CT-032

$\Rightarrow$ Relative Frequency $= \dfrac{\text{Frequency}(f_i)}{\Sigma f_i} \times 100$

$\Rightarrow$ Cumulative Frequency = Total freq. at some point/interval. (Basically it's running total of frequency)

$\Rightarrow$ Arithmetic Mean $\Rightarrow$ $GD = \dfrac{\sum_{i=1}^{n} (f_i \times x_i)}{\sum_{i=1}^{n} f_i}$ | $UG = \dfrac{\sum_{i=1}^{n} (x_i)}{n}$

$\Rightarrow$ Geometric Mean $\Rightarrow$ $GD = AL\left(\dfrac{\sum_{i=1}^{n} f_i \times (\log x_i)}{\sum_{i=1}^{n}(f_i)}\right)$ | $UG = \text{Antilog}\left(\dfrac{\sum_{i=1}^{n}(\log x_i)}{n}\right)$

$\Rightarrow$ Harmonic Mean $\Rightarrow$ $GD = \dfrac{\sum_{i=1}^{n} f_i}{\sum_{i=1}^{n}\left(f_i \times \frac{1}{x_i}\right)}$ | $UG = \dfrac{n}{\sum_{i=1}^{n}\left(\frac{1}{x_i}\right)}$

$\Rightarrow$ Quartile $= i\left(\dfrac{n+1}{4}\right)^{th}$ value } ungrouped

$\Rightarrow$ Decile $= i\left(\dfrac{n+1}{10}\right)^{th}$ value , Percentile $= i\left(\dfrac{n+1}{100}\right)^{th}$ value.

Grouped data $\Rightarrow$ Quartile $= l + \dfrac{h}{f}\left(i\dfrac{\Sigma f}{4} - C.F\right)$

$P_i = l + \dfrac{h}{f}\left(i\dfrac{\Sigma f}{100} - C.F\right)$

$D_i = l + \dfrac{h}{f}\left(i\dfrac{\Sigma f}{10} - C.F\right)$

$M = l + \dfrac{h}{f}\left(i\dfrac{\Sigma f}{2} - C.F\right)$
(Median).

where;
h = width of class boundary
l = lowest bound.
$i\dfrac{\Sigma f}{4}$ = group highlighter.

-1-032

⇒ Mean of GD: Mean = $\frac{\sum_{i=1}^{n}(f_i \times x_i)}{\sum_{i=1}^{n} f_i}$ ⇒ UG = $\frac{u_1+u_2+u_3}{3}$

- Median of GD: $M = l + \frac{h}{f}\left(\frac{\sum f}{2} - C.F\right)$ ⇒ UG = Q_2, D_5, P_{50}

⇒ Mode of GD: Mode = $l + \left(\frac{f_m - f_1}{2f_m - f_1 - f_2} \times h\right)$ f_m = high freq, f_1 = uske upar wala, f_2 = uske niche wala.

UG ⇒ most appeard.

$Q_3 - \cancel{Q}_2 = Q_2 - Q_1$ Sy

∵ $MD(\bar{x}) = \frac{\sum|X_i - \bar{X}|}{n}$ or $\frac{\sum[f_i|x_i - \bar{x}|]}{\sum f_i}$.

∵ $\sum(X_i - \bar{X}) = 0$

∵ $C.O.M.D(\bar{x}) = \frac{MD(\bar{x})}{\bar{x}} \times 100$.

∵ $y = 2\cancel{x} + 3x \Rightarrow \bar{y} = 2 + 3\bar{x}$.

∵ $\sqrt{\beta_1} = \frac{\mu_3}{(\mu_2)^{3/2}}$ (skewness)

∵ Co-efficient of Range = $\left(\frac{X_m - X_0}{X_m + X_0}\right) \times 100$

∵ $\beta_2 = \frac{\mu_4}{(\mu_2)^2}$

∵ Co-efficient of Quartile Deviation = $\left(\frac{Q_3 - Q_1}{Q_3 + Q_1}\right) \times 100$

∵ Interquartile Range = $Q_3 - Q_1$

∵ Quartile Deviation = $\frac{Q_3 - Q_1}{2}$.

