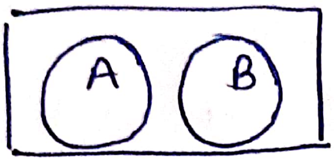
Important Points

- Probability – "chances"

- Exclusive or Disjoint Events –



Sample spaces have no point in common

For Exp – Tossing a coin –

A : getting a head $= \{H\}$

B : getting a tail $= \{T\}$

$$A \cap B = \phi$$

- Exhaustive Events –

"Total number of all possible outcomes of Random Experiment"

$$\text{Probability of an Event} = \frac{\text{number of favourable cases}}{\text{Total number of cases.}}$$

$$P(A) = \frac{n(A)}{n(S)}$$

Probability lies between 0 to 1

$$0 \leq P(A) \leq 1$$

Addition Law of probability

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

For mutually disjoint / Exclusive events -

$$P(A \cap B) = 0$$

Hence $P(A \cup B) = P(A) + P(B)$

Multiplication theorem of probability

$$P(A \cap B) = P(A) \times P(B/A)$$

OR

$$P(A \cap B) = P(B) \times P(A/B)$$

For Independent events

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

If A & B are Independent events - then -

(i) A & $\bar{B}$ are Independent events.

(ii) $\bar{A}$ & B are Independent Events

(iii) $\bar{A}$ & $\bar{B}$ are Independent events.

conditional probability -

$$P\left(\frac{A}{B}\right) = \frac{n(A \cap B)}{n(B)}$$

↳ Probability of A when B has already occured.

Probability Spaces

Three tuple (S,F,P) components are -

(1) S → (Sample space)

(2) F → (Event space)

(3) P → (Probability measure) $P : F \rightarrow R$

Baye's theorem - (learn with proof)

If E_1, E_2 & $E_3 \ldots$ & E_n are mutually Exclusive and exhaustive Events with $P(E_i) \neq 0$ $(i = 1, 2, \ldots n)$ of a Random Experiment then for any arbitary Event A of Sample space of Experiment with $P(A) \geq 0$

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

Random variables/variate (X)

A function X assigning to every element $w \in S$ to a real number, $X(w)$ is called a Random variable.

Types

Discreate — finite number of distinct values.

Ex. no. of Employees.

continous — defined a function in interval.

$x \in (0,1)$

Theorems on Random variable –

① $\{w : X(w) < a\} \in F$ (Event space)

② If X_1 & X_2 are Random variable & c is constant then cX_1, cX_2, $X_1 + X_2$, $X_1 X_2$ also Random variables.

Distribution function

$$F(x) = P(X \leq x) \quad \forall \; -\infty \leq x \leq \infty$$

Probability mass function –

$$p(x) = \begin{cases} P(X = x_i) = P_i & x = x_i \\ 0 & x \neq x_i \end{cases}$$

conditions –

① $P(x_i) \geq 0 \; \forall i$ ② $\sum_{i=1}^{n} p(x_i) = 1$

Distribution function properties -

1. If F is distribution function of Random variable X and $a < b$ then,

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

$$P(a < X < b) = F(b) - F(a) - P(X = b)$$

$$P(a \leq x \leq b) = P(X = a) + F(b) - F(a) - P(X = b)$$

$$P(a \leq x \leq b) = F(b) - F(a) + P(X = a)$$

2. If $0 \leq F(x) \leq 1$ and $x < y$ then $F(x) \leq F(y)$
i.e., Distribution function is increasing function.

Probability density function -

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

Conditions/Properties -

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

$$P(\alpha \leq X \leq \beta) = \int_{\alpha}^{\beta} f(x)\,dx$$

Expectation value of Random variable -

$$E(X) = x_1 p_1 + x_2 p_2 + x_3 p_3 + \dots + x_n p_n.$$

$$= \sum_{i=1}^{n} x_i p_i(x)$$

$\rightarrow E(X) = \bar{X} = \mu$ {Mean}

For Random discreate variable -

$$E(g(x)) = \Sigma g(x) p(x) dx$$

<u>variance</u> -

$$\text{vare}(X) = E(X^2) - (E(X))^2 = \sigma^2$$

SD

$$\boxed{SD = \sqrt{var(X)} = \sigma}$$

$$\text{varience}(X) = E\{(X - E(X))^2\}$$

$\boxed{\text{M.M. Imp}}$

<u>chebyshev's Inequality</u> -

$X \rightarrow$ Random variable $\mu \rightarrow$ mean.

