x - \bar{x} , y = y - \bar{y}$$

$x$	$y$	$X$	$Y$	$x^2$	$y^2$	$XY$
21	60	-32	-32	1296	1024	1152
23	71	-34	-21	1156	441	714
30	72	-27	-20	729	400	840
54	83	-3	-9	9	81	27
57	110	0	18	0	824	0
58	84	1	-8	1	64	-8
72	100	15	8	225	64	120
78	92	21	0	441	0	0
87	113	30	21	900	441	630
90	135	33	43	1089	1899	1419
				5842	4688	4594

$$H = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}} = \frac{4594}{5233.29}$$

$$H = 0.8778$$

$\bar{x}$	3	7	5	4	6	8	2	7
$\bar{y}$	7	12	8	9	8	10	13	5

$$\det \begin{pmatrix} a & b \\ b & a \end{pmatrix} = 10 \cdot 3 - 10 \cdot 3 = 0$$

Ans: 0.8778

CEPES REVISION  
Lecture 5

OP 153 3 5 3 2 1 7 2 1 2 1 0 3  
2 1 2 1 3 2 1 6 2 0 1 1 0 3 3 1

# # Spearman's Rank Correlation Coefficient

$$r = 1 - \frac{6 \sum d^2}{n(n^2-1)}$$

for equal ranks

$$d = R_1 - R_2$$

$$r = 1 - \frac{6[\sum d^2 + \frac{1}{12}(m_1^3 - m_1) + \frac{1}{12}(m_2^3 - m_2)]}{n(n^2-1)}$$

Q. Obtain the rank correlation coeff for the following data:

x	68	64	75	50	64	80	75	40	55	64
y	62	58	68	45	81	60	68	48	50	70

x	y	R <sub>1</sub>	R <sub>2</sub>	d	d <sup>2</sup>
68	62	4	5	-1	1
64	58	6	7	-1	1
75	68	2.5	3.5	-1	1
50	45	9	10	-1	1
64	81	6	1	5	25
80	60	1	6	-5	25
75	68	2.5	3.5	-1	1
40	48	10	9	1	1
55	58	8	8	0	0
64	70	6	2	4	16
					72

If same, ranking then takes average of their ranks

$$n = 10$$

~~$\sum d^2 = 72$~~

$$m_1 = 2 \quad (75)$$

$$m_2 = 2 \quad (68)$$

$$m_3 = 3 \quad (64)$$

$$\mu = 1 - \frac{1}{6} \left[ \sum d^2 + \frac{1}{2} (m_1^3 - m_1) + \frac{1}{2} (m_2^3 - m_2) + \frac{1}{2} \left( \frac{m_3^3 - m_3}{m_3^2 - m_3} \right) \right]$$

$$n(n^2 - 1)$$

$$= 0.545$$

# Regression  $\rightarrow$  relationship

line of Regression  $X$  on  $Y$ .

$$x - \bar{x} = b_{xy} (y - \bar{y})$$

coefficient of regression

Line of Regression  $Y$  on  $X$

$$y - \bar{y} = b_{yx} (x - \bar{x})$$

$$b_{xy} = \text{or } \frac{\sigma_x}{\sigma_y} = \frac{N \sum xy - \sum x \sum y}{N \sum y^2 - (\sum y)^2}$$

