[Ch :- 2]

Measures of Central Tendency

* Arethmetic mean = A.M = $\bar{x}$.

$\bar{x} = \frac{\Sigma x_i}{n}$

$\bar{x} = \frac{\Sigma f_i x_i}{n}$, where $n = \Sigma f_i$.

$\bar{x} = A + \frac{\Sigma f_i d_i}{n} \times C$

where, $d_i = \frac{x_i - A}{c}$ = Deviation

A = Assumed mean.

C = class interval.

* Properties of A.M. :-

1. If all Observations are equal then A.m. = the Same Observation.

2. * the Sum of Deviations taken from A.M. is always zero, i.e. $\Sigma(x - \bar{x}) = 0$.

3. The Sum of squares of the Deviations is minimum when it is taken from A.m. i.e. $\Sigma(x - \bar{x})^2 \leq \Sigma(x - A)^2$. where, A = any number.

4. If observations are At equal distances then, A.m. = $\frac{\text{Smallest} + \text{Largest}}{2}$.

5. A.m. is Affected by the change of origin a

the change of scale.
i.e. If $y = a + bx$,
then $\bar{y} = a + b \cdot \bar{x}$.

6. When frequency (f_i) are multiplied or Divided by a constant number then arithmetic mean (A.M.) remains unchange.

7. A.M. of first n natural numbers $= \frac{n+1}{2}$.

⇒ The A.M. of first n odd natural numbers $= n$.
⇒ The A.M. of first n even natural numbers $= n+1$.

8. When observations are having different importance in the data, In place ~~ob~~ of simple mean, we use weighted mean.
i.e. $\bar{x}_w = \frac{\Sigma w_i x_i}{\Sigma w_i}$.

✱ Combine mean

$$\bar{x}_c = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{n_1 + n_2}$$

✱ Median :—— (M).

→ It is a positional average.
→ It is not affected by xtreme observations (અંતિમ મૂલ્યો).
→ After arranging observations into Ascending order (ચડતા ક્રમમાં), the observation which falls exactly in the middle is known as median of the data.

C.f. = Cumulative frequency.

→ for ungroup data,

$$M = \left(\frac{n+1}{2}\right)^{th} \text{ observation.}$$

→ when classes are given,

$$m = L + \frac{\left(\frac{n}{2}\right) - cf}{f} \times C.$$

where, L = Lower Limit of median class:
(A class having $(n/2)^{th}$ observation is called median class.)
f = frequency of the median class.
cf = cf of previous class.
C = class interval.

Partition Values.

	For ungrouped data.	for grouped
Quartiles.	$Q_j = j\left(\frac{n+1}{4}\right)^{th}$ obs^n where j = 1, 2, 3.	$Q_j = L + \frac{j\left(\frac{n}{4}\right) - Cf}{f} \times C.$
Deciles.	$D_j = j\left(\frac{n+1}{10}\right)^{th}$ obs^n where j = 1, 2, 3, ..., 9.	$D_j = L + \frac{j\left(\frac{n}{10}\right) - Cf}{f} \times C.$
percentiles	$P_j = j\left(\frac{n+1}{100}\right)^{th}$ obs^n where, j = 1, 2, 3, ... 99.	$P_j = L + \frac{j\left(\frac{n}{100}\right) - Cf}{f} \times C.$

[illegible]
[illegible]
[illegible]

વિષમતા - Skewness.
સંમિત - Symmetrical.

Mode :- Z.

→ The most frequently occured observation of the data is known as mode of the data.

→ For ungroup data, mode can be obtain just by an observation.

→ When classis are given,

$$Z = L + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times C.$$

→ In the following situations we can not use the above formula :-

(1) when largest freqency is at
- more then one classes.
- first class.
- Last class.

(2) when class intervals are unequal

$$\bar{x} - Z = 3(\bar{x} - M)$$

or

$$Z = 3M - 2\bar{x}$$

Some Imp Points Regarding :- Mode

→ Mode is not unique. It is ill define measure.

→ mode is unstable measure.

→ mode is affected by change of origin & change of scale.
i.e; If $y = a + bx$, then $Z_y = a + b \cdot Z_x$.

Geometric mean :-

$$G.M = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \ldots \cdot x_n}$$

or

$$\left[x_1 \cdot x_2 \cdot x_3 \cdot \ldots \cdot x_n\right]^{1/n}$$

$$G.M. = \text{Antilog}\left[\frac{\Sigma \log x_i}{n}\right]$$

when frequency (f_i) are given,

$$G.M. = \left[x_1^{f_1} \cdot x_2^{f_2} \cdot x_3^{f_3} \cdot \ldots \cdot x_k^{f_k}\right]^{1/N}$$

where, $n = \Sigma f_i$

Combined G.M. ;

$$G.M_C = \left[G_1^{n_1} \times G_2^{n_2}\right]^{\frac{1}{n_1+n_2}}$$

Harmonic mean :-

→ H.M. has very limited use, generally it is use to find Average speed, Average price per unit etc.

$$H.M. = \frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\ldots+\frac{1}{x_n}} = \frac{n}{\Sigma\left(\frac{1}{x_i}\right)}$$

$\boxed{\frac{n}{\Sigma}}$ ને A.M. કહે છે.

