Unit 2: Test of statistical significance

Hypothesis

A hypothesis is defined as: it is a presumption or quantitative statement made for the population parameter which may be true or false. It should be kept in mind that hypothesis is merely assumption relating to the population parameter and can be tested by using sample information.

Null hypothesis: The hypothesis whose validity is going to be tested is known as null hypothesis. It is denoted by H_o. It is formulated on the assumption that there is no significant difference between the sample statistic and population parametric value. It is also known as hypothesis of no difference.

For example: If we are interested to test whether the average time taken by a new drug to be relief is 2 hours, then we formulate the H_o as,

$H_o: \mu = 2$ hours i.e. the average time taken to be relief is 2 hours.

Alternative hypothesis: The hypothesis which is complementary to H_o is known as alternative hypothesis. It is denoted by H_1 and is also known as hypothesis of difference. The alternative hypothesis for the above assumption may be any one of the following three statements:

For example: The alternative hypothesis for the second example above may be,

$H_1: \mu \neq 2$ hours i.e. the average time taken to be relief is not 2 hours.

$H_1: \mu > 2$ hours i.e. the average time taken to be relief is greater than 2 hours.

$H_1: \mu < 2$ hours i.e. the average time taken to be relief is less than 2 hours.

Although alternative hypothesis can be set up in above three ways, one and only one alternative hypothesis can be set up at a time.

Errors in testing of hypothesis

The following table gives a summary of possible results of any hypothesis test:

Nature of statement	Decision	
	Accept H_o	Reject H_o
H_o is true	Correct decision Probability = $(1-\alpha)$	Type I error Probability = α
H_o is false	Type II error Probability = β	Correct decision Probability = $(1-\beta)$

Type I error

In testing of hypothesis, a type I error occurs when the null hypothesis is rejected though it is true. The probability of type I error is denoted by α (alpha).

α = P (type I error) = P (rejecting H_o/H_o is true)

A type I error can also be referred to as an error of the first kind.

Type II error

In testing of hypothesis test, a type II error occurs when the null hypothesis is accepted, when it is false. The probability of a type II error is denoted by β (beta). The exact probability of a type II error is generally unknown.

β = P (type II error) = P (accepting H_o/ H_o is false)

A type II error can also be referred to as an error of the second kind.

There is an intuitively appealing inverse relationship between the probabilities of the two types of errors. As α increases; β decreases and vice versa i.e. smaller the risk of one, higher the risk of other. The only one way to reduce α and β simultaneously is to increase the sample size.

p value

The p value for a hypothesis test is the probability of obtaining a value of test statistic as extreme as or more extreme in the direction of alternative hypothesis than the one actually computed, when null hypothesis is true.

The decision rules for accepting or rejecting null hypothesis in the p value approach are as follows:

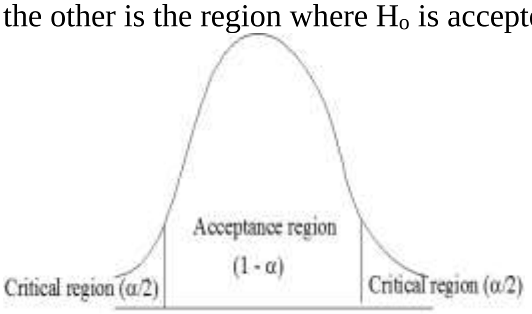
If p value $\geq \alpha$, H_o is accepted i.e. the result is not significant.

If p value $< \alpha$, H_o is rejected i.e. the result is significant.

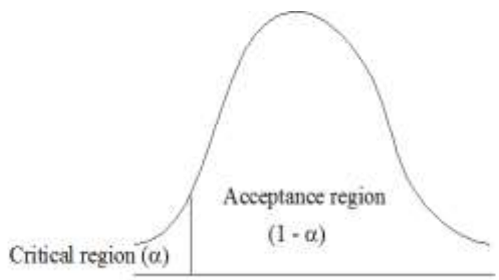
Conventionally a p value of 0.05 is generally regarded as sufficiently small to reject the null hypothesis.

Critical region and acceptance region

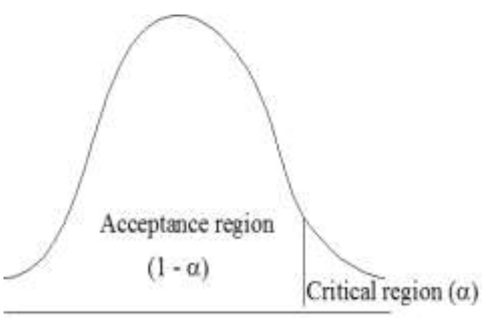
The critical region or rejection region is a set of values of the test statistics for which the null hypothesis is rejected. The size of this region is determined by the probability α of the sample point falling in the critical region when H_o is true. The sample space for the test statistic is partitioned into two regions; the critical region which will lead us to reject the null hypothesis and the other is the region where H_o is accepted is called the acceptance region.



