Formulas:

UNIT - 1

Mean $(\bar{x}) = \frac{\Sigma x_i}{n}$

Median $= \frac{N}{2}$

Mean deviation $= \frac{\Sigma f_i |x_i - \text{Median}|}{\Sigma f_i}$

Mean deviation $= \frac{\Sigma |x_i - \bar{x}|}{n}$

Mean $(\bar{x}) = \frac{\Sigma x_i f_i}{\Sigma f_i}$

Mean deviation $= \frac{\Sigma f_i |x_i - \bar{x}|}{\Sigma f_i}$

For Continuous:-

Mean $= A + \frac{\Sigma f_i d_i}{\Sigma f_i} \times h$ $\quad \therefore d_i = \frac{x_i - A}{h}$

Mean deviation $= \frac{\Sigma f_i |x_i - \bar{x}|}{\Sigma f_i}$

Median $= l + \frac{h}{f}\left(\frac{N}{2} - c\right)$

Mean deviation $= \frac{\Sigma f_i |x_i - \text{Median}|}{\Sigma f_i}$

Variance $(\sigma^2) = \frac{\Sigma (x_i - \bar{x})^2}{n}$

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

Descrete data:-

Variance $(\sigma^2) = \frac{\Sigma f_i (x_i - \bar{x})^2}{\Sigma f_i}$

Standard Deviation (S.D) $= \sqrt{\sigma^2}$

(or)

σ

Continuous Data:-

Variance $(\sigma^2) = \frac{h^2}{N^2}\left[N \Sigma f_i d_i^2 - (\Sigma f_i d_i)^2\right]$

* Cumulative frequency $= \frac{\sigma_1}{\bar{x}} \times 100$ and $\frac{\sigma_2}{\bar{x}} \times 100$

Karl pearlson's Skewness $(S_{KP}) = \frac{\text{Mean} - \text{Mode}}{\sigma}$

Coefficient of Corelation $(r) = \frac{\Sigma xy}{\sqrt{\Sigma x^2}\sqrt{\Sigma y^2}}$ (or) $\frac{\Sigma xy - \frac{\Sigma x \Sigma y}{N}}{\sqrt{\Sigma x^2 - \frac{(\Sigma x)^2}{N}}\sqrt{\Sigma y^2 - \frac{(\Sigma y)^2}{N}}}$

* Rank corelation $= 1 - \frac{6\Sigma D^2}{N(N^2-1)}$

Repeated Ranks $= 1 - 6\left[\frac{\Sigma D^2 + \frac{1}{12}(m^3 - m) + \frac{1}{12}(m^3 - m) + \cdots}{N(N^2-1)}\right]$

Straight line is

$\Sigma y = Na + b\Sigma x$

$\Sigma xy = a\Sigma x + b\Sigma x^2$

→ Regression Coefficient y on x

$$b_{yx} = \frac{\Sigma dx dy - \frac{(\Sigma dx \Sigma dy)}{N}}{\Sigma dx^2 - \frac{(\Sigma dx)^2}{N}}$$

→ Regression Coefficient x on y

$$b_{xy} = \frac{\Sigma dx dy - \frac{(\Sigma dx \Sigma dy)}{N}}{\Sigma d y^2 - \frac{(\Sigma dy)^2}{N}}$$

→ Regression Eq^n y on x

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

→ Regression Equation x on y

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

UNIT - II

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

$$P(E) = \frac{n(e)}{n(S)}$$

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

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

$$P\left(\frac{E_1}{E_2}\right) = \frac{P(E_1 \cap E_2)}{P(E_2)} \text{ (or) } \frac{n(E_1 \cap E_2)}{n(E_2)}$$

$$P\left(\frac{E_2}{E_1}\right) = \frac{P(E_1 \cap E_2)}{P(E_1)} \text{ (or) } \frac{n(E_1 \cap E_2)}{n(E_1)}$$

* $P(A/A) = 1$
* $P(A - B) = P(A) - P(A \cap B)$
* $P(A/B) = \frac{P(A \cap B)}{P(B)}$
* $P(E_1 \cap E_2) = P(E_1) P(E_2 | E_1)$

$$P\left(\frac{E_K}{A}\right) = \frac{P(E_K) P(A/E_K)}{P(E_1)P(A/E_1) + P(E_2)P(A/E_2) + \cdots + P(E_n)P(A/E_n)}$$

* $E(X) = \sum_{i=1}^{n} P_i X_i$
* $E(x^r) = \sum_{i=1}^{n} P_i x_i^r$

$$E(xy) = E(x) \cdot E(y)$$

$$E(x + y) = E(x) + E(y)$$

Variance (σ^2) (or) $V(x) = E(x^2) - [E(x)]^2$

Standard deviation (S.D) $= \sqrt{E(x^2) - [E(x)]^2}$

UNIT-III

* $\text{Mean}(\mu) = np$
* $\text{Variance}(x) = npq$

$SD = \sqrt{npq}$

* $\text{Mean} = \frac{\Sigma f_i x_i}{\Sigma f_i}$ (or)