→ H.M. is ~~[illegible]~~ Reciprocal of A.M of reciprocal of

$\boxed{1/x}$

the observation.

→ when frequency (f_i) are given,

$$H.M. = \frac{n}{\frac{f_1}{x_1} + \frac{f_2}{x_2} + \frac{f_3}{x_3} \dots + \frac{f_k}{x_k}} = \frac{n}{\Sigma\left(\frac{f_i}{x_i}\right)}$$

where, $n = \Sigma f_i$.

* Combined H.M. :- [નાના અવલોકનોની કિંમતની અસર વધુ થાય (∵ વ્યસ્ત છે)]

$$H.M_c = \frac{n_1 + n_2}{\frac{n_1}{H_1} + \frac{n_2}{H_2}}$$

Note :-

1. In General,

$$AM \geq GM \geq HM.$$

2. If observations are disting (different)

$$\therefore AM > GM > HM.$$

3. If all observations are equal,

$$AM = GM = HM.$$

4. For Two possitive numbers,

$$GM = \sqrt{AM \times HM}$$

or

$$(GM)^2 = AM \times HM$$

GM is a Geometric mean of AM and HM.

$\rightarrow$ $\boxed{\bar{x} = \frac{2n+1}{3}}$;

$\rightarrow$ જ્યારે પ્રથમ n પ્રાકૃતિક સંખ્યા x અને f તેને આપેલ હોય ત્યારે ઉપરનું સુત્ર મુકવું.

where data

x	f
1	1
2	2
3	3
$\vdots$	$\vdots$
n	n

then,

$$Am = \frac{2n+1}{3}$$

PAGE NO.
DATE 17 / 5 / 23

Ch:- 3. Measures of Dispersion.

☆ Range (R) :- (વિસ્તાર)

$$R = H - L$$ [Absolute measure
નિરપેક્ષ માપ]

Coefficient of Range (વિસ્તારાંક) $= \dfrac{H-L}{H+L} \times 100$ [Relative measure
સાપેક્ષ માપ]

☆ Quartile Deviation (Q.D.) :- (ચતુર્થક વિચલન)

$$Q.D. = \frac{Q_3 - Q_1}{2}$$ [Absolute measure]

Coefficient of Q.D. $= \dfrac{Q.D.}{M} \times 100$

Or

$= \dfrac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$ [when distribution is symmetrical]

☆ Mean Deviation (M.D.) :- (સરેરાશ વિચલન)

M.D. from A.M. $= \dfrac{\Sigma|x-\bar{x}|}{n}$ or $\dfrac{\Sigma f \cdot |x-\bar{x}|}{n}$

M.D. from M $= \dfrac{\Sigma|x-M|}{n}$ or $\dfrac{\Sigma f \cdot |x-M|}{n}$

M.D. from Z $= \dfrac{\Sigma|x-Z|}{n}$ or $\dfrac{\Sigma f \cdot |x-Z|}{n}$

Absolute measures

Coeff. of MD from A.M. $= \dfrac{MD}{\bar{x}} \times 100$

$$\text{Coeff}^{t} \text{ of MD}_{\text{from } M} = \frac{MD}{M} \times 100$$

$$\text{Coeff}^{t} \text{ of MD}_{\text{from } Z} = \frac{MD}{Z} \times 100$$

Relative measure.

Standard Deviation (S or S.D. or σ)

$$S = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}} = \sqrt{\frac{\Sigma x^2}{n} - (\bar{x})^2}$$

or

$$S = \sqrt{\frac{\Sigma f(x-\bar{x})^2}{n}} = \sqrt{\frac{\Sigma f x^2}{n} - (\bar{x})^2}$$

or

$$S = \sqrt{\frac{\Sigma f_i d_i^2}{n} - \left(\frac{\Sigma f_i d_i}{n}\right)^2} \times C$$

where, $d_i = \dfrac{x_i - A}{C}$

Absolute measure

And, variance $= (S.D.)^2$

Relative measure of S.D. is called coefficient of Variation (C.V.)

$$C.V. = \frac{S}{\bar{x}} \times 100$$

More C.V. $\Leftrightarrow$ Less consistency

DATE / /

Some important points for all four measures of Dispersion :—

1. If all observations are equal, all measures of Dispersion will be zero (0).

2. All measures of Dispersion are always non-negative (~~-~~). (0, +).

3. All measures of Dispersion are independent of the change of origin but affected by the change of scale.

$$\text{If, } y = a + b.x$$

then,

(1) $R_y = |b| \cdot R_x$
(2) $QD_y = |b| \cdot QD_x$
(3) $MD_y = |b| \cdot MD_x$
(4) $SD_y = |b| \cdot SD_x$

$\therefore$ Variance of $y = b^2 \cdot$ variance of x.

Short cut :- 4. If $Ax + By + C = 0$

then Dispersion measure of $y = \left|\frac{A}{B}\right| \cdot$ (Dispersion measure of x)

i.e. Measure of $y = \left|\frac{\text{Coeff}^t \text{ of } x}{\text{Coeff}^t \text{ of } y}\right| \cdot$ (measure of x).

Linear correlation & Linear Regression.