Two tailed test



Left tailed test



Right tailed test

Level of significance

The level of significance of a statistical hypothesis test is a fixed probability of wrongly rejecting the null hypothesis H_o, if it is in fact true. The maximum size of type I error that we tolerate in decision process due to use of sample information is called level of significance. The level of significance is usually denoted by α.

Level of significance = P(type I error) =P (rejecting H_o when it is true) = α

There is no single standard or universal level of significance for testing of hypothesis. Level of significance (α) can be set as 1%, 2%, 5%, 10%. But usually, the significance level is chosen to be α = 5% unless it is mentioned. α = 5% means we are likely to reject 5 cases although it is true in reality when the decision is repeated 100 times under the same condition using same decision rule.

One-tailed Test

If the direction of the difference is specified in any hypothesis test, we call it one tailed test. Identifying the words for comparative study such as high, low, increase, decrease, at least, at most, greater than, superior, inferior, taller than, smaller than, only, etc we use one tailed test. After identifying one tailed test, the right and left tailed test can be identified as,

If the sample statistic < population parameter, we use left tailed test.

If the sample statistic > population parameter, we use right tailed test.

If the 1st sample statistic < 2nd sample statistic, we use left tailed test.

If the 1st sample statistic > 2nd sample statistic, we use right tailed test.

Two tailed test

A two-sided test is a statistical hypothesis test in which the values for which we can reject the null hypothesis, H_o are located in both tails of the probability distribution**.** In other words, the critical region for a two-sided test is the set of values less than a first critical value of the test and the set of values greater than a second critical value of the test. A two-sided test is also referred to as a two-tailed test of significance.

The choice between a one-sided test and a two-sided test is determined by the purpose of the investigation.

Steps to be followed in testing of hypothesis

1. Set up null and alternative hypothesis.
2. Select suitable test statistic (T).

   $$\text{Test statistic} = \frac{\text{Statistic - Parameter}}{\text{Standard error of the statistic}} \text{ or } \frac{\text{Difference}}{\text{Standard error}}$$
3. Set the level of significance α and the degree of freedom where necessary.
4. Obtain the critical value of the test statistics (C).
5. Give decision: Is $T \leq C$?

   If yes, accept null hypothesis (fail to reject the null hypothesis)

   If no, reject null hypothesis, i.e. accept alternative one.

Large sample test (n > 30)

Large sample test (z-test)

Test of significance of single mean

The steps in testing are as follows:

1. **Formulate null (H_o) and alternative (H_1) hypothesis**

Null hypothesis: H_o: $\mu = \mu_o$, there is no significant difference between sample mean and population mean or the samples have drawn from the population with mean μ_o and variance σ^2 or the population mean has a specified value μ_o.
Alternative hypothesis: H_1: $\mu \neq \mu_o$, there is a significant difference between sample mean and population mean or the samples have not been drawn from the population with mean μ_o and variance σ^2 or the population mean has not a specified value μ_o.
or H_1: $\mu > \mu_o$, the population mean is greater than μ_o (right tailed)
or H_1: $\mu < \mu_o$, the population mean is less than μ_o (left tailed)

2. **Compute the test statistic**

Under H_o, the test statistic is,

$$z = \frac{\text{Difference}}{\text{Standard error}} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

Where $\bar{x}$ = sample mean, μ = population mean, σ = population standard deviation,
n = sample size
If the population standard deviation is unknown, then we use its estimate provided by the sample variance s^2 then,

$$z = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \sim N(0,1) \qquad (\hat{\sigma} = s \text{ for large samples})$$

3. **Level of significance**

Chose a appropriate level of significance, α can be taken as 0.01, 0.05. 0.1

4. **Critical value**

Obtain the critical value or tabulated value of z from the table depending whether the alternative hypothesis is one tailed or two tailed.

5. **Decision**

If $z_{cal} \leq z_{tab}$, null hypothesis is accepted, otherwise reject the null hypothesis.

Exercise

1. The mean height obtained from a random sample of size 100 is 64 inches. The standard deviation of the distribution of height of the population is known to be 3 inches. Test the statement that the mean height of the population is 67 inches at 5% level of significance. Also set up 99% confidence limits of the mean heights of the population. (10; 63.2, 64.8)
2. A research team is willing to assume the systolic blood pressures in a certain population of males are approximately normally distributed with a standard deviation of 16. A simple random sample of 64 males from the population had a mean systolic blood pressure reading of 133. At the 0.05 level of significance, do these data provide sufficient evidence for us to conclude that population mean is greater than 130? (1.5)

