16

Statistical Inference: Hypothesis Testing

16.1 INTRODUCTION

Hypothesis testing is a very important phase of statistical inference. It is a procedure which enables us to decide on the basis of information obtained from sample data whether to accept or reject a statement or an assumption about the value of a population parameter. Such a statement or assumption which may or may not be true, is called a *statistical hypothesis*. We *accept* the hypothesis as being true, when it is supported by the sample data. We *reject* the hypothesis when the sample data fail to support it.

It is important to understand what we mean by the terms *reject* and *accept* in hypothesis testing. The *rejection* of a hypothesis is to declare it false. The *acceptance* of a hypothesis is to conclude that there is not sufficient evidence to reject it. Acceptance does not necessarily mean that the hypothesis is true.

The basic concepts associated with hypothesis testing are discussed below:

16.1.1. Null and Alternative Hypothesis. A *null hypothesis*, generally denoted by the symbol H_0, is any hypothesis which is to be tested for possible rejection under the assumption that it is true. Today the term is used for any hypothesis that is being tested. The word *null* in the term *null hypothesis* implies that usually H_0 is the hypothesis of no effect. A null hypothesis should always be precise such as "the given coin is unbiased" or "a drug is ineffective in curing a particular disease" or "the difference between the two teaching methods is null or zero." The hypothesis is usually assigned a numerical value. For example, suppose we think that the average height of students in all colleges is 62″. This statement is taken as a hypothesis and is written symbolically as $H_0 : \mu = 62''$. In other words, we hypothesize that $\mu = 62''$.

An *alternative hypothesis* is any other hypothesis which we accept when the null hypothesis H_0 is rejected. It is customarily denoted by H_1 or H_A. A null hypothesis H_0 is thus tested against an alternative hypothesis H_1. For example, if our null hypothesis is $H_0 : \mu = 62''$, then our alternative hypothesis may be $H_1 : \mu \neq 62''$ or $H_1 : \mu > 62''$ or $H_1 : \mu < 62''$.

16.1.2. Simple and Composite Hypotheses. A *simple hypothesis* is one in which *all* parameters of the distribution are specified. For example, if the heights of college students are normally distributed with $\sigma^2 = 4$, the hypothesis that its mean μ is, say, 62", that is $H : \mu = 62$, we have stated a simple hypothesis, as the mean and variance together specify a normal distribution completely. A simple hypothesis, in general, states that $\theta = \theta_0$ where θ_0 is the specified value of a parameter θ, (θ may represent $\mu, p, \mu_1 - \mu_2$. etc.).

A hypothesis which is not simple (i.e. in which not all of the parameters are specified) is called a *composite hypothesis*. For instance, if we hypothesize that $H : \mu > 62$ (and $\sigma^2 = 4$) or $H : \mu = 62$ and $\sigma^2 < 4$, the hypothesis becomes a composite hypothesis because we cannot know the exact distribution of the population in either case. Obviously, the parameters $\mu > 62''$ and $\sigma^2 < 4$ have more than one value and no specified values are being assigned. The general form of a composite hypothesis is $\theta \leq \theta_0$ or $\theta \geq \theta_0$, that is the parameter θ does not exceed or does not fall short of a specified value θ_0. The concept of simple and composite hypotheses applies to both null hypothesis and alternative hypothesis.

Hypotheses may also be classified as *exact* and *inexact*. A hypothesis is said to be an *exact hypothesis* if it selects a unique value for the parameter such as $H : \mu = 62$ or $p = 0.5$. A hypothesis is called an *inexact hypothesis* when it indicates more than one possible values for the parameter such as $H : \mu \neq 62$ or $H : p > 0.5$. A simple hypothesis must be an exact one while an exact hypothesis is not necessarily a simple hypothesis. An inexact hypothesis is a composite hypothesis.

16.1.3. Test-statistic. A sample statistic which provides a basis for testing a null hypothesis, is called a *test-statistic*. Every test-statistic has a probability (sampling) distribution which gives the probability of obtaining a specified value of the test-statistic when the null hypothesis is true. It is important to remember that a test-statistic does not *prove* the hypothesis to be correct but it furnishes an *evidence* against the hypothesis. The sampling distributions of the most commonly used test-statistics are normal, t, chi-square or F.

16.1.4. Acceptance and Rejection Regions. All possible values which a test-statistic may assume can be divided into two mutually exclusive groups: one group consisting of values which appear to be consistent with the null hypothesis, and the other having values which are unlikely to occur if H_0 is true. The first group is called the *acceptance region* and the second set of values is known as the *rejection region* for a test. The rejection region is also called the *critical region*. The value(s) that separates the critical region from the acceptance region, is called the *critical value*(s). The critical value which can be in the same units as the parameter or in the standardized units, is to be decided by the experimenter keeping in view the degree of confidence he (she) is willing to have in the null hypothesis.

