

# Chapter(11)

## Tow-Samples Tests of Hypothesis (Examples)

# Test of hypothesis about the difference between Two population means Independent samples

This lecture by HesaAlawwad

**Step (1): State the Null ( $H_0$ ) and alternate ( $H_1$ ) hypothesis**

**Case1:**

$$\begin{array}{ll} H_0 : \mu_1 \leq \mu_2 & \text{Or} \quad H_0 : \mu_1 - \mu_2 \leq 0 \\ H_1 : \mu_1 > \mu_2 & H_1 : \mu_1 - \mu_2 > 0 \end{array}$$

**Case2:**

$$\begin{array}{ll} H_0 : \mu_1 \geq \mu_2 & \text{Or} \quad H_0 : \mu_1 - \mu_2 \geq 0 \\ H_1 : \mu_1 < \mu_2 & H_1 : \mu_1 - \mu_2 < 0 \end{array}$$

**Case3:**

$$\begin{array}{ll} H_0 : \mu_1 = \mu_2 & \text{Or} \quad H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 \neq \mu_2 & H_1 : \mu_1 - \mu_2 \neq 0 \end{array}$$

**Step (2): Select a level of significance.**

**Step (3): Select the Test Statistic (computed value)**

I. with known population standard deviation  $\sigma$ .

$$\text{Test statistic} = Z_c = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

II.: with unknown population standard deviation,

(  $n_1 + n_2 - 2 \geq 30$  &  $\sigma_1 = \sigma_2$  ).

$$\text{Test statistic } Z_c = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{\bar{X}_1 - \bar{X}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

III. with unknown population standard deviation,

(  $n_1 + n_2 - 2 < 30$  &  $\sigma_1 = \sigma_2$  ).

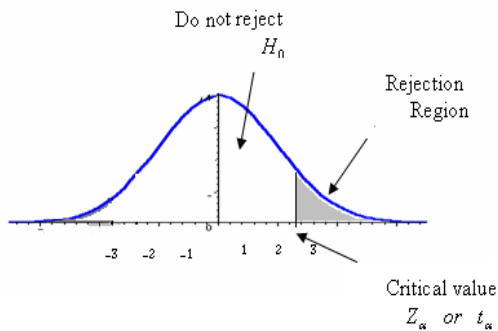
$$\text{Test statistic} = t_c = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{\bar{X}_1 - \bar{X}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

#### Step (4): Selected the Critical value

Types of test	Z	t
The one – tailed test (Right)	$Z_{\alpha}$	$t_{(\alpha, n_1+n_2-2)}$
The one – tailed test (left)	$-Z_{\alpha}$	$-t_{(\alpha, n_1+n_2-2)}$
The two – tailed test	$\pm Z_{\alpha/2}$	$\pm t_{(\frac{\alpha}{2}, n_1+n_2-2)}$

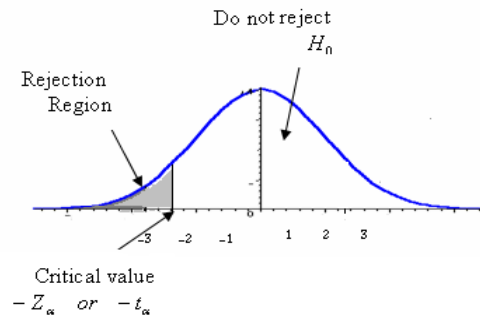
#### Step (5): Formulate the Decision Rule and Make a Decision