3. A random sample of 900 children was found to have a mean fat fold thickness at triceps of 3.4 mm with a standard deviation of 2.3 mm. Can it be reasonably regarded as a representative sample of population having a mean thickness of 3.2 mm? (2.61)
4. In a sample of 49 adolescents who served as the subjects in an immunologic study, one variable of interest was the diameter of skin test reaction to an antigen. The sample mean and standard deviation were 21 and 11 mm erythema respectively. Can it be concluded from these data that the population mean is less than 30 at 5% level of significance? (5.73)
5. A pharmaceutical firm maintains that the mean time for a drug to take effect is 24 minutes. In a sample of 400 trials, the mean time is 26 minutes with a standard deviation of 4 minutes. Test the hypothesis that the mean time is 24 minutes against the alternative hypothesis that is not equal to 24 minutes. Use a level of significance of 0.01. (10)
6. 50 smokers were questioned about the number of hours 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. If the sample mean is 7.5 and the standard deviation is 0.5, what can you conclude? (2.83)

Test of significance difference between two means

The steps in testing are as follows:

Formulate null and alternative hypothesis

H_o: $\mu_1 = \mu_2$, there is no significant difference between two population means, or the two samples have been drawn from the same population or two population means are equal.

H_1: $\mu_1 \neq \mu_2$, there is a significant difference between two population means, or the two samples have not been drawn from the same population or two population means are not equal.

or H_1: $\mu_1 > \mu_2$, the mean of first population is greater than the mean of second population (right tailed)

or H_1: $\mu_1 < \mu_2$, the mean of first population is less than the mean of second population (left tailed).

Compute the test statistic

Under H_o, the test statistic is,

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Here the population variance is assumed to be known. If they are unknown, then their estimate is given by the corresponding sample variance s_1^2 and s_2^2 (for large samples $\hat{\sigma}_1 = s_1$ and $\hat{\sigma}_2 = s_2$). Hence the test statistic z is given by,

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

When $\sigma_1^2 = \sigma_2^2 = \sigma^2$ (say), then

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

If the common variance σ^2 is unknown then it is estimated by the relation,

$$\hat{\sigma}^2 = \frac{n_1 s_1^2 + n_1 s_2^2}{n_1 + n_2}$$

The other steps like level of significance, critical value and decision making is same as that we have done in test of significance of single mean.

Exercise

1. Suppose a randomized clinical trial is being conducted to compare two drugs for the treatment of hypertension. In a group of 100 hypertensive patients half are allocated to receive drug A and half drug B. An appropriate measure of efficacy is determined to be the systolic blood pressure reading (mm/Hg) after three months of drug use. The results obtained are as follows:

	n	$\bar{x}$	s
Drug A	50	145	9.9
Drug B	50	135	10.0

   Can you conclude that drug B results in lower systolic blood pressure in patients with hypertension than does drug A at 5% level of significance? (5.02)
2. In a group of 196 adults in the age group of 45 to 53 years, belonging to a social class I. The mean serum cholesterol was 180mg% with a standard deviation of 42mg%. In other comparable groups of 144 adults belonging to social class II the mean serum cholesterol was 150 mg% with a standard deviation of 48mg%. At the 1% level of significance, is the difference in the cholesterol level of the two classes statistically significant?
3. In a nutritional study, 100 children were given a usual diet and vitamins A and D tablets. After 6 months, their average weight was 30 kg with a standard deviation of 2 kg while the average weight of the second comparable group of 100 children who were taking usual diet only was 29 kg with a SD of 1.8 kg. Can we say that vitamins A and D were responsible for this difference? (3.78)
4. If 60 MPH students are found to have a mean height of 63. 60 inches and 50 M. Pharm. students a mean height of 69.51 inches. Would you conclude that the MPH students are taller than M.Pharm. students? Assume the standard deviation of height of post graduate students to be 2.48 inches. (12.44)

Test of significance of sample proportion

The steps in testing are as follows:

1. Formulate null and alternative hypothesis

H_o: $P = P_o$, there is no significant difference between sample proportion and population proportion or the samples have been drawn from the population with proportion P_o, or the population proportion has specified value P_o.

H_1: $P \neq P_o$, there is a significant difference between sample proportion and population proportion or the samples have not been drawn from the population with proportion P_o, or the population proportion has not specified value P_o.
or H_1: $P > P_o$, the population proportion is greater than P_o (right tailed)
or H_1: $P < P_o$, the population proportion is less than P_o (left tailed)

2. Compute the test statistic

Under H_o, the test statistic is,

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

Where P = population proportion of success
Q = population proportion of failure = 1 – P.

3. The other steps like level of significance, critical value and decision making is same as that we have done in test of significance of single mean.