16.1.5. Type I and Type II Errors. When we perform a hypothesis test, we derive the evidence from the sample in the form of a test-statistic. There is a possibility that the sample evidence may lead us to make a wrong decision. We may reject a null hypothesis H_0, when it is, in fact, true or we may accept a null hypothesis H_0, when it is actually false. The former type is called an *error of the first kind* or a *Type I-error*, while the latter, an *error of the second kind* or a *Type II-error*. The decision and the corresponding two types of error may be displayed in a tabular form as below:

True Situation	DECISION	
	Accept H_0	Reject H_0 (or accept H_1)
H_0 is true	Correct decision (No error)	Wrong decision (Type-I error)
H_0 is false	Wrong decision (Type-II error)	Correct decision (No error)

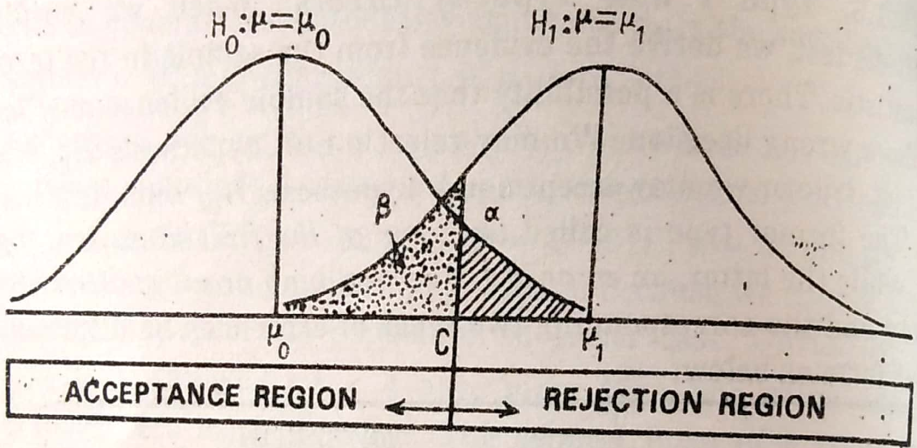
A legal analogy will help in understanding the difference between Type I and Type II errors. In a court trial, the supposition of Law is that the accused (the defendant) is innocent. This supposition of innocence may be regarded as a kind of null hypothesis H_0 that is to be rejected or accepted. After having heard the evidence presented during the trial, the judge arrives at a decision. Suppose the accused is, in fact, innocent (i.e. H_0 is true), but the finding of the judge is guilty. The judge has rejected a true null hypothesis and in so doing, has made a Type I error. If, on the other hand, the accused is, in fact, guilty (*i.e.* H_0 is false) and the finding of the judge is innocent, the judge has accepted a false null hypothesis and by accepting a false hypothesis, he has committed a Type II error.

The probability of making a Type I error is conventionally denoted by α (*alpha*) and that of committing a Type II error is indicated by β (*beta*). Thus α is the probability of rejecting H_0 when H_0 is true and β is the probability of accepting H_0 when H_0 is false (*i.e.* H_1 is true). In symbols, we may write

$$\alpha = P\,(\text{Type I error}) = P\,(\text{reject } H_0 \,/\, H_0 \text{ is true}),$$

$$\beta = P\,(\text{Type II error}) = P\,(\text{accept } H_0 \,/\, H_0 \text{ is false}).$$

Let us consider two distributions: one under the null hypothesis $H_0 : \mu = \mu_0$ (*i.e.* distribution assuming H_0 is true) and the other under alternative hypothesis $H_1 : \mu = \mu_1$ (*i.e.* distribution assuming H_1 is true).



