

# Chapter 4

## Tests of Hypothesis

### 4.1 statistical hypothesis

A statistical hypothesis is a statement concerning one population or more.

#### 4.1.1 The Null and The Alternative Hypothesis:

The structure of hypothesis testing will be formulated with the use of the term null hypothesis. This refers to any hypothesis we wish to test that called  $H_0$ . The rejection of  $H_0$  leads to the acceptance of an alternative hypothesis denoted by  $H_1$ . A null hypothesis concerning a population parameter,  $\theta$  will always be stated so as to specify an exact value of the parameter, whereas the alternative hypothesis allows for the possibility of several values. We usually test the null hypothesis:  $H_0 : \theta = \theta_0$  against one of the following alternative hypothesis:

$$H_1 : \left\{ \begin{array}{l} \theta \neq \theta_0 \\ \theta > \theta_0 \\ \theta < \theta_0 \end{array} \right\}$$

### 4.1.2 Types of Errors:

#### (i) Type One Error I:

Rejection of the null hypothesis when it is true is called a type I error. The probability of committing a type I error also called the level of significance which is denoted by  $\alpha$ . Sometimes  $\alpha$  is called the size of the critical region or the size of the test.

#### (ii) Type Two Error II:

Acceptance of the null hypothesis when it is false is called a type II error, which is denoted by  $\beta$ .

#### **Possible situations in testing a statistical hypothesis**

	$H_0$ is true	$H_0$ is false
Accept $H_0$	Correct decision	Type II error, $\beta$
Reject $H_0$	Type I error, $\alpha$	Correct decision

**Type I error:** rejecting  $H_0$  when  $H_0$  is true.

**Type II error:** accepting  $H_0$  when  $H_0$  is false.

**P (Type I error)** = P (rejecting  $H_0$  |  $H_0$  is true) =  $\alpha$ .

**P (Type II error)** = P (accepting  $H_0$  |  $H_0$  is false) =  $\beta$ .

Ideally we like to use a test procedure for which both the type I and type II errors are small. It is noticed that a reduction in  $\beta$  is always possible by increasing the size of the critical region,  $\alpha$ .

For a fixed sample size, decrease in the probability of one error will usually result in an increase in the probability of the other error.

Fortunately the probability of committing both types of errors can be reduced by increasing the sample size.

### **(iii) One – Tailed and Two – Tailed test:**

A test of any statistical hypothesis where the alternative is one – sided such as:

$$H_0 : \theta = \theta_0 \quad \text{vs} \quad H_1 : \theta > \theta_0 \\ \text{or} \quad H_1 : \theta < \theta_0$$

is called a one – tailed test. The critical region for the alternative hypothesis  $\theta > \theta_0$  lies entirely in the right tail of the distribution while the critical region for the alternative hypothesis  $\theta < \theta_0$  lies entirely in the left tail.

$$H_1 : \theta > \theta_0$$

$$H_1 : \theta < \theta_0$$

A test of any statistical hypothesis where the alternative is two – sided, such as:

$$H_0 : \theta = \theta_0 \quad \text{vs} \quad H_1 : \theta \neq \theta_0$$

is called two – tailed test since the critical region is split into two parts having equal probabilities placed in each tail of the distribution of the test statistic.

## 4.2 P – Values in Decision Making:

A p – value is the lowest level (of significance) at which the observed value of the test statistic is significant.

$$P\text{-value} = 2P(Z > |Z_{obs}|) \text{ when } H_1 \text{ is as follows: } H_1 : \theta \neq \theta_0$$

$$p\text{-value} = P(Z > Z_{obs}) \text{ when } H_1 \text{ is as follows: } H_1 : \theta > \theta_0$$

$$p\text{-value} = P(Z < Z_{obs}) \text{ when } H_1 \text{ is as follows: } H_1 : \theta < \theta_0$$

$H_0$  is rejected if  $\alpha \geq P\text{-value}$  otherwise  $H_0$  is accepted.

### Ex (1):

$$H_0 : \mu = 10 \quad \text{vs} \quad H_1 : \mu \neq 10, \quad \alpha = 0.05 \Rightarrow Z = 1.87$$

$$\begin{aligned} P\text{-value} &= 2P(Z > |1.87|) = 2P(Z > 1.87) = 2[1 - P(Z \leq 1.87)] \\ &= 2[1 - 0.9693] = 2(0.0307) = 0.0614 \end{aligned}$$

Since  $P\text{-value} > \alpha$  then  $H_0$  is accepted.