**Case1:** Reject  $H_0$  if  $Z_c > Z_{\alpha}$  ,  $t_c > t_{v,\alpha}$



**Case2:** Reject  $H_0$  if  $|Z_c| > Z_{\alpha}$  ,  $|t_c| > t_{v,\alpha}$

This means  $Z_c < -Z_{\alpha}$  ,  $t_c < -t_{v,\alpha}$

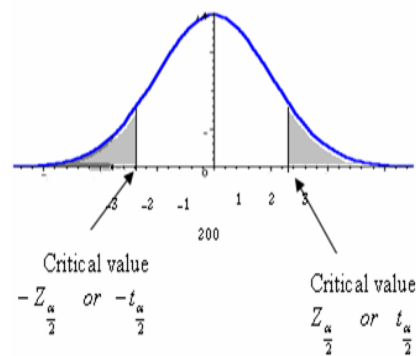


**Case3:** Reject  $H_0$  if  $|Z_c| > Z_{\frac{\alpha}{2}}$  ,  $|t_c| > t_{v,\frac{\alpha}{2}}$  which means:

$$Z_c > Z_{\frac{\alpha}{2}} , t_c > t_{v,\frac{\alpha}{2}}$$

or

$$Z_c < -Z_{\frac{\alpha}{2}} , t_c < -t_{v,\frac{\alpha}{2}}$$



### Example (1)

Customers at Food Town Super Markets have a choice when paying for their groceries. They may check out and pay using the standard cashier – assisted checkout, or they may use the new Fast Lane procedure. In the standard procedure a Food-Town employee puts it on a short conveyor where another employee puts it in a bag and then into the grocery cart. In the Fast Lane procedure the customer scans each item, bags it, and places the bags in the cart themselves.

The Fast Lane procedure is designed to reduce the time a customer spends in the checkout line.

The Fast Lane facility was recently installed at the Byrne RoadFoodTownlocation. The store manger would like to know if the mean checkout time using the standard checkout method **is longer** than using the Fast Lane. She gathered the following sample information. The time is measured from when the customer enters the line until their bags are in the cart. Hence the time includes both waiting in line and checking out. With **0.01 significant level** test the hypothesis for difference between two means for the following:

Customer type	Sample mean	Population $\sigma$	Sample size
Standard	5.50 minutes	0.40 minutes	50
Fast Lane	5.30 minutes	0.30 minutes	100

**Compute the p-value?**

**Solution:**

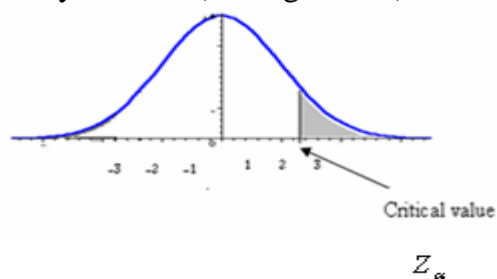
**Step 1-state the null hypothesis**

$$H_0 : \mu_1 \leq \mu_2$$

$$H_1 : \mu_1 > \mu_2$$

This one-tailed test (right)

The key word is (is longer than)



**Step 2: Select the level of significance.**

$\alpha = 0.01$  as stated in the problem

**Step 3: Select the test statistic.**

Use Z-distribution since the assumptions are met

$$\bar{X}_1 = 5.50 \quad , \quad \sigma_1 = 0.40 \quad , \quad n_1 = 50$$

$$\bar{X}_2 = 5.30 \quad , \quad \sigma_2 = 0.30 \quad , \quad n_2 = 100$$

$$\text{Then } Z_c = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{5.5 - 5.3}{\sqrt{\frac{0.40^2}{50} + \frac{0.30^2}{100}}} = \frac{0.2}{0.064} = 3.13$$

**Step 4: Formulate the decision rule. (Critical value)**

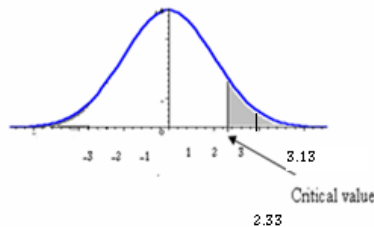
$$Z_{0.01} = 2.33$$

$$0.5 - 0.01 = 0.4900$$

Reject  $H_0$  if  $Z_c > 2.33$

**Step 5: Make a decision and interpret the result.**

$$Z_c = 3.13 > 2.33$$



The decision is to Reject  $H_0$  at significant level 0.01 which means that we got significant result that the difference of 0.20 minutes between the mean checkout time using the standard method and using the Fast Lane it is not due to chance but it is significantly different, we conclude that the Fast Lane methods is faster.

