3·1. INTRODUCTION

Statistical inference is that branch of statistics which is concerned with using probability concept to deal with uncertainty in decision-making. The field of statistical inference has had a fruitful development since the latter half of the 19th century.

It refers to the process of selecting and using a sample statistic to draw inference about a population parameter based on a subset of it—the sample drawn from the population. Statistical inference treats two different classes of problems :

1. *Hypothesis testing,** *i.e.,* to test some hypothesis about parent population from which the sample is drawn.

2. *Estimation, i.e.,* to use the 'statistics' obtained from the sample as estimate of the unknown 'parameter' of the population from which the sample is drawn.

In both these cases the particular problem at hand is structured in such a way that inferences about relevant population values can be made from sample data.

Hypothesis Testing. Hypothesis testing begins with an assumption, called a *Hypothesis,* that we make about a population parameter. A hypothesis is a supposition made as a basis for reasoning. According to Prof. Morris Hamburg, "A hypothesis in statistics is simply a quantitative statement about a population." Palmer O Johnson has beautifully described hypothesis as "islands in the uncharted seas of thought to be used as bases for consolidation and recuperation as we advance into the unknown."

There can be several types of hypotheses. For example, a coin may be tossed 200 times and we may get heads 80 times and tails 120 times. We may now be interested in testing the hypothesis that the coin is unbiased. To take another example we may study the average weight of the 100 students of a particular college and may get the result as 110 lb. We may now be interested in testing the hypothesis that the sample has been drawn from a population with average weight 115 lb. Similarly, we may be interested in testing the hypothesis that the variable in the population are uncorrelated.

3·2. PROCEDURE OF TESTING HYPOTHESIS

The procedure of testing hypotheses is briefly described below :

1. **Set up a hypothesis.** The first thing in hypothesis testing is to set up a hypothesis about a population parameter. Then we collect sample data produce sample statistics, and use this information to decide how likely it is that our hypothesized population parameter is correct. Say, we assume a certain

value for a population mean. To test the validity of our assumption, we gather sample data and determine the difference between the hypothesized value and the actual value of the sample mean. Then we judge whether the difference is significant. The smaller the difference, the greater the likelihood that our hypothesized value for the mean is correct. The larger the difference, the smaller the likelihood.

The conventional approach to hypothesis testing is not to construct a single hypothesis about the population parameter, but rather to set up two different hypotheses. These hypotheses must be so constructed that if one hypothesis is accepted, the other is rejected and *vice versa*.

The two hypotheses in a statistical test are normally referred to as :

(*i*) Null hypothesis, and

(*ii*) Alternative hypothesis.

The null hypothesis is a very useful tool in testing the significance of difference. In its simplest form the hypothesis asserts that there is no real difference in the sample and the population in the particular matter under consideration (hence the word "null" which means invalid, void, or amounting to nothing) and that the difference found is accidental and unimportant arising out of fluctuations of sampling. The null hypothesis is akin to the legal principle that a man is innocent until he is proved guilty. It constitutes a challenge; and the function of the experiment is to give the facts a chance to refute (or fail to refute) this challenge. For example, if we want to find out whether extra coaching has benefited the students or not, we shall set up a null hypothesis that "extra coaching has not benefited the students". Similarly, if we want to find out whether a particular drug is effective in curing malaria we will take the null hypothesis that "the drug is not effective in curing malaria". The rejection of the null hypothesis indicates that the differences have statistical significance and the acceptance of the null hypothesis indicates that the differences are due to chance. Since many practical problems aim at establishment of statistical significance of differences, rejection of the null hypothesis may thus indicate success in statistical project.

As against the null hypothesis, the alternative hypothesis specifies those values that the researcher believes to hold true, and, of course, he hopes that the sample data lead to acceptance of this hypothesis as true. The alternative hypothesis may embrace the whole range of values rather than single point. Now a days, it is usually accepted common practice not to associate any special meaning two the null or alternative hypothesis but merely to let these terms represent to different assumptions about the population parameter. However, for statistical convenience it will make a difference as to which hypothesis is called the null hypothesis and which is called the alternative.