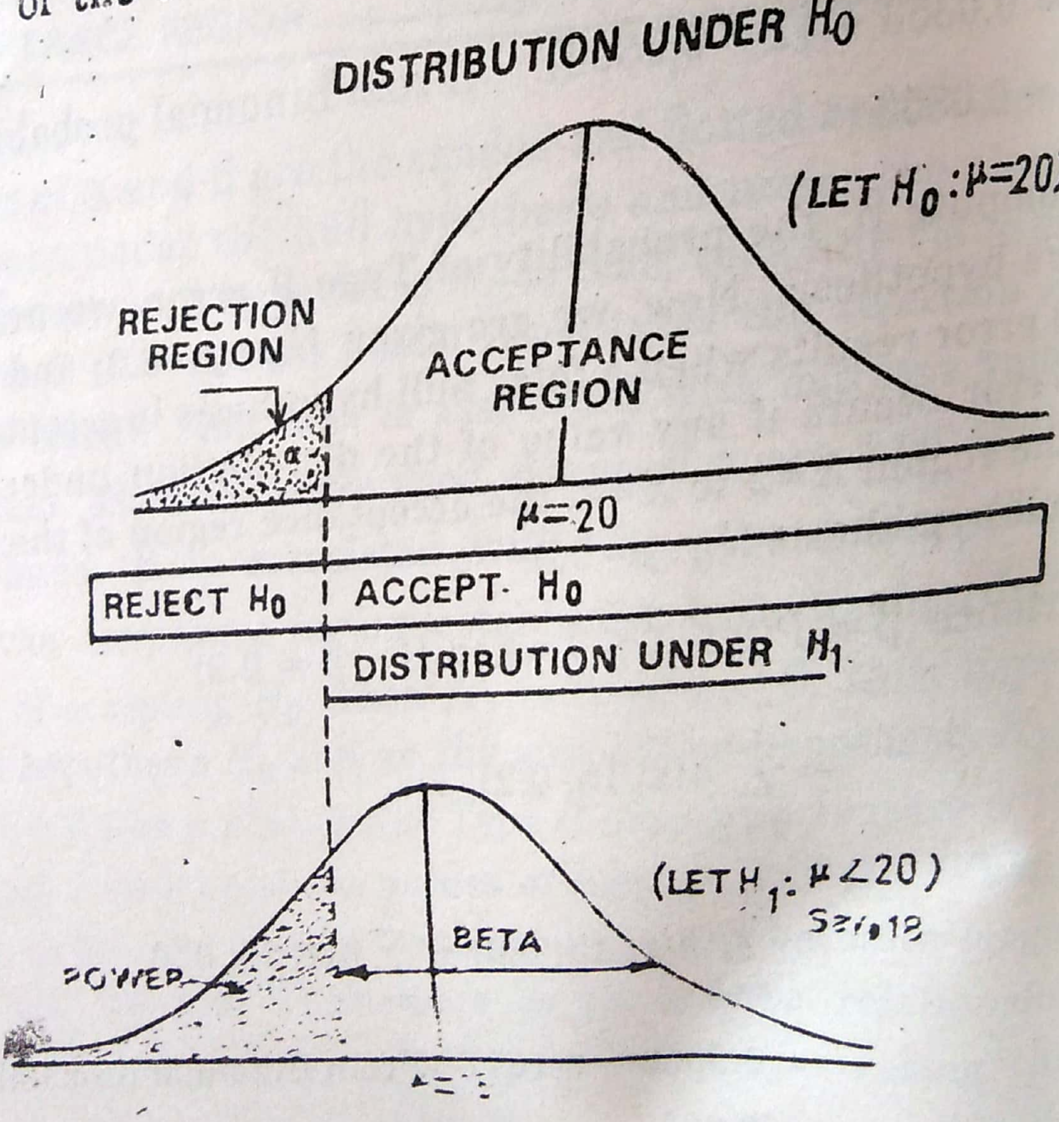
The probabilities of α and β are the shaded and dotted areas respectively of the distributions under the null hypothesis and under the alternative hypothesis. When our null hypothesis H_0 is true, then any value greater than or equal to C (the critical point) constitutes the rejection region equal to α (one–sided). That is α is associated with extreme values of the μ_0-distribution. The commonly used values of α are 0.05 and 0.01. On the other hand, β is associated with the area under the μ_1-distribution in the *acceptance region* established from μ_0-distribution. The probability of accepting H_0 when H_1 is true, *i.e.* β, thus depends both on the null hypothesis H_0 and on the alternative hypothesis H_1. In order to determine β (the probability of Type II error) we require α (the probability of Type I error) and the values of both μ_0 and μ_1. When α becomes smaller, β tends to become larger and when α becomes larger, β tends to become smaller. Thus there is an inverse relationship between α and β. We can decrease both α and β by increasing the sample size.

$= 0.7869 - 0.00$...

16.1.6. The Power of a Test with respect to a specified alterna[tive] hypothesis, is the probability of rejecting a null hypothesis when [it is] actually false. The *power* is the complement of β, the probability [of] committing a Type II error. It is therefore numerically equivalent to [one] minus β. Symbolically,

$$\text{Power} = P\,(\text{reject } H_0 / H_0 \text{ is false})$$
$$= 1 - \beta$$

To represent α, β and power of a test graphically, we show th[e] distributions of the test-statistic under both hypotheses H_0 and H_1 a[s] below:



15 Statistical Inference: Estimation

15.1 INTRODUCTION

The process of drawing inferences about a population on the basis of information contained in a sample taken from the population is called *Statistical Inference*. Statistical inference is traditionally divided into two major areas: *estimation of parameters* and *testing of hypothesis*.

Estimation is a procedure by which we obtain an estimate of the true but unknown value of a population parameter by using the sample observations $X_1, X_2, \ldots, X_n$ from the population. For example, we may estimate the mean and the variance of a population by computing the mean and the variance of a sample drawn form the population.

Testing of hypothesis is a procedure which enables us to decide on the basis of information obtained by sampling whether to accept or reject any specified statement or hypothesis regarding the value of the parameter in a statistical problem.

We shall discuss estimation in this chapter and we shall deal with testing of hypothesis in the next chapter.

15.2 ESTIMATES AND ESTIMATORS

An *estimate* is a numerical value of the unknown parameter obtained by applying a rule or a formula, called an *estimator*, to a sample $X_1, X_2, \ldots, X_n$ of size n, taken from the population. In other words, an *estimator* stands for the rule or method that is used to estimate a parameter whereas an *estimate* stands for the numerical value obtained by substituting the sample observations in the rule or the formula. For instance, if $X_1, X_2, \ldots, X_n$ is a random sample of size n from a population