The P-value for the test statistic is:

$$p\text{-value} = P(Z > 3.13) = 0.5 - 0.4991 = 0.0009$$

$$p\text{-value} = 0.0009 < \alpha = 0.01$$

$\therefore$  Reject the  $H_0$ .

**Example (2)**

A manufacturer suspects a difference in the quality of the spare parts he receives from two suppliers. He obtains the following data on service life of random samples of parts from two suppliers,

For suppliers A:  $n_1 = 50$  ,  $\bar{X}_1 = 150$  ,  $S_1 = 10$

For suppliers B:  $n_2 = 100$  ,  $\bar{X}_2 = 153$  ,  $S_2 = 5$

Test whether the difference between the 2 samples is statistically significant at the 1% level of significance. ( $\sigma_1 = \sigma_2$ )

Solution:

**Step 1: state the hypothesis:**

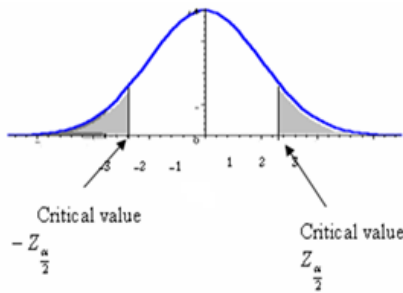
$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

Test is two-tailed test (key word is difference between 2 samples)

**Step 2- Select the level of significance.**

$\alpha = 0.01$  as stated in the problem



### Step 3: Select the test statistic.

Use Z-distribution since the assumptions are met

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = \frac{49 \times 100 + 99 \times 25}{50 + 100 - 2} = \frac{4900 + 2475}{148} = \frac{7375}{148} = 49.8311$$

$$S_p = 7.0591$$

$$\begin{aligned} Z_c &= \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{\bar{X}_1 - \bar{X}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{150 - 153}{7.0591 \sqrt{\frac{1}{50} + \frac{1}{100}}} = \frac{-3}{7.0591 \sqrt{0.02 + 0.01}} = \frac{-3}{7.0591 \sqrt{0.03}} \\ &= \frac{-3}{(7.0591)(0.1732)} \\ &= \frac{-3}{1.2227} = -2.4536 \end{aligned}$$

### Step 4: Formulate the decision rule. (Critical value)

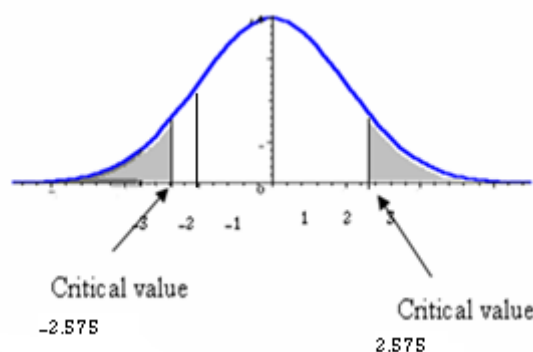
$$Z_{\frac{\alpha}{2}} = Z_{\frac{0.01}{2}} = Z_{0.005} = \pm 2.58$$

$$0.5 - 0.005 = 0.495$$

Reject  $H_0$  if  $Z_c > 2.575$  Or  $Z_c < -2.575$

### Step 5: Make a decision and interpret the result.

Reject  $H_0$  if  $Z_c > Z_{\frac{\alpha}{2}}$  Or  $Z_c < -Z_{\frac{\alpha}{2}}$



Since  $-2.575 < -2.4536 < 2.575$  so we don't reject  $H_0$  at significant level 0.01 and conclude that there is no significant difference between the two samples.

$$P\text{-value} = 2P(Z > 2.4536) = 2(0.5 - 0.49286) = 2(0.00714) = 0.01428$$

$$P\text{-value} = 0.01428 > \alpha = 0.01$$

$\therefore$  Don't reject  $H_0$ .

