Chapter: **Hypothesis Testing**

**Hypothesis**

An assumption we make about a population parameter. A **statistical hypothesis test** is a method of making statistical decisions using experimental data.

**Test of significance or test of a hypothesis**

Significance test or test of a hypothesis is a statistical procedure of reaching a decision.

**Null hypothesis**

The statistical hypothesis which is picked up for the test is known as the null hypothesis. The null hypothesis is usually denoted by .



**Example**

Let us consider a normal distribution with mean and variance. Then the hypothesis that the normal distribution has specified mean 270 i., e., is known as null hypothesis.



**Alternative hypothesis**

Any hypothesis other than the null hypothesis is known as alternative hypothesis. It is usually denoted by or .



**Example**

Let us consider a normal distribution with mean and variance. We are going to test the hypothesis,. Then the alternative hypothesis may be



Or



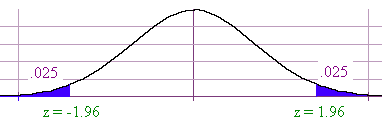
Or



**Rejection Regions**

Suppose that.  We can draw the appropriate picture and find the score for and.  We call the outside regions the rejection regions.





We call the blue areas the *rejection region* since if the value of falls in these regions, we can say that the null hypothesis is very unlikely so we can reject the null hypothesis.



**Example**

50 smokers were questioned about the number of hours they sleep each day.  We want to test the hypothesis that the smokers need less sleep than the general public which needs an average of 7.7 hours of sleep.  We follow the steps below:

1. Compute a rejection region for a significance level of 0.05.
2. If the sample mean is 7.5 and the standard deviation is 0.5, what can you conclude?

**Solution**

First, we write down the null and alternative hypotheses



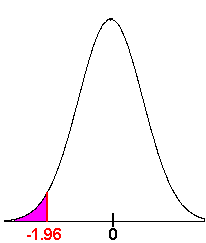
This is a left tailed test.  The -score that corresponds to 0.05 is -1.96.  The critical region is the area that lies to the left of -1.96.  If the z-value is less than -1.96 there we will reject the null hypothesis and accept the alternative hypothesis.  If it is greater than -1.96, we will fail to reject the null hypothesis and say that the test was not statistically significant.



We have



Since -2.83 is to the left of -1.96, it is in the critical region.



Hence we reject the null hypothesis and accept the alternative hypothesis.  We can conclude that smokers need less sleep.

**Test**

A body of rules which leads to the decision regarding acceptance or rejection of the hypothesis is called a test. The statistic which is usually used to test the parameter of a population is known as test statistic.

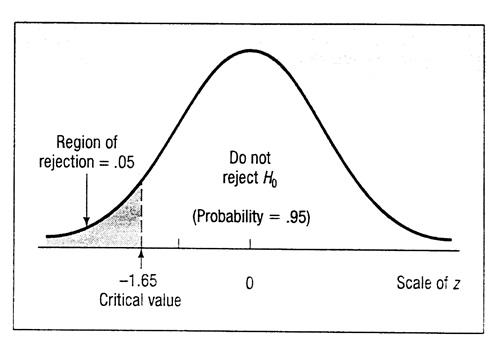
Test may be classified as

* One tailed test
* Two tailed test

**One tailed test**

A test for which the entire rejection region lies in only of two tails either in the right tail or in the left tail of the sampling distribution of the test statistic is called one tailed. If we are interested to test the hypothesis vs then we should use right tailed test. If we test vs then we should use right tailed test.



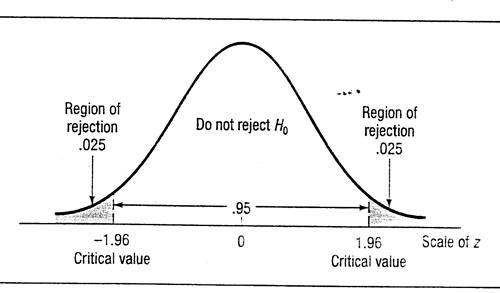
 **Fig:** Sampling Distribution of the Statistic z, Right and Left-Tailed Test, 0.05 Level of Significance.