☆ Correlation :-

☆ Karl - Pearsons Product moment method :—

→ r = correlation coefficient.

$$\therefore r = \frac{cov(x, y)}{s_x \cdot s_y}$$

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

$$= \frac{\Sigma xy - n\bar{x}\bar{y}}{n}$$

→ Standard Deviation = S.D. of $x = s_x = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}$

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

→ S.D. of $y = s_y = \sqrt{\frac{\Sigma(y-\bar{y})^2}{n}} = \sqrt{\frac{\Sigma y^2}{n} - (\bar{y})^2}$

(1) $$r = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma(x-\bar{x})^2} \cdot \sqrt{\Sigma(y-\bar{y})^2}}$$

(2) $$r = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{n\Sigma x^2 - (\Sigma x)^2} \cdot \sqrt{n\Sigma y^2 - (\Sigma y)^2}}$$

(3) $$r = \frac{n\Sigma uv - (\Sigma u)(\Sigma v)}{\sqrt{n\Sigma u^2 - (\Sigma u)^2} \cdot \sqrt{n\Sigma v^2 - (\Sigma v)^2}}$$

where, $u = x - A$ or $\frac{x-A}{c_x}$

$v = y - B$ or $\frac{y-B}{c_y}$.

(4) $r = \dfrac{\Sigma(x-\bar{x})(y-\bar{y})}{n \cdot S_x \cdot S_y}$

(5) $r = \dfrac{\Sigma xy - n\bar{x}\bar{y}}{n \cdot S_x \cdot S_y}$

★ Properties of r :-

1) $-1 \leq r \leq 1$.

2) r is a unit free measure. (r is a pure number and Hence, it is a relative measure.)

3) $r(x,y) = r(y,x)$.

4) r is independent of the change of origin & the change of scale.

5) (i) $r(-x,y) = -r(x,y)$.
(ii) $r(x,-y) = -r(x,y)$
(iii) $r(-x,-y) = r(x,y)$.

★ → 1) If $r_{xy} = 0.8$,
(a) $r(3x+2, 5y-7) = \underline{0.8}$

(b) $r\left(\dfrac{5x-7}{10}, \dfrac{2y+10}{10}\right) = \underline{0.8}$

(c) $r\left(\dfrac{3-5x}{2}, \dfrac{10+3y}{7}\right) = \underline{-0.8}$

(d) $r\left(\dfrac{5x}{3}, \dfrac{5-y}{2}\right) = \underline{-0.8}$

(e) $r\left(\frac{10-3x}{2}, \frac{5-3y}{2}\right) = \underline{0.8}$

★ $r(x,y) = -0.6.$

$u = 3x + 5,\ v = -2y + 7$ find $r(u,v)$.

$\therefore\ r(u,v) = \boxed{0.6}$

★ $r_{xy} = 0.5$, find r_{uv}.

$3u - 2x = 11$

$\therefore\ 3u = 11 + 2x$

$\therefore\ u = \frac{11 + 2x}{3}$

$5v + 7y - 8 = 0$

$\therefore\ 5v = 8 - 7y$

$\therefore\ v = \frac{8 - 7y}{5}$

$\therefore\ r_{uv} = -0.5$

★ Note :—

* If $y = a + bx\ (b > 0)$ then $r = +1$.
* If $y = a + bx\ (b < 0)$ then $r = -1$.
* ~~If~~

Eg :- (i) $y = 10 + 3x$ then $r = \underline{+1}$.

(ii) $y = -7 + 1.5x$ then $r = \underline{+1}$

(iii) $2x + 3y = 15$ then $r = \underline{-1}$.

$\therefore\ 3y = 15 - 2x$

(iv) $-y = 5x + 2$ then $r = \underline{-1}$.

★ Probbable Error [P.E.] :—

→ $$P.E. = \frac{0.6745\,(1 - r^2)}{\sqrt{n}}$$

→ (1) Probbable limits = $(r - P.E.)$ to $(r + P.E.)$

→ (2) Significance of Correlation in Population

પ્રથમ n પ્રાકૃતિક સંખ્યાઓનો સરવાળો $= \frac{n(n+1)}{2}$

(2) C.f. = correction factor

be checked.

→ * If $\left(\frac{r}{P.E.}\right) < 1$ → Not significant.

* If $\left(\frac{r}{P.E.}\right) > 6$ → Significant.

* otherwise, sample is insufficient.

☆ Coefficient of Determination (R^2) :—

→ If $r = 0.8$, then $R^2 = r^2 = 0.64$.

$\therefore$ 64% = Explained variation.

(100 − 64) = 36% = Unexplained variation.

$$\therefore R^2 = \frac{\text{Explained Variation}}{\text{Total variation}}$$

☆ Coefficient of non Determination :—

→ $= 1 - R^2 = \frac{\text{Unexplained Variation}}{\text{Total Variation}}$

☆ Spearman's Rank Correlation method :—

→ (1) when all observations are different,

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

(2) when some observations are repeated,

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

where, $d = R_x - R_y$ = Difference of Ranks.

Σd^2 = sum of squares of difference in the Ranks.

$$C.f. = \Sigma\left(\frac{m^3 - m}{12}\right)$$

m = number of times a particular observation is repeated.

$\Rightarrow$ implies/તેથી.

