06/02/09

# RESEARCH METHODOLOGY

By Prof. Khan

Reference book : C.R. Kothari (Revised Edition)

MARKS BREAKUP:

A] Theory — 30%.
B] Numericals based on the tools — 30%.
C] Practical approach / performing Research — 40%.

A] THEORY:

Syllabus:

Ch. 1 :- Complete
Ch. 2 :- Complete
Ch. 3 :- Research Design (Pg. 31 to 40)
Ch. 5 : Pg. 69 to 75
Ch. 14 : Report Writing (Pg. 344 to 358)

C] Practical Approach:

1. Identify the project — Applied
2. Data Collection — Collection
3. Analysis — Steps
4. Interpretation / Generalisation — Conclusion
5. Report Writing — Presentation

(PPT to explain the work)

# CH. 1 : Research

- Meaning
- Objectives
- Types
  - (i.) Descriptive v/s Analytical
  - (ii.) Applied v/s Fundamental
  - (iii) Quantitative v/s Qualitative
  - (iv.) Free, laboratory, historical

## STEPS:

1.) Formulation of research problem
2.) Extensive Literature Study
3.) Development of working hypothesis
4.) Preparing the research time
* 5.) Determining the sample design
* 6.) Collection of design / data
7.) Execution of the project
* 8.) Analysis of data
* 9.) Hypothesis testing
10.) Generalisation & Interpretation (conclusion)
11.) Report Writing

(* - Numerical's - can be studied from any book)

17/02/09

## I.] Estimation:

Point Estimation – Exact value is determined
Interval Estimation – Location of interval | range

Test of Hypothesis – Make a statement regarding population parameters & then do research to accept or reject the statement.

'NULL' Hypothesis – $H_0$ – Hypothesis believed to be true.

(Population parameter = $\mu$ = population or mean)

## II.] Sample Size:

1. Large sample test ($n \geq 30$)
   - Difference of proporation
   - Single mean
   - Difference mean
   - Single proportion

2. Small sample test ($n \leq 30$)
   - Simple
   - Different
   - Paired T test

3. Chi Square distribution
   - Goodness of Fit
   - Independence of attributes

## III.] Procedure for testing of hypothesis:

Step 1. Define $H_0$ ie: Null hypothesis & $H_1$: Alternate hypothesis

Step 2 Fix level of significance : $\alpha$ (Alpha) : Maximum probability of committing type I error is level of significance.

$$P(\text{Type I error}) \leq \alpha$$

| | Fact ($H_0$) | |
|---|---|---|
| | True | False |
| Accept | ✓ | Type II error |
| Reject | Type I error | ✓ |

Step 3 Find test Statistics under $H_0$;

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

Step 4 Find Critical value & decision criteria

Step 5 Compare & conclude

# A] LARGE SAMPLE TEST:

1) When to use?

- if sample is large ( > 30)
- Parent population tends to normal

2) Critical Values :

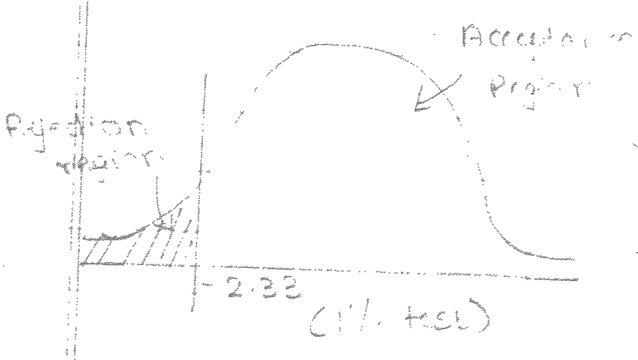
Z test is used (Normal test)