**Two tailed test**

A test for which the rejection region is divided equally between two tails of the sampling distributions of the test statistics is called a two tailed test. If we are interested to test the

vs then we should use two tailed test.





**Fig:** Regions of Non-rejection and Rejection for a Two-Tailed Test, 0.05 Level of Significance.

**Critical value**

The value of the standard statistic ( or ) beyond which we reject the null hypothesis; the boundary between the acceptance and rejection regions.



**Error**

When using probability to decide whether a statistical test provides evidence for or against our predictions, there is always a chance of driving the wrong conclusions. Even when choosing a probability level of 95%, there is always a 5% chance that one rejects the null hypothesis when it was actually correct. This is called ***Type I error***, represented by the Greek letter. The probability of a type I error (α) is called the ***significance level***.



It is possible to err in the opposite way if one fails to reject the null hypothesis when it is, in fact, incorrect. This is called ***Type II error***, represented by the Greek letter . These two errors are represented in the following chart.



| **Table: Types of error** | | |
| --- | --- | --- |
| **Type of decision** | **H0 true** | **H0 false** |
| Reject H0 | Type I error () | Correct decision (1-) |
| Accept H0 | Correct decision (1-) | Type II error () |

A related concept is power, which is the probability of rejecting the null hypothesis when it is actually false. Power is simply 1 minus the Type II error rate, and is usually expressed as.



When choosing the probability level of a test, it is possible to control the risk of committing a Type I error by choosing an appropriate.



**P-values**

There is another way to interpret the test statistic.  In hypothesis testing, we make a yes or no decision without discussing borderline cases.  For example with α = 0.06, a two tailed test will indicate rejection of for a test statistic of or for , but is much stronger evidence than .  To show this difference we write the *p-value* which is the *lowest significance level such that we will still reject Ho*.  For a two tailed test, we use twice the table value to find p, and for a one tailed test, we use the table value.



**Example:**

Suppose that we want to test the hypothesis with a significance level of 0.05 that the climate has changed since industrialization.  Suppose that the mean temperature throughout history is 50 degrees.  During the last 40 years, the mean temperature has been 51 degrees with a standard deviation of 2 degrees.  What can we conclude?

We have

        vs



We compute the z score:



The table gives us 0.9992 .

So that

        p = (1 - 0.9992)(2) = 0 .002

Since  0.002 < 0.05

We can conclude that there has been a change in temperature.

Note that small p-values will result in a rejection of and large p-values will result in failing to reject .



**Steps in Hypothesis Testing**

**Step 1**

## Identify the null hypothesis and the alternate hypothesis .



**Step 2**

Choose . The value should be small, usually less than 10%. It is important to consider the consequences of both types of errors.



**Step 3**

Select the test statistic and determine its value from the sample data. This value is called the observed value of the test statistic. Remember that a t statistic is usually appropriate for a small number of samples; for larger number of samples, a z statistic can work well if data are normally distributed.

**Step 4**

Compare the observed value of the statistic to the critical value obtained for the chosen.



**Step 5**

Make a decision.

If the test statistic falls in the critical region:Reject in favour of .



If the test statistic does not fall in the critical region:Conclude that there is not enough evidence to reject .



**Important tests of significance**

The important tests of significance in statistics can be classified broadly as

1. Normal test
3. test



c) test



d) test.



**Test of significance about mean**

Here we consider the following cases:

* + Comparison of a sample mean with an assigned population mean.
  + Comparison of two independent sample means.
  + Comparison of two correlated sample means.
  + Comparison of independent sample means.



**Comparison of a sample mean with an assigned population mean**

**Case 1:** is known or estimated from a large sample .



**Case 2** is unknown and the sample is small.



**Case 1 Case 2** 

**Hypothesis Hypothesis**



**Test statistic Test statistic**

with .