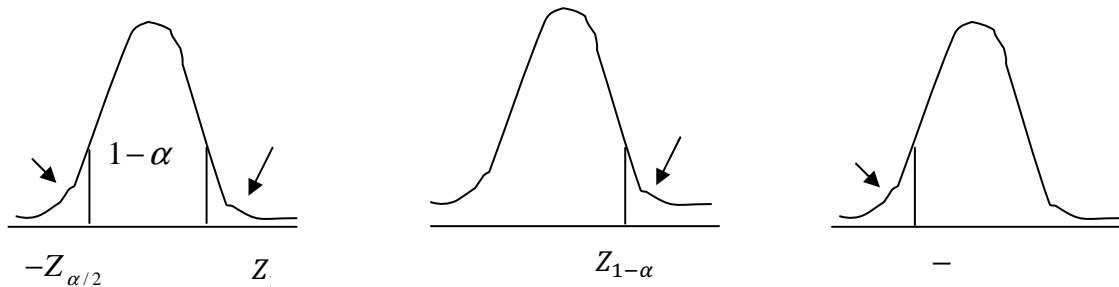
## 4.3 Testing Hypothesis Concerning one Population

We have the following steps to test any hypothesis:

1. Stating the null hypothesis  $H_0 : \theta = \theta_0$ .
2. Choosing an appropriate alternative hypothesis from one of the alternatives  $H_1 : \theta < \theta_0$  or  $\theta > \theta_0$  or  $\theta \neq \theta_0$ .

3. determining the significance level of size  $\alpha = 0.01, 0.025, 0.05$  or  $0.1$

4. Determining the rejection or critical region (R.R.) and the acceptance region (A.R.) of  $H_0$ .



5. Selecting the appropriate test statistic and establish the critical region. If the decision is to be based on a  $p$  – value it is not necessary to state the critical region.

6. Computing the value of the test statistic from the sample data.

7. Decision rule:

rejecting  $H_0$  if the value of the test statistic in the critical (rejection) region or if the  $p$  –value  $\leq \alpha$  also.

accepting  $H_0$  if the value of the test statistic in the acceptance region or if the  $p$  –value  $> \alpha$  also.

### **4.3.1 Testing hypothesis for a Single population Mean**

There are different cases to study the population mean  $\mu$  that is:

#### **(i) Large Samples ( $n > 30$ )**

1-data needed:  $n, \alpha, \mu_0, \sigma$  or  $S, \bar{X}$ .

-type of population(normal or not normal)

- the variance of population is known  $\sigma^2$  or unknown then using  $S$

2- state the hypothesis  $H_0: \mu = \mu_0$  and  $H_1(\mu < \mu_0$  or  $\mu > \mu_0$  or  $\mu \neq \mu_0)$

3- the statistic:

- if  $\sigma$  is known 
$$Z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

- if  $\sigma$  is unknown 
$$Z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$
, where  $Z \sim N(0,1)$

4- Determining the rejection or critical region (R.R.) and the acceptance region (A.R.) of  $H_0$ , that is:

i)if  $H_1: \mu > \mu_0$ , reject  $H_0$  if  $Z > Z_{1-\alpha}$

ii)if  $H_1: \mu < \mu_0$ , reject  $H_0$  if  $Z < Z_\alpha$  ( $Z_\alpha$  always negative)

iii) if  $H_1: \mu \neq \mu_0$ , reject  $H_0$  if  $Z > Z_{\frac{1-\alpha}{2}}$  or  $Z < Z_{\frac{\alpha}{2}}$  (where  $Z_{\frac{\alpha}{2}} = -Z_{\frac{1-\alpha}{2}}$ )

5-Decision rule:

if the rejection rule is hold then we reject  $H_0$  and accept  $H_1$

if the rejection rule isn't hold then we accept  $H_0$  and reject  $H_1$

## **EX(2)**

We take a random sample of 36 apples from an apples farms and found that the average weight of the apple is 190 gm. Assuming that the apples weight in this farm has normal distribution with standard deviation 18

gm. Test that the average weight of apple in this farm less than 200 gm at  $\alpha = 0.01$ .

