26/oct

WEEK : 08

Hypothesis Testing:

→ H_0 — Null Hypothesis

$P(H_0 \text{ is true}) = 1-\alpha$

i.e. 90%, 95%, 99%

↓ state of the art

→ H_A — Alternative hypothesis

$P(H_A \text{ is true}) = \alpha$

i.e. 10%, 5%, 1%

↓ Research hypothesis

Type I error:

$\alpha = P(\text{Guilty} \mid \text{Innocent})$

~~β~~ Type II error

$\beta = P(\text{Innocent} \mid \text{Guilty})$

Type 2 error more serious.

eg If healthy person is declared covid then this is not risky (Type 1)

but if covid patient declared clear than this is serious (type 2)

→ Assume data is normally distributed.

If not normally distributed we will make according to it.

Hypothesis Testing for Mean, Variance, proportion, chi-square testing (important)

~~Variance Testing~~ 4 methods.

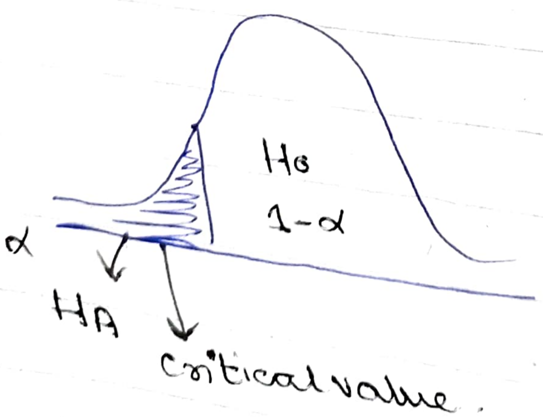
1) Define H_0
2) Define H_A
3) α (probability of H_A is true)

If α value not given then assume 5%

4) Test Statistic / calculations

$$Z_{cal} = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \quad \text{OR}$$

$$t_{cal} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$



$Z \sim N(0,1)$

→ Assume data is normally distributed
→ Jise test krna hai use Alternative hypothesis longy.

~~Ex. 04~~ P-value approach.

Easy way for calculation.

If P-value < α
then Reject H_0

P value more then 5% then we do not accept alternative value.

$\Rightarrow \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \Rightarrow \dfrac{2450-2600}{600/\sqrt{50}} = \dfrac{-150}{84.86}$

$\Rightarrow z = -1.76$

$\Rightarrow 0.0392 = \phi(z)$

Two sided test so $\alpha/2 = 5/2 = 2.5$

$0.039 \not< 0.025$

Accept H_0

Testing for difference of Two Population Mean ($\mu_1 - \mu_2$)

1) $H_0 : \mu_1 = \mu_2$

   $H_0 : \mu_1 - \mu_2 = 0$

2) $H_A : \mu_1 < \mu_2$ (or) $\mu_1 > \mu_2$

   (or) $H_A : \mu_1 \neq \mu_2$

3) α (e.g 10% 5% or 1%)

4) Test Statistic

From Ho

i) $$Z_{cal} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Population parameters

ii) $$Z_{cal} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$ $n > 30$

When variance is same $n < 30$

iii) $$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

iv) $$t_{cal} = \frac{\bar{x}_d - \mu_D}{s_d / \sqrt{n}}$$

iv) $\cancel{\bar{x} = 96.92}$

$\cancel{\bar{x}_2 = 83.03}$

$\mu_g < \mu_p$

$\mu_1 < \mu_2$ left

Ex book

Sample 1

$s_1 = 19.34$

$\bar{X}_1 = 96.93$

$n_1 = 14$

Sample 2

$s_2 = 28.48$

$\bar{X}_2 = 83.03$

$n_2 = 12$

$$s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$$

$$= \sqrt{\frac{(14-1)(19.34)^2 + (12-1)(28.48)^2}{14+12-2}}$$

18

$$= \sqrt{\frac{4862.46 + 8922.21}{24}}$$

$$= \sqrt{574.36}$$

$s_p = 23.96$ 18.95

Case 3

$$t_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\overset{0}{\mu_1 - \mu_2})}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