$\left.\begin{array}{l} \Sigma = + \\ E = - \end{array}\right\}$ E ઓપરેટર Error

☆ Note :-

(1) $\Sigma d = \Sigma(R_x - R_y) = \boxed{0}$

(2) If for each pair of observation $R_x = R_y$, then each $d = 0$.

$$\therefore \Sigma d^2 = 0 \Rightarrow \boxed{r = 1}$$

(Perfect aggreement between two Judges.)

(3) If the Ranks of x and y are exact in ~~Revers~~ Reverse orders then,

$$\therefore \boxed{r = -1}$$

(Perfect disagreement between two Judges.)

Notes :-

(i) If $c = \underline{m}$ then $r = \underline{+1}$

(ii) If $c = \underline{m/2}$ then $r = \underline{0}$.

(iii) If $c = \underline{0}$ then $r = \underline{-1}$.

☆ When x & y are co-related ;

1. $V(x+y) = V(x) + V(y) + 2\cdot Cov(x,y)$
2. $V(x-y) = V(x) + V(y) - 2\cdot Cov(x,y)$.

☆ When x & y are unco-related (Independent) ;

1. $V(x+y) = V(x) + V(y)$
2. $V(x-y) = V(x) + V(y)$

☆ Linear Regression :-

Regression line of y on x	Regression line of x on y
$y = a + b_{yx} \cdot x$	$x = c + b_{xy} \cdot y$
$b_{yx} = \dfrac{\Sigma (x-\bar{x})(y-\bar{y})}{\Sigma (x-\bar{x})^2}$	$b_{xy} = \dfrac{\Sigma (x-\bar{x})(y-\bar{y})}{\Sigma (y-\bar{y})^2}$
$= \dfrac{n\Sigma xy - (\Sigma x)(\Sigma y)}{n\Sigma x^2 - (\Sigma x)^2}$	$= \dfrac{n\Sigma xy - (\Sigma x)(\Sigma y)}{n\Sigma y^2 - (\Sigma y)^2}$

$$= \frac{n\Sigma uv - (\Sigma u)(\Sigma v)}{n\Sigma u^2 - (\Sigma u)^2}$$

Here, $u = x - A$ &

$v = y - B$.

$$= \frac{n\Sigma uv - (\Sigma u)(\Sigma v)}{n\Sigma v^2 - (\Sigma v)^2}$$

Here, $u = x - A$ &

$v = y - B$.

$$= \frac{n\Sigma uv - (\Sigma u)(\Sigma v)}{n\Sigma u^2 - (\Sigma u)^2} \times \frac{cy}{cx}$$

Here, $u = \frac{x - A}{cx}$, $v = \frac{y - B}{cy}$

$$= \frac{n\Sigma uv - (\Sigma u)(\Sigma v)}{n\Sigma v^2 - (\Sigma v)^2} \times \frac{cx}{cy}$$

Short sum :-

$$b_{yx} = r \cdot \frac{Sy}{Sx}$$

$$b_{xy} = r \cdot \frac{Sx}{Sy}$$

$$= \frac{\mathrm{cov}(x, y)}{Sx^2}$$

$$= \frac{\mathrm{cov}(x, y)}{Sy^2}$$

$$a = \bar{y} - b_{yx} \cdot \bar{x}$$

$$c = \bar{x} - b_{xy} \cdot \bar{y}$$

☆ Notes :-

(1) * Another form of y on x line :-

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

* Another form of x on y line :-

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

(2) The constants a & c of two regression lines are intercepts of the lines.

(3) Two regression coefficients are also known as slopes of the regression lines.

Meaning of regression coefficient :-

The probable change in the value of dependent variable due to one unit change in the value of independent variable is called regression coefficient.

$\therefore$ b_{yx} = The change in y due to 1 unit change in x.
b_{xy} = The change in x due to 1 unit change in y.

Properties of Regression coefficients :-

(1). $b_{yx} \cdot b_{xy} = r^2$
Or
$\therefore r = \pm\sqrt{b_{yx} \cdot b_{xy}}$

[i.e. r is Geometric mean of two regression coefficient

(2) The Signs of r, b_{yx} & b_{xy} are always ~~sa~~ the same

(3) If one of the regression coefficient is >1 then another regression coefficient must be <1, So that their product remain from 0 to 1.

(4) If there is perfect correlation ($r = \pm 1$). then Two regression coefficients are reciprocal (ulat) of each other.

(5). Regression coefficients are independent of the change of origin but dependent on (effected by) the change of Scale.

{ i.e. ; * If $u = x - A$ & $v = y - B$.
$b_{yx} = b$ ...

If $u = \frac{x-A}{cx}$ & $v = \frac{y-B}{cy}$

$b_{yx} = b_{vu} \times \frac{cy}{cx}$ & $b_{xy} = b_{uv} \times \frac{cx}{cy}$

Some Important Points for Two Regression lines :-

(1) If there is perfect corelation ($r = \pm 1$), then two regression lines are identical / coinside.

(2) Smaller the angle between two regression lines ~~called~~, higher is the degree of corelation.

(3) If $r = 0$ then two regression lines are perpendicular (परस्पर लंब - 90°).

(4) Two regression lines always intersect at a point $(\bar{x}, \bar{y})$.

Notes ! — imp.

→ When regression equations are in the form of $Ax + By + C = 0$.

→ from y on x line : $b_{yx} = -\frac{\text{Coefficient of } x}{\text{Coefficient of } y}$.