**Example (3)** Owens Lawn Care, Inc., manufactures and assembles lawnmowers that are shipped to dealers throughout the United States and Canada. Two different procedures have been proposed for mounting the engine on the frame of the lawnmower. The question is: Is there a difference in the mean time to mount the engines on the frames of the lawnmowers? The first procedure was developed by longtime Owens employee Herb Welles (designated as procedure 1), and the other procedure was developed by Owens Vice President of Engineering William Atkins (designated as procedure 2). To evaluate the two methods, it was decided to conduct a time and motion study.

A sample of five employees was timed using the Welles method and six using the Atkins method. The results, in minutes, are shown below. **Is there a difference** in the mean mounting times? Use the 0.10 significance level. ( $\sigma_1 = \sigma_2$ )

Methods	Sample mean	Sample S	Sample size
Welles Method	4 minutes	2.92 minutes	5
Atkins Method2	5 minutes	2.09 minutes	6

**Solution:**

**Step 1—state the null hypothesis**

$$H_0 : \mu_1 = \mu_2$$

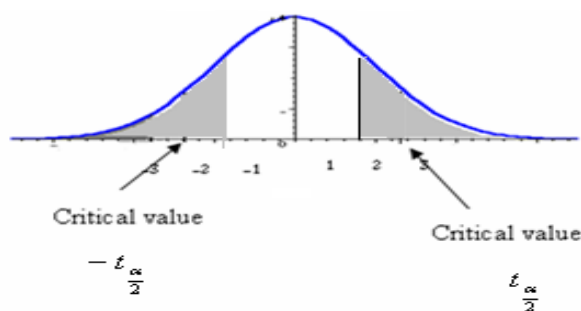
$$H_1 : \mu_1 \neq \mu_2$$

This two-tailed test

The key word is (Is there a difference)

**Step 2: Select the level of significance.**

$\alpha = 0.10$  as stated in the problem



**Step 3: Select the test statistic.**

Use t-distribution since the assumptions are met

$$\bar{X}_1 = 4, \quad S_1 = 2.92, \quad n_1 = 5$$

$$\bar{X}_2 = 5, \quad S_2 = 2.09, \quad n_2 = 6$$

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = \frac{4 \times 8.53 + 5 \times 4.37}{9} = \frac{34.12 + 21.85}{9} = \frac{55.97}{9} = 6.22$$

$$S_p = 2.49$$

$$t_c = \frac{\bar{X}_1 - \bar{X}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{4 - 5}{(2.49) \sqrt{\frac{1}{5} + \frac{1}{6}}} = \frac{-1}{(2.49)(0.609)} = \frac{-1}{1.5146} = -0.662$$

**Step 4: Formulate the decision rule.**

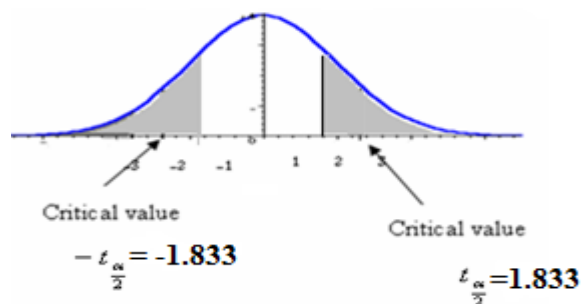
$$\pm t\left(\frac{\alpha}{2}, n_1 + n_2 - 2\right) = \pm t\left(\frac{0.10}{2}, 5 + 6 - 2\right) = \pm t(0.05, 9) = \pm 1.833$$