Where, and



**Example**

The mean life time of a sample of 100 light tubes produced by a company is found to be 1570 hours with standard deviation of 80 hours. Test the hypothesis that the mean life time of the tubes produced by the company is 1600 hours (consider 5% significance level).

**Solution**

* 1. **Hypothesis**

We consider the following hypothesis

vs



**2. Significance level**

Given that the significance level



**3. Test statistic**

In order to test the hypothesis we consider the following test statistic



We have, , , .



**4. Critical value**

The critical value or tabulated value is 1.96.

**5. Making decision**

Since the calculate value is greater than the tabulated value, so we reject the null hypothesis. So that, the mean life time of the tubes produced by the company is not 1600 hours.

**Example**

A sample of 400 male students is found to have a mean height 67.47 inches. Can it be reasonably regarded as a sample from a large population with mean height 67.39 inches and standard deviation 1.30 inches? Test at 5% level of significance.

**Solution**

1. **Hypothesis**

We consider the following hypothesis

vs



**2. Significance level**

Given that the significance level



**3. Test statistic**

In order to test the hypothesis we consider the following test statistic



We have, , , .



**4. Critical value**

The critical value or tabulated value is 1.96.

**5. Making decision**

Since the calculate value is less than the tabulated value, so we accept the null hypothesis. So that, we may conclude that the given sample (with mean height 67.47 inches) can be regarded to have been taken from a population with mean height 67.39 inches and standard deviation 1.30 inches at 5% level of significance.

**Example**

Suppose that it is known from experience that the standard deviation of the weight of 8 ounces packages of cookies made by a certain bakery is 0.16 ounces. To check its production is under control on a given day, the true average of the packages is 8 ounces, they select a random sample of 40 packages and find their mean weight is 8.122 ounces. Test whether the production is under control or not at 5% level of significance.

**Solution**

1. **Hypothesis**

We consider the following hypothesis

vs



**2. Significance level**

Given that the significance level



**3. Test statistic**

In order to test the hypothesis we consider the following test statistic



We have, , , .



**4. Critical value**

The critical value or tabulated value is 1.96.

**5. Making decision**

Since the calculate value is greater than the tabulated value, so we reject the null hypothesis. So that, the production is not under control.

**Case 2: is unknown and the sample is small**



**Example**

A random sample of 10 boys had the following I. Q’s:

70, 120, 110, 101, 88, 83, 95, 98, 107, 100

Do these data support the assumption of a population mean I. Q. of 100?

**Solution**

1. **Hypothesis**

We consider the following hypothesis

vs



**2. Significance level**

Given that the significance level



**3. Test statistic**

In order to test the hypothesis we consider the following test statistic



We have, ,



, .



**4. Critical value**

The critical value or tabulated value of for is 2.2622.



**5. Making decision**

Since the calculate value is less than the tabulated value, so we accept the null hypothesis. So that, the data support that the population mean I. Q is 100.

**Example**

Is the temperature required to damage a computer on the average less than 110 degrees?  Because of the price of testing, twenty computers were tested to see what minimum temperature will damage the computer.  The damaging temperature averaged 109 degrees with a standard deviation of 3 degrees.  (Use α = 0.05)

**Solution**

1. **Hypothesis**

We consider the following hypothesis

vs



**2. Significance level**

Given that the significance level



**3. Test statistic**

In order to test the hypothesis we consider the following test statistic



Here, , and



**4. Critical value**

This is a one tailed test, so we can go to our t-table with 19 degrees of freedom to find the critical value or tabulated value. The critical value is 1.73.

**5. Making decision**

Since the calculate value is less than the tabulated value, so we accept the null hypothesis and conclude that there is insufficient evidence to suggest that the temperature required to damage a computer on the average less than 110 degrees.

**Example**

The specimen of copper wires drawn from a large lot has the following breaking strength (in Kg. weigth):

578, 572, 570, 568, 572, 578, 570, 572, 596, 544

Test whether the mean breaking strength of the lot may be taking to be 578 Kg. weights by using 10% level of significance.