**Solu.**

$n = 36, \alpha = 0.01, \mu_0 = 200, \sigma = 18, \bar{X} = 190$   
 $H_0: \mu = 200$  against  $H_1: \mu < 200$

$$Z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = Z = \frac{190 - 200}{\frac{18}{\sqrt{36}}} = -3.33$$

$$Z_{\alpha} = Z_{0.01} = -2.33$$

Since  $Z = -3.33 < Z_{\alpha} = Z_{0.01} = -2.33$ , thus we reject  $H_0$  and accept  $H_1$

i.e., the average weight of the apples in this farm is less than 200gm at  $\alpha = 0.01$ .

**Ex (3):**

A random sample of **100** recorded deaths in the United States during the past year showed an average life span of **71.8** years with a standard deviation of **8.9** years. Dose this seem to indicate that the average life span today is greater than **70** years? Use a **0.05** level of significance.

**Solu.**

$n = 100, \alpha = 0.05, \mu_0 = 70, S = 8.9, \bar{X} = 71.8$

$H_0: \mu = 70$  against  $H_1: \mu > 70$

$$Z = \frac{\bar{x} - \mu_0}{\frac{S}{\sqrt{n}}} = Z = \frac{71.9 - 70}{\frac{8.9}{\sqrt{100}}} = 2.02$$

$$Z_{1-\alpha} = Z_{0.95} = 1.645$$

Since  $Z = 2.02 > Z_{1-\alpha} = Z_{0.95} = 1.645$ , thus we reject  $H_0$  and accept  $H_1$   
i.e., the average life span is greater than 70 years at  $\alpha = 0.05$ .

**or**

$$P\text{-value} = P(Z > 2.02) = 1 - P(Z \leq 2.02) = 1 - 0.9783 = 0.0217$$

reject  $H_0$  since  $\alpha > p\text{-value}$

#### **Ex (4):**

A manufacturer of sports equipment has developed a new synthetic fishing line that he claims has a standard deviation of **0.5** kilogram. Test the hypothesis that  $\mu = 8$  kilograms against the alternative that  $\mu \neq 8$  kilograms if a random sample of **50** lines is tested and found to have a mean breaking strength of **7.8** kilograms. Use a **0.01** level of significance.

#### **Solution:**



$$H_0 : \mu = 8 \quad \text{vs} \quad H_1 : \mu \neq 8 \quad , \quad \alpha = 0.01$$

$$n = 50, \bar{X} = 7.8, \sigma = 0.5$$

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{7.8 - 8}{0.5/\sqrt{50}} = -2.82$$

$$Z_{1-\alpha/2} = Z_{0.995} = 2.575 \quad \text{and} \quad -Z_{1-\alpha/2} = -Z_{0.995} = -2.575$$

$$R.R. : Z > 2.575 \quad \text{or} \quad Z < -2.575$$

Since  $Z = -2.83 \in R.R. \rightarrow$  we reject  $H_0$  at  $\alpha = 0.01$

**or**

$$\begin{aligned} p\text{-value} &= 2P(Z > |-2.83|) = 2P(Z > 2.83) \\ &= 2(1 - P(Z \leq 2.83)) = 2(1 - 0.9977) \\ &= 2(0.0023) = 0.0046 \end{aligned}$$

$H_0$  is rejected since  $p\text{-value} \leq \alpha$

### **(ii) Small Samples ( $n < 30$ )**

In small samples we have three cases for the statistic as follows:

1) Normal population and  $\sigma$  is known, then the statistic is:

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

Which has standard normal distribution, when  $H_0$  is true.

2) Normal population and  $\sigma$  is unknown, then the statistic is:

$$T = \frac{\bar{x} - \mu_0}{\frac{S}{\sqrt{n}}}$$

Which has t-distribution with  $(n-1)$  degrees of freedom, when  $H_0$  is true.