$N(q+p)^n$

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

* $P(X=0) = n_{C_r} p^r q^{n-r}$

→ $z = \frac{x-\mu}{\sigma}$

$P(x_1 \leq X \leq x_2) = |A(z_2) - A(z_1)|$

$P(x_1 \leq X \leq x_2) = (A(z_2) + A(z_1))$

→ $P(z \leq z_1) = 1 - P(z > z_1)$

- $z_1 \geq 0 \quad = 1 - [0.5 - A(z_1)]$

$0.5 + A(z_1)$

- $z_1 \leq 0 ; \quad 1 - [0.5 + A(z_1)]$

$= 0.5 - A(z_1)$

→ $P(z \geq z_1)$

$z_1 \geq 0, \quad P(z \geq z_1) = 0.5 - A(z_1)$

$z_1 \leq 0, \quad P(z \leq z_1) = 0.5 + A(z_1)$

UNIT-IV

Correction factor $= \left(\frac{N-n}{N-1}\right)$

Control limit $= z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$

Null Hypothesis :- $\mu = \mu_0$

LOS :- $\alpha = 0.5$ (or) $\alpha = 1\%$

Acceptance Region :-

$z = \frac{t - E(t)}{\text{S.D expect of } t}$

$z = \frac{\text{Observed value} - \text{Expected value}}{\text{standard expect value}}$

Level of significance	1%	5%	10%
Critical value of TTT	2.58	1.96	1.645
RTT	2.33	1.645	1.28
LTT	-2.33	-1.645	1.28

→ Single Proportion $= z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}$

Two Sample Proportion $= z = \frac{P_1 - P_2}{\sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}}}$

$\boxed{Q = 1 - P}$

Pooling $z = \frac{P_1 - P_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

Substitution $z = \frac{P_1 - P_2}{\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}}$

- Test Statistics $z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$

- Single mean-Large Sample

$$z = \begin{cases} \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} & \text{if } \sigma \text{ known} \\ \frac{\bar{x} - \mu}{s/\sqrt{n}} & \text{if } \sigma \text{ not known} \end{cases}$$

UNIT - V

t-distribution :- $t = \dfrac{\bar{x}-\mu}{s\sqrt{n-1}}$

$s^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$ (when SD not given)

Difference of mean:- $t = \dfrac{\bar{x}-\bar{y}}{\sqrt{\dfrac{s^2}{n_1}+\dfrac{s^2}{n_2}}}$ (or) $\dfrac{\bar{x}-\bar{y}}{s\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$

$\bar{x} = \dfrac{\Sigma x_i}{n_1}$ $\bar{y} = \dfrac{\Sigma y_i}{n_2}$ $\nu = n_1+n_2-2$

$|\bar{x}-\bar{y}| = |t| \cdot s\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}$

if $\sigma_1 = \sigma_2 = \sigma$

$s^2 = \dfrac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2}$ (or) $s^2 = \dfrac{\Sigma(x_i-\bar{x})^2 + \Sigma(y_i-\bar{y})^2}{n_1+n_2-2}$

Paired Sample test:- $t = \dfrac{\bar{d}-\mu}{s/\sqrt{n}}$ (or) $t = \dfrac{\bar{d}}{s/\sqrt{n}}$ $(\mu = 0)$

where $\bar{d} = \dfrac{\Sigma d_i}{n}$

$s = \sqrt{\dfrac{\Sigma(d_i-\bar{d})}{n-1}}$ $s^2 = \dfrac{\Sigma d^2 - n(\bar{d})^2}{n-1}$ (or) $s^2 = \dfrac{\Sigma(d_i-\bar{d})^2}{n-1}$

F-test:-

$F = \dfrac{S_1^2}{S_2^2}$ $S_1^2 > S_2^2$

$S_1^2 = \dfrac{\Sigma(x_i-\bar{x})^2}{n_1-1}$, $S_2^2 = \dfrac{\Sigma(x_i-\bar{x})^2}{n_2-2}$

$S_1^2 = \dfrac{n_1 s_1^2}{n_1-1}$, $S_2^2 = \dfrac{n_2 s_2^2}{n_2-1}$

Chi- Square test (χ^2):-

$\chi^2 = \sum_{i=1}^{n}\left(\dfrac{O_i-E_i}{E_i}\right)^2$

$E_i = \dfrac{\text{Row Table} \times \text{column Table total}}{\text{Grand (total) total}}$

d.f = (no. rows - 1) (no. columns - 1)