**Solution**

1. **Hypothesis**

We consider the following hypothesis

vs



**2. Significance level**

Given that the significance level



**3. Test statistic**

In order to test the hypothesis we consider the following test statistic



We have, ,



, .



**4. Critical value**

The critical value or tabulated value of for is 1.8331.



**5. Making decision**

Since the calculate value is less than the tabulated value, so we accept the null hypothesis at 10% level of significance. Hence we may conclude that the mean breaking strength of copper wires lot may be taken as 578 Kg. weight.

**Home work:**

**Book: By D. Lind.**

**Page 331, Exercise: 1, 2, 3.**

**Page 338, Example.**

**Page 343, Self review 10.5 and Exercise: 21, 22.**

**Book: By Paul Newbold.**

**Page 322, Example 9.3**

**Page 323, Exercise 9.5**

**Page 327, Exercise 9.15, 9.16, 9.17**

**Page 364, Exercise 9.63**

**Comparison of two independent sample means**

Suppose we want to test two independent sample means are equal. Then we have to test the hypothesis vs .



We have the following three cases

**Case 1:** Variance known or samples are large .



**Case 2:** Small samples and variances (unknown) assumed equal.

**Case 3:** Small samples and variances (unknown) assumed not equal.

**Case 1: Variance known or samples are large**



**Hypothesis**

vs



**Test statistic**



**Example**

Intelligence test given to two groups of boys and girls gave the following information:

|  | **Mean score** | **Standard deviation** | **Number** |
| --- | --- | --- | --- |
| **Girls** | 75 | 10 | 50 |
| **Boys** | 70 | 12 | 100 |

Is the difference in the mean scores of boys and girls statistically significant? Use .



**Solution**

1. **Hypothesis**

We consider the following hypothesis

vs



**2. Significance level**

Given that the significance level .



**3. Test statistic**

In order to test the hypothesis we consider the following test statistic



We have, , ,



, ,



**4. Critical value**

The critical value or tabulated value at 1% level of significance is 2.58.

**5. Making decision**

Since the calculate value is greater than the tabulated value, so we reject the null hypothesis. Hence there is a difference between the mean score of boys and girls.

**Example**

Suppose you are working as a purchase manager for a company. The following information has been supplied to you by two manufactures of electric bulbs:

|  | **Company A** | **Company B** |
| --- | --- | --- |
| **Mean life (in hours)** | 1300 | 1248 |
| **Standard deviation (in hours)** | 82 | 93 |
| **Sample size** | 100 | 100 |

Which brand of bulbs are you going to purchase if you desire to take a risk of 5%?

**Solution**

1. **Hypothesis**

We consider the following hypothesis

vs



**2. Significance level**

Given that the significance level .



**3. Test statistic**

In order to test the hypothesis we consider the following test statistic



We have, , ,



, ,



**4. Critical value**

The critical value or tabulated value at 5% level of significance is 1.96.

**5. Making decision**

Since the calculate value is greater than the tabulated value, so we reject the null hypothesis. Hence the qualities of two brands of bulbs are not same.

**Case 2: Small samples and variances (unknown) assumed equal**

**Hypothesis**

vs



**Test statistic**

where, ,



with



**Example**

The gain in weights (in lbs) of pigs fed on two diets A and B are given below:

|  | **Gain in weight** |
| --- | --- |
| **Diet A** | 25, 32, 30, 34, 24, 14, 32, 24, 30, 31, 35, 25 |
| **Diet B** | 44, 34, 22, 10, 47, 31, 40, 30, 32, 35, 18, 21, 35, 29, 22 |

Test if the two diets differ significantly as regards their effect on increase in weight. Use 5% level of significance.

**Solution**

1. **Hypothesis**

We consider the following hypothesis

vs



**2. Significance level**

Given that the significance level .