$\sigma^2 \rightarrow$ variance $K \rightarrow$ +ve number

$$P\{|X - \mu| \geq K\sigma\} \leq 1/K^2$$

or

$$P\{|X - \mu| < K\sigma\} \leq 1 - \frac{1}{K^2}$$

UNIT - II

STATISTICS

Mean → Average value {Expectation value}

median → middle value.

mode → Repeating number (most)

Moments -

$$\mu_r' = E(X-A)^r$$

Moment about **origin**

① $\mu_0' = 1$.

② $\mu_1' = E(X) = \mu$

→ 1st moment about origin is mean.

③ $\mu_2' = E(X)^2$

Moment about **mean** → $\mu_r = E(X-\bar{X})^r$

① $\mu_0' = 1$

② $\mu_1 = 0$

Ist moment about mean is zero (0)

③ $\mu_2 = E(X^2) - (E(X))^2$

2nd moment about mean is variance.

Shape of any distribution can be described by its various moments -

1) mean → central tendency.
2) Second moment → varience (width)
3) third moment → Skewness (assymetric lining to Either left or Right)
4) fourth moment → Kurtosis (degree of central peaknes or Equivalently) (The flatness of outer tales)

Pearson's β and γ cofficients -

$$\boxed{\beta_1 = \frac{\mu_3^2}{\mu_2^3}}$$

$$\boxed{\gamma_1 = \frac{\mu_3}{\sigma^3} = \pm\sqrt{\beta_1}}$$

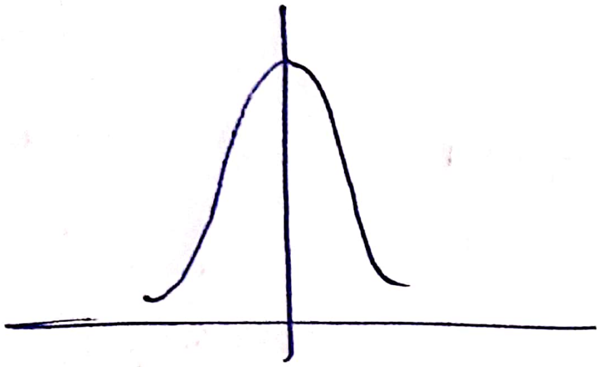
$$\boxed{\beta_2 = \frac{\mu_4}{\mu_2^2}}$$

$$\boxed{\gamma_2 = \beta_2 - 3}$$

Skewness - "lack of symmetry"

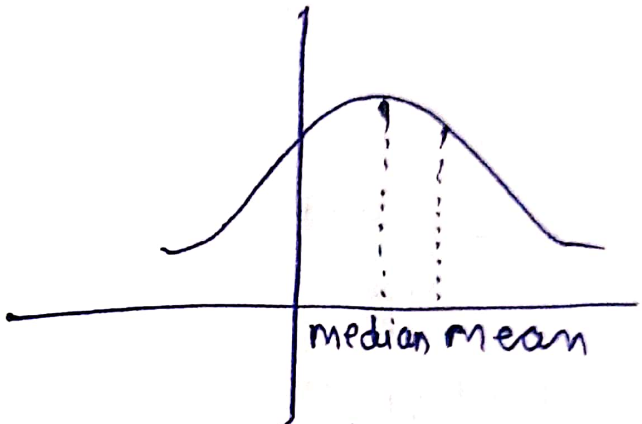
↳ study to take an idea about shape of curve.