The null and alternative hypotheses are distinguished by the use of two different symbols, H_0 representing the null hypothesis and H_a the alternative hypothesis. Thus a psychologist who wishes to test whether or not a certain class of people have a mean *I.Q.* higher than 100 might establish the following null and alternative hypotheses :

$$H_0 \;:\mu = 100 \text{ (null hypothesis)}$$

H_a : $\mu \neq 100$ (alternative hypothesis)*

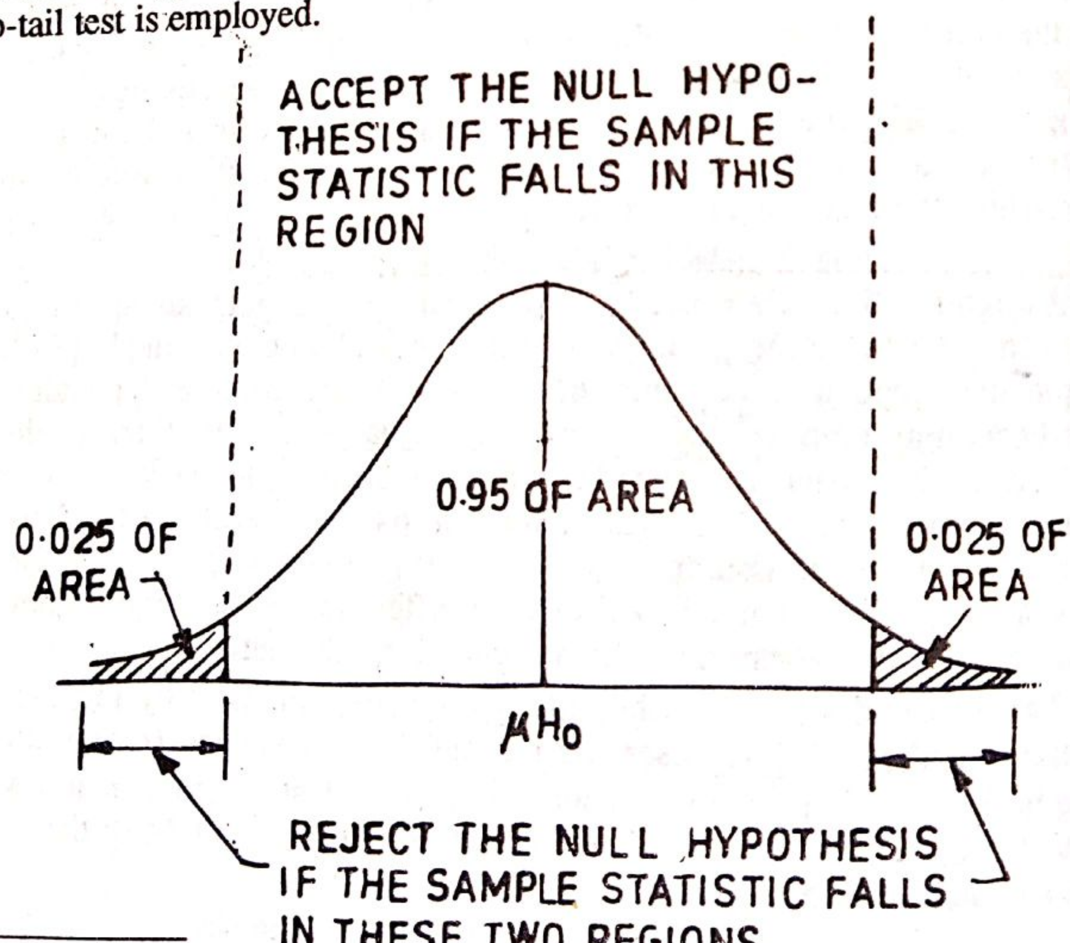
Or, if he is interested in testing the differences between the mean *I.Q.* of two groups, this psychologist may like to establish the null hypothesis that the two groups have equal means ($\mu_1 - \mu_2 = 0$) and the alternative hypothesis that their means are not equal ($\mu_1 - \mu_2 \neq 0$)

H_0 : $\mu_1 - \mu_2 = 0$ (null hypothesis)

H_a : $\mu_1 - \mu_2 \neq 0$ (alternative hypothesis)

2. *Set up a suitable significance level.* Having set up the hypothesis, the next step is to test the validity of H_0 against that of H_a at ascertain level of significance. The confidence with which an experimenter rejects—or retains—a null hypothesis depends upon the significance level adopted. The significance level is customarily expressed as a percentage, such as 5 per cent, is the probability of rejecting the null hypothesis if it is true. When the hypothesis in question is accepted at the 5 per cent level, the statistician is running the risk that, in the long run, he will be making the wrong decision about 5 per cent of the time. By rejecting the hypothesis at the same level he runs the risk of rejecting a true hypothesis in 5 out of every 100 occasions. By testing at the 1 per cent level he seeks to reduce the chance of making a false judgment but some element of risk remains (1 out of 100 occasions) that he will make the wrong decision, *i.e.*, he may accept where he ought to have rejected or *vice versa*.

The following diagram illustrates the region in which we would accept or reject the null hypothesis when it is being tested at 5% level of significance and a two-tail test is employed.