Exercise

1. 20 people were attacked by a disease and only 18 survived. Will you reject the hypothesis that the survival rate, if attacked by the disease is 85% in favor of the hypothesis that it is more at 5% level of significance? (1.39)
2. A commonly prescribed drug on the market for relieving pain is believed to be only 60% effective. Experimental results with a new drug administered to a random sample of 100 adults suffering from the pain showed that 70% received relief. Is this sufficient evidence to conclude that the new drug is superior to the one commonly prescribed? Use a 0.05 level of significance. (2)
3. A patented medicine claimed that it is effective in curing 90% of the patients suffering from malaria. From a sample of 200 patients using this medicine, it was found that only 170 were cured. Determine whether the claim is right or wrong. Take 1% level of significance. (2.38)
4. In a large village, out of 300 men, 165 were found to be smokers. Does this data support the hypothesis that majority of men in the village are smokers?
5. A random sample of 600 persons selected from a certain large city gives the result that females are 53%. Is there any reason to doubt the hypothesis that males and females are in equal numbers in the city?
6. It was found that 66 percent of a sample of 670 infants had completed hepatitis B series. Can you conclude on the basis of these data that in the sampled population more than 60 percent have completed the series? Let α = 5%. (3.17)
7. In Far western development region of Nepal, 70% farmers lie below the poverty line. A random sample of 100 farmers is taken from the region and it is found that 75 farmers lie below the poverty line. Test the hypothesis that there is no significance difference between the sample and population proportion of poor farmers.
8. In a maternity hospital, 510 female babies were born out of 1000 babies. A random sample of 150 babies is drawn and it is found that 78 were the female babies. At 5% level of significance, test the hypothesis that there is no difference between the sample proportion and population proportion of female babies born.

Test of significance difference of two proportions

The steps in testing are as follows:

1. Null and alternative hypothesis:

H_0: $P_1 = P_2 = P$ (say) i.e. two population proportions are same or the two samples have been drawn from the same proportion (percentage) or there is no significant difference between two population proportions.

H_1: $P_1 \neq P_2$ i.e, two population proportions are not equal or the two samples have not drawn from the same proportion (percentage) or there is a significant difference between two population proportions.

Or

H_1: $P_1 > P_2$ i.e, the population proportion of first group is greater than second group

Or

H_1: $P_1 < P_2$ i.e, the population proportion of first group is less than second group

2. Test statistic:

Under H_0, the test statistic is

$$Z = \frac{p_1 - p_2}{\sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}}}$$

Where p_1 = observed sample proportion of success from the first population = $\frac{X_1}{n_1}$

p_2 = observed sample proportion of success from the second population = $\frac{X_2}{n_2}$

P_1 = population proportion of success in the first group (attribute)

P_2 = population proportion of success in the second group (attribute)

$Q_1 = 1 - P_1$ = population proportion of failure in the first group

$Q_2 = 1 - P_2$ = population proportion of failure in the second group

n_1 = size of the first sample

n_2 = size of the second sample

When population proportion of success remain unknown then we use,

$\hat{P} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{X_1 + X_2}{n_1 + n_2}$ and $\hat{Q} = 1 - \hat{P}$. Thus

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

3. Other steps are same as before.

Exercise

1. In a survey, 246 town school children and 349 village school children were examined. Out of 246 children, 36 suffered from conductive hearing loss and among 349 children 61

suffered with hearing loss. Does this data present any evidence that conductive hearing loss is as common among town children as among village children? (0.94)

2. In a study of obesity, the following results were obtained from the sample of male and female between the ages of 20 and 75 years

Sample	Sample size	Number
Male	150	21
Female	200	48

Can we conclude from these data in the sampled population there is a difference in the proportions who are over weight? Let $\alpha = 0.05$. (-2.33)

3. In a random sample of 500 men in a large city, 350 were found to be smokers. After the tax on tobacco has been heavily increased, only 300 were found to be smoker in another random sample of 500 men in the same city. Is the heavy taxation on tobacco an effective instrument to decrease the smoking habit or the men?
4. A random sample of 144 families drawn from a certain metropolitan city showed that 48 families have two or more children. Another random sample of 150 families drawn from another municipality showed that 70 families have two or more children. Test whether the difference between the two proportions is significant or not.
5. A random samples of 600 and 1000 men from two cities, 400 and 600 men are found to be literate. Do the data indicate at 5% level of significance that the population are significantly different in the percentage of literacy?

Assumption of t test

1. The population from which the samples are drawn is normally distributed.
2. The sample size is less than or equal to 30 (i.e. $n \leq 30$)
3. The samples are drawn by random sampling method and are independent.
4. The population standard deviation is unknown.

Degree of freedom

Degree of freedom refers to the number of values in a sample that can be chose freely. When a statistic is used to estimate a parameter, the number of degrees of freedom (df) available depends upon the restriction imposed. It is denoted by nu (υ).The degree of freedom is not always $(n - 1)$, it may vary with the types of the problems under study and the restriction imposed. In some cases, degree of freedom is also determined as the sample size minus the number of population parameters that are estimated from the sample observations.

Application of t test

The t distribution has a number of applications in testing of hypothesis. Some mostly used test is:

1. test of significance of single mean, population variance being unknown.
2. test of significance difference between two means, the population variance being unknown but equal.
3. test of significance of an observed sample correlation coefficient.
4. test of significance of an observed partial and multiple correlation coefficient.

Test of significance of single mean

The steps in testing are as follows:

1. **Formulate null and alternative hypothesis**

Null hypothesis: H_o: $\mu = \mu_o$, there is no significant difference between sample mean and population mean or the samples have drawn from the population with mean μ_o and variance σ^2 or the population mean has a specified value μ_o.