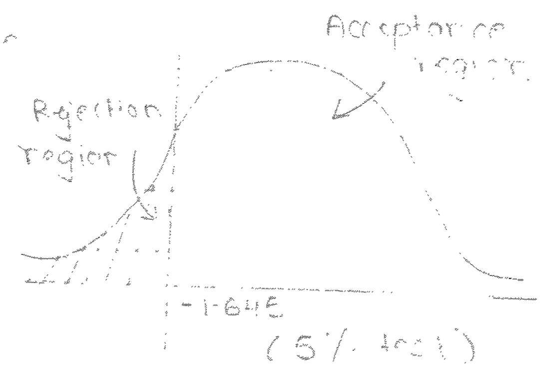
| | Signs of Alternate hypothesis | Level of Significance 1% | Level of Significance 5% |
|---|---|---|---|
| i) left tailed | $\leq$ | $-2.33$ | $-1.645$ |
| ii) Right tailed | $\geq$ | $+2.33$ | $+1.645$ |
| iii) two tailed | $\neq$ | $\pm 2.58$ | $\pm 1.96$ |

(i) Left tailed test:

$H_0 : \mu = 30\%$

$H_1 : \mu < 30\%$



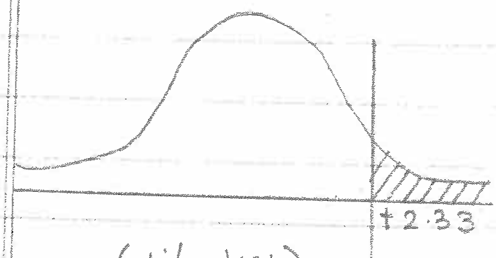

Calculate 'Z' cal

Z cal is $\leq -2.33$, Ho is rejection

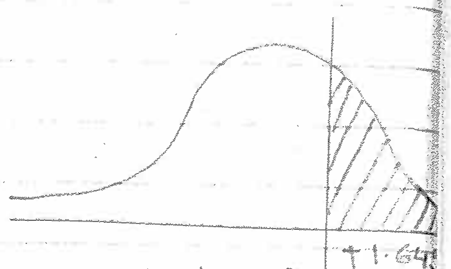
($H_1$ → Alternate Hypothesis + Level of Significance = Decision Criteria)

2. <u>Right tailed test</u>:

$$H_0 : \mu = 30\%$$
$$H_1 : \mu \geqslant 30\%$$



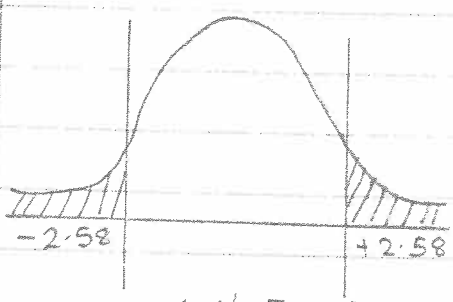

(1% test)



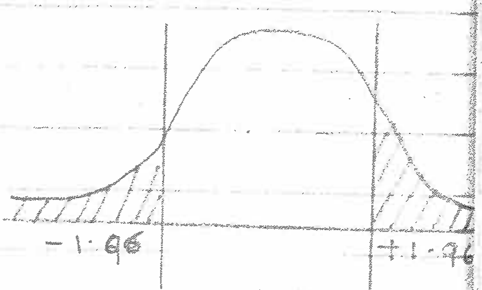

(5% test)

3.) <u>Two tailed test</u>:

$$H_0 : \mu = 30\%$$
$$H_1 : \mu \neq 30\%$$



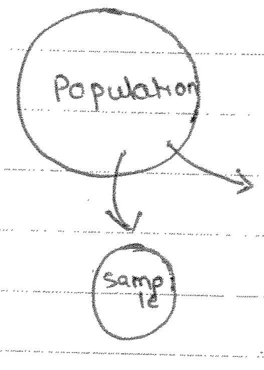

(1% Test)




(5% Test)

## LST I : TEST FOR SINGLE MEAN

When ?




$\sigma$ - population std. deviation

$(H_0 : \mu = \mu_0)$?

$\bar{x} \rightarrow$ single mean

to check whether between $\mu_0$ & $\bar{x}$, is there any significant difference.

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

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

ex 1: Vegetable oil company:

1 tin = 5 kgs. Acceptable

Sample = n = 200 tins

Average wt. = $\bar{x}$ = 4.96 kgs.