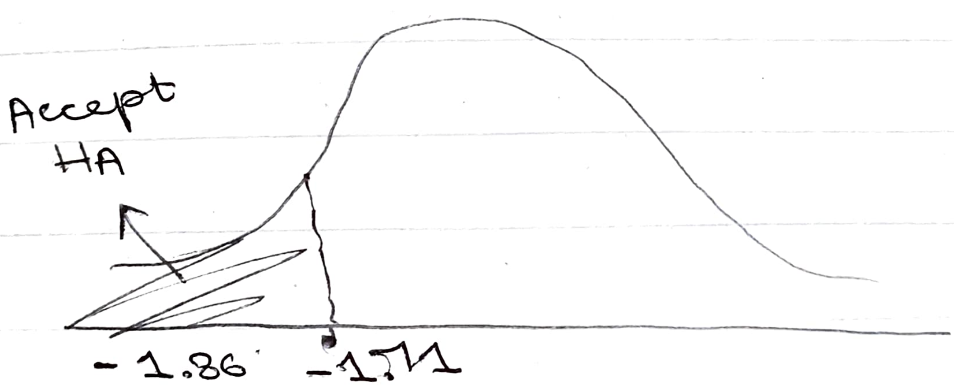
$$\Rightarrow t_{cal} = \frac{(96.93-83.03)-0}{\cancel{23.96}\; 18.95\sqrt{\frac{1}{14}+\frac{1}{12}}}$$

$$= \frac{13.9}{18.95\;\cancel{23.96}(0.393)} = \frac{13.9}{\cancel{9.42}\;7.44}$$

$$\Rightarrow t_{cal} = \cancel{1.47}\; 1.86$$

$$t_{(\alpha, n_1+n_2-2)} = t_{(0.05,\, 14+12-2)}$$

$$= t_{(0.05,24)} = 1.71$$



9/Nov

Hypothesis Testing For Variance

1) H_0 $\sigma^2 = \sigma_0^2$

2) H_A $\sigma^2 < \sigma_0^2$

or H_A $\sigma^2 > \sigma_0^2$

or H_A $\sigma^2 \neq \sigma_0^2$

3) Level of Significance (α)

4) Test Statistic

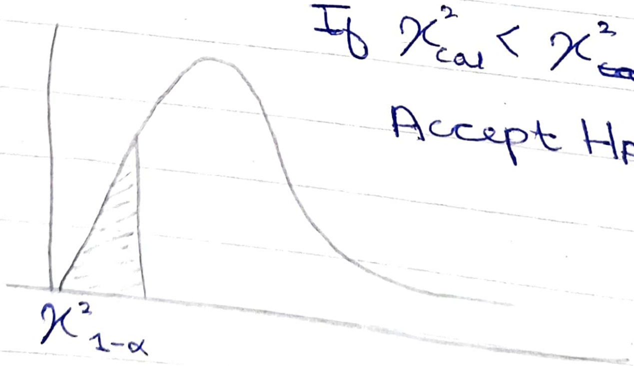
$$\chi^2_{cal} = \frac{(n-1)s^2}{\sigma_0^2}$$

5) Critical Region (The area where H_A is true)