Alternative hypothesis: H_1: $\mu \neq \mu_o$, there is a significant difference between sample mean and population mean or the samples have not been drawn from the population with mean μ_o and variance σ^2 or the population mean has not a specified value μ_o.

or H_1: $\mu > \mu_o$, the population mean is greater than μ_o (right tailed)

or H_1: $\mu < \mu_o$, the population mean is less than μ_o (left tailed)

2. **Compute the test statistic**

Under H_o the test statistic is,

$$t = \frac{\text{Difference}}{\text{Standard error}} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \sim t_{n-1}$$

Where, $\bar{x}$ = sample mean $= \frac{\sum x}{n}$, μ = population mean, s = sample standard deviation, n = sample size

s^2 is an unbiased estimate of the population variance σ^2 and is calculated by the formula,

$$s^2 = \frac{1}{n-1}\sum\left(x-\bar{x}\right)^2$$

This formula to calculate s^2 is appropriate when $\bar{x}$ is in whole number. When the value of s^2 is in fraction, the computation of s^2 by the above formula is quite tedious. In such case we use,

$$s^2 = \frac{1}{n-1}\left[\sum x^2 - \frac{\left(\sum x\right)^2}{n}\right]$$

When the number of observations are large, then we calculate $\bar{x}$ and s^2 by taking deviation of the items from any arbitrary value i.e. d = x – a, where a is called assumed mean.

$$\bar{x} = a + \frac{\sum d}{n},\ s^2 = \frac{1}{n-1}\left[\sum d^2 - \frac{\left(\sum d\right)^2}{n}\right]$$

When the biased estimate of population variance σ^2 i.e. s^2 is given, then the test statistic is,

$$t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n-1}}}$$

3. **3.** Chose the appropriate level of significance α and obtain the appropriate degree of freedom $\upsilon = n-1$.
4. **Critical value**: Obtain the critical or tabulated value of t from the table.
5. **Decision**: If $t_{cal} \leq t_{tab}$, null hypothesis is accepted, otherwise reject the null hypothesis.

Exercise

1. The average tablet potency of a batch of tablets based on assay of 10 randomly selected tablets is 98.88 mg with a standard deviation of 0.954. Test whether the samples have been from a batch with mean tablet potency 100 mg or not. (3.52)
2. Nine laboratory animals were infected with a certain bacterium and then immunosuppressed. The mean number of organisms later recovered from tissue specimen was 6.5 with standard deviation of 0.6. Can one conclude from these data that the population mean is greater than 6? Use 5% level of significance. (2.36)
3. Mean hemoglobin level of 25 pre school children was observed to be 10.6 gm/dl with a standard deviation of 1.15 gm/dl. Is it significantly different from population mean value of 11.0 gm/dl? (1.70)
4. The researcher claims that the average age of patients coming in the Hospital is 14 years. To test this claim, the Director of the Hospital took a sample of 25 patients. He found that the sample mean was 12.5 years and standard deviation of 2.8 years. Test this claim at 5% level of significance, assuming that the ages of the patients are approximately normally distributed.
5. The following are the systolic blood pressure (mm Hg) of 12 patients undergoing drug therapy for hypertension:
   183, 152, 178, 157, 194, 163, 144, 114, 178, 152, 118, 158
   Can we conclude on the basis of these data that the population mean is less than 165? Let $\alpha = 0.01$. (1.05)

6. The following are the basal pulse (beats/min) collected on a sample of 8 patients with hypertension: 5.1, 3.8, 8.2, 5.8, 7.0, 9.3, 2.5 and 6.5. Can we conclude that on the basis of these data that the population mean is greater than 8.25 at $\alpha = 0.01$? (2.82)
7. The following data are oxygen uptakes (milliliters) during incubation of a random sample of 15 cell suspensions: 14.0, 14.1, 14.5, 13.2, 11.2, 14.0, 14.1, 12.2, 11.1, 13.7, 13.2, 16.0, 12. 8, 14.4, 12.9. Do the data provide sufficient evidence at 0.05 level of significance that the population mean is not 12 ml? (4.33)

Test of significance difference between two means

The steps in testing are as follows:

1. **Formulate null and alternative hypothesis**

$H_o: \mu_1 = \mu_2$, there is no significant difference between two sample means or the samples have been drawn from normal population with same mean or two population mean do not differ significantly.

$H_1: \mu_1 \neq \mu_2$, the samples have not drawn from normal population with same mean.

or $H_1: \mu_1 > \mu_2$, mean of the first population is greater than mean of the second population (right tailed test)

or $H_1: \mu_1 < \mu_2$ mean of the first population is lesser than mean of the second population (left tailed test)

After identifying one tailed test, right tailed or left tailed can be identified as follows:

(i) If 1st sample statistic < 2nd sample statistic, we use left tailed test.

(ii) If 1st sample statistic > 2nd sample statistic, we use right tailed test.