→ from x on y line : $b_{xy} = -\frac{\text{coefficient of } y}{\text{coefficient of } x}$.

6. → $2x + 3y + 50 = 0$.

→ (b). $\therefore b_{yx} = -\frac{2}{3} = -\frac{\text{coefficient of } x}{\text{coefficient of } y}$.

7. → $4y - 5x = 15$ (y on x)

$$b_{yx} = \frac{-(-5)}{4} = 1.25$$

$r = 0.45$, $b_{xy} = ?$

C.y. - Current year
·P.y. - Past year
B.y. - Base years

Fixed base Index no. & chain base Index no.

1. Fixed base Index number of C.y. $= \frac{\text{Price of C.y.}}{\text{" of B.y.}} \times 100 = \frac{P_1}{P_0} \times 100 = I = P_1$

2. Link ~~of~~ Relative of C.y. $= \frac{\text{Price of C.y.}}{\text{Price of p.y.}} \times 100 = \frac{P_n}{P_{n-1}} \times 100$

3. C.B.I.N. of C.y. $= \frac{\text{L.R. of C.y.} \times \text{CBIN of P.y.}}{100}$

4. C.B.I.N. of C.y. $= \frac{\text{FBIN of C.y.}}{\text{FBIN of P.y.}} \times 100.$

5. F.B.IN of C.y. $= \frac{\text{CBIN of C.y.} \times \text{FBIN of P.y.}}{100}$

18. → $= \frac{\text{FBIN of C.y.}}{\text{FBIN of P.y.}} \times 100$

19. → $\frac{\text{CBIN of Cy} \times \text{FBIN of P.y.}}{100}$

20. → CB $= \frac{\text{L.R. of C.y.} \times \text{CBIN of P.y.}}{100}$

$$K \,\&\, P = B \pm \sqrt{B^2 - f^2}$$

Quantity index શોધવા માટે price index માં P ના બદલે q લેવું.

Value = price × quantity

Value index $= \frac{\sum p_1 q_1}{\sum p_0 q_0} \times 100$

$= \frac{V_1}{V_0} \times 100$

Ch:-6. Probability And Random Variable

★ Results :-

1. $P(A) = \frac{m}{n}$

2. $P(\phi) = 0$ and $P(S) = 1$.

3. $0 \leq P(A) \leq 1$.

4. $P(A') = 1 - P(A)$ i.e. $P(A) + P(A') = 1$

5. for two events A and B,
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
[Addition Rule of Probability for two events]
→ if A and B are mutually exclusive events $(A \cap B = \phi)$ then,
$$P(A \cup B) = P(A) + P(B).$$
→ if A and B are mutually exclusive and exhaustive events $(A \cap B = \phi$ & $A \cup B = S)$ then,
$$P(A) + P(B) = 1.$$

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

7. $P(A \cup B) = 1 - P(A' \cap B')$
i.e. $P(A' \cap B') = 1 - P(A \cup B)$.

8. $P(A' \cup B') = 1 - P(A \cap B)$.
i.e. $P(A \cap B) = 1 - P(A' \cup B')$.

PAGE NO.
DATE / /

9. $P(A-B) = P(A \cap B') = P(A) - P(A \cap B)$

& $P(B-A) = P(B \cap A') = P(B) - P(A \cap B)$.

10. $P(A \cap B') = 1 - P(A' \cup B)$

i.e. $P(A' \cup B) = 1 - P(A \cap B')$

→ $P(A \cup B') = 1 - P(A' \cap B)$

11. For 3 events A, B & C ;

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

[Additional Rule of Prob. for 3 events]

→ if A, B, C are mutually exclusive and exhaustive events then,

$$P(A) + P(B) + P(C) = 1.$$

12. $P(A \cup B \cup C) = 1 - P(A' \cap B' \cap C')$
i.e. $P(A' \cap B' \cap C') = 1 - P(A \cup B \cup C)$

13. If A and B are two independent events then,
(i) $P(A \cap B) = P(A) \times P(B)$.
(ii) $P(A \cap B') = P(A) \cdot P(B')$.
(iii) $P(A' \cap B) = P(A') \cdot P(B)$.
(iv) $P(A' \cap B') = P(A') \cdot P(B')$

i.e. A & B', A' & B and A' & B' are also indep.

1. <u>Conditional Probability</u> :-

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \; ; \; P(B) \neq 0.$$

$$P(B/A) = \frac{P(A \cap B)}{P(A)} \; ; \; P(A) \neq 0.$$

i.e. $P(A \cap B) = P(A) \cdot P(B/A)$.
& $P(A \cap B) = P(B) \cdot P(A/B)$.

multiplication Rule of Prob. for two events (composite event or compound event)

Non leap (Normal)	leap
365 days	366 days.
52 weeks	52 weeks.
$\times$ 7	$\times$ 7
364 days	364 days
1 Extra	2 extra.
{mon, T, W, Th, F, S, Sun}	{(m, T), (T, W), (W, Th), (Th, F), (F, S), (S, Sun), (Sun, m)}

Or

એક જગ્યા બધાંને છોડેલો
3 person છે તો 4,

Short cut :-

② $\frac{3}{2!}$

$$\frac{4 \times 2 \times 6!}{8!}$$

Ch:-7 Random variable

Definition of Expectation :-

Suppose $x_1, x_2, x_3, \ldots x_n$ are n different values of random variable (r.v.) x and $p_1, p_2, p_3, \ldots, p_n$ are their respective probabilities then Expectation of x is the sum of product of different values of x with their corresponding probabilities. and it is denoted by $E(x)$.