$$\therefore = \frac{\sum xy}{\sum y^2}$$

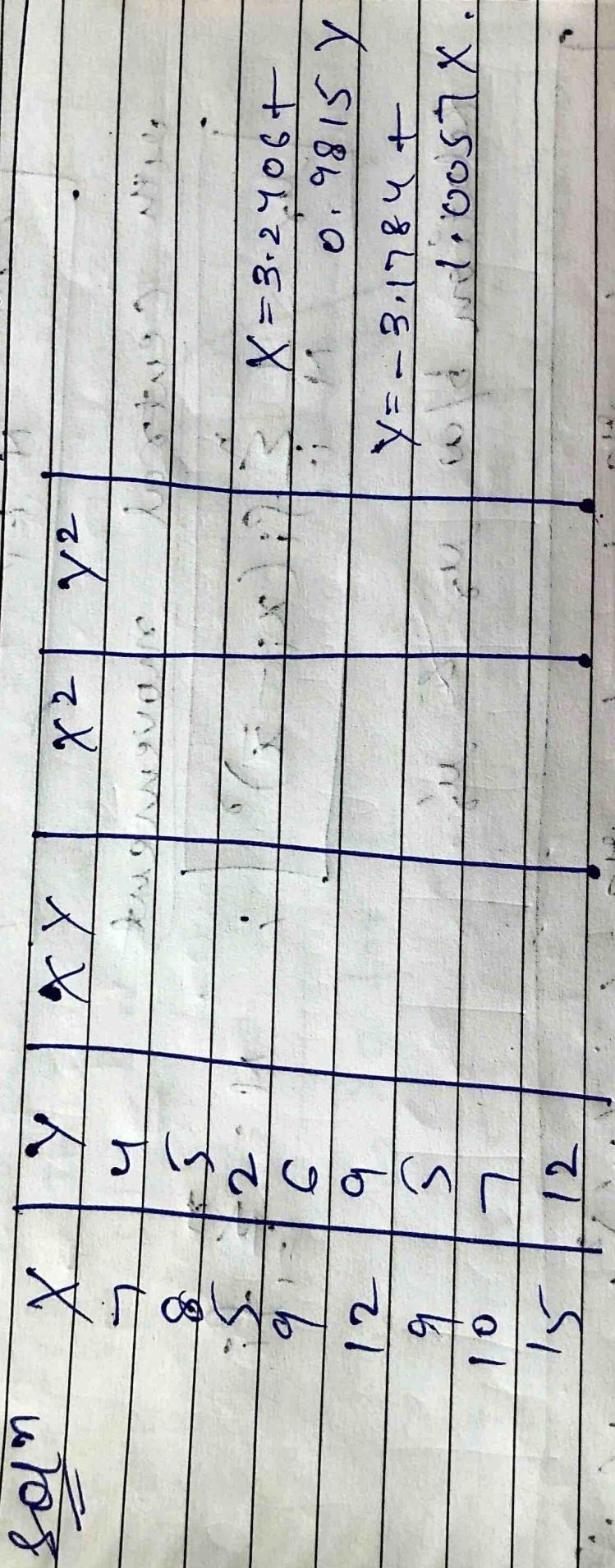
$$\sigma^2 = b_{yx} \cdot b_{xy} \quad : \quad x = x - \bar{x}$$

$$y = y - \bar{y}$$

$$b_{yx} = \text{or } \frac{\sigma_y}{\sigma_x} = \frac{N \sum xy - \sum x \sum y}{N \sum x^2 - (\sum x)^2} = (\sum x)^2$$

from the following data obtain the two regression lines:-

Capital Employed (Rs. in Lakh)	7	8	5	9	12	9	10	15
Sales Volume (Rs. in Lakh)	4.5	5.2	2.6	9.5	15.7	12.5	11.0	17.2



## Moments

$\text{ith moment about any pt. } x=a.$

$$1. \Gamma \quad \mu'_r = \frac{1}{N} \sum_{i=1}^n f_i (x_i - a)^r ; \quad N = \sum_{i=1}^n f_i$$

$\text{ith central movement}$

$$\mu_r = \frac{1}{N} \sum_{i=1}^n f_i (x_i - \bar{x})^r ; \quad N = \sum_{i=1}^n f_i$$

relation b/w  $\mu_r$  &  $\mu'_r$

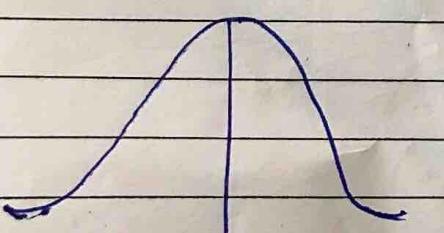
$$\mu_r = \mu'_r - r! C_1 (\mu'_{r-1} - \mu'_1) + r! C_2 (\mu'_{r-2} - \mu'_2)^2 + \dots + (-1)^r (r-1)! (\mu'_1)$$

$\Rightarrow$  Pearson's  $\beta$  &  $\gamma$  co-eff.

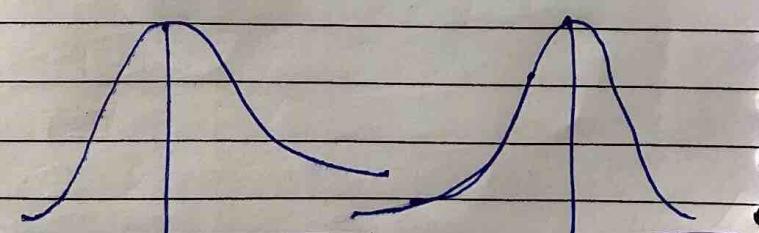
$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}, \quad \gamma_1 = \sqrt{\beta_1}, \quad \beta_2 = \frac{\mu_4}{\mu_2^2}$$