2. **Compute the test statistic**

Under H_o the test statistic is,

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{S^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim t_{n_1+n_2-2}$$

Where $\bar{x}_1 = \text{mean of first sample} = \frac{\sum x_1}{n_1}$

$\bar{x}_2 = \text{mean of second sample} = \frac{\sum x_2}{n_2}$

$S^2 = \frac{1}{n_1 + n_2 - 2}\left[\sum\left(x_1 - \bar{x}_1\right)^2 + \sum\left(x_2 - \bar{x}_2\right)^2\right]$, if we have to calculate S.

or $S^2 = \frac{1}{n_1 + n_2 - 2}\left[\sum x_1{}^2 - \frac{\left(\sum x_1\right)^2}{n_1} + \sum x_2{}^2 - \frac{\left(\sum x_2\right)^2}{n_2}\right]$

When the sample variance are given, S^2 is calculated by the formula,

$$S^2 = \frac{n_1 s_1{}^2 + n_2 s_2{}^2}{n_1 + n_2 - 2}$$

3. **Level of significance and degree of freedom**: Fix a level of significance and degree of freedom $(n_1 + n_2 - 2)$.
4. **Critical value**: Obtain the tabulated value of t at α % level of significance for $(n_1 + n_2 - 2)$ d.f for one or two tailed test.

Note: If $(n_1 + n_2 - 2) > 30$, use the tabulated value of z at α% level of significance.

5. Decision: If $t_{cal} \le t_{tab}$, null hypothesis is accepted, otherwise reject the null hypothesis.

Exercise

1. Two drug manufacturing companies produce headache remedies. Each company claims that its drug brings faster-acting relief. A consumer protection agency wants to test if one drug brings relief faster than the other. An experiment was performed to compare the mean lengths of time required for bodily absorption of both drugs. The length of time in minutes for the drugs to reach a specified level in the blood was recorded for both drugs. The sample size, means and standard deviations of the two samples are recorded as follows

	n	$\bar{x}$	s
Drug A	12	10.1	4.2
Drug B	12	8.9	3.8

Use a 5% level of significance to test the hypothesis that there is no difference in the mean time required for bodily absorption of these two drugs.

2. A random sample of 10 patients at the emergency room of hospital A revealed an average waiting time before getting medical attention to be 31 minutes with a standard deviation of 4.63 minutes. A similar study of 8 patients at hospital B revealed an average waiting time of 24 minutes with a standard deviation of 3.87 minutes. Using 1% level of significance, can we conclude that there is a significant difference in the waiting time for patients in these hospitals before getting medical attention.

3. High blood pressure is a leading cause of strokes. Medical researchers are constantly seeking ways to treat patients suffering from this condition. A specialist in hypertension claims that regular aerobic exercise can reduce high blood pressure just as successfully as drugs, with none of the adverse side effects. To test the claim, 50 patients who suffer from high blood pressure were chosen to participate in an experiment. For 60 days half of the sample exercised three times per week for one hour; the other half took the standard medication. The percentage reduction in blood pressure was recorded for each individual and the resulting data are shown in the accompanying table. Can we conclude at the 1% significance level that exercise is at least as effective as medication in reducing hypertension?

	Exercise	Medication
Mean	14.31	13.28
Standard deviation	1.63	1.82

4. A group of seven week old chicken reared on a high protein diet weigh 12, 15, 11, 16, 14, 14 and 16 ounces, a second group of 5 chickens similarly treated except that they receive low protein diet weigh 8, 10, 14, 10 and 13 ounces. Test whether there is significant evidence that additional protein has increased the weight of the chicken. (2.42)

5. Chest circumference in cm of 10 normal children and 10 malnourished children aged one year are given below:
Normal group: 42, 46, 50, 48, 50, 52, 41, 49, 51, 56
Malnourished group: 38, 41, 36, 35, 30, 42, 31, 29, 31, 35
Test the statistical significance of the difference in chest circumference on an average between two groups. (6.73)

6. A drug formulated to be dissolved more rapidly by substituting lactose for part of lipoidal lubricant in the regular release product. The formulator is convinced that this dissolution for six tablets of each product (minutes) is:
   Original product: 25, 22, 29, 30, 26, 24
   Modified product: 18, 23, 24, 22, 19, 16
   Test whether the new formulation, cause faster dissolution of the drug. (3.23)

Paired t test

In testing the significance difference for dependent samples paired t test is applied for the cases like before and after, pre and post, with and without etc. The steps in testing are as follows:

1. **Formulate null and alternative hypothesis**

Null hypothesis: H_o: $\mu_x = \mu_y$ or $d = 0$, i.e. there is no significant difference between the observation before and after treatment or treatment is not effective

Alternative hypothesis: H_1: $\mu_x \neq \mu_y$ or $d \neq 0$, i.e. there is a significant difference between the observation before and after treatment or treatment is effective.

or H_1: $\mu_x > \mu_y$ or $d > 0$, the treatment is effective or there is negative impact in the observation after treatment.

or H_1: $\mu_x < \mu_y$ or $d < 0$, the treatment is effective or there is positive impact in the observation after treatment.