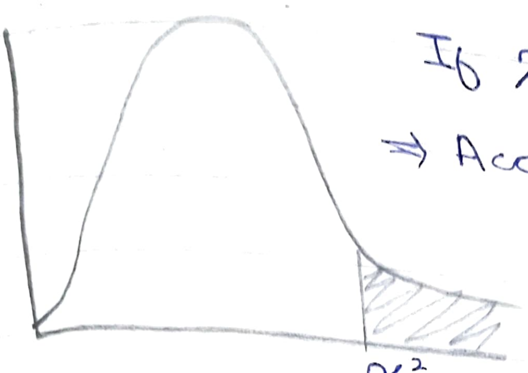
a) If H_A $\sigma^2 < \sigma_0^2$

If $\chi^2_{cal} < \chi^2_{\cancel{cal}\,tab}$

Accept H_A



b) If H_A $\sigma^2 > \sigma_0^2$



If $\chi^2_{cal} > \chi^2_{tab}$

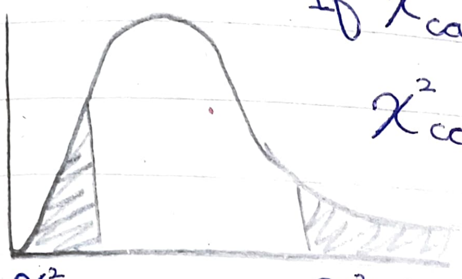
$\Rightarrow$ Accept H_A

c) If H_A $\sigma^2 \neq \sigma_0^2$

If $\chi^2_{cal} < \chi^2_{1-\frac{\alpha}{2}}$ OR

$\chi^2_{cal} > \chi^2_{\frac{\alpha}{2}}$

$\Rightarrow$ Accept H_A



Example 10.12:

$\Rightarrow n = 10 \quad S = 1.2 \quad \sigma_0 = 0.9$

$$\chi^2_{cal} = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(10-1)(1.2)^2}{(0.9)^2}$$

$$= 16.0$$

we will check this calculated value from table. approximately where it is lying.

$P \approx 0.7$ (lying b/w 10% and 5%)

Hypothesis Testing for σ_1^2/σ_2^2 OR σ_2^2/σ_1^2

1) $H_0 : \sigma_1^2 = \sigma_2^2 \Rightarrow H_0 : \dfrac{\sigma_1^2}{\sigma_2^2} = 1$

2) $H_A : \sigma_1^2 < \sigma_2^2$ OR

$H_A \quad \sigma_1^2 > \sigma_2^2$ OR

$H_A \quad \sigma_1^2 \neq \sigma_2^2$

3) Define α

4) $f_{cal} = \dfrac{S_1^2}{S_2^2}$ OR $f_{cal} = \dfrac{S_2^2}{S_1^2}$

(whichever is larger)

$\sigma_1^2 < \sigma_2^2$

5) $f_{cal} < f_{1-\alpha}(V_1, V_2) \Rightarrow$ Accept H_A

$\sigma_1^2 > \sigma_2^2$

$f_{cal} > f_{\alpha}(V_1, V_2) \Rightarrow$ Accept H_A

$\sigma_1^2 \neq \sigma_2^2$

$$\left.\begin{array}{l} f_{cal} < f_{1-\frac{\alpha}{2}}(V_1, V_2) \text{ OR} \\ \\ f_{cal} > f_{\frac{\alpha}{2}}(V_1, V_2) \end{array}\right\} \text{Accept } H_A$$

Example 10.13

$n_1 = 12$ $\quad v_1 = 11$

$n_2 = 10$ $\quad v_2 = 9$

$\alpha = 10\%$ $\quad \frac{\alpha}{2} = 5\%$

$$f_{\frac{\alpha}{2}}(v_1, v_2) = f_{0.05}(11, 9) = 3.11$$

$$f_{1-\frac{\alpha}{2}}(v_1, v_2) = f_{0.95}(11, 9)$$

$$= \frac{1}{f_{0.05}(9, 11)} = \frac{1}{2.90}$$

$$= 0.34$$

Expected Frequency

$$e_f = \frac{R_i \times C_j}{G}$$

$O_{i,j}$	$e_{i,j}$	$(O_i - e_i)^2 / e_i$
182	200.9	
213	209.9	
203	187.2	
154	135.1	
135	141.1	
110	125.8	

$$\chi^2_{cal} = \sum_{i=1}^{r \times c} \left(\frac{(O_i - e_i)^2}{e_i} \right) \quad \text{here } r = 2,\ c = 3$$

$$= 7.85$$

Critical Region

$$\chi^2_{tab} = \chi^2_{\alpha}((r-1)(c-1))$$

$$= \chi^2_{0.05}\,\underline{(2-1)(3-1)}$$

$$= 2$$

[If any value in table/matrix is less than 5 then we merge two values called Yaches correction].