PROBABILITY AND STATISTICS (PMA303)
Lecture-[29]
(Testing of hypothesis for difference between two population means)
Para. & Non-para., Hypothesis Testing: (Unit VI-VII)

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Dr. Rajanish Rai

Assistant Professor

Department of Mathematics

Thapar Institute of Engineering and Technology, Patiala

many decision involve a comparison of two population means :

- For example, state government is interested to introduced the NEW POLICIES in schools by replacing the offline mode with the online computerized mode. To determine whether policies are significant or not.

μ_1

$\overline{x_1}$

p_1

n_1

μ_2

$\overline{x_2}$

p_2

n_2
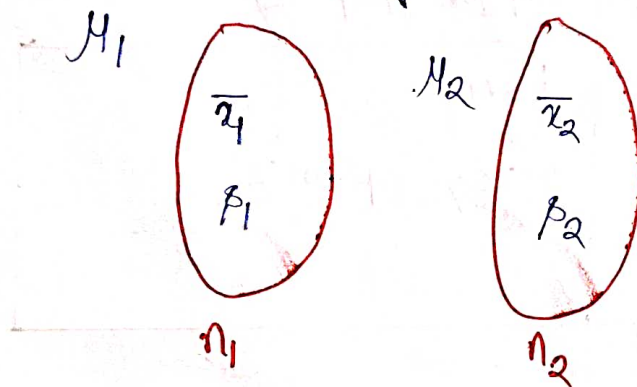
Large — sample statistical test for $\mu_1 - \mu_2$ —:

(i) Define the Hypothesis —:

Null hypothesis —:

$H_0 : \mu_1 - \mu_2 = D$, where

D is some specified difference that you wish to test.

$\left[\begin{array}{c} \text{in most cases} \\ D = 0 \end{array}\right]$

Alternative hypothesis —:

One failed test	Two-failed test
$H_1 : \mu_1 - \mu_2 > D$ or $H_1 : \mu_1 - \mu_2 < D$	$H_1 : \mu_1 - \mu_2 \neq D$.

② Test statistics -:

$$X \sim N(M_1, \sigma_1^2) \qquad Y \sim N(M_2, \sigma_2^2)$$
$$\bar{X} \sim N\left(M_1, \frac{\sigma_1^2}{n_1}\right) \qquad \bar{Y} \sim N\left(M_2, \frac{\sigma_2^2}{n_2}\right)$$
$$\bar{X} - \bar{Y} \sim N\left(M_1 - M_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - D}{\sqrt{\frac{\beta_1^2}{n_1} + \frac{\beta_2^2}{n_2}}} = \frac{\bar{x}_1 - \bar{x}_2 - (M_1 - M_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$
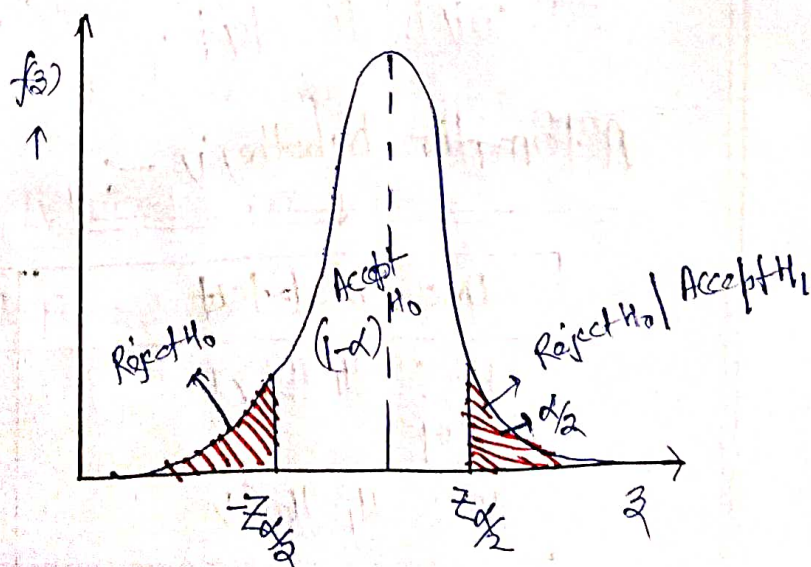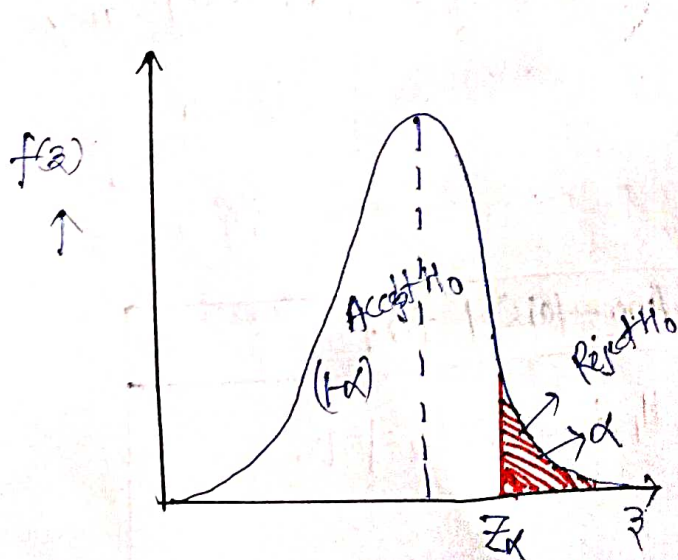
or   p-value