**Step 5: Make a decision and interpret the result.**

Reject  $H_0$  if  $t_c > t_{\alpha/2}$  or  $t_c < -t_{\alpha/2}$

Or

$$t_c < -t_{\alpha/2} \text{ or } t_c > t_{\alpha/2}$$



The decision is not to reject the null hypothesis, because 0.66 falls in **the region between -1.833 and 1.833.**



# Test of hypothesis about the difference between Two population means

## (Dependent or paired samples)

**Purpose is to compare** means of two non- independent samples. Two ways to obtain related samples:

**1- Repeated measures designs:** for, measurements on the same individuals before and after a treatment.

e.g., Memory retention in quiet and noisy environment (on same observations)

Measure heart rate before & after exercise (on same observations)

**2- Matched samples :** each Population in one sample matched, on specific variable(s), with a Population in another sample

**We will use the five steps Hypotheses – testing procedure:**

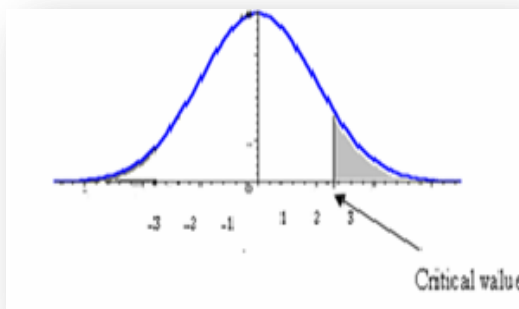
**Step (1): State the null and alternate hypotheses:**

Case1: One-tailed test (Right)

Upper tail critical (when  $X_1 > X_2$ )

$H_0: \mu_d \leq 0$

$H_1: \mu_d > 0$

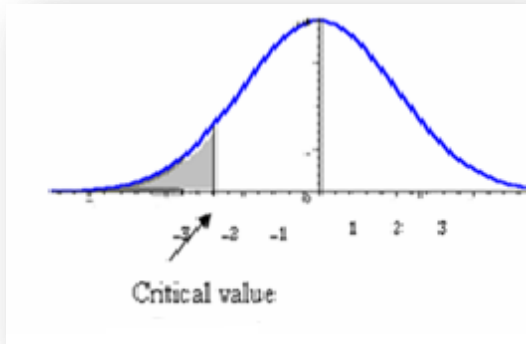


**Case2: One-tailed test (Left)**

Lower tail critical (when  $X_1 < X_2$ )

$H_0: \mu_d \geq 0$

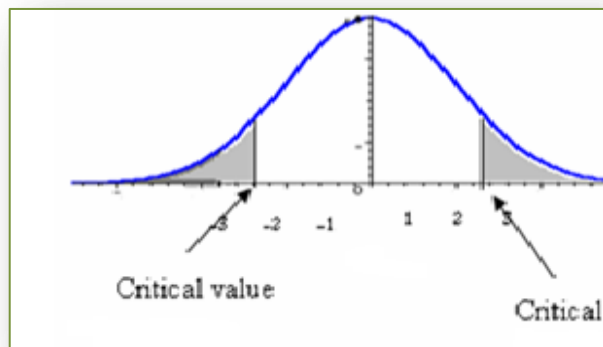
$H_1: \mu_d < 0$



**Case3: (Two-tailed test)**

$$H_0: \mu_d = 0$$

$$H_1: \mu_d \neq 0$$



**Step (2): Select the level of significance ( $\alpha$ )**

**Step (3): The test statistic**

$$t = \frac{\bar{d}}{S_d / \sqrt{n}}$$

$d_i$  : The **difference** between the paired or related observations.

$$d_i = (X_{i1} - X_{i2})$$

$\bar{d}$  : The mean of the difference between the paired or related observations.

$S_d$  : The standard deviation of the paired or related observations.

$$S_d = \sqrt{\frac{\sum (d - \bar{d})^2}{n - 1}} = \sqrt{\frac{\sum d^2 - \frac{(\sum d)^2}{n}}{n - 1}}$$

$n$  : Number of paired differences.