Symmetrical

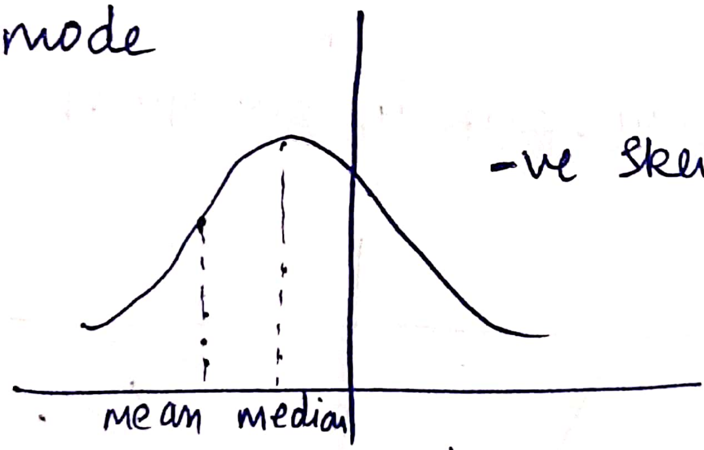


mean = = median = = mode

+ve skewed



mode



mode

-ve skewed

If,

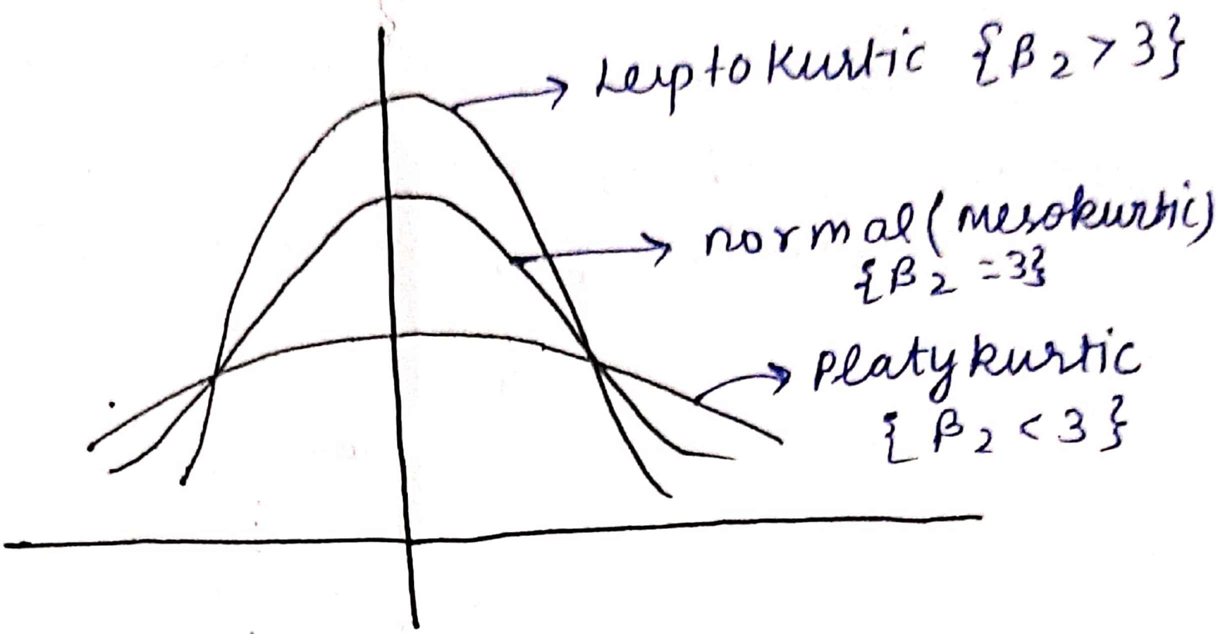
$\beta_1 = 0$, curve will be symmetrical.

$\beta_1 > 0$ curve will be positive skewed.

$\beta_1 < 0$, curve will be negetive skewed

Kurtosis :-

↳ flatness or peakness of frequency curve.



Probability distribution

- Discreate
  - Binomial
  - Possion
- Continous
  - ↓ Normal

Moment Genrating function -

$$M_a(t) = E(e^{t(x-a)})$$

For Discreate Random variable

$$M_a(t) = \sum p(x) e^{t(x-a)}$$

$p(x) \rightarrow$ Probability mass function.

For continous Random variable

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

$f(x) \rightarrow$ Probability density function

Binomial distribution / Bernoulli's distribution -

→ Probability of Success or failure outcome, occurance or Non occurance yes or No when Repeated multiple times.

$$P(X=x) = \begin{cases} p(x) = {}^nC_x \cdot p^x q^{n-x} & x = 0, 1, 2 \ldots \\ 0 & \text{otherwise} \end{cases}$$

$p \rightarrow$ Probability of success.