For one-tailed test	For two tailed test
$p\text{-value} = P(Z > z)$ or $P(Z < -z)$	$p\text{-value} = P(Z > z) + P(Z < -z)$

③ Rejection region -:

Here α is the level of significance

One tailed test	Two-tailed test
$Z > Z_\alpha$ or $Z < -Z_\alpha$	$Z > Z_{\alpha/2}$ or $Z < -Z_{\alpha/2}$

④ **Conclusion —:**

- Reject H_0 and Conclude that H_1 is true.
- Accept (do not reject) H_0 is true.

Question —: An investigation of relative merits of two kinds of FLASHLIGHT BATTERIES showed that a random sample of 100 batteries of brand A lasted on the average of 36.5 hours with a S.D. of 1.8 hours, while a random sample of 80 batteries of brand B lasted on the average of 36.8 hours with a S.D. of 1.5 hours. Use 5% level of significance ($Z_\alpha = 1.96$); test whether the observed difference between the average life times is significant.

Solution —: Given that

Sample	Size (n)	mean ($\bar{x}$)	S.D. (s)
A	100	36.5	1.8
B	80	36.8	1.5

(i) **Define the Hypothesis —:**

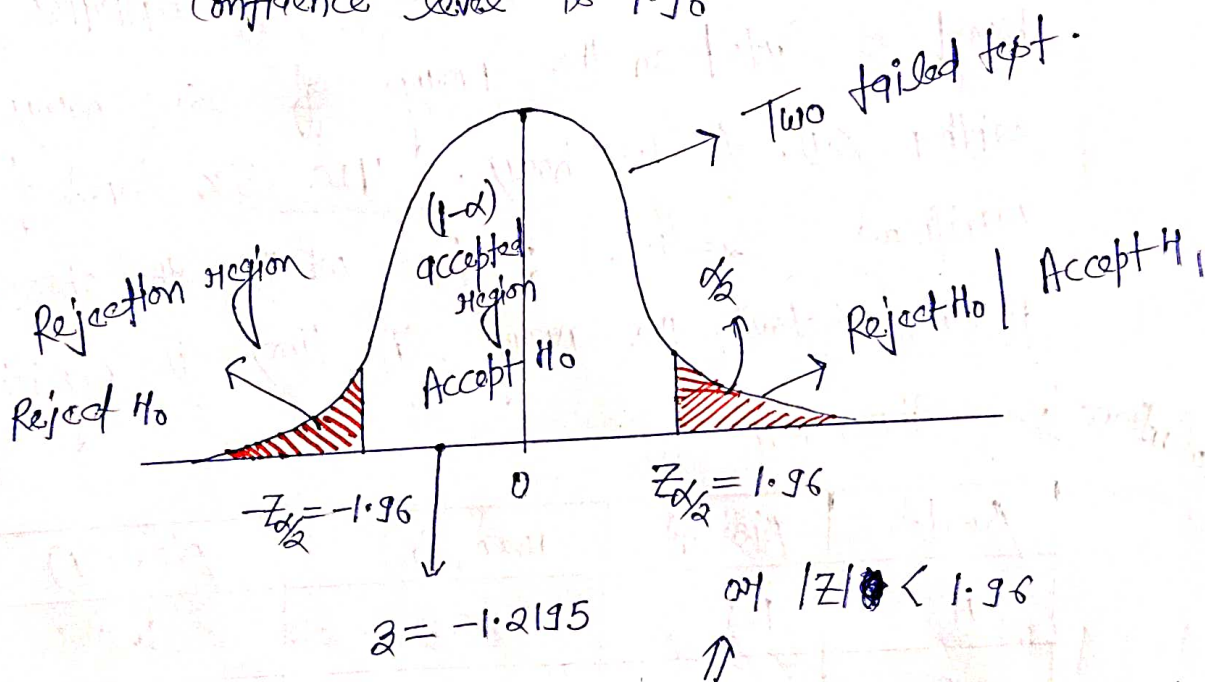Null Hypothesis : H_0 : $M_1 - M_2 = 0$

Vs

∅ Alternative hypothesis H_1 : $M_1 - M_2 \neq 0$ (Two tailed test)

② <u>Test - statistics-:</u>

$$Z = \dfrac{\overline{x_1} - \overline{x_2}}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} = \dfrac{36.5 - 36.8}{\sqrt{\dfrac{(1.8)^2}{100} + \dfrac{(1.5)^2}{80}}}$$