Std. dev. = $s = \sigma$ = 0.22 kgs.

Soln.

i.) $H_0 : \mu = 5$

$H_1 : \mu \neq 5$

ii.) $\alpha$ = 5% Level of Significance (L.S.)

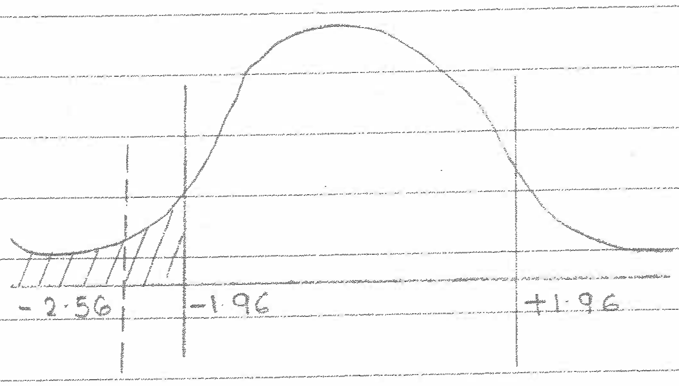
iii.) $\therefore \quad Z = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$

$$\therefore Z = \frac{4.96 - 5}{0.22/\sqrt{200}}$$

$$\therefore Z = -2.56$$

(As the sign of alternate hypothesis $H_1$ is $\neq$, this is a two tailed sum).

iv.)




-2.56 lies in the rejected zone

v.) $\therefore$ $H_0$ is rejected & $H_1$ is accepted.

The difference in sample mean & population mean is significant.

⇒ RIGHT TAILED: (Not recommended)

(Not recommended as alternative ie $H_1$ as the average wt. given is less than 5 ie 4.96

i.) $H_0 : \mu = 5$

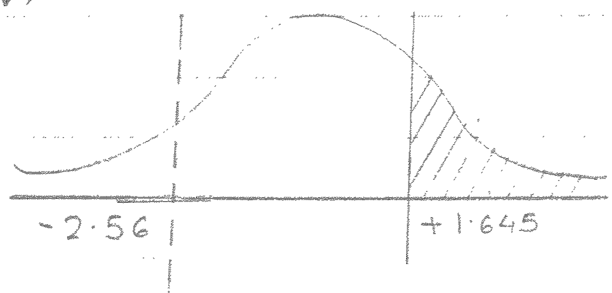
$H_1 : \mu > 5$

ii.) 5% L.S.

iii.) $$Z = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$= 2.56$$

iv)




$-2.56$ is lying in the acceptance zone

iv.) $H_0$ is accepted ie. weight of tin is 5 kgs. & difference is not significant.

⇒ LEFT TAILED:   Best

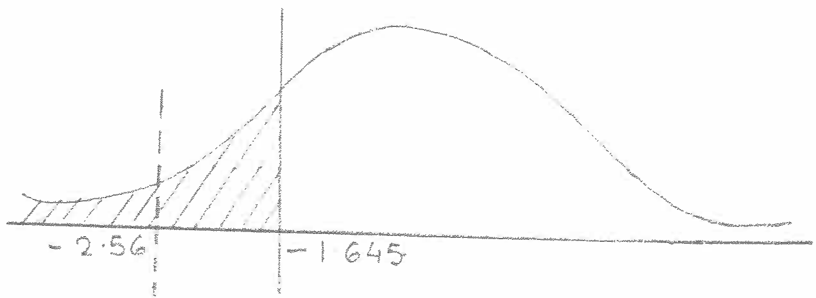
i.) $H_0 : \mu = 5$
$H_1 : \mu \leq 5$

ii.) L.S. $= 5\%$

iii) $$Z = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$= -2.56$$

iv)



