Test of Hypothesis

Introduction:

Hypothesis tests are widely used in business and industry for making decisions. Probability and sampling theory plays an ever increasing role in constructing the criteria on which business decisions are made.

For example, in order to increase consumers awareness of a product or service, it might be necessary to compare the effectiveness of different types of advertising campaigns.

In order to offer more profitable investment opportunities to its customers, an investment firm might wish to compare the profitability of different types of investment of portfolios.

If a sample contains 30 or more $(n \geq 30)$ observations, then it is called large sample. A study of test statistic for large samples is called large sample statistic

In this chapter, we will be studying only normal test (Z-test) for large sample ($n \geq 30$) and also three important test statistic such as students t-test, F-test, and χ^2-test for small sample ($n < 30$).

Introduction:

In most of the situations, it is very difficult to study the whole population. The value of the population parameter is usually unknown, and one objective of sampling is to estimate its value. The choice of appropriate statistics depends on which population parameter is of interest to statistician. Any inference drawn about the population is based on sample statistic.

Now the question arises, whether the sample statistic is a represtative value of respective population parameter or whether there is any significant difference between the parameter and statistic to some extent. This ~~pa~~ matter can be ensured by the test of hypothesis.

For example, a reputed toothpaste producer claims that average weight of its big size toothpaste is 140 gm. The customer association of Bangladesh (CAB) can verify ~~this~~ claim by the following steps.

(I) Collecting a random sample of toothpaste

(II) Determining the average weight of toothpastes

& (III) Performing a test of hypothesis about the mean weight.

W

Definition of Hypothesis Testing:

The process that enables a decision-maker to draw an inference about population characteristics by analyzing the difference between the value obtained from sample and the hypothesized value of parameter is called hypothesis testing.

Concepts of Hypothesis Testing:

The process of hypothesis testing can be compared with criminal jury trial. In a jury trial it is assumed that the criminal is innocent, and the jury will decide that a person is guilty only if there is a very strong evidence against the presumption of innocence. This criminal jury trial process for choosing between guilt and innocence possesses

① Rigorous procedures for presenting and evaluating evidence.

(II) A judge to enforce the rules.

(III) A decision process that assumes innocence unless there is evidence to prove guilt beyond a reasonable doubt.

It is to be noted that this process will fail to convict a number of people who are, in fact, guilty.

But if a person's innocence is rejected and the person is found guilty, we have a strong belief that the person is guilty. In testing the hypothesis, the sample collected from population acts as evidence of trial.

Hypothesis & Statistical Hypothesis :

Every moment, we use different types of statements regarding different events of our life. Statistics is not concerned with all types of statements. Statistics usually deals with statements, which have some relation with uncertainity.

Thus, any statement about any aspect of a phenomenon is considered as hypothesis.

Tomorrow will be a sunny day, he/she is the president of some association, the firm is not running well, etc. are examples of the statements about tomorrow's weather, chief of an association, state of the firm, respectively. Hence, these are sample hypothesis.

However, in attempting to take decision regarding some characteristics of population on the basis of sample, it is necessary to make some assumption regarding parameters of the population, such assumptions which may or may not be true, are called statistical hypothesis.

Hence, Statistical Hypothesis is a claim or statement (belief or assumption) about unknown feature (distribution or parameter) of a population.

For example, average monthly sale of a store is TK. 10,000, average hourly production of a machine is 200 units, proportion of defective products produced by a certain machine is less than other and so on.

Definition of Hypothesis :

Any statement about any phenomenon is termed as hypothesis.

Definition of Statistical Hypothesis :

Statistical hypothesis is a statement about population characteristic that can be tested on the basis of sample data.

For example, a pious person will go to the heaven is not a statistical hypothesis since it can not be proved with the statistical data. It does not mean that it has no value. It is a faith.

In statistical tests of significance, two mutually exclusive hypotheses are to be used: these are null hypothesis and alternative hypothesis.

Parametric Hypothesis :

Any hypothesis about the parameter of a population distribution is known as parametric hypothesis.

For example, Suppose, the average age of all students in a college is 20 years. The hypothesis about their mean value is called parametric hypothesis.

Non-Parametric Hypothesis :

Any hypothesis about a population distribution is called a non-parametric hypothesis.

For example, a company's owner claim that the weight of his produced commodity follows the normal distribution. The hypothesis about the weight of his produced commodity is called non-parametric hypothesis.