$$\boxed{Z = -1.2195}$$

③ <u>Rejection region -:</u> In two failed test, critical value of 95%.
Confidence level is 1.96

Two failed test.

Rejection region
Reject H_0

$(1-\alpha)$
accepted region
Accept H_0

$\frac{d}{2}$
Reject H_0 / Accept H_1

$-Z_{d/2} = -1.96$ 0 $Z_{d/2} = 1.96$

$2 = -1.2195$

or $|Z| < 1.96$

④ <u>Conclusion-:</u> Since $-1.2195 > -1.96$, which fail to reject H_0. Hence we conclude that, H_0 is accepted, i.e., there is no <u>significant</u> difference between the average life times of brand. A and B batteries.

Remark :—

Revisit the previous example and suppose that an investigation will be concerned only if they detected a _difference of more than 0.2 hours_ in the lives of two batteries. Based on your Confidence interval 95%, Is the statistical significance in this case? Explain.
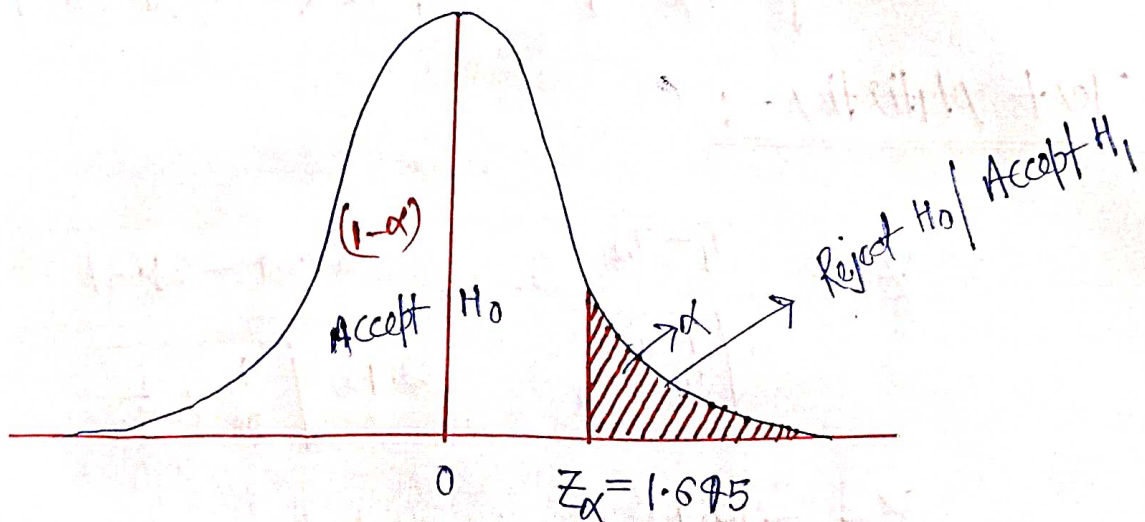
(i) $H_0 : M_1 - M_2 = 0.2$ (Null Hypothesis)

$H_1 : M_1 - M_2 > 0.2$ (Alternative hypothesis)
$$\overset{\shortparallel}{D}$$
One tailed test

For 95% C.I $\Rightarrow Z_\alpha = 1.645$

(ii) $Z = \dfrac{\overline{x}_1 - \overline{x}_2 - D}{\sqrt{\dfrac{\beta_1^2}{n_1} + \dfrac{\beta_2^2}{n_2}}}$ (do it).