**3. Test statistic**

In order to test the hypothesis we consider the following test statistic

where, ,



We have, , ,



, ,



**4. Critical value**

For the critical value or tabulated value at 5% level of significance is 2.0595.



**5. Making decision**

Since the calculate value is less than the tabulated value, so we accept the null hypothesis. Hence there is no difference between two diets as regards their effect on increase in weight.

**Example**

The purpose of a study is to investigate the nature of lung destruction in cigarette smokers. We have the following data:

|  | **Lung destructive index scores** |
| --- | --- |
| **Non smokers** | 18.1, 6.0, 10.8, 11.0, 7.7, 17.9, 8.5, 13.0, 18.9 |
| **Smokers** | 16.6, 13.9, 11.3, 26.5, 17.4, 15.3, 15.8, 12.3, 18.6, 12.0, 24.1, 16.5, 21.8, 16.3, 23.4, 18.8 |

Make a conclusion on the basis of the data about the lung damage.

**Solution**

1. **Hypothesis**

We consider the following hypothesis

vs



**2. Significance level**

Given that the significance level .



**3. Test statistic**

In order to test the hypothesis we consider the following test statistic

where, ,



We have, , ,



, ,



**4. Critical value**

For the critical value or tabulated value at 5% level of significance is 2.0687.



**5. Making decision**

Since the calculate value is greater than the tabulated value, so we reject the null hypothesis. Hence we may conclude that as measured by the index used in the study, smoker have greater damage than non smokers.

**Case 3: Small samples and variances (unknown) assumed not equal**

**Hypothesis**

vs



**Test statistic**

where, with



**Example**

A company is to choose from two brands of tyres A and B, for their vehicles. An experiment is conducted in which 10 tyres of brand A and 12 tyres of brand B are used. They are run under similar conditions until they were out. The experimental results are as follows:

| **Brand A** | **Brand B** |
| --- | --- |
| miles | miles |
| miles | miles |
|  |  |

Which of these brands do you consider superior (equality of variances are not assumed)? Use 5% level of significance.

**Solution**

1. **Hypothesis**

We consider the following hypothesis

vs



**2. Significance level**

Given that the significance level .



**3. Test statistic**

In order to test the hypothesis we consider the following test statistic



We have, , ,



, ,



**4. Critical value**

For the critical value or tabulated value at 5% level of significance is 2.093.



**5. Making decision**

Since the calculate value is greater than the tabulated value, so we reject the null hypothesis. Hence, Brand B is superior to Brand A.

**Home work**

**Page 361, Exercise 1, 2**

**Page 369, Self review 11.3**

**Page 370, Exercise 15, 17, 18**

**Book: By D. Lind**

**Page 345, Exercise: 9.35, 9.36, 9.37**

**Page 366, 9.81**

**Book : By Paul Newbold.**

**Comparison of two correlated sample means**

**Hypothesis**

**vs**



**Test statistic**

with .



Where, and



**Example**

In order to compare two drugs A and B, a group of 20 rats were used in an experiment. These rats were grouped into 10 pairs such that animals in each pair are alike in respect of age, sex and weight. The treatments are allotted at random to the animals in each pair, the allotments for any pair being independent of that for any other pair. The experimental results ate as follows:

| **Treatment** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **A** | 30 | 27 | 20 | 40 | 16 | 34 | 42 | 20 | 16 | 19 |
| **B** | 25 | 28 | 18 | 44 | 18 | 35 | 44 | 23 | 12 | 23 |

Test the hypothesis that the two drugs are same at 5% level of significance.

**Solution**

1. **Hypothesis**

We consider the following hypothesis

vs



**2. Significance level**

Given that the significance level .



**3. Test statistic**

In order to test the hypothesis we consider the following test statistic

with .