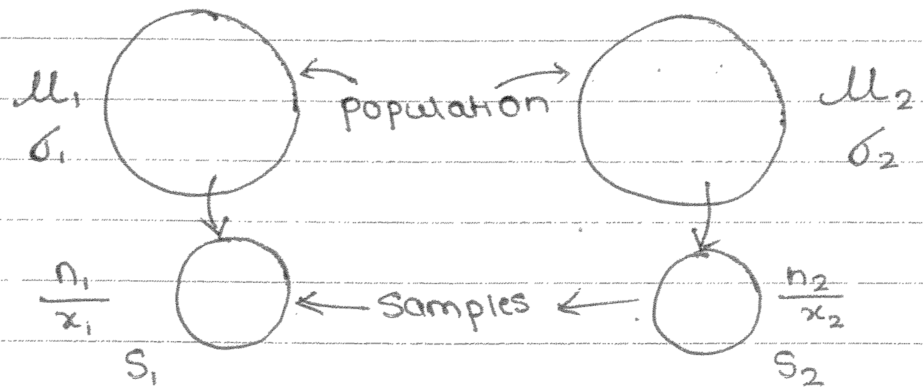

-2.56 lies in the rejection zone

v) ∴ $H_0$ is rejected & $H_1$ is accepted

12/03/09

# LST II : TEST FOR DIFFERENCE OF MEAN

Two large samples




$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \ (\leq, \neq, \geq) \ \mu_2$

Under $H_0$ test statistics :-

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Ex. 1: There are two groups consisting of 121 girls and 81 boys. Their intelligence test results are :

| | Girls | Boys |
|---|---|---|
| | 121 | 81 |
| Mean | 84 | 81 |
| Std. Dev. | 10 | 12 |

(i) Is the difference in intelligence significant?

Sol$^n$: (i) $H_0 : \mu_1 = \mu_2$

(means there is no significant difference & whatever difference is there, is due to sampling)

$H_1 : \mu_1 > \mu_2$

ii) L.S. = 5%

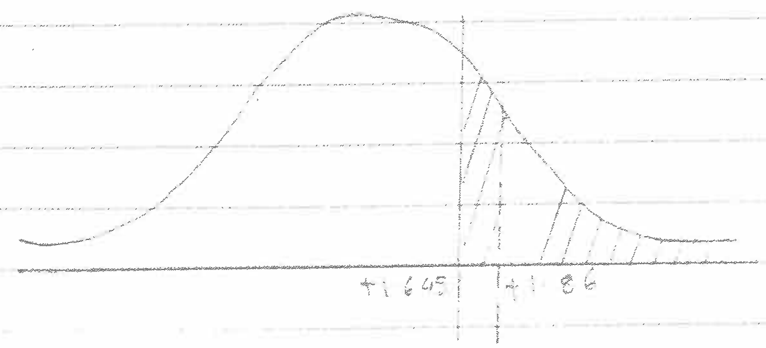
iii) Under $H_0$ test of statistics:

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$\therefore Z = \frac{84 - 81}{\sqrt{\frac{100}{121} + \frac{144}{81}}}$$

$$\therefore Z = 1.86$$

iv)




v) $H_0$ is rejected.

ie: there is significant difference in the performance of girls & boys. $\therefore H_1$ is accepted

ie: girls intelligence is more than boys.

Ex. 2: Random samples drawn from 2 places.
Adult males & their height

| | Place A | Place B |
|---|---|---|
| Avg. | 68.5 | 68.58 |
| Std. Dev. | 2.5 | 3.0 |
| Size | 1200 | 1500 |

Test @ 5% L.S.

Test whether mean height is same at both places.

Sol^n^: i) $H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$ (ie: there is significant diff.)

ii) L.S = 5%

iii) $$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

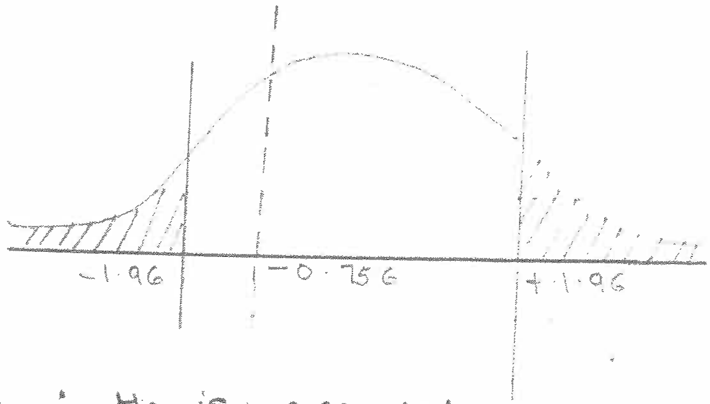
$$\therefore Z = \frac{68.5 - 68.58}{\sqrt{\frac{(2.5)^2}{1200} + \frac{(3)^2}{1500}}}$$