2. **Compute the test statistic**

Under H_o the test statistic is,

$$t = \frac{\bar{d}}{\frac{s}{\sqrt{n}}}$$

Where, $d = X - Y$ or $Y - X$ i.e. difference between set of observations

$\bar{d} = \frac{\sum d}{n}$ = mean of the difference

s = standard deviation of the distribution of the difference between the paired observation

$$s = \sqrt{\frac{1}{n-1}\sum\left(d-\bar{d}\right)^2} \text{ or } s = \sqrt{\frac{1}{n-1}\left[\sum d^2 - \frac{\left(\sum d\right)^2}{n}\right]}$$

3. Obtain a critical value of t at α% level of significance for $(n - 1)$ degree of freedom and give decision as before.

Exercise

1. The following pulmonary function data were collected on children with muscular dystrophy before and after a period of respiratory therapy.

	Forced vital capacity (liters)									
Before	74	65	84	89	84	65	78	86	83	82
After	79	78	100	92	104	70	81	84	85	90

   Test whether one should conclude that the therapy is effective at 5% level of significance. (3.33)

2. A researcher wanted to find the effect of a special diet on systolic blood pressure. She selected a sample of seven adults and put them on this dietary plan for three months. The following table gives the systolic blood pressure of these seven adults before and after the completion of this plan.

Before	210	180	195	220	231	199	224
After	193	186	186	223	220	183	233

Using the 1% significance level, can we conclude that the dietary plan is effective in reducing blood pressure? (1.23)

3. A certain stimulus administered to each of the 12 patients resulted in the following increase in blood pressure: 5, 2, 8, –1, 3, 0, –2, 1, 5, 0, 4 and 6. Can it be concluded that the stimulus will, in general be accompanied by an increase in blood pressure? (2.89)
4. A certain diet newly introduced to each of the 12 pigs resulted in the following increase in body weight.
   6, 3, 8, –2, 3, 0, –1, 1, 6, 0, 5 and 4
   Can you conclude that the diet is effective in increasing the weight of pigs?
5. Two laboratories carry out independent estimates of particular chemical in medicine produced by a certain firm. A sample is taken from each batch, halved and separate halves sent to the two laboratories. The following data is obtained.
   No. of samples 10
   Mean value of difference of estimates 0.6
   Sum of the squares of difference (from their mean) 20
   Is the difference significant?

Normality test

In statistics, normality tests are used to determine whether a data set is well-modeled by a normal distribution or not, or to compute how likely an underlying random variable is to be normally distributed.

Graphical methods

An informal approach to testing normality is to compare a histogram of the residuals to a normal probability curve. The actual distribution of the residuals (the histogram) should be bell-shaped and resemble the normal distribution. This might be difficult to see if the sample is small. In this case one might proceed by regressing the measured residuals against a normal distribution with the same mean and variance as the sample. If the regression produces an approximately straight line, then the residuals can safely be assumed to be normally distributed.

A more formal graphical tool is the normal probability plot, a quantile- quantile plot against the standard normal distribution. Here the correlation coefficient of the data (the goodness of fit of the best fit line) gives a measure of how well the data is modeled by a normal distribution. These also have the benefit that outliers stick out, and that they can be used for communication with non-statisticians more easily than numbers.

Histograms:

The histogram for each sample has a reference normal distribution curve for a normal distribution with the same mean and variance as the sample. This provides a reference for detecting gross nonnormality when the sample sizes are large.

Boxplots:
Suspected outliers appear in a boxplot as individual points o or x outside the box. If these appear on both sides of the box, they also suggest the possibility of a heavy-tailed distribution. If they appear on only one side, they also suggest the possibility of a skewed distribution. Skewness is also suggested if the mean (+) does not lie on or near the central line of the boxplot, or if the central line of the boxplot does not evenly divide the box. Examples of these plots will help illustrate the various situations.

Normal probability plot:
The normal probability plot may be the single most valuable graphical aid in diagnosing how a population distribution appears to differ from a normal distribution.
For values sampled from a normal distribution, the normal probability plot, (normal Q-Q plot) has the points all lying on or near the straight line drawn through the middle half of the points. Scattered points lying away from the line are suspected outliers. Examples of these plots will help illustrate the various situations.

Frequency tests
Normality tests include D'Agostino's K-squared test, the Jarque–Bera test, the Anderson–Darling test, the Cramér–von-Mises criterion, the Lilliefors test for normality (itself an adaptation of the Kolmogorov–Smirnov test), the Shapiro–Wilk test, the Pearson's chi-square test, and the Shapiro–Francia test for normality.
Historically, the third and fourth standardized moments (skewness and kurtosis) were some of the earliest test for normality; other early test statistics include the ratio of the mean absolute deviation to the standard deviation and of the range to the standard deviation.