**Step (4): The critical value:**

Types of test	critical value ( t )
The one – tailed test (Right)	$t_{(\alpha, n-1)}$
The one – tailed test (left)	$t_{(-\alpha, n-1)}$
The two – tailed test	$\pm t_{(\frac{\alpha}{2}, n-1)}$

**Step (5): Formulate the Decision Rule**

**Case1:** Reject  $H_0$  if  $t_c > t_{(\alpha, n-1)}$

**Case2:** Reject  $H_0$  if  $t_c > -t_{(\alpha, n-1)}$

**Case3:** Reject  $H_0$  if  $t_c > t_{(\alpha/2, n-1)}$  or  $t_c < -t_{(\alpha/2, n-1)}$

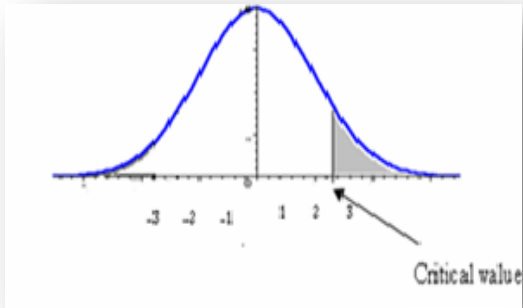
**Example (4)**

Advertisements by SylphFitnessCenter claim that completing its course will result in losing weight. A random sample of eight recent participants showed the following weights before and after completing the course. At the 0.01 significance level, can we conclude the students lost weight (in pounds)  
Note: 1 kg = 2.20 pounds

No	Before	After
1	155	154
2	228	207
3	141	147
4	162	157
5	211	196
6	164	150
7	184	170
8	172	165

**Solution:****Step (1): State the null and alternate hypotheses**

$$H_0 : \mu_d \leq 0 \quad H_1 : \mu_d > 0$$



**Step (2): Select the level of significance ( $\alpha=0.01$ )**

**Step (3): The test statistic**

No	Before B	After A	d (B-A)	$d^2$
1	155	154	1	1
2	228	207	21	441
3	141	147	-6	36
4	162	157	5	25
5	211	196	15	225
6	164	150	14	196
7	184	170	14	196
8	172	165	7	49
Total			71	1169

$$\begin{aligned}
 S_d &= \sqrt{\frac{\sum d^2 - \frac{(\sum d)^2}{n}}{n-1}} = \sqrt{\frac{1169 - \frac{(71)^2}{8}}{7}} = \sqrt{\frac{1169 - 630.125}{7}} = \sqrt{\frac{538.875}{7}} \\
 &= \sqrt{76.9821} = 8.774 \\
 \bar{d} &= \frac{\sum d}{n} = \frac{71}{8} = 8.875 \\
 t_c &= \frac{\bar{d}}{S_d / \sqrt{n}} = \frac{8.875}{8.775 / \sqrt{8}} = \frac{8.875}{8.775 / 2.8284} = \frac{8.875}{3.10246} = 2.861
 \end{aligned}$$

**Step (4): The critical value**

In one –tailed test (Right)

$$t_{(\alpha, n-1)} = t_{(0.01, 7)} = 2.998$$

**Reject  $H_0$  if**

$$t_c > t_{0.01,7}$$

**Step (5): Formulate the Decision Rule**

Do not reject  $H_0$ . We cannot conclude that the students lost weight

# Two-Samples Tests of Hypothesis

## Test of hypothesis about the difference between Two population Proportions

We investigate whether two samples came from populations with an equal proportion of successes.

