→ Testing of Single mean (Large Sample) :-

Step 1 :- Understanding the problem.
Identification of random variable.

Step 2 :- Stating the hypothesis

Null hypothesis $H_0 : \mu = \mu_0$

vs

Alternate hypothesis $H_1 : \mu > \mu_0$

$\mu < \mu_0$

$\mu \neq \mu_0$

Step 3 :- Test statistic

$$z = \frac{\bar{x} - \mu_0}{S.E(\bar{x})}$$

Sampling distribution
→ drawing random samples from the population.

where $S.E(\bar{x})$ = Standard error($\bar{x}$)
= Standard deviation of $\bar{x}$
$= \sigma/\sqrt{n}$

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

Step 4 :- Decision Rule

α = level of significance = P(Type I error)

β = P(Type II error)

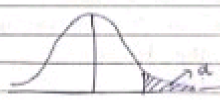
Eg :-

	Raining	Not Raining
umbrella	✓	×
No umbrella	×	✓

α = 5% or 1%

	Accept H_0	Reject H_0
True H_0	✓	× → Type I error
False H_0	× → Type II error	✓

case 1 : $H_1 : \mu > \mu_0$ (Right tail test)

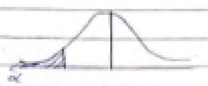


At α%, accept H_0 when

→ Critical / Rejection region

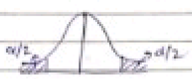
Accept H_0 when $Z_{calculated} < Z_\alpha$

case 2 :- $H_1 : \mu < \mu_0$ (Left tail test)



Accept H_0 when $z > z_\alpha$

case 3 :- $H_1 : \mu \neq \mu_0$



$|z| < Z_{\frac{\alpha}{2}}$

Step 5 :- Numerical Computation

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

$\bar{x} = \frac{\Sigma x}{n}$, $\hat{\sigma}$ (unknown) → when popⁿ S.D is not known we replaced with its unbiased estimator

$$\hat{\sigma} = s = \sqrt{\frac{1}{n}\sum_i (x_i - \bar{x})^2}$$

Conclusion, Accepting H_0 based on Step 4.

Test of significance concerning difference of mean (Large sample).

* $X_1 \sim N(\mu_1, \sigma_1^2)$
  $X_2 \sim N(\mu_2, \sigma_2^2)$

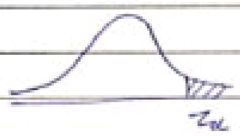
* $H_0 : \mu_1 = \mu_2$
  $H_1 : \mu_1 > \mu_2$
  $\mu_1 < \mu_2$
  $\mu_1 \neq \mu_2$

* Test statistic

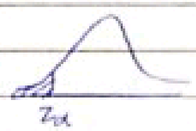
$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

* Decision rule,

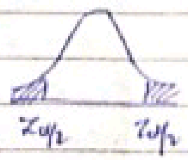
At $\alpha\%$, we accept H_0 when

→ 

$z < z_\alpha$

→ 

$z > z_\alpha$

→ 

$|z| < z_{\alpha/2}$

Page No.
Date / /

* Numerical computation,

$$\bar{x}_1 = \frac{\sum_i x_{1i}}{n_1} \qquad \bar{x}_2 = \frac{\sum_i x_{2i}}{n_2}$$