Rejection rules for  $H_0$  will be as follows:

i) If  $H_1 : \mu > \mu_0$  then , Reject  $H_0$  if  $T > t_{n-1, 1-\alpha}$

ii) if  $H_1 : \mu < \mu_0$  then , Reject  $H_0$  if  $T < t_{(n-1), \alpha}$  ( where  $t_{(n-1), \alpha} = -t_{n-1, 1-\alpha}$  )

iii) if  $H_1 : \mu \neq \mu_0$  then , Reject  $H_0$  if  $T > t_{n-1, \frac{\alpha}{2}}$  or  $T < t_{n-1, \frac{\alpha}{2}}$

$$\text{where } t_{n-1, \frac{\alpha}{2}} = - t_{n-1, 1-\frac{\alpha}{2}}$$

3)if the population is not Normal, it will be study at the end of this course.

### Ex (5):

If a random sample of **12** homes with a mean  $\bar{X} = 42$  included in a planned study indicates that vacuum cleaners expend an average of **42** kilowatt – hours per year with standard deviation of **11.9** kilowatt hours dose this suggest at the **0.05** level of significance that vacuum cleaners expend on the average less than **46** kilowatt hours annually, assume the population of kilowatt - hours to be normal?

### Solu.

$$H_0 : \mu = 46 \quad \text{vs} \quad H_1 : \mu < 46, \quad \alpha = 0.05$$

$$n = 12, \bar{X} = 42, S = 11.9$$

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{42 - 46}{11.9/\sqrt{12}} = -1.16$$

$$v = n - 1 = 11, -t_{1-\alpha} = -t_{0.95} = -1.796$$

$$R.R.: T < -1.796$$

Since  $T = -1.16 \in A.R. \rightarrow$  we accept  $H_0$  at  $\alpha = 0.05$

accept  $H_0$  since the value of  $t$  is in the acceptance region (A.R.)

### 4.3.2 Testing hypothesis for a Single population Variance

We will always assume that the population is normally distributed. (there is no results for this by using computer).

#### Steps for this test

1- Data :  $\sigma_0^2, n, S^2, \alpha,$

2- hypothesis :  $H_0: \sigma^2 = \sigma_0^2$

$$H_1: \begin{cases} \sigma^2 < \sigma_0^2 \\ \sigma^2 > \sigma_0^2 \\ \sigma^2 \neq \sigma_0^2 \end{cases}$$

3- the statistic:  $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$

4- the table value:  $\begin{cases} \chi_{(n-1), \alpha}^2 & \text{when } H_1: \sigma^2 < \sigma_0^2 \\ \chi_{(n-1), 1-\alpha}^2 & \text{when } H_1: \sigma^2 > \sigma_0^2 \\ \chi_{(n-1), \frac{\alpha}{2}}^2, \chi_{(n-1), (1-\frac{\alpha}{2})}^2 & \text{when } H_1: \sigma^2 \neq \sigma_0^2 \end{cases}$

5- the decision: we reject  $H_0$  if

$$\begin{cases} \chi^2 < \chi_{(n-1),\alpha}^2 & \text{at } H_1: \sigma^2 < \sigma_0^2 \\ \chi^2 > \chi_{(n-1),1-\alpha}^2 & \text{at } H_1: \sigma^2 > \sigma_0^2 \\ \chi^2 < \chi_{(n-1),\frac{\alpha}{2}}^2 \text{ or } \chi^2 > \chi_{(n-1),(1-\frac{\alpha}{2})}^2 & \text{at } H_1: \sigma^2 \neq \sigma_0^2 \end{cases}$$

### Ex (6):

Chemically evaluated irrigation water samples from 14 Qatif wells. The percent of Na cations (الأيون الموجب) in the water was measured:

43, 47, 40, 45, 45, 48, 47, 47, 46, 52, 50, 50, 51, 49

Assuming a normal distribution, test whether the variance of the total Na cations is less than 10 at  $\alpha = 0.05$ .

### Solu.

1- Data :  $\sigma_o^2 = 10, n = 14, \alpha = 0.05, S^2 = 10.5934$

2- hypothesis :  $H_0: \sigma^2 = 10$

$$H_1: \sigma^2 < 10$$

3- the statistic:  $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{(13)(10.5934)}{10} = 13.77$

4- the table value:  $\chi_{(n-1),\alpha}^2 = \chi_{13,0.05}^2 = 5.892$

5- the decision: we reject  $H_0$  if  $\chi^2 < \chi_{(n-1),\alpha}^2 \rightarrow 13.77 \not< 5.892$

Thus, we accept  $H_0$  and reject  $H_1$ , i.e., we will reject the assumption that the variance is less than 10.