∵ σ^2 = Variance = $\frac{\sum(X_i - \bar{X})^2}{n}$ or $\frac{\sum[f_i(x_i - \bar{x})^2]}{\sum f_i}$

∵ Co-efficient of Variance = $\frac{\sigma}{\bar{X}} \times 100$.

∵ σ & σ^2 always +ve.

∵ σ & σ^2 of constant is zero.

∵ $V(ax) = a^2 V(x)$. ∵ $\sigma^2 = \mu^2$

Relation between μ & μ'

$\mu_1 = 0$

$\mu_2 = \mu_2' - \mu_1'\mu_1'$

$\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2\mu_1'\mu_1'^2$

$\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6\mu_1'^2 + \mu_2' - 3\mu_1'^3\mu_1'$

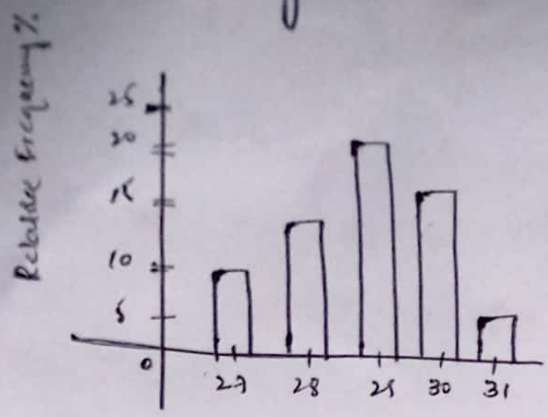
$\mu_r = \frac{\sum(X_i - \bar{X})^r}{n}$ OR $\mu_r = \frac{\sum[f_i(x_i - \bar{x})^r]}{\sum f_i}$

Moment About Origin:

$\mu_r' = \frac{\sum(X_i)^r}{n}$ or $\mu_r' = \frac{\sum f_i(X_i)^r}{f_i}$

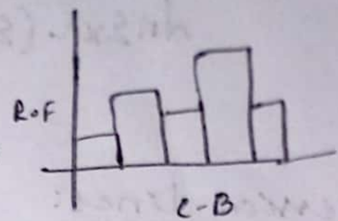
∵ $\mu_1 = 0$,

Relative Frequency:

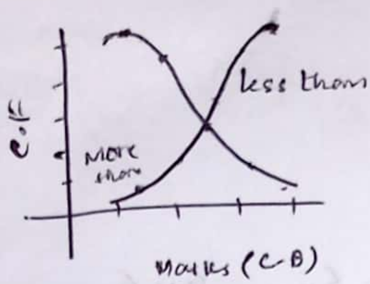


(Bar Graph)

We can also use Histogram:

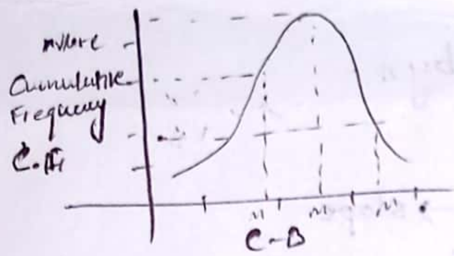


Cumulative Frequency:

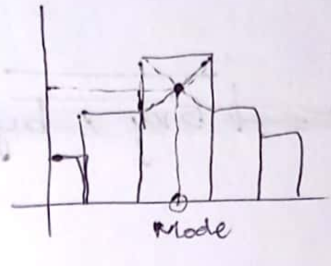


(O-give Curve)

Mean, Median, Mode



Mark the value of group highlighter on y-axis (C.F) Then find ~~median~~ median.



Correlation Co-efficient:

$$\gamma = \frac{n\sum xy - \sum x \sum y}{\sqrt{n\sum x^2 - (\sum x)^2}\sqrt{n\sum y^2 - (\sum y)^2}}$$

<u>Regression Lines:</u>

① $Y - \bar{Y} = b_{yx}(X - \bar{X})$

② $X - \bar{X} = b_{xy}(Y - \bar{Y})$.

<u>Regression Co-efficients:</u> (slopes).

$$b_{yx} = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}, \quad b_{xy} = \frac{n\sum xy - \sum x \sum y}{n\sum y^2 - (\sum y)^2}.$$

$\therefore \gamma = \sqrt{b_{xy} \times b_{yx}} \qquad \therefore \gamma^2 = b_{xy} \times b_{yx}$.

