MAT2001
Statistics for Engineers

Module 5
Hypothesis Testing I

Introduction

Population and Sample

Every statistical investigation aims at collecting information about some aggregate or collection of individuals or of their attributes, rather than the individuals themselves. In statistical language, such a collection is called *a population* or *universe*. For example, we have the population of products turned out by a machine, of lives of electric bulbs manufactured by a company etc. A population is finite or infinite, according as the number of elements is finite or infinite. In most situations, the population may be considered infinitely large. A finite subset of a population is called *a sample* and the process of selection of such samples is called *sampling*. The basic objective of the theory of sampling is to draw inference about the population using the information of the sample.

Parameters and Statistics

Generally in statistical investigations, our ultimate interest will lie in one or more characteristics possessed by the members of the population. If there is only one characteristic of importance, it can be assumed to be a variable x. If x_i be the value of x for the ith member of the sample, then $(x_1, x_2, \ldots, x_n)$ are referred to as sample observations. Our primary interest will be to know the values of different statistical measures such as mean and variance of the population distribution of x. Statistical measures, calculated on the basis of population values of x are called *parameters*. Corresponding measures computed on the basis of sample observations are called *statistics*.

$\theta_0 \rightarrow$ Parmeter

$\theta \rightarrow$ Statistic

Sampling Distribution

If the number of samples is large, the values of the statistic may be classified in the form of a frequency table. The probability distribution of the statistic that would be obtained if the number of samples, each of same size were infinitely large is called the *sampling distribution* of the statistic. If we adopt random sampling technique that is the most popular and frequently used method of sampling [the discussion of which is beyond the scope of this book], the nature of the sampling distribution of a statistic can be obtained theoretically, using the theory of probability, provided the nature of the population distribution is known.

Standard Error

Like any other distribution, a sampling distribution will have its mean, standard deviation and moments of higher order. The standard deviation of the sampling distribution of a statistic is of particular importance in tests of Hypotheses and is called *the standard error* of the statistic.

Statistical Hypotheses

When we attempt to make decisions about the population on the basis of sample information, we have to make assumptions or guesses about the nature of the population involved or about the value of some parameter of the population. Such assumptions, which may or may not be true, are called *statistical hypotheses*.

Null and Alternative Hypotheses

Very often, we set up a hypothesis which assumes that there is no significant difference between the sample statistic and the corresponding population parameter or between two sample statistics. Such a hypothesis of no difference is called a *null hypothesis* and is denoted by H_0. A hypothesis that is different from (or complementary to) the null hypothesis is called an *alternative hypothesis* and is denoted by H_1. A procedure for deciding whether to accept or to reject a null hypothesis (and hence to reject or to accept the alternative hypothesis respectively) is called *the test of hypothesis.*

Significant Difference and Test of Significance

If θ_0 is a parameter of the population and θ is the corresponding sample statistic, usually there will be some difference between θ_0 and θ since θ is based on sample observations and is different for different samples. Such a difference which is caused due to sampling fluctuations is called *insignificant difference*. The difference that arises due to the reason that either the sampling procedure is not purely random or that the sample has not been drawn from the given population is known as *significant difference*. This procedure of testing whether the difference between θ_0 and θ is significant or not is called *the test of significance.*