$$\gamma_2 = \beta_2 - 3$$

# Skewness  $\rightarrow$  measure of asymmetry.



Mean = Median = Mode

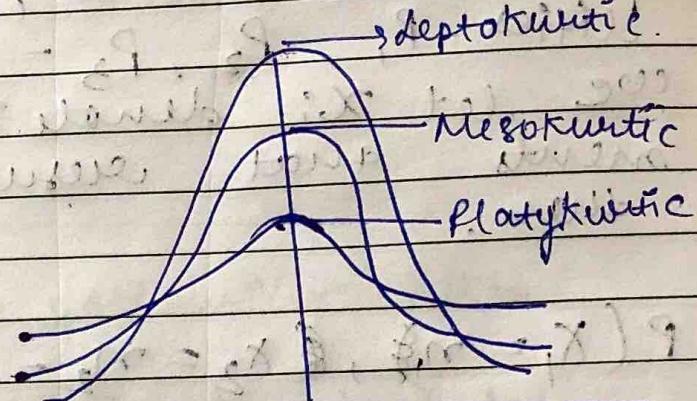


Mode < Median < Mean

Mean < Median < Mode

ESTD 1971

no. skewed + very skewed - very skewed

Karl Pearson's coeff. of skewness =  $\frac{\text{mean} - \text{mode}}{3 \times \text{S.D.}}$ Measure of skewness =  $B_2$ Kurtosis  $\rightarrow$  normally distributed $B_2 = 3$  or  $\gamma_2 = 0$  for M $B_2 > 3$  or  $\gamma_2 > 0$  for L $B_2 < 3$  or  $\gamma_2 < 0$  for P

mean = mode = median.

Q1. Calculate the first 4 moments for the following frequency distribution about the mean &amp; comment upon these.

$x$	-4	-3	-2	-1	0	1	2	3	4	5
$f$	3	4	5	7	12	7	5	4	3	1

$$\mu_1 = 0, \quad \mu_2 = 44.4, \quad \mu_3 = 0, \quad \mu_4 = 47.16$$

$$B_1 = 0, \quad B_2 = 2.39 < 3$$

## Multidimensional Distribution

A joint distribution, which arises when a sequence of  $n$  independent and identical experiments is performed. Suppose that each experiment can result in any one of  $r$  possible outcomes, with respective probabilities  $P_1, P_2, P_3, \dots, P_r, \sum_{i=1}^r p_i = 1$ . If we let  $x_i$  denote the no. of experiments that result in outcome no.  $i$  then.

$$P(x_1 = n_1, x_2 = n_2, \dots, x_r = n_r)$$

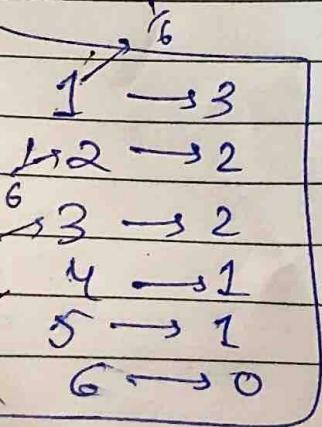
$$\frac{n_1! n_2! \dots n_r!}{n_1! n_2! \dots n_r!} p_1^{n_1} p_2^{n_2} p_3^{n_3} \dots p_r^{n_r}$$

$$\sum_{i=1}^r n_i = n$$

Q1. If a fair die is rolled 9 times. find the prob. that one appears 3 times, 2 & 3 appears twice each, 4 & 5 one each and 6 not at all.

$$\text{Soln. Req. Probability} = \frac{9!}{3! 2! 2! 1! 1! 0!} \left(\frac{1}{6}\right)^3 \left(\frac{1}{6}\right)^2$$

$$\left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right) \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^0$$



$$n_4 = 1$$

$$n_5 = 1$$

$$n_6 = 0$$

$\Rightarrow \frac{9!}{3!2!2!1!1!0!} \times \binom{1}{5}^9$

$\Rightarrow \frac{7 \times 5}{3 \times 6^5} = \frac{35}{23328}$

Q1. calculate  $\sigma$  for the following data.

$x$	9	8	7	6	5	4	3	2	1	1
$y$	15	16	14	13	11	12	10	8	9	11

Q2. Find rank correlation

$x$	56	42	72	36	63	47	55	49	38
$y$	147	125	160	118	149	128	150	145	115

Q3.  $X$  is normally distributed  $N(30, 5)$

= find probability.

(a)  $x \geq 45$

(b)  $26 \leq x \leq 40$

(c)  $|x - 30| > 5$

~~A~~ ~~Q~~ here,  $\bar{x} = 5$   
 $\bar{y} = 12$ .

$X$	$Y$	$X^2$	$Y^2$	$XY$
4	3	16	9	12
3	4	9	16	12
2	2	4	4	4
1	1	1	1	1
0	-1	0	1	0
-1	0	1	0	0
-2	-2	4	4	4
-3	-4	9	16	12
-4	-3	16	9	12
		60	60	57

$$\text{S.D. } \sigma = \frac{\sqrt{57}}{\sqrt{60 \times 60}} \approx \frac{\sqrt{57}}{\sqrt{3600}} = \frac{\sqrt{57}}{60} \approx 0.9$$