$$\therefore E(x) = x_1 \cdot p_1 + x_2 \cdot p_2 + x_3 \cdot p_3 + \ldots + x_n \cdot p_n$$

$$\therefore E(x) = \Sigma x_i \cdot p_i$$

It is also known as mean or Average of x. and it is also denoted by μ.

$$\therefore \text{mean} = \text{Average} = \mu = E(x) = \Sigma x_i \cdot p_i$$

Properties of Expectation :-

1. $E(k) = k$. eg:- $E(5) = 5$, $E(-8/3) = -8/3$.

2. $E(ax) = aE(x)$ eg:- $E(7x) = 7 \cdot E(x)$.
$$E\left(\frac{-9x}{5}\right) = -\frac{9}{5} E(x).$$

3. $E(ax \pm b) = a E(x) \pm b$. eg:- $E(3x+5) = 3E(x)+5$
$$E(y) = 7E(x) + 2.$$
$$E(2x - 7) = 2E(x) - 7.$$

4. For two variables x & y,

$$E(ax \pm by) = aE(x) \pm bE(y).$$

Σ σ ← સિગ્મા
Δ δ ← ડેલ્ટા

Function → વિધેય

PAGE NO.
DATE / /

Eg :- $E(3x + 5y) = 3E(x) + 5E(y)$.

5. For two independent variables x & y,
$$E(x \cdot y) = E(x) \cdot E(y).$$

6. $E(x - \mu) = 0$. $\left[\because \Sigma(x_i - \bar{x}) = 0\right]$

7. If $g(x)$ is a function of x then,
$$E[g(x)] = \Sigma\, g(x) \cdot p(x).$$

Eg :- $E(x^2) = \Sigma x^2 \cdot p(x) = \Sigma x_i^2 \cdot p_i$

$E(x^3) = \Sigma x^3 \cdot p(x) = \Sigma x_i^3 \cdot p_i$

Definition of variance :-

Variance is the average of squares of deviation taken from mean of a r.v. x and it is denoted by $V(x)$ or σ^2.

i.e.
$$V(x) = E(x - \mu)^2 = E[x - E(x)]^2$$

or $V(x) = E(x^2) - [E(x)]^2 = E(x^2) - \mu^2$

$\boxed{\mu = E(x)}$

Properties of variance :-

1. $V(k) = 0$ Eg :- $V(5) = 0$

2. $V(ax) = a^2 \cdot V(x)$ Eg :- $E(5x) = 5 \cdot E(x)$
$V(5x) = 25 \cdot V(x)$
$V(-3x) = 9\, V(x)$

3. $V(ax \pm b) = a^2 \cdot V(x)$ Eg :- $E(3x - 5) = 3E(x) - 5$
$V(3x - 5) = 9\, V(x)$
$V(7 - 9x) = 81\, V(x)$

PAGE NO.
DATE / /

4. For two independent variables x & y,

$$V(ax \pm by) = a^2 \cdot V(x) + b^2 \cdot V(y).$$

Eg :- $V(3x-5y) = 9\,V(x) + 25\,V(y)$.
$V(2x+7y) = 4\,V(x) + 49\,V(y)$.

5. Positive square root of variance is known as Standard deviation.
i.e. $S.D. = \sqrt{V(x)}$.

→ If $y = a \pm bx$ then
SD of $y = |b| \cdot$ SD of x.

Eg :- If SD of $x = 10$ then find SD of $7x-5$.
$\therefore$ SD of $7x-5 = 7 \times$ SD of x
$= 7 \times 10$
$= 70$.

Eg :- If SD of $x = 10$ then find SD of $5-7x$.
$\therefore$ SD of $5-7x = |-7| \cdot$ SD of x.
$= 7 \times 10$
$= 70$

Ch:- 8. Theoretical Distribution

Binomial Distribution :-

$$P(x) = nC_x p^x q^{n-x} \quad ; x = 0, 1, 2, 3, \dots, n$$
$$0 < p, q < 1.$$
$$p + q = 1.$$

where, n = no. of trials.
x = no. of successes.
p = Prob. of successes.
q = Prob. of failure.

Binomial Distribution is also denoted by $\boxed{X \sim B(n, p)}$

[X follows B.D. with parameters n & p].

Total Probability = 1.

$\therefore P(0) + P(1) + P(2) + \dots + P(n) = 1$

$$\therefore \sum_{x=0}^{n} P(x) = 1.$$

$$\therefore \sum_{x=0}^{n} nC_x p^x q^{n-x} = 1.$$

$$\therefore (q + p)^n = 1.$$

i.e. $\boxed{\sum_{x=0}^{n} nC_x p^x q^{n-x} = (q + p)^n}$

Eg:- $n=2$ $\quad \sum_{x=0}^{2} nC_x p^x q^{n-x} = (q+p)^2$

$$\therefore 2C_0 p^0 q^{2-0} + 2C_1 p^1 q^{2-1} + 2C_2 p^2 q^{2-2} = q^2 + 2pq + p^2$$