### 4.3.3 Testing hypothesis for a Single population Proportion

We will assume that the sample size is large ( $n > 30$ ) and the proportion sample is  $r = \frac{a}{n}$ .

### Steps for this test

1- Data :  $P_o, n, \alpha, r = \frac{a}{n}$

2- hypothesis :  $H_0: P = P_0$

$$H_1: \begin{cases} P < P_0 \\ P > P_0 \\ P \neq P_0 \end{cases}$$

3- the statistic:  $Z = \frac{r - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}}$

4- the table value:  $\begin{cases} Z_\alpha & \text{when } H_1: P < P_0 \\ Z_{1-\alpha} & \text{when } H_1: P > P_0 \\ Z_{1-\frac{\alpha}{2}}, Z_{\frac{\alpha}{2}} & \text{when } H_1: P \neq P_0 \end{cases}$

5- the decision: we reject  $H_0$  if

$$\begin{cases} Z < Z_\alpha & \text{when } H_1: P < P_0 \\ Z > Z_{1-\alpha} & \text{when } H_1: P > P_0 \\ Z > Z_{1-\frac{\alpha}{2}} \text{ or } Z < Z_{\frac{\alpha}{2}} & \text{when } H_1: P \neq P_0 \end{cases}$$

### Ex (7):

A random sample of size 200 unit from a factory production and found that it contained 8% defective units. Can we say that the defective proportion units in the factor production is more than 7% at  $\alpha = 0.05$ .

### Solu.

1- Data :  $P_o = 0.07, n = 200, \alpha = 0.05, r = 0.08$

2- hypothesis :  $H_0: P = 0.07$

$$H_1: P > 0.07$$

3- the statistic:  $Z = \frac{r - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}} = \frac{0.08 - 0.07}{\sqrt{\frac{0.07(1-0.07)}{200}}} = 0.554274$

4- the table value  $Z_{1-\alpha} = Z_{0.95} = 1.645$

5- the decision: we accept  $H_0$  since  $Z \not> Z_{1-\alpha} (0.554274 < 1.645)$

#### 4.4 Testing Hypothesis Concerning two independent Normal Populations

##### 4.4.1 testing for the two populations means

###### (i) Large Samples ( $n_1, n_2 > 30$ )

1- data needed:  $n_1, \mu_1, \sigma_1$  (or  $S_1$ ),  $\bar{X}_1$  and  $n_2, \mu_2, \sigma$  (or  $S_2$ ),  $\bar{X}_2$

2- the hypothesis:  $H_0: \mu_1 = \mu_2 \rightarrow \mu_1 - \mu_2 = 0$

$$H_1: \begin{cases} \mu_1 < \mu_2 \rightarrow \mu_1 - \mu_2 < 0 \\ \mu_1 > \mu_2 \rightarrow \mu_1 - \mu_2 > 0 \\ \mu_1 \neq \mu_2 \rightarrow \mu_1 - \mu_2 \neq 0 \end{cases}$$

3- the statistic:

- if  $\sigma_1, \sigma_2$  is known, then use  $Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$

- if  $\sigma_1, \sigma_2$  is unknown, then use  $Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$ , where  $Z \sim N(0,1)$

4- Determining the rejection of  $H_0$ , that is:

i)if  $H_1: \mu_1 > \mu_2 \rightarrow \mu_1 - \mu_2 > 0$  ,reject  $H_0$  if  $Z > Z_{1-\alpha}$

ii) if  $H_1: \mu_1 < \mu_2 \rightarrow \mu_1 - \mu_2 < 0$  ,reject  $H_0$  if  $Z < Z_\alpha$  ( $Z_\alpha$  always negative)

iii) if  $H_1: \mu_1 \neq \mu_2 \rightarrow \mu_1 - \mu_2 \neq 0$  , reject  $H_0$  if  $Z > Z_{1-\frac{\alpha}{z}}$  or  $Z < Z_{\frac{\alpha}{z}}$

### (ii) Small Samples( $n_1, n_2 < 30$ )

We will follow the same steps in the case of large samples except that we use t- distribution if populations variances are unknown and equal, thus:

3- the statistic:

- if  $\sigma_1, \sigma_2$  is known, then use  $Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$