Critical Region

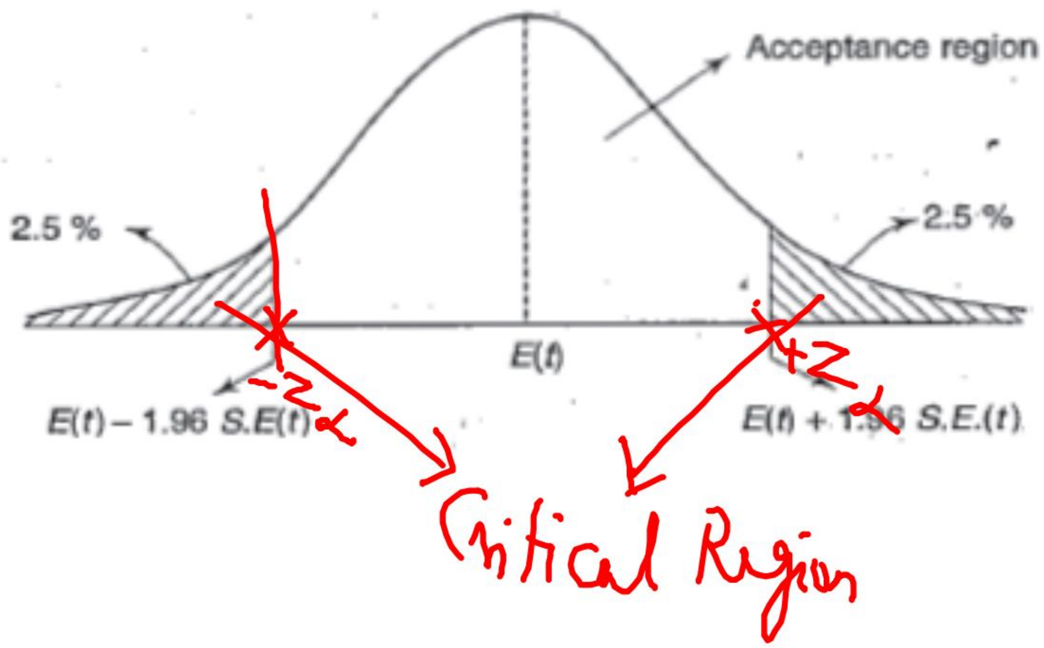
If we are prepared to reject a null hypothesis when it is true or if we are prepared to accept that the difference between a sample statistic and the corresponding parameter is significant, when the sample statistic lies in a certain region or interval, then that region is called the *critical region* or *region of rejection*. The region complementary to the critical region is called *the region of acceptance*.

In the case of large samples, the sampling distributions of many statistics tend to become normal distributions. If 't' is a statistic in large samples, then t follows a normal distribution with mean $E(t)$, which is the corresponding population parameter, and S.D. equal to S.E. (t). Hence $\boxed{Z = \frac{t - E(t)}{S \cdot E \cdot (t)}}$ is a standard normal variate i.e., Z (called *the test statistic*) follows a normal distribution with mean zero and S.D unity.

$$Z = \frac{X - \mu}{\sigma}$$

Critical Region

It is known from the study of normal distribution, that the area under the standard normal curve between $t = -1.96$ and $t = +1.96$ is 0.95. Equivalently the area under the general normal curve of 't' between $[E(t) - 1.96 \text{ S.E. } (t)]$ and $[E(t) + 1.96 \text{ S.E. } (t)]$ is 0.95. In other words, 95 per cent of the values of t will lie between $[E(t) \mp 1.96 \text{ S.E. } (t)]$ or only 5 per cent of values of t will lie outside this interval.



Level of Significance

The probability 'α' that a random value of the statistic lies in the critical region is called *the level of significance* and is usually expressed as a percentage.

Note *The level of significance can also be defined as the maximum probability with which we are prepared to reject* H_0 *when it is true. In other words, the total area of the region of rejection expressed as a percentage is called the level of significance.*

Critical Values or Significant Values

The value of the test statistic z for which the critical region and acceptance region are separated is called *the critical value* or *the significant value* of z and denoted by z_α, when α is the level of significance. It is clear that the value of z_α depends not only on α but also on the nature of alternative hypothesis.

Types of Errors

The level of significance is fixed by the investigator and as such it may be fixed at a higher level by his wrong judgement. Due to this, the region of rejection becomes larger and the probability of rejecting a null hypothesis, when it is true, becomes grater. The error committed in rejecting H_0, when it is really true, is called *Type I error*. This is similar to a good product being rejected by the consumer and hence Type I error is also known as *producer's risk*. The error committed in accepting H_0, when it is false, is called *Type II error*. As this error is similar to that of accepting a product of inferior quality, it is also known as *consumer's risk*.

The probabilities of committing Type I and II errors are denoted by α and β respectively. It is to be noted that the probability α of committing Type I error is the level of significance.

One-Tailed and Two-Tailed Tests

If θ_0 is a population parameter and θ is the corresponding sample statistic and if we set up the null hypothesis $H_0 : \theta = \theta_0$, then the alternative hypothesis which is complementary to H_0 can be any one of the following:

(i) $H_1 : \theta \neq \theta_0$, i.e. $\theta > \theta_0$ or $\theta < \theta_0$

(ii) $H_1 : \theta > \theta_0$

(iii) $H_1 : \theta < \theta_0$.