$\because$ Curve (Straight line) => $Y = \underbrace{a}_{\text{intercept}} + \overbrace{b}^{\text{slope}}X$.

$\because$ Parabola second degree => $Y = a + bx + cx^2$.

<u>Rank Correlation:</u>

$$\rho = 1 - \frac{6\sum d^2}{n(n^2-1)} \quad \text{(Non-repeating)}.$$

$$\rho = 1 - \frac{6(\sum d^2 + F)}{n(n^2-1)} \quad \text{(Repeating)}.$$

$$F = \frac{m(m^2-1)}{12}, \quad m = \text{no. of repeatition}.$$

$$\text{Probability} = \frac{\text{Favourable Outcomes}}{\text{All Possible Outcomes}}.$$

⇒ Permutation ⇒ Arrangement

⇒ Arrangement in a line = $n!$

" " circle = $(n-1)!$

⇒ Distinct Permutation = $\frac{n!}{n_1! n_2! \cdots}$

⇒ Arrangement of some Obj (r) from Total (n)

→ $^nP_r = \frac{n!}{(n-r)!}$ (repetition not allowed)

→ $^nP_r = (n)^r$ (rep. allowed)

Combination = Selection.

$$^nC_r = \frac{n!}{r!(n-r)!}$$

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

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

$$P(A \cap B) = P(A)\,P(B).$$

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

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

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

$$P(A \cap B \cap C') = P(A \cap B) - P(A \cap B \cap C)$$

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

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

Statistical Inference

Case I: σ unknown & $n \geq 30$

$$\bar{X} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

Case II: Same. ($n<30$)

Case III: σ unknown & $n \geq 30$

$$\bar{X} - Z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \bar{X} - Z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$$

Case IV: σ unknown & $n<30$

$$\bar{X} - t_{(\frac{\alpha}{2}, n-1)} \frac{s}{\sqrt{n}} < \mu < \bar{X} - t_{(\frac{\alpha}{2}, n-1)} \frac{s}{\sqrt{n}}$$

$\because$ $s = \sqrt{\dfrac{\Sigma (X_i - \bar{X})^2}{n}}$

$\because$ $V = s^2$.

$\because$ $Z = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ $\quad \because$ $t = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}}$

$\because$ $X^2 = \dfrac{(n-1)s^2}{\sigma^2}$

$\because$ $X^2_{cal} = X_{(1-\alpha, n-1)} \rightarrow <$ left tailed

$= X_{(\alpha, n-1)} \longrightarrow >$ Right tailed.

$\because$ $X^2 = \left(\dfrac{(O_i - e_i)^2}{e_i}\right) \Sigma$. $\quad \because X^2_{(\alpha, (r-1)(c-1))}$

$\because$ $e_i = \dfrac{\text{row total} \times \text{col. total}}{\text{Total}}$.

Reject H_0 if:

$\neq$: $Z_{cal} > Z_{tab}$ OR $Z_{cal} < -Z_{tab}$

$<$: $Z_{cal} < -Z_{tab}$

$>$: $Z_{cal} > Z_{tab}$

Same for t and X^2 distribution

Test for Independence (Table)

$$X^2_{cal} > X^2_{tab}$$

Right Tailed/left Tailed (One tail):

Z_α, $t_{(\alpha, n-1)}$

For two tailed:

$Z_{\frac{\alpha}{2}}$, $t_{(\frac{\alpha}{2}, n-1)}$

CDF of Continuous:

∵ $F(x) = P(-\infty \leq \boxed{X \leq X_i})$.

Extract this only.

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

> Agar que me limit 0-3 he to 1-3 direct find nhi karskte. Phele 0-1 phir 0-3 then dono ko minus.

∵ $f(x) \rightarrow$ PDF given hoga or $F(X)$ nekalna hoga (CDF) to $\int_0^{x} f(x)\, dx$ se karege, 0 ki jaga given lower limit aegi.

∵ $F(X) \longrightarrow$ CDF given hoga or $f(x)$ nikalna hoga to der. apply krege.

∵ $F(X)$ given hoga to direct Probability nikal skte hen by putting upper limit.

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

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