Skewness - বিষমতা
Skewed - বিষম

PAGE NO.
DATE / /

$\therefore q^2 + 2pq + p^2 = q^2 + 2qp + p^2.$

Properties of Binomial Distribution :

1. This is the distribution of discrete random variable.

2. n and p are the parameters.
   [i.e. It is bi-parametric distribution]

3. Mean $= np$

4. variance $= npq$ $\therefore$ SD $= \sqrt{npq}$

5. $q = \frac{v}{m}$, $p = 1 - q$, $n = \frac{m}{p}$

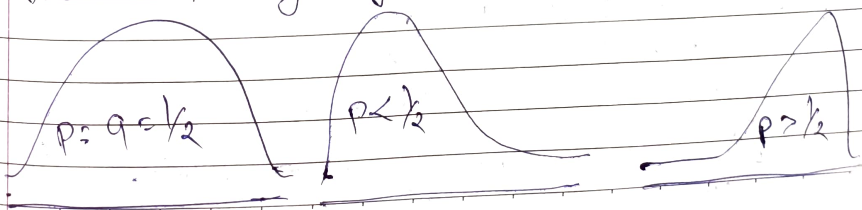
6. Mean is always greater than variance. [$np > npq$.]

7. Maximum value of variance $= \frac{n}{4}$

[when $p = q = 1/2$ then variance is max.]

8. If $n = 1$ then it is known as Bernoulli distribution.

9. If $p = q = 1/2$ then it is symmetrical dist., if $p < 1/2$ then it is positively skewed dist. and if $p > 1/2$ then it is negatively skewed dist.



integer - ycoi's
fraction - aiyoi's

10. If $(n+1)\cdot p$ is integer then it is bimodal dist. [i.e. Two modes] And Two modes are $(n+1)\cdot p$ & $(n+1)\cdot p-1$.

→ If $(n+1)\cdot p$ is in fraction then it is uni-modal dist. And the mode is integer part of $(n+1)\cdot p$.

→ Eg If $X \sim B(17, 1/3)$ find mean & mode.

Here $n=17$, $p=1/3$

$\therefore$ mean $= np = 17 \times \frac{1}{3} = 5.67$

Now, $(n+1)\cdot p = (18)(1/3) = 6$

$\therefore$ Two modes are 6 & 5

→ If $x \sim B(17, 1/4)$ find variance & mode.

Here $n=17$, $p=1/4$, $q=3/4$.

Variance $= npq = 17 \times \frac{1}{4} \times \frac{3}{4} = \frac{51}{16}$

Now,

$(n+1)\cdot p = 18 \times \frac{1}{4} = 4.5$

$\therefore$ mode is 4.

11. If $X \sim B(n_1, p)$ and
$Y \sim B(n_2, p)$ then
$X+Y \sim B(n_1+n_2, p)$.
[i.e. Sum of two independent binomial variate is also binomial variate]

PAGE NO.
DATE / /

11. If $p = q = \frac{1}{2}$ then $P(x) = \frac{nC_x}{2^n}$.

eg. if $x \sim B(6, 0.8)$ and
$y \sim B(9, 0.8)$
then $x + y \sim B(15, 0.8)$.

P. same હોવા જોઈએ.

12. If n is very large $(n \to \infty)$ and p is very small $(p \to 0)$ then B.D. follows to Poisson dist.

13. If n is very large $(n \to \infty)$ and p is not very small $(p \nrightarrow 0)$ then B.D. follows to normal dist.

Poisson distribution

→ $$P(x) = \frac{e^{-m} \cdot m^{x}}{x!} \; ; \; x = 0, 1, 2, 3, \ldots$$

m = mean (np)
($m > 0$)
$e = 2.7183$
x = no. of successes.

→ It is denoted by $X \sim P(m)$.

Total prob. = 1.
i.e. $P(0) + P(1) + P(2) + \ldots = 1$.
i.e. $$\sum_{x=0}^{\infty} P(x) = 1$$

$$\therefore \sum_{x=0}^{\infty} \frac{e^{-m} \cdot m^{x}}{x!} = 1$$

PAGE NO.
DATE / /

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

$\rightarrow \boxed{P(0)} = \frac{e^{-m} \cdot m^0}{0!} = \frac{e^{-m} \cdot 1}{1} = \boxed{e^{-m}}$

$\therefore \boxed{P(0) = e^{-m}}$

$\rightarrow P(1) = \frac{e^{-m} \cdot m^1}{1!} = e^{-m} \cdot m$ $\therefore \boxed{P(1) = e^{-m} \cdot m}$

i.e. $\boxed{P(1) = P(0) \times m}$

$\rightarrow P(2) = \frac{e^{-m} \cdot m^2}{2!} = \frac{e^{-m} \cdot m^2}{2}$ i.e. $\boxed{P(2) = P(1) \times \frac{m}{2}}$

$\rightarrow P(3) = \frac{e^{-m} \cdot m^3}{3!} = \frac{e^{-m} \cdot m^2}{2} \times \frac{m}{3}$

i.e. $\boxed{P(3) = P(2) \times \frac{m}{3}}$

Similarly $P(4) = P(3) \times \frac{m}{4}$

⋮

$P(9) = P(8) \times \frac{m}{9}$

⋮

$$\boxed{P(k+1) = P(k) \times \frac{m}{k+1}}$$

DATE / /

Now,

$p(0) + p(1) + p(2) + p(3) + \ldots$

$= e^{-m} + e^{-m} \cdot m + \frac{e^{-m} \cdot m^2}{2} + \frac{e^{-m} \cdot m^3}{6} + \frac{e^{-m} \cdot m^4}{24}$