Null Hypothesis :

The approach of statistical hypothesis testing starts with a statement complement to the original claim. The hypothesis about the parameter of a population such as mean μ, the variance σ^2, or the proportion π which is formulated for sole purpose of rejecting or nullifying it, is called null hypothesis. Hence, null hypothesis is a statement about no difference between the parameter and statistic. Null hypothesis is denoted by H_0.

For example, if we want to decide whether a given coin is not fair, we formulate the null hypothesis that the coin is fair, i.e,

$H_0 : \pi = 0.5$,

Definition of Null Hypothesis :

The hypothesis that is formulated for its possible rejection using sample data is called null hypothesis. Null hypothesis is denoted by H_0.

Alternative Hypothesis :

The alternative hypothesis is a logical opposite statement of null hypothesis. If null hypothesis is rejected (or actually it is false), then some alternative form of parameter should be true.

Thus, any hypothesis that differs from a given null hypothesis is called an alternative hypothesis. Alternative hypothesis is denoted by H_1 or H_A.

For example, if the null hypothesis about population mean is $H_0 : \mu_0 = 55$, an alternative might be $H_1 : \mu_1 \neq 55$ or $\mu_1 < 55$ or $\mu_1 > 55$.

Definition of Alternative Hypothesis :

The hypothesis, which is true if the null hypothesis is false is called alternative hypothesis. Alternative hypothesis indicates the type of test (left, right, or two-tailed) It is denoted by H_1 or H_A.

Simple Hypothesis :

A hypothesis is said to be a simple hypothesis if it completely specifies the distribution of the population from which the sample has been taken or considered. In this case, the information about all the parameters of population distribution is known.

For example, if a coin is tossed 50 times (that means n of binomial distribution is 50) to determine if the coin is fair one, the null hypothesis to be formulated as $H_0: \pi = 0.50$, which is a simple hypothesis because it specifies the population distribution completely.

Again, for testing $H_0: \mu = \mu_0$ of a normal distribution if the population variance σ^2 is known, it is a simple hypothesis.

Definition of Simple Hypothesis :

The hypothesis, which completly specifies all the parameters of the related population, is called simple hypothesis

Composite Hypothesis :

On the other hand, if a hypothesis does not specify the population distribution completely, it is called a composite hypothesis. In the above coin example, if $n = 50$ is not specified, $H_0 : \pi = 0.50$, would be a composite hypothesis, because there is a number of distributions all with $\pi = 0.50$.

In this case, n is called a nuisance parameter.

Similarly, even if we know $n = 50$ but the hypothesis to be tested is defined as $H_0 : \pi \neq 0.50$ or $H_0 : \pi > 0.50$, it would be composite hypothesis, because the value of π is not specified by a single value.

Again, for testing $H_0 : \mu = \mu_0$ of a normal distribution if the information about σ^2 is not given, it is a composite hypothesis.

In this case, the parameter σ^2 is known as nuisance parameter.

Definition of Composite Hypothesis:

The hypothesis, which does not completely specify the parameters, is called a composite hypothesis.

Errors in Decision-making :

In any decision-making, if the decision is not correct, the decision-maker may commit error in two mutually exclusive ways, termed as type-I error and type-II error.

Decision	State of nature	
	H_0 is True	H_0 is False
Reject H_0	Type-I error	Correct decision
Fail to reject H_0 or Accept H_0	Correct decision	Type-II error

Definition of Type-I error:

The error of rejecting the null hypothesis when it is in fact true is called type-I error. That means, type-I error occurs when null hypothesis is wrongly rejected. A type-I error is also known as first kind of error.

Definition of Type-II error:

The error of accepting the null hypothesis when it is false is called type-II error. Type-II error occurs when null hypothesis is accepted wrongly. A type-II error is also known as second kind of error.

A type-I error is often considered to be more serious, and therefore more important to avoid than a type-II error. The hypothesis test procedure is therefore adjusted so that there is a guaranteed 'low' probability of rejecting the null hypothesis wrongly. This probability is never 0. While the exact probability of type-II error is generally unknown.

If we do not reject the null hypothesis, it may still be false (a type-II error) as the sample may not be representative enough to identify the falseness of the null hypothesis (especially if the truth is very close to hypothesis)

For any given set of data, type-I & type-II errors are inversely related; the smaller the risk of one, the higher the risk of the other.