Kolmogorov-Smirnov test
The Kolmogorov-Smirnov test can be applied to test whether data follow any specified distribution, not just the normal distribution. As a general test, it may not be as powerful as a test specifically designed to test for normality. Moreover, the Kolmogorov-Smirnov test becomes a conservative test (and thus loses power) if the mean and/or variance is not specified beforehand, but must be calculated from the sample data. And the Kolmogorov-Smirnov test will not indicate the type of nonnormality, say whether the distribution appears to be skewed or heavy-tailed. Examination of the calculated skewness and kurtosis, and of the histogram, boxplot, and normal probability plot for the data may provide clues as to why the data failed the Kolmogorov-Smirnov test.

The Kolmogorov-Smirnov one-sample test for normality is based on the maximum difference between the sample cumulative distribution and the hypothesized cumulative distribution. If the D statistic is significant, then the hypothesis that the respective distribution is normal should be rejected. For many software programs, the probability values that are reported are based on those tabulated by Massey (1951); those probability values are valid when the mean and standard deviation of the normal distribution are known a-priori and not estimated from the data. However, usually those parameters are computed from the actual data. In that case, the test for normality involves a complex conditional hypothesis ("how likely is it to obtain a D statistic of this magnitude or greater, contingent upon the mean and standard deviation computed from the data"), and the Lilliefors probabilities should be interpreted (Lilliefors, 1967). Note that in recent years, the Shapiro-Wilks' W test has become the preferred test of normality because of its good power properties as compared to a wide range of alternative tests.

Kolmogorov Smirnov test

Null and alternative hypothesis
H_o: This distribution follows a normal distribution.
H_1: This distribution does not follow a normal distribution.

Test statistic
The K – S statistic is the maximum absolute deviation of expected relative frequency (F_e) and the observed relative frequency (F_o). It is denoted by Dn.
$\therefore$ K – S statistic: $D_n = \max |F_e - F_o|$

Critical value and level of significance
The K – S test is always a one tailed test for a given level of significance α.
The critical values for D_n can be tabulated by using table.

Decision:
If $Dn_{cal} \leq Dn_{tab}$, H_o is accepted otherwise rejected.

Example: Below if the table of observed frequencies along with the frequency to the observed under a normal distribution.

(a) Calculate the K – S statistics.
(b) Can we conclude that this distribution does infact follow a normal distribution? Use 0.10 level of significance.

Test score	51 – 60	61 – 70	71 – 80	81 – 90	91 – 100
Observed frequency	30	100	440	500	130
Expected frequency	40	170	500	390	100

Solution:
H_o: This distribution follows a normal distribution.
H_1: This distribution does not follow a normal distribution.

Test score	Observed frequency	Observed Cumulative frequency	Observed relative frequency (F_o)	Expected frequency	Expected cumulative frequency	Expected relative frequency (F_e)	$D = \|F_e - F_o\|$
51 – 60	30	30	$\frac{30}{1200} = 0.025$	40	40	$\frac{40}{1200} = 0.033$	0.008
61 – 70	100	130	$\frac{130}{1200} = 0.108$	170	210	$\frac{210}{1200} = 0.175$	0.067

71 – 80	440	570	$\frac{570}{1200} = 0.475$	500	710	$\frac{710}{1200} = 0.592$	0.117
81 – 90	500	1070	$\frac{1070}{1200} = 0.891$	390	1100	$\frac{1100}{1200} = 0.92$	0.029
91 – 100	130	1200	$\frac{1200}{1200} = 1$	100	1200	$\frac{1200}{1200} = 1$	0

a) K – S statistic: $D_n = \max |F_e - F_o| = 0.117$

b) The tabulated value of D_n for n = 5 and $\alpha = 0.1$ is 0.510

Since $D_{cal} < D_{tab}$, H_o is accepted which means that the distribution follows a normal distribution.

Table

Critical values of D_n in the Kolmogorov Smirnov Goodness of Fit test

Sample size (n)	Level of significance for D_n = Maximum |Fe – Fo|				
	0.20	0.15	0.10	0.05	0.01
1	.900	.925	.950	.975	.995
2	.684	.726	.776	.842	.929
3	.565	.597	.642	.708	.828
4	.494	.525	.564	.624	.733
5	.446	.474	.510	.565	.669
6	.410	.436	.470	.521	.618
7	.381	.405	.438	.486	.577
8	.358	.381	.411	.457	.543
9	.339	.360	.388	.432	.514
10	.322	.342	.368	.410	.490
11	.307	.326	.352	.391	.468
12	.295	.313	.338	.375	.450
13	.284	.302	.325	.361	.433
14	.274	.292	.314	.349	.418
15	.266	.283	.304	.338	.404
16	.258	.274	.295	.328	.392
17	.250	.266	.286	.318	.381
18	.244	.259	.278	.309	.371
19	.237	.252	.272	.301	.363
20	.231	.246	.264	.294	.356
25	.21	.22	.24	.27	.32
30	.19	.20	.22	.24	.29
35	.18	.19	.21	.23	.27
Over 35	$\frac{1.07}{\sqrt{n}}$	$\frac{1.14}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$