- if  $\sigma_1, \sigma_2$  is unknown (but equal), then use  $T = \frac{\bar{X}_1 - \bar{X}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ ,

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

4- Determining the rejection of  $H_0$ , that is:

i)if  $H_1: \mu_1 > \mu_2 \rightarrow \mu_1 - \mu_2 > 0$  ,reject  $H_0$  if  $T > t_{1-\alpha, (n_1+n_2-2)}$

ii) if  $H_1: \mu_1 < \mu_2 \rightarrow \mu_1 - \mu_2 < 0$  ,reject  $H_0$  if  $T < t_{\alpha, (n_1+n_2-2)}$

iii)if  $H_1: \mu_1 \neq \mu_2 \rightarrow \mu_1 - \mu_2 \neq 0$  , reject  $H_0$  if  $T > t_{1-\frac{\alpha}{z}, (n_1+n_2-2)}$  or  $T < t_{\frac{\alpha}{z}, (n_1+n_2-2)}$

5-Decision rule:

☒ if the rejection rule is hold then we reject  $H_0$  and accept  $H_1$

☒ if the rejection rule isn't hold then we accept  $H_0$  and reject  $H_1$

### Ex (8):

To compare the income level in two cities, a randomly two samples are selected . the first sample of size 50 family from the first city has average income 64 thousand dollar per year and variance 6 thousand dollar and a sample of size 60 family from the second city has an average income 66 thousand dollar with variance 5 thousand dollar. is there exist a significant difference between the average income for families in the two cities at  $\alpha = 0.05$ .

### Solu.

1- Data:

$$n_1 = 50 \quad \bar{X}_1 = 64 \quad S_1^2 = 6$$

$$n_2 = 60 \quad \bar{X}_2 = 66 \quad S_2^2 = 5 \quad \alpha = 0.05$$

2-Hypothesis:

$$H_0: \mu_1 = \mu_2 \rightarrow \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 \neq \mu_2 \rightarrow \mu_1 - \mu_2 \neq 0$$

3-the statistic:

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{64 - 66}{\sqrt{\frac{6}{50} + \frac{5}{60}}} = -4.435$$

4-reject  $H_0$  if  $Z > Z_{1-\frac{\alpha}{z}}$  or  $Z < Z_{\frac{\alpha}{z}}$



$$Z_{1-\frac{\alpha}{z}} = Z_{0.975} = 1.96 \quad \text{and} \quad Z_{\frac{\alpha}{z}} = Z_{0.025} = -1.96$$

Thus, reject  $H_0$  since  $Z < Z_{\frac{\alpha}{2}}$ , i.e., there is a significant difference between the average of family income in the two cities.

#### **4.4.2 Testing for the two populations Proportions**

If we have two independent samples of size  $n_1$  and  $n_2$  with proportions  $r_1$  and  $r_2$  respectively. Thus, we will use the following steps:

1-data needed:  $n_1, r_1 = \frac{a_1}{n_1}$  and  $n_2, r_2 = \frac{a_2}{n_2}$

2- the hypothesis:  $H_0: P_1 = P_2 \rightarrow P_1 - P_2 = 0$

$$H_1: \begin{cases} P_1 < P_2 \rightarrow P_1 - P_2 < 0 \\ P_1 > P_2 \rightarrow P_1 - P_2 > 0 \\ P_1 \neq P_2 \rightarrow P_1 - P_2 \neq 0 \end{cases}$$

3- the statistic:

$$Z = \frac{r_1 - r_2}{\sqrt{\hat{r}(1-\hat{r})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \text{where} \quad \hat{r} = \frac{n_1 r_1 + n_2 r_2}{n_1 + n_2} = \frac{a_1 + a_2}{n_1 + n_2}$$

4- Determining the rejection of  $H_0$ , that is:

i) if  $H_1: P_1 > P_2 \rightarrow P_1 - P_2 > 0$ , reject  $H_0$  if  $Z > Z_{1-\alpha}$

ii) if  $H_1: P_1 < P_2 \rightarrow P_1 - P_2 < 0$ , reject  $H_0$  if  $Z < Z_{\alpha}$

iii) if  $H_1: P_1 \neq P_2 \rightarrow P_1 - P_2 \neq 0$ , reject  $H_0$  if  $Z > Z_{1-\frac{\alpha}{z}}$  or  $Z < Z_{\frac{\alpha}{z}}$