$= e^{-m} \left[1 + m + \frac{m^2}{2} + \frac{m^3}{6} + \frac{m^4}{24} + \frac{m^5}{120} + \ldots \right]$

$$\boxed{\frac{\log m}{\log n} = \log_n m}$$

DATE / /

Normal distribution

If a continuous random variable x possesses following probability density function (pdf) then x is known as normal variate and its dist. is known as normal dist.

$$\underset{pmf}{P(x)} = \underset{pdf}{f(x)} = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty,\ -\infty < \mu < \infty,\ \sigma > 0$$

Or

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

where, x = Normal variable

μ = mean

σ = standard deviation

Parameters

$e = 2.7183$

$\pi = \frac{22}{7} = 3.1416$

it is denoted by $x \sim N(\mu, \sigma^2)$

[x follows normal dist. with mean μ & variance σ^2.]

eg:- $x \sim N(50, 25)$ i.e. $\mu = 50$, $\sigma^2 = 25$

$\therefore \sigma = 5$

integration (m key) $\int$ Sum

PAGE NO.

DATE / /

Total Prob. = 1 $\quad \Sigma \quad \int$

discrete continuous

continuous

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

$$\therefore \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1.$$

Standard Normal dist :-

If we take $z = \frac{x-\mu}{\sigma}$ then z is known as Standard normal variate and its dist. is known as standard normal dist.

$$f(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}z^2} = \frac{1}{\sqrt{2\pi}} \cdot e^{-z^2/2}$$

$$; -\infty < z < \infty$$

where, z = standard normal variate

mean = 0

SD = 1 $\qquad$ Here, (mean & varian

variance = 1.

$e = 2.7183$

$\pi = \frac{22}{7} = 3.1416$

PAGE NO.
DATE / /

It is denoted by $Z \sim N(0,1)$

[Z follows normal dist. with mean 0 & Variance 1]

[for standard normal dist., there is no Parameter]

Total Prob. = 1

$$\therefore \int_{-\infty}^{+\infty} f(z)\,dz = 1$$

$$\therefore \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}z^2}\,dz = 1$$

* Properties :-

1. it is a dist. of continuous random variable.

2. Here μ & σ are the Parameters.
[for standard normal dist, there is no Parameters]

3. mean = μ & SD = σ
[for Z, mean = 0 & SD = 1]

4. It is always Symmetrical.

5. Skewness is zero.

6. Mean = Medium = Mode.

DATE / /

7. $Q_1 = \mu - 0.675\sigma$
& $Q_3 = \mu + 0.675\sigma$

8. $Q_3 + Q_1 = 2\mu$
$\therefore \mu = \frac{Q_3 + Q_1}{2}$
$\therefore m = \frac{Q_3 + Q_1}{2}$

9. $QD = \frac{2}{3}\sigma = 0.675\sigma$
$QD = \frac{Q_3 - Q_1}{2}$

10. $MD = \frac{4}{5}\sigma = 0.8\sigma$.
$MD = \frac{\Sigma |x_i - \bar{x}|}{n}$

11. $QD : MD : SD = 10 : 12 : 15$

12. $\mu - \sigma$ and $\mu + \sigma$ are known as points of inflexion.
[for z, -1 & 1 are inflexion]

13. $\phi(a)$ = Cumulative dist. function for z.
$\phi(a) = P(z \le a) = P(-\infty \le z \le a)$

Eg. Find $\phi(1)$.
$\phi(1) = P(z \le 1)$
$= \cancel{0.8413} = 0.5000 + 0.3413$
$= 0.8413$

- $\mu \pm \sigma$ (i.e. $-1 < z < 1$) $\doteq 68.26\%$
- $\mu \pm 2\sigma$ (i.e. $-2 < z < 2$) $= 95.45\%$
- $\mu \pm 3\sigma$ (i.e. $-3 < z < 3$) $\doteq 99.73\%$
- $\mu \pm 1.96\sigma$ (i.e. $-1.96 < z < 1.96$) $= 95\%$
- $\mu \pm 2.58\sigma$ (i.e. $-2.58 < z < 2.58$) $= 99\%$

14. $F(a)$ = Cumulative dist. function for x.

$$F(a) = P(x \le a) = P(-\infty \le x \le a).$$

Eg:- if $\mu = 100$ & $\sigma = 15$ then find $F(130)$.

$$F(130) = P(x \le 130)$$
$$= P\left(z \le \frac{130-100}{15}\right)$$
$$= P(z \le 2)$$
$$= 0.4772 + 0.5000$$
$$= 0.9772$$

15. if $x \sim N(\mu_1, \sigma_1^2)$ and
$y \sim N(\mu_2, \sigma_2^2)$
then $x + y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.
& $x - y \sim N$ $(\mu_1 - \mu_2, \sigma_1^2, \sigma_2^2)$

16. if $\mu = 50$ & $\sigma = 10$ find $P(x = 70) = 0$.

(Perfect value ની Prob. હંમેશાં 0 જ થાય.)