**Step (1): State the Null ( $H_0$ ) and alternate ( $H_1$ ) hypothesis**

$$\text{Case 1: } \begin{array}{l} H_0 : \pi_1 \geq \pi_2 \\ H_1 : \pi_1 < \pi_2 \end{array} \quad \text{OR} \quad \begin{array}{l} H_0 : \pi_1 - \pi_2 \geq 0 \\ H_1 : \pi_1 - \pi_2 < 0 \end{array}$$

$$\text{Case 2: } \begin{array}{l} H_0 : \pi_1 \leq \pi_2 \\ H_1 : \pi_1 > \pi_2 \end{array} \quad \text{OR} \quad \begin{array}{l} H_0 : \pi_1 - \pi_2 \leq 0 \\ H_1 : \pi_1 - \pi_2 > 0 \end{array}$$

$$\text{Case 3: } \begin{array}{l} H_0 : \pi_1 = \pi_2 \\ H_1 : \pi_1 \neq \pi_2 \end{array} \quad \text{OR} \quad \begin{array}{l} H_0 : \pi_1 - \pi_2 = 0 \\ H_1 : \pi_1 - \pi_2 \neq 0 \end{array}$$

**Step (2): Select a level of significance.**

**Step (3): Select the Test Statistic (computed value)**

The value of the test statistic is computed from the following formula.

$$Z_c = \frac{P_1 - P_2}{\sqrt{\frac{P_p(1-P_p)}{n_1} + \frac{P_p(1-P_p)}{n_2}}}$$

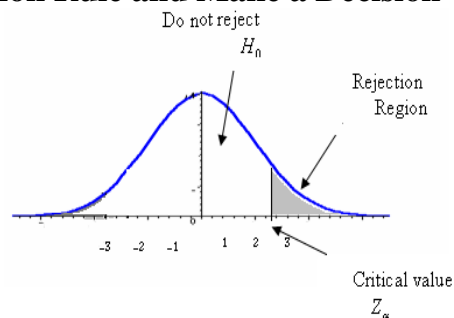
The two samples are pooled using the following formula.

$$P_p = \frac{X_1 + X_2}{n_1 + n_2}$$

**Step (4): Selected the Critical value**

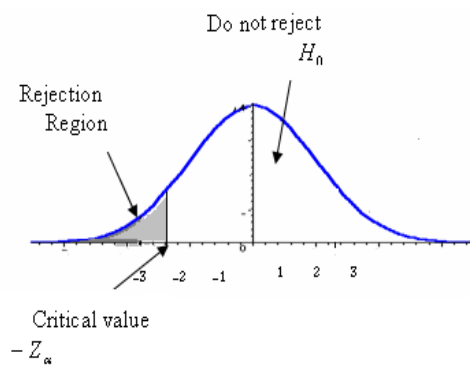
The one – tailed test (Right)	$Z_\alpha$
The one – tailed test (left)	$-Z_\alpha$
The two – tailed test	$\pm Z_{\alpha/2}$

**Step (5): Formulate the Decision Rule and Make a Decision**



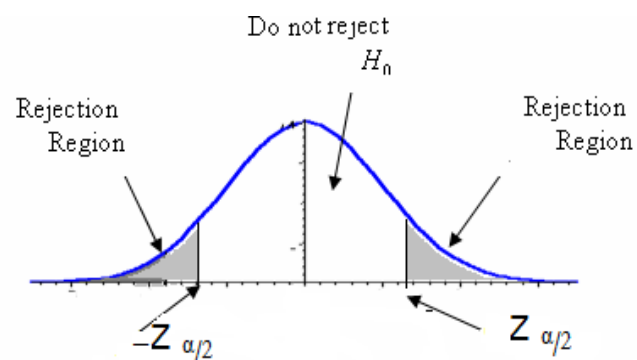
**Case1:** Reject  $H_0$  if  $Z_c > Z_\alpha$