#### **Ex (9):**

Two machine A and B, a random sample of size 300 units from machine A with defective proportion 8% and another sample of size 200 units from

machine B with defective proportion 4%. The manager think that the defective proportion from machine A is differ from the defective proportion from machine B, is he right?. use  $\alpha = 0.05$

**Solu.**

1-data needed:  $n_1 = 300, r_1 = 0.08$  and  $n_2 = 200, r_2 = 0.04, \alpha = 0.05$

2- the hypothesis:  $H_0: P_1 = P_2 \rightarrow P_1 - P_2 = 0$

$$H_1: P_1 \neq P_2 \rightarrow P_1 - P_2 \neq 0$$

3- the statistic:

$$Z = \frac{0.08 - 0.04}{\sqrt{0.064(1 - 0.064)\left(\frac{1}{300} + \frac{1}{200}\right)}} = 0.895$$

$$\text{where } \hat{r} = \frac{n_1 r_1 + n_2 r_2}{n_1 + n_2} = \frac{300(0.08) + 200(0.04)}{500} = 0.064$$

4- reject  $H_0$  if  $Z < Z_{\frac{\alpha}{2}} = Z_{0.025} = -1.96$  or  $Z > Z_{1-\frac{\alpha}{2}} = 1.96$

Thus, we accept  $H_0$  and reject  $H_1$  that says there is a difference between the defective proportions from machines A and B.

**4.4.3 Testing for the two populations Variances**

We have two independent random samples of size  $n_1$  and  $n_2$ , with variances  $s_1^2$  and  $s_2^2$  respectively. We will assume that each population has Normal distribution. we will use the following steps:

1-data needed:  $n_1, s_1^2$  and  $n_2, s_2^2, \alpha$

2- the hypothesis:  $H_0: \sigma_1^2 = \sigma_2^2$

$$H_1: \begin{cases} \sigma_1^2 < \sigma_2^2 \\ \sigma_1^2 > \sigma_2^2 \\ \sigma_1^2 \neq \sigma_2^2 \end{cases}$$

3- the statistic:  $F = s_1^2/s_2^2$

4- Determining the rejection of  $H_0$ , that is:

i) if  $H_1: \sigma_1^2 > \sigma_2^2$ , reject  $H_0$  if  $F > F_{1-\alpha, n_1-1, n_2-1}$

ii) if  $H_1: \sigma_1^2 < \sigma_2^2$ , reject  $H_0$  if  $F < F_{\alpha, n_1-1, n_2-1} = \frac{1}{F_{1-\alpha, n_1-1, n_2-1}}$

iii) if  $H_1: \sigma_1^2 \neq \sigma_2^2$ , reject  $H_0$  if  $F > F_{1-\frac{\alpha}{2}, n_1-1, n_2-1}$  or  $F < F_{\frac{\alpha}{2}, n_1-1, n_2-1} = \frac{1}{F_{1-\frac{\alpha}{2}, n_1-1, n_2-1}}$

### Ex (10):

For the data given in the following table, test whether the variance of bacteria counts of pasteurized milk is less than the variance of for counts for the raw milk. Use  $\alpha = 0.01$

	Sample size	Mean	Standard deviation
Raw Milk	32	31291.547	70521.846
Pasteurized Milk	23	784.838	2358.172

### Solu.

1-data needed:

$$n_1 = 32, s_1 = 70521.846 \text{ and } n_2 = 23, s_2 = 2358.172, \alpha = 0.01$$

2- the hypothesis:  $H_0: \sigma_1^2 = \sigma_2^2$

$$H_1: \sigma_1^2 > \sigma_2^2$$

3- the statistic:  $F = \frac{s_1^2}{s_2^2} = \left( \frac{70521.846}{2358.172} \right)^2 = 894.327$

4-  $F_{1-\alpha, n_1-1, n_2-1} = F_{0.99, 31, 22} = 2.67$

5- We reject  $H_0$  and accept  $H_1$ , since,

$F = 894.327 > F_{0.99, 31, 22} = 2.67$ . Thus, we conclude that the variance of bacteria counts of pasteurized milk is less than the variance of for counts for the raw milk. Use  $\alpha = 0.01$