---

* Alternative hypothesis is any admissible hypothesis ... under test

The above diagram illustrates how to interpret a 5% level of significance. It may be noted that 2·5 per cent of the area under the curve is located in each tail.

3. *Setting a test criterion.* The third step in hypotheses testing procedure is to construct a test criterion. This involves selecting an appropriate probability distribution for the particular test, that is, a probability distribution which can properly be applied. Some probability distributions that are commonly used in testing procedures are t, F and χ^2. Test criteria must employ an appropriate probability distribution; for example, if only small sample information is available, the use of the normal distribution would be inappropriate.

4. *Doing computations.* Having taken the first three steps, we have completely designed a statistical test. We now proceed to the fourth step—performance of various computations—from a random sample of size n, necessary for the test. These calculations include the testing statistic and the standard error of the testing statistic.

5. *Making decisions.* Finally, as a fifth step, we may draw statistical conclusions and take decisions. A statistical conclusion or statistical decision is a decision either to reject or to accept the null hypothesis. The decision will depend on whether the computed value of the test criterion falls in the region of rejection or the region of acceptance If the hypothesis is being tested at 5% level and the observed set of results has probabilities less than 5 per cent, we consider the difference between the sample statistics and the hypothetical parameter significant. In other words, we think that the sample result is so rare that it cannot be explained by chance variation alone. We then decide to reject H_0 and state : "the null hypothesis is false", or "the sample observations are not consistent with the null hypothesis" (the rejection of H_0 automatically leads to acceptance of H_a).

On the other hand, if at 5% level of significance the observed set of results has probability more than 5 per cent we give reason that the difference between the sample result and the hypothetical parameter can be explained by chance variations and, therefore, is not significant statistically. Consequently, we decide not to reject H_0 and state : "The sample observations are not inconsistent with the null hypothesis." If the probability is about 5 per cent, the wisest course may be to resolve judgment and draw another sample, if possible.

The reader might have noted above that the rejection statement is much stronger than the acceptance statement. In other words, if the null hypothesis is not rejected, the statistician does not then categorically conclude that the hypothesis is true. The difference in attitudes arises essentially from the fact that, in logic, it is always easier to prove something false than to prove it true.

It should be clearly noted that the practical "managerial decision" is outside the responsibility of the statistician. He does not make the decision; he purely provides information on the basis of which the businessman or administrator can be assisted in making his decisions.

Two Types of Errors in Testing of Hypothesis

When a statistical hypothesis is tested there are four possibilities :

1. The hypothesis is true but our test rejects it. (Type I error)
2. The hypothesis is false but our test accepts it. (Type II error)

3. The hypothesis is true and our test accepts it. (Correct decision)
4. The hypothesis is false and our test rejects it. (Correct decision)

Obviously, the first two possibilities lead to errors.

In a statistical hypothesis testing experiment, a Type I error is committed by rejecting the null hypothesis when it is true. The probability of committing a Type I error is denoted by a α (pronounced as alpha), where

$$\alpha = \text{Prob. (Type I error)}$$
$$= \text{Prob. (Rejecting } H_0/H_a \text{ is true)}$$

On the other hand, a Type II error is committed by not rejecting (*i.e.*, accepting) the null hypothesis when it is false. The probability of committing a Type II error is denoted by β (pronounced as beta), where

$$\beta = \text{Prob. (Type II error)}$$
$$= \text{Prob. (Not rejecting or accepting } H_0/H_a \text{ is false)}$$

The distinction between these two types of errors can be made by an example. Assume that the difference between two population means is actually zero. If our test of significance when applied to the sample means leads us to believe that the difference in population means is significant, we make a Type I error. On the other hand, suppose there is true difference between the two population means. Now if our test of significance leads to the judgment "not significant", we commit a Type II error. We thus find ourselves in the situation which is described by the following table :

	Accept H_0	Reject H_a
H_0 is true	Correct decision	Type I error
H_0 is false	Type II error	Correct decision

While testing hypothesis the aim is to reduce both the types of error, *i.e.*, Type I and Type II. But due to fixed sample size, it is not possible to control both the errors simultaneously. There is a *trade-off* between these types of errors : the probability of making one type of error can only be reduced if we are willing to increase the probability of making the other type of error. In order to get a low β, we will have to put up with a high α. To deal with this trade-off in business situations, managers decide the appropriate level of significance by examining the costs or penalties attached to both types of errors.

It is more dangerous to accept a false hypothesis (Type II error) than to reject a correct one (Type I error). Hence we keep the probability of committing Type I error at a certain level, called the level of significance. The level of significance (also known as the size of the rejection region or size of the critical region or simply size of the test) is traditionally denoted by the Greek letter α. In most