$$= \frac{-0.08}{\sqrt{0.0112}}$$

$$= \frac{-0.08}{0.1058}$$

$$= \underline{\underline{-0.756}}$$

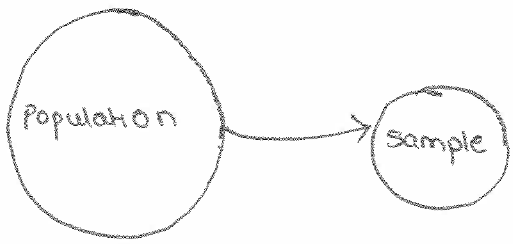
iv.)




v.) ∴ Ho is accepted.

## LST III : TEST FOR SINGLE PROPORTION

$$Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}$$




Let $n = 30$, $x = 20$

$\therefore p = 20/30 = 0.67$ ....> sample populati

$\therefore q = 1 - p = 0.33$

$H_0 : p = p_0$ ....> ($p_0$ can also be wri as $P$

$H_1 : p \, (\leq, \neq, >) \, p_0$

Ex. 1: A manufacturer claims that 90% components confirm to the specification.

A random sampling of 200 samples shows that only 164 meets the standard. Test the claim at 5% L.S.

Soln: (i.) $P_0 = 0.9$

$\therefore H_0 : P = P_0$

$\therefore H_0 : P = 0.9$

$$p = \frac{164}{200} = 0.82$$

$\therefore Q = 1 - P = 1 - 0.9 = 0.1$

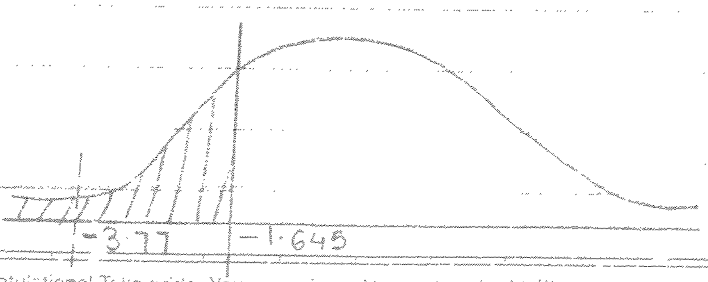
ii.) Test at 5% L.S.

iii.) $$Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}$$

$$\therefore Z = \frac{0.82 - 0.9}{\sqrt{\frac{0.9 \times 0.1}{200}}}$$

$$= \underline{\underline{-3.77}}$$

iv.)

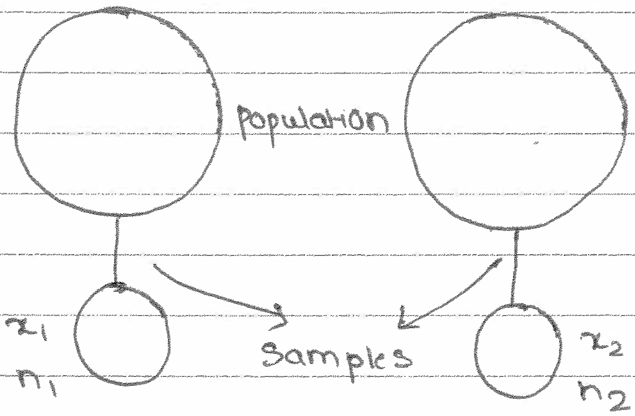


5.) $H_0$ is rejected.

$\therefore$ $H_1$ is accepted.

$\therefore$ Suppliers products are less than 90% who confirm the specification.