$q \rightarrow$ Probability of failure.

FOR Binomial distribution

M.G.F

$M_x(t) = (e^t p + q)^n$

$p \to$ Probability of Success.
$q \to$ Probability of failure

Mean & variance -

(μ) Mean $= np$

$n \to$ no. of times Experiment Performed.

(σ^2) variance $= npq$

$S.D = \sqrt{\text{variance}} = \sqrt{\sigma^2}$
$= \sqrt{npq}$

$$\boxed{\mu_3 = npq\,(q-p)}$$

$$\boxed{\mu_4 = npq\,(1+3(n-2)pq)}$$

$$\boxed{\beta_1 = \frac{(q-p)^2}{npq}} \qquad \boxed{\gamma_1 = \frac{1-2p}{\sqrt{npq}}}$$

* If ~~p~~ $p > \frac{1}{2}$ then Skewness -ve.
* If $p < \frac{1}{2}$ then Skewness +ve
* If $p = \frac{1}{2}$ then Skewness is Symmetrical

$$\boxed{\beta_2 = 3 + \frac{(1-6pq)}{npq}} \qquad \boxed{\gamma_2 = \frac{1-6pq}{npq}}$$

Poisson Distribution -

* Limiting case of Binomial distribution
  $n >>>$, $p <<<<$, $np = \text{constant}$
  $\boxed{np = m}$

Proof ★★★

$$\boxed{P(r) = \frac{m^r}{r!} e^{-r}}$$

mean $= np = m$.

variance $= npq = m$.

$\boxed{\mu_3 = m}$; $\boxed{\mu_4 = 3m^2 + m}$

$\boxed{\gamma_1 = \frac{1}{\sqrt{m}}}$; $\boxed{\beta_2 = 3 + \frac{1}{m}}$

Normal distribution

$$Z = \frac{x - np}{\sqrt{npq}} = \frac{x - \mu}{\sigma}$$

$\mu \rightarrow$ mean

$\sigma \rightarrow$ st. deviation

In normal distribution

mean = median = mode.

$$y = P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

$$= \frac{1}{\sigma\sqrt{2\pi}} e^{-z^2/2}$$

$y = P(x) \longrightarrow$ Normal curve.

moment about mean -

$$\mu_{2n+1} = 0$$

$$\mu_{2n} = [(2n-1)(2n-3) \cdots 5 \cdot 3 \cdot 1]\sigma^{2n}$$

$\boxed{\mu_2 = \sigma^2}$ $\boxed{\beta_1 = 0}$

$\boxed{\mu_4 = 3 \cdot \sigma^4}$ $\boxed{\beta_2 = 3}$

Correlation -

numerical measure of correlation
→ cofficient of correlation.

$$\underbrace{\text{for } r}_{\text{Karl pearson's cofficient of correlation}} = \frac{\Sigma XY}{n\sigma_x\sigma_y} = \frac{cov(x,y)}{\sigma_x\sigma_y}$$

Karl pearson's
cofficient of correlation

$$cov(x,y) = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{n}$$

$$\sigma_x = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}} \qquad \sigma_y = \sqrt{\frac{\Sigma(y-\bar{y})^2}{n}}$$

$X \rightarrow$ Deviation from mean $\bar{x} = x - \bar{x}$
$Y \rightarrow$ Deviation from mean $\bar{y} = y - \bar{y}$

NOTE - Cofficient of correlation lies between $-1 \leq \rho \leq 1$

Method of calculation

Direct

Step deviation

Direct -

[illegible]

$$\boxed{\rho = \frac{\Sigma XY}{\sqrt{\Sigma X^2 \Sigma Y^2}}}$$

$$\boxed{X = x - \bar{x}} \qquad \boxed{Y = y - \bar{y}}$$

Step deviation method -

$$\boxed{\rho = \frac{n \Sigma d_x \Sigma d_y - \Sigma d_x \Sigma d_y}{\sqrt{[n \Sigma d_x^2 - (\Sigma d_x)^2][n \Sigma d_y^2 - (\Sigma d_y)^2]}}}$$

$$d_x = \frac{x - a}{h} \qquad d_y = \frac{y - b}{k}$$