$(1-\alpha)$

Accept H_0

Reject H_0 / Accept H_1

0 $Z_\alpha = 1.645$

If 60 B. Tech Computer science student are found to have mean height of 63.60 inches and 50 B. Tech mechanical students a mean height of 69.51 inches. Would you conclude that mechanical students are taller than Computer science students? Assume that the standard deviation of height of under-graduate students to be 2.48 inches. Use 5% level of significance $\left(Z_\alpha = 1.645, \text{ one failed test} \right)$.

Solution —:

Given that

$$\overline{x_1} = 63.60 \qquad n_1 = 60$$

$$\overline{x_2} = 69.51 \qquad n_2 = 50$$

large sample $(n > 30)$.

$$6 = 2.48$$

(i) **Define the Hypothesis —:**

Null hypothesis H_0 : $\quad M_1 - M_2 = 0$

Alternative hypothesis H_1 : $\quad M_1 - M_2 < 0 \quad \left(\text{One failed test} \right)$.
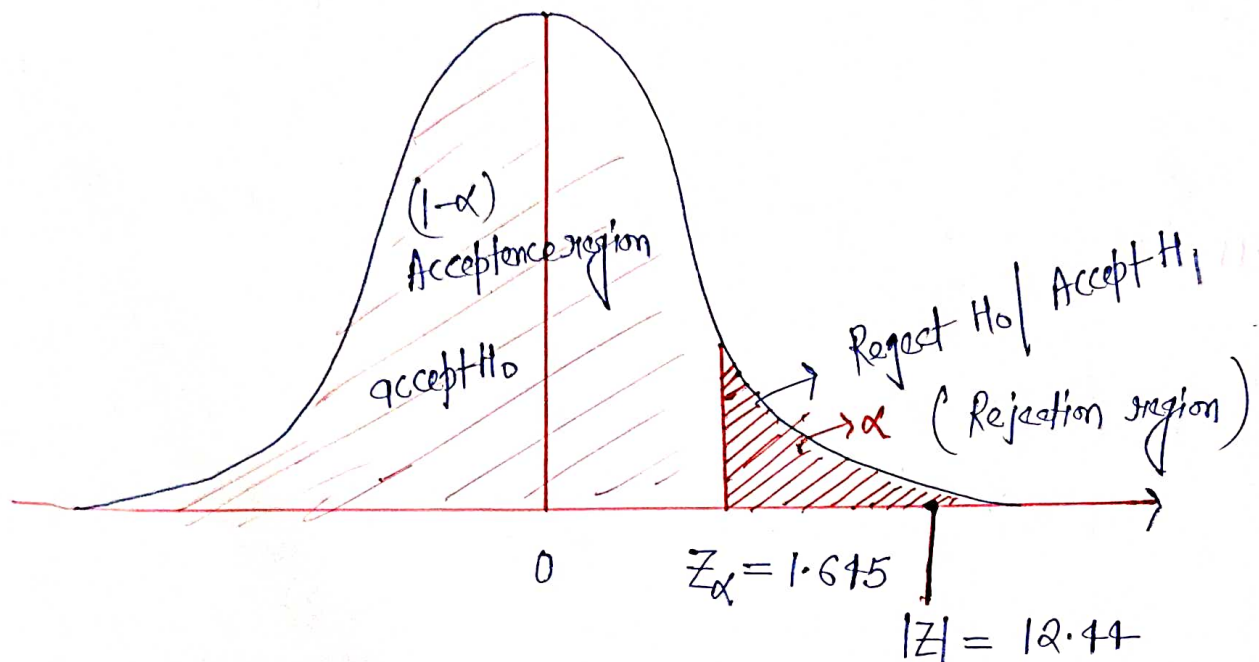
(ii) **Test statistics —:**

$$Z = \frac{\overline{x_1} - \overline{x_2}}{6\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{63.60 - 69.51}{2.48\sqrt{\frac{1}{60} + \frac{1}{50}}}$$

$$Z = -12.44 \quad i.e., \quad |Z| = 12.44.$$

(3) Rejection region -:

For one tail, 95% confidence interval, the critical value $Z_\alpha = 1.645$.

$|Z| = 12.44$

(4) Conclusion -:

Since $|Z| > Z_\alpha = 1.645$, so H_0 is rejected, and we conclude that H_1 is accepted. That is yes, we can conclude that there is highly significant difference between their mean height, i.e., machenical students are taller than Computer science students. #