H_1 given in (i) is called a two tailed alternative hypothesis, whereas H_1 given in (ii) is called a right-tailed alternative hypothesis and H_1 given in (iii) is called a left-tailed alternative hypothesis.

When H_0 is tested while H_1 is a one-tailed alternative (right or left), the test of hypothesis is called *a one-tailed test*.

When H_0 is tested while H_1 is two-tailed alternative, the test of hypothesis is called *a two-tailed test*.

Table for *Z*-Test

The critical values for some standard LOS's are given in the following table both for two-tailed and one-tailed tests

Nature of test | LOS $\alpha\%$	1% (.01)	2% (.02)	5% (.05)	10% (.1)
Two-tailed	$\lvert z_\alpha \rvert = 2.58$	$\lvert z_\alpha \rvert = 2.33$	$\lvert z_\alpha \rvert = 1.96$	$\lvert z_\alpha \rvert = 1.645$
Right-tailed	$z_\alpha = 2.33$	$z_\alpha = 2.055$	$z_\alpha = 1.645$	$z_\alpha = 1.28$
Left-tailed	$z_\alpha = -2.33$	$z_\alpha = -2.055$	$z_\alpha = -1.645$	$z_\alpha = -1.28$

Procedure for Testing Hypothesis

1. Null hypothesis H_0 is defined.
2. Alternative hypothesis H_1 is also defined after a careful study of the problem and also the nature of the test (whether one-tailed or two tailed) is decided.
3. LOS 'α' is fixed or taken from the problem if specified and z_α is noted.
4. The test-statistic $z = \dfrac{t - E(t)}{S \cdot E \cdot (t)}$ is computed.
5. Comparison is made between $|z|$ and z_α. If $|z| < z_\alpha$, H_0 is accepted or H_1 is rejected, i.e. it is concluded that the difference between t and $E(t)$ is not significant at α % L.O.S.

On the other hand, if $|z| > z_\alpha$, H_0 is rejected or H_1 is accepted, i.e. it is concluded that the difference between t and $E(t)$ is significant at α % L.O.S.

Procedure for Testing Hypothesis

① . Null Hypothesis:

$H_0 : \theta = \theta_0$

(or)

$H_0 : \theta_1 = \theta_2$

② . Alternative Hypothesis:

* $H_1 : \theta \neq \theta_0 \ (\theta_1 \neq \theta_2)$ → Two-Tailed Test
* $H_1 : \theta > \theta_0$ → Right One-Tailed Test
* $H_1 : \theta < \theta_0$ → Left One-Tailed Test

③ . LOS $= \alpha\% = 1\%$ (or) 5%

$= 0.01$ (or) 0.05

$Z_{\alpha\%} = Z_{Tab}$

④ . Test Statistic: $Z_{Cal} = \dfrac{t - E(t)}{S.E(t)}$

⑤ . Comparison and Conclusion:

(i). $|Z_{Cal}| < |Z_{Tab}|$

Then, H_0 is accepted (or) H_1 is rejected.

(ii). $|Z_{cal}| > |Z_{Tab}|$

Then, H_0 is rejected (or) H_1 is accepted.

—— × ——

Types of Sampling Theory

1. Large Sample:

Size of the Sample (n) is greater than or equal to 30.

2. Small Sample:

Size of the Sample (n) is less than 30.

Large Sample Tests (*Z*-Tests)

Tests of Significance for Large Samples

It is generally agreed that, if the size of the sample exceeds 30, it should be regarded as a large sample. The tests of significance used for large samples are different from the ones used for small samples for the reason that the following assumptions made for large samples do not hold for small samples

1. The sampling distribution of a statistic is approximately normal, irrespective of whether the distribution of the population is normal or not.
2. Sample statistics are sufficiently close to the corresponding population parameters and hence may be used to calculate the standard error of the sampling distribution.

Z-Test for Single Proportion

Test I

Test of significance of the difference between sample proportion and population proportion.

The test statistic $z = \dfrac{p - P}{\sqrt{\dfrac{PQ}{n}}}$.

Note 1. *If* P *is not known, we assume that* p *is nearly equal to* P *and hence* S.E. (p) *is taken as* $\sqrt{\dfrac{pq}{n}}$. *Thus* $z = \dfrac{p - P}{\sqrt{\dfrac{pq}{n}}}$.