$$\sigma_1^2 = s_1^2 = \frac{1}{n_1} \sum_i (x_{1i} - \bar{x}_1)^2$$

$$\sigma_2^2 = s_2^2 = \frac{1}{n_2} \sum_i (x_{2i} - \bar{x}_2)^2$$

Testing of hypothesis for single proportion

→ $X \sim B(n, p)$

$X \sim N(\mu = nP, \sigma^2 = nPQ)$

→ deals with characteristic

→ $H_0 : P = P_0$

vs

$H_1 : P > P_0$

$P < P_0$

$P \neq P_0$

→ $z = \dfrac{p - E(p)}{S.E(p)} \sim N(0,1)$

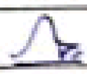
where, $E(p) = P$

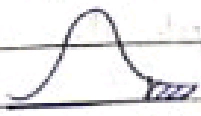
$S.E(p) = \sqrt{\dfrac{PQ}{n}}$

$z = \dfrac{p - P}{\sqrt{\dfrac{PQ}{n}}} \sim N(0,1)$

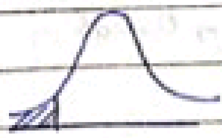
where $p = \dfrac{x}{n}$

→ At α%, accept H_0 when

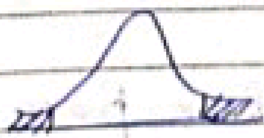




$z < z_\alpha$



$z > z_\alpha$



$|z| < z_{\alpha/2}$

Test of Significance for difference of proportions.

→ $X_1 \sim B(n_1, P_1)$

$X_2 \sim B(n_2, P_2)$

→ $H_0 : P_1 = P_2$

vs

$H_1 : P_1 > P_2$

$P_1 < P_2$

$P_1 \neq P_2$

Page No.

Date

→ $$z = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim N(0,1)$$

where Pooled proportion $= \dfrac{x_1 + x_2}{n_1 + n_2} = P$

or $P = \dfrac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$

and $Q = 1 - P$

Testing equality of Variances:-

→ $X_1 \sim N(\mu_1, \sigma_1^2)$
$X_2 \sim N(\mu_2, \sigma_2^2)$

→ $H_0 :- \sigma_1^2 = \sigma_2^2$

vs

$H_1 :- \sigma_1^2 \neq \sigma_2^2$

→ $F = \dfrac{S_1^2}{S_2^2} \sim F_{n_1-1, n_2-1, \alpha}$

Note :- Numerator > Denominator

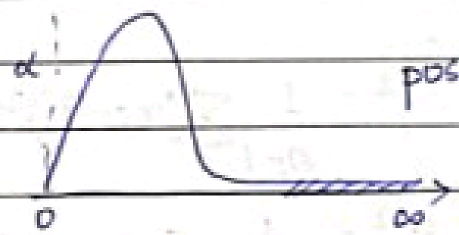
$\dfrac{S_2^2}{S_1^2} \sim F_{n_2-1, n_1-1, \alpha}$

F - Snedecor's distribution

→ Continous prob. distribution

→ pdf is expressed in terms of β function.

→ $0 < \alpha < \infty$

→ 

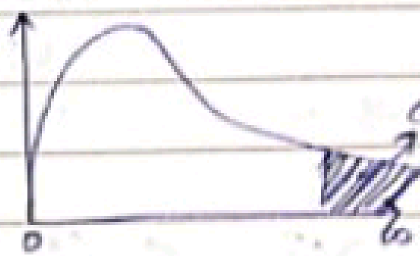
positively skewnes

→ At α%, accept H_0 when $F < F_{n_1-1, n_2-1, \alpha}$

$$s^2 = \frac{1}{n-1} \sum_{i} (x_i - \bar{x})^2$$

Chi. Square Distribution (χ^2)

→ Non-parametric test
→ Distribution free test
→ Positively skewed
→ Deals with attributes
→ Applications :-
- ★ Goodness of fit
- ★ Independence of attributes
- ★ Homogenity of data



Testing of Hypothesis for goodness of fit :-

→ H_0 :- Distribution is a good fit.

vs

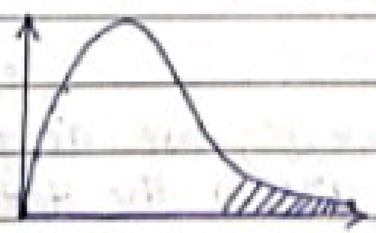
H_1 :- Distribution is not a good fit.