Where, and



We have,

| **Treatment** | **A (x)** | **B (y)** |  |  |
| --- | --- | --- | --- | --- |
| 1 | 30 | 25 | 5 | 25 |
| 2 | 27 | 28 | -1 | 1 |
| 3 | 20 | 18 | 2 | 4 |
| 4 | 40 | 44 | -4 | 16 |
| 5 | 16 | 18 | -2 | 4 |
| 6 | 34 | 35 | -1 | 1 |
| 7 | 42 | 44 | -2 | 4 |
| 8 | 20 | 23 | -3 | 9 |
| 9 | 16 | 12 | 4 | 16 |
| 10 | 19 | 23 | -4 | 16 |
| Total |  |  | -6 | 96 |

So that we get, and



.



**4. Critical value**

For the critical value or tabulated value at 5% level of significance is 2.2622.



**5. Making decision**

Since the calculate value is less than the tabulated value, so we accept the null hypothesis. Hence, two drugs A and B are same.

**Example**

Memory capacity of 9 students was tested before and after training. State at 5% level of significance whether the training was effective from the following scores:

| Student | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Before | 10 | 15 | 9 | 3 | 7 | 12 | 16 | 17 | 4 |
| After | 12 | 17 | 8 | 5 | 6 | 11 | 18 | 20 | 3 |

**Solution**

1. **Hypothesis**

We consider the following hypothesis

vs



**2. Significance level**

Given that the significance level .



**3. Test statistic**

In order to test the hypothesis we consider the following test statistic

with .



Where, and



We have,

| **Student** | **Before (x)** | **After (y)** |  |  |
| --- | --- | --- | --- | --- |
| 1 | 10 | 12 | -2 | 4 |
| 2 | 15 | 17 | -2 | 4 |
| 3 | 9 | 8 | 1 | 1 |
| 4 | 3 | 5 | -2 | 4 |
| 5 | 7 | 6 | 1 | 1 |
| 6 | 12 | 11 | 1 | 1 |
| 7 | 16 | 18 | -2 | 4 |
| 8 | 17 | 20 | -3 | 9 |
| 9 | 4 | 3 | 1 | 1 |
| Total |  |  | -7 | 29 |

So that we get, and



.



**4. Critical value**

For the critical value or tabulated value at 5% level of significance is 1.8595.



**5. Making decision**

Since the calculate value is less than the tabulated value, so we accept the null hypothesis. Hence, we may conclude that the difference in score before and after training is insignificant. So that we can infer that the training was not effective.

**Example**

Ten persons were appointed for the post Officer in an office. Their performance was noted by giving a test and the marks were recorded out of 100. They were given 3 months training and a test was held and marks were recorded out of 100.

| Employees | A | B | C | D | E | F | G | H | I | J |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Before training | 80 | 76 | 92 | 60 | 70 | 56 | 74 | 56 | 70 | 56 |
| After training | 84 | 70 | 96 | 80 | 70 | 52 | 84 | 72 | 72 | 50 |

By applying test can it be concluded that the employees have benefited by the training?



**Solution**

1. **Hypothesis**

We consider the following hypothesis

vs



**2. Significance level**

Given that the significance level .



**3. Test statistic**

In order to test the hypothesis we consider the following test statistic

with .



Where, and



We have,

| **Employees** | **Before training (x)** | **After training (y)** |  |  |
| --- | --- | --- | --- | --- |
| A | 80 | 84 | -4 | 16 |
| B | 76 | 70 | 6 | 36 |
| C | 92 | 96 | -4 | 16 |
| D | 60 | 80 | -20 | 400 |
| E | 70 | 70 | 0 | 0 |
| F | 56 | 52 | 4 | 16 |
| G | 74 | 84 | -10 | 100 |
| H | 56 | 72 | -16 | 256 |
| I | 70 | 72 | -2 | 4 |
| J | 56 | 50 | 6 | 36 |
| Total |  |  | -40 | 880 |

So that we get, and



.



**4. Critical value**

For the critical value or tabulated value at 5% level of significance is 2.2622.



**5. Making decision**

Since the calculate value is less than the tabulated value, so we accept the null hypothesis. Hence, we may conclude that the employees have not benefited by the training.