2. *95 per cent confidence limits for* P *are then given by* $\dfrac{|P - p|}{\sqrt{\dfrac{Pq}{n}}} \le 1.96$, *i.e. they are* $\left(p - 1.96\sqrt{\dfrac{pq}{n}},\ p + 1.96\sqrt{\dfrac{pq}{n}}\right)$.

Z-Test for Difference of Proportions

Test 2

Test of significance of the difference between two sample proportions.

Let p_1 and p_2 be the proportions of successes in two large samples of size n_1 and n_2 respectively drawn from the same population or from two population with the same proportion P.

The test statistic $$z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}.$$

Note:

If P is not known, an unbiased estimate of P based on both samples, given by

$$\hat{P} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2},$$ is used in the place of P.

Z-Test for Single Mean

Test 3

Test of significance of the difference between sample mean and population mean.

The test statistic $z = \dfrac{\bar{X} - \mu}{\sigma / \sqrt{n}}$.

Note 1. *If σ is not known, the sample S.D. 's' can be used in its place, as s is nearly equal to σ when n is large.*

2. 95% confidence limits for μ are given by $\dfrac{|\mu - \bar{X}|}{\sigma / \sqrt{n}} \leq 1.96$, i.e. $\left(\bar{X} - 1.96 \dfrac{\sigma}{\sqrt{n}}, \bar{X} + 1.96 \dfrac{\sigma}{\sqrt{n}}\right)$, if σ is known. If σ is not known, then the 95% confidence interval is $\left(\bar{X} - \dfrac{1.96\, s}{\sqrt{n}}, \bar{X} + \dfrac{1.96\, s}{\sqrt{n}}\right)$

Z-Test for Difference of Means

Test 4

Test of significance of the difference between the means of two samples.

Let $\overline{X}_1$ and $\overline{X}_2$ be the means of two large samples of sizes n_1 and n_2 drawn from two populations (normal or non-normal) with the same mean μ and variances σ_1^2 and σ_2^2 respectively.

The test statistic $$z = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \qquad (1)$$

Z-Test for Difference of Means

Note 1. If the samples are drawn from the same population, i.e. if $\sigma_1 = \sigma_2 = \sigma$ then

$$z = \frac{\bar{X}_1 - \bar{X}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad (2)$$

2. If σ_1 and σ_2 are not known and $\sigma_1 \neq \sigma_2$, σ_1 and σ_2 can be approximated by the sample S.D.'s σ_1 and σ_2. Hence, in such a situation,

$$z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \qquad (3) \text{ [from (1)]}$$

3. If σ_1 and σ_2 are equal and not known, then $\sigma_1 = \sigma_2 = \sigma$ is approximated by $\sigma^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2}$. Hence, in such a situation,

$$z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}, \quad \text{from (2)}$$

i.e.

$$z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_2} + \frac{s_2^2}{n_1}}} \qquad (4)$$

4. The difference in the denominators of the values of z given in (3) and (4) may be noted.

Example:

The fatality rate of typhoid patients is believed to be 17.26 per cent. In a certain year 640 patients suffering from typhoid were treated in a metropolitan hospital and only 63 patients died. Can you consider the hospital efficient?

Solution:

$H_0 : p = P$, i.e. the hospital is not efficient. $H_1 : p < P$.

One-tailed (left-tailed) test is to be used

Let us assume that LOS = 1%. $\therefore$ $z_\alpha = -2.33$

$$z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}, \text{ where } p = \frac{63}{640} = 0.0984 \text{ and}$$

$$P = 0.1726 \quad \text{and hence} \quad Q = 0.8274.$$

$$z = \frac{0.0984 - 0.1726}{\sqrt{\frac{0.1726 \times 0.8274}{640}}} = -4.96$$

$\therefore$ $\quad |z| > |z_\alpha|$

$\therefore$ The difference between p and P is significant. i.e., H_0 is rejected and H_1 is accepted.

i.e. The hospital is efficient in bringing down the fatality rate of typhoid patients.

Exercise:

A salesman in a departmental store claims that at most 60 percent of the shoppers entering the store leaves without making a purchase. A random sample of 50 shoppers showed that 35 of them left without making a purchase. Are these sample results consistent with the claim of the salesman? Use a level of significance of 0.05.

Example:

In a large city A, 20 per cent of a random sample of 900 school boys had a slight physical defect. In another large city B, 18.5 percent of a random sample of 1600 school boys had the same defect. Is the difference between the proportions significant?