Although it is desirable to keep the both types of errors at minimum level, but unfortunately in practice it is not possible. Hence the probability of type I error is kept fixed (to be considered at the beginning of testing procedure) and then try to minimize the probability of type-II error.

Scanned with CamScanner

Level of Significance :

In testing a given hypothesis, the maximum probability with which we would be willing to take risk of rejecting a hypothesis when it should be accepted, is called the level of significance of the test. This probability is denoted by α, generally specified before any sample is drawn so that the results obtained will not influence the choice of the decision-maker.

- Since type-I is the more serious error (usually) that is the one we concentrate on. We usually pick it to be very small such as 0.05, 0.01 or in some cases 0.001. It is to be noted here that α is not a type-I error, α is the probability of committing a type-I error and β is the probability of committing a type-II error.

Definition of Level of Significance :-

The probability of committing a type-I error is called the level of significance. In other words, it is the total area under the critical region.

Symbolically, $\alpha = P(\text{Reject } H_0 \mid H_0 \text{ is true})$

It is also known as size of a test.

$1-\alpha = P(\text{Accept } H_0 / H_0 \text{ is true})$ is called the confidence interval.

Scanned with CamScanner

Definition of Power of a Test:

The complement of the probability of type-II error is called the power of a test. That means, the probability of rejecting a false null hypothesis is the power of a test. In other words, the probability of correct decision is the power of a test. Symbolically

Power of a test $1-\beta = 1 - P(\text{Accept } H_0 / H_0 \text{ is false})$

$$= P(\text{Reject } H_0 / H_0 \text{ is false})$$

Interpretation of Level of Significance :

Generally, a significance level of 0.05 or 0.10 is considered, although other values are also used. Thus, if the level of significance is 0.05, it will mean that there are about 5 samples out of 100 that would direct to reject the hypothesis when it should be actually accepted.

So $(1-0.05) = 0.95$ is the probability of accepting null hypothesis when it is true, i.e., there is 95% confidence in taking the right decision.

In such case it is said that the hypothesis has been rejected at 5% level of significance, which again means that the probability of wrong decision is 0.05.

One-tailed & two-tailed test:

A one-tailed test is a test, which is concerned about possible deviation of the value of the parameter in only one direction from the specified value defined in the null hypothesis, while a two-tailed test is a test, which is concerned about the possible deviation of the parametric value in both directions.

These are also called a one-sided alternative or two-sided alternative.

In this case, the parameter can take any value other than the value specified by null hypothesis.

For example, $H_0: \mu = 0$ against $H_1: \mu \neq 0$ is a two-tailed test, while $H_0: \mu = 0$, against $H_1: \mu > 0$ or $H_0: \mu = 0$, against $H_1: \mu < 0$ is a one-tailed test.

A left-tailed test:

When the rejection region is in the left tail of the distribution of the test statistic, the test is called a left-tailed test. If the null hypothesis is $H_0: \mu_0 = 0$, then the alternative hypothesis will be $H_1: \mu < 0$.

A right-tailed test:

When the rejection region is in the right tail of the distribution of the test statistic, the test is called a right-tailed test. If the null hypothesis is $H_0: \mu_0 = 0$ then the alternative hypothesis will be $H_1: \mu > 0$.

Alternative hypotheses defined in the above two tests are called one-sided alternatives.

A two-tailed test:

When the rejection region is equally divided in the left and right tails of the distribution of the test statistic, the test is called a two-tailed test. The alternative hypotheses defined in this test is called two-sided alternatives. If the null hypothesis is $H_0: \mu_0 = 0$, then the two sided alternative hypothesis is defined by $H_1: \mu_0 \neq 0$.

- In the test of significance, the decision whether the use of a one-tailed or a two-tailed test is appropriate depends on the objective of the test, i.e., it is chosen on the basis of the direction of claimed deviation of the parameter.

Suppose that a manufacturer of ballpoint pen, whose machine produces on average 1000 pens per hour, is planning to purchase a new machine.

The authority will not buy a new machine unless it is definitely proved as superior. For this he would test the hypothesis that the new machine is no better than the existing machine or similar to the machine now available. in the market against the alternative that the new machine is superior with respect to the hourly average production of the existing machine.

In other words, it is required to test the null hypothesis $H_0 : \mu = 1000$ against the alternative $H_1 : \mu > 1000$ and buy the new machine only if the null hypothesis is rejected. Such an alternative test will result in a one_tailed test with the critical region in the right tail.