**Case2:** Reject  $H_0$  if  $|Z_c| > Z_\alpha$  or  $Z_c < -Z_\alpha$



**Case3:** Reject  $H_0$  if  $|Z_c| > Z_{\frac{\alpha}{2}}$  which mean

$$Z_c > Z_{\frac{\alpha}{2}} \quad \text{or} \quad Z_c < -Z_{\frac{\alpha}{2}}$$



### Example (5)

Manila Perfume Company recently developed a new fragrance that it plans to market under the name Heavenly. A number of market studies indicate that Heavenly has very good market potential. The Sales Department at Minnelli is particularly interested in whether there is a difference in the proportions of younger and older women who would purchase Heavenly if it were marketed. There are two independent populations; a population consisting of the 100 younger women, 19 of them liked the Heavenly and a population consisting of the 200 older women. 62 of them liked Heavenly. Each sampled woman will be asked to smell Heavenly and indicate whether she likes the fragrance well enough to purchase a bottle. (With significant level 0.05)

#### Solution:

**Step 1: State the null and alternate hypotheses.**

$$H_0: \pi_1 = \pi_2$$

$$H_1: \pi_1 \neq \pi_2$$

**Step 2: State the level of significance.**

The .05 significance level is stated in the problem.

**Step 3: Find the appropriate test statistic.**

We will use the z-distribution

let  $P_1$  = young women

$$P_1 = \frac{X_1}{n_1} = \frac{19}{100} = 0.19$$

let  $P_2 = \text{Older women}$

$$P_2 = \frac{X_2}{n_2} = \frac{62}{200} = 0.31$$

$$P_p = \frac{X_1 + X_2}{n_1 + n_2} = \frac{19 + 62}{100 + 200} = \frac{81}{300} = 0.27$$

$$Z_c = \frac{P_1 - P_2}{\sqrt{\frac{P_p(1-P_p)}{n_1} + \frac{P_p(1-P_p)}{n_2}}} = \frac{0.19 - 0.31}{\sqrt{\frac{0.27(1-0.27)}{100} + \frac{0.27(1-0.27)}{200}}} = -2.21$$

**Step 4: Selected the Critical value.**

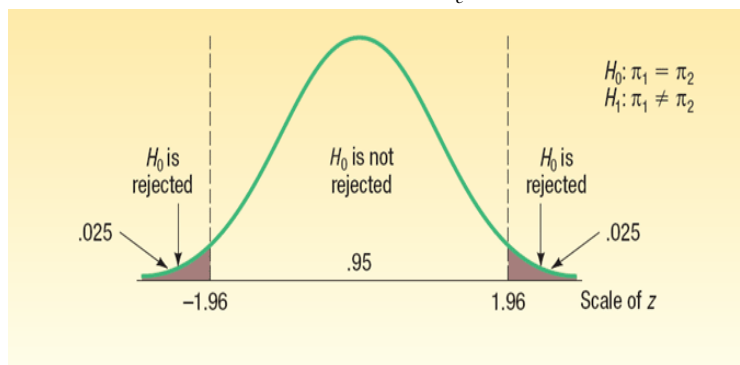
$$\pm Z_{\frac{\alpha}{2}} = \pm Z_{\frac{0.05}{2}} = \pm Z_{0.025}$$

$$0.5 - 0.025 = 0.475$$

$$\pm Z_{0.025} = \pm 1.96$$

**Step 5: State the decision rule**

$$\begin{array}{l} \text{Reject } H_0 \text{ if } Z_c > Z_{\frac{\alpha}{2}} \quad \text{Or} \quad Z_c < -Z_{\frac{\alpha}{2}} \\ Z_c > 1.96 \quad \text{Or} \quad Z_c < -1.96 \end{array}$$



Whereas  $Z_c < -1.96$ ,  $-2.21 < -1.96$

Then Reject  $H_0$  at significant level 0.05

The computed value of 2.21 is in the area of rejection. Therefore, the null hypothesis is rejected at the .05 significance level. To put it another way, we reject the null hypothesis that the proportion of young women who would purchase Heavenly is equal to the proportion of older women who would purchase Heavenly.