Solution:

$$p_1 = 0.2, \quad p_2 = 0.185, \quad n_1 = 900 \quad \text{and} \quad n_2 = 1600$$

$$H_0 : p_1 = p_2$$

$$H_1 : p_1 \neq p_2$$

Two tailed test is to be used.

Let L.O.S. be 5% $\therefore z_\alpha = 1.96$

$$z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad (1)$$

Since P, the population proportion, is not given, we estimate it as $\hat{P} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{180 + 296}{900 + 1600} = 0.1904.$

Using in (1), we have

$$z = \frac{0.2 - 0.185}{\sqrt{0.1904 \times 0.8096 \times \left(\frac{1}{900} + \frac{1}{1600}\right)}} = 0.92$$

$|z| \leq z_\alpha$. Therefore The difference between p_1 and p_2 is not significant at 5 per cent level.

Exercise:

Before an increase in excise duty on tea, 800 people out of a sample of 1000 were consumers of tea. After the increase in duty, 800 people were consumers of tea in a sample of 1200 persons. Find whether there is significant decrease in the consumption of tea after the increase in duty.

Example:

A sample of 100 students is taken from a large population. The mean height of the students in this sample is 160 cm. Can it be reasonably regarded that, in the population, the mean height is 165 cm, and the S.D. is 10 cm?

Solution:

A sample of 100 students is taken from a large population. The mean height of the students in this sample is 160 cm. Can it be reasonably regarded that, in the population, the mean height is 165 cm, and the S.D. is 10 cm?

$$\bar{x} = 160, \quad n = 100, \quad \mu = 165 \quad \text{and} \quad \sigma = 10.$$

$H_0 : \bar{x} = \mu$ (i.e. the difference between $\bar{x}$ and μ is not significant)

$H_1 : \bar{x} \neq \mu.$

Two-tailed test is to be used.

Let LOS be 1% $\therefore z_\alpha = 2.58$

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{160 - 165}{10 / \sqrt{100}} = -5$$

Now $\quad |z| > z_\alpha.$

$\therefore$ The difference between $\bar{x}$ and μ is significant at 1% level.

i.e. H_0 is rejected.

i.e. it is not statistically correct to assume that $\mu = 165$.

Exercise:

The mean breaking strength of the cables supplied by a manufacturer is 1800 with a S.D. of 100. By a new technique in the manufacturing process, it is claimed that the breaking strength of the cable has increased. In order to test this claim, a sample of 50 cables is tested and it is found that the mean breaking strength is 1850. Can we support the claim at 1 per cent level of significance?

Example:

A simple sample of heights of 6400 English men has a mean of 170 cm and a S.D. of 6.4 cm, while a simple sample of heights of 1600 Americans has a mean of 172 cm and a S.D. of 6.3 cm. Do the data indicate that Americans are, on the average, taller than the Englishmen?

Solution:

$$n_1 = 6400,\ \bar{x}_1 = 170 \quad \text{and} \quad s_1 = 6.4$$

$$n_2 = 1600,\ \bar{x}_2 = 172 \quad \text{and} \quad s_2 = 6.3$$

$$H_0 : \mu_1 = \mu_2 \quad \text{or} \quad \bar{x}_1 = \bar{x}_2,$$

i.e. the samples have been drawn from two different populations with the same mean.

$$H_1 : \bar{x}_1 < \bar{x}_2 \quad \text{or} \quad \mu_1 < \mu_2.$$

Left-tailed test is to be used.

Let LOS be 1%. $\therefore z_\alpha = -2.33$

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

[$\because \sigma_1 \simeq s_1$ and $\sigma_2 \simeq s_2$. Refer to Note 2 under Test 4]

$$= \frac{170 - 172}{\sqrt{\frac{(6.4)^2}{6400} + \frac{(6.3)^2}{1600}}} = -11.32$$

Now $\quad |z| > |z_\alpha|$

$\therefore$ The difference between $\bar{x}_1$ and $\bar{x}_2$ (or μ_1 and μ_2) is significant at 1% level.

i.e. H_0 is rejected and H_1 is accepted.

i.e. The Americans are, on the average, taller than the Englishmen.

Exercise:

In a random sample of size 500, the mean is found to be 20. In another independent sample of size 400, the mean is 15. Could the samples have been drawn from the same population with S.D. 4?