## LST IV : TEST OF DIFFERENCE OF PROPORTIO




$H_0$ : $P_1 = P_2$

$H_1$ : $P_1 \ (\leq, \neq, \geq) \ P_2$

Under $H_0$ test statistics

$$Z = \frac{p_1 - p_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where $p$ = proportion of combined sample.

$$p = \frac{x_1 + x_2}{n_1 + n_2} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

Ex. 1: In a large city in a random sample of 1100 students, 20% of the students have some physical defects.
In another large city 200 out of 900 school boys had the same defect.
Do you think the % is less in the former city.

Soln. $p_1 = 0.2$ (20%)
$p_2 = 0.22$ (22.22% → 200/900)

(i) $H_0 : p_1 = p_2$
$H_1 : p_1 \leq p_2$

$$\therefore \quad p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

$$= \frac{(1100)(0.2) + (900)(0.22)}{1100 + 900}$$

$$= \frac{220 + 200}{2000}$$

$$= 0.21$$

$$\therefore \quad q = 1 - p = 1 - 0.21 = 0.79$$

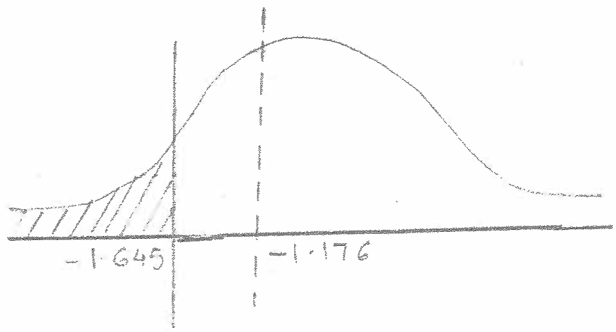
iii) $$Z = \frac{p_1 - p_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{0.2 - 0.22}{\sqrt{(0.21)(0.79)\left(\frac{1}{1100} + \frac{1}{900}\right)}}$$

$$= \frac{-0.02}{0.0187}$$

$$= \underline{\underline{-1.176}}$$

iv.)




v.) ∴ $H_1$ is rejected & $H_0$ is accepted.
∴ There no significant difference in the no. of students having physical defects in the two cities.

Ex. 2: A machine produce 20 defective articles in a batch of 400.
After treatment & improvement it produc 10 defective pieces in a batch of 300.
Has the machine improved?

Soln:

$p_1 = 0.05 \; (20/400)$
$p_2 = 0.033 \; (10/300)$

$n_1 = 400$

$n_2 = 300$

$$\therefore p = \frac{400(0.05) + 300(0.033)}{400 + 300}$$

$$= \frac{20 + 10}{700}$$

$$= \frac{30}{700}$$

$$= 0.0428$$

(i) $H_0 : p_1 = p_2$

$H_1 : p_1 < p_2$

$\therefore q_1 = 1 - p = 1 - 0.0428$

$= 0.957$

ii) 5% L.S.

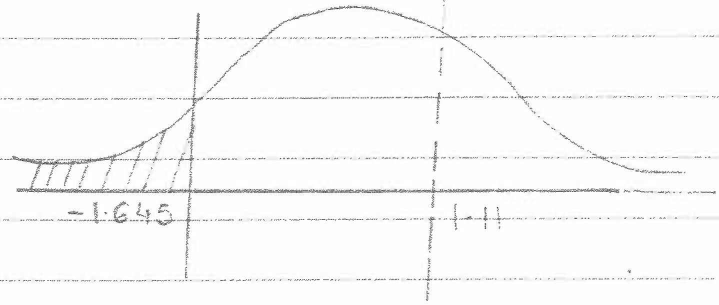
iii) $$Z = \frac{p_1 - p_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{0.05 - 0.033}{\sqrt{(0.042)(0.957)\left(\frac{1}{400} + \frac{1}{300}\right)}}$$

$$= \frac{0.017}{0.0153}$$

$$= \underline{1.11}$$

iv)




v) ∴ $H_0$ is accepted & $H_1$ is rejected

∴ No. of defective protectors before after repair are same.