→ $$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_{n-k}$$ → (degree's of freedom)

n - no. of observations

k - no. of constraints

→ At $\alpha\%$, we accept H_0 when



$\chi^2 < \chi^2_{n-k, \alpha}$

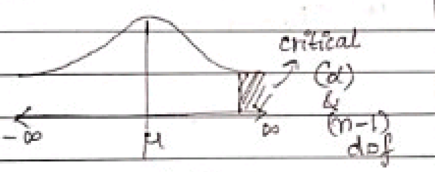
(table values)

Testing of Hypothesis for single mean (Small Sample) :-

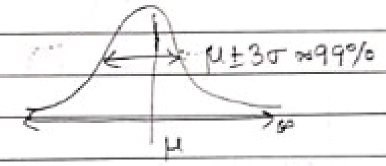
$$\frac{\bar{x} - E(\bar{x})}{S.E(\bar{x})} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

t - distribution / Student's t distribution.

→ Continous prob. distribution.
→ Similar to normal distribution



t- curve



normal dist^n

Degrees of freedom : The no. of values in the final calculation of a statistic, that are free to vary.

→ $X \sim N(\mu, \sigma^2)$
Samples are drawn randomly.

→ $H_0 : \mu = \mu_0$ vs $H_1 : \mu > \mu_0$
$\mu < \mu_0$
$\mu \neq \mu_0$

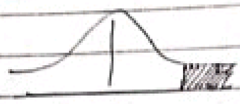
→ Test Statistic,

$$\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim t_{\alpha, n-1}$$

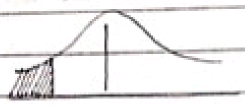
$$\hat{\sigma} = S = \sqrt{\frac{1}{n-1} \sum_i (x - \bar{x})^2}$$

In calc^n, $\sigma = s$
$s = S$

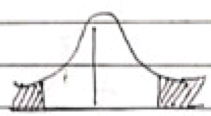
→ Decision rule,
At $\alpha\%$, accept H_0 when



$t < t_{\alpha, n-1}$



$t > t_{\alpha, n-1}$



$|t| < t_{\alpha, n-1}$

Testing of difference of mean for small samples :-

→ Comparative study of 2 population mean

$X_1 \sim N(\mu_1, \sigma_1^2)$

$X_2 \sim N(\mu_2, \sigma_2^2)$

→ Samples are independent

→ H_0 :- $\mu_1 - \mu_2 = S$

vs

H_1 :- $\mu_1 - \mu_2 \neq S$

$\mu_1 - \mu_2 < S$

$\mu_1 - \mu_2 > S$

→ Decision rule,

Accept H_0, when

$|t| < t_{\alpha/2, n_1+n_2-2}$

$t > t_{\alpha, n_1+n_2-2}$

$t < t_{\alpha, n_1+n_2-2}$

→ $$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{\alpha, n_1+n_2-2}$$

S^2 - Pooled variance

$$= \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$$

$$S^2 = \frac{1}{n-1}\sum_1 (x-\bar{x})^2, \quad s^2 = \frac{1}{n}\sum_1 (x-\bar{x})^2$$

→ Also

$$\Rightarrow \boxed{(n-1)S^2 = ns^2}$$

Paired t test :-

→ $X_1 \sim N(\mu_1, \sigma_1^2)$
$X_2 \sim N(\mu_2, \sigma_2^2)$

→ Samples are dependent.

$$d_i = X_1 - X_2$$

$$d \sim N(\mu_d, \sigma_d^2)$$

→ H_0 :- $\mu_1 - \mu_2 = 0$ or $\mu_d = 0$ → Accept H_0 when

vs

H_1 :- $\mu_1 \neq \mu_2$ or $\mu_d \neq 0$ $|t| < t_{\alpha/2, n-1}$

$\mu_1 - \mu_2 < 0$ or $\mu_d < 0$ $t > t_{\alpha, n-1}$

$\mu_1 - \mu_2 > 0$ or $\mu_d > 0$ $t < t_{\alpha, n-1}$

→ $$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} \sim t_{\alpha, n-1}$$

$$\bar{d} = \frac{\sum_i d}{n} \qquad s_d = \frac{1}{(n-1)} \sum_i (d_i - \bar{d})^2$$