CH # 16

Hypothesis testing

Hypothesis testing is a very important phase of statistical inferences. It is a procedure which enable 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 statement or assumption which may or may not be true is called statistical hypothesis.

Null Hyppothesise

A null hypothesis is generally denoted by 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.

or

That purpose That no statistical significane exists in a set of given observation.
it may in form of equality ($=$, $\leq$ or $\geq$)

Example:-

we think average height of student in all colleges is 62" so written as

"$H_0 : \mu = 62$"

Alternative Hypothesise

An alternative is any other hypothesise which we accept when the null hypothesis H_0 is rejected.
it is a statment in which there is some statistical relationship between two variables.

Example:

If null hypothesise is "$H_0 : \mu = 62$" than our alternative hypothesise "$H_1 : \mu \neq 62$ or $H_1 : \mu > 62$" $H_1 : \mu < 62$.

It is denoted by H_1.

Simple Hypothesis

Simple Hypothesis is one which all parameters of the distribution are speified.

Composite Hypothesis:-

A hpothesis which is not simple (in which not all of The parameters are speified is called composite Hypothesis.

Exact ~~Hp~~ Hypothesis:-

A hypothesis is said to be an exact hypothesis if it select a unique value for the parameters if it select a unique value for parameters

$H: \mu = 62$ or $p = 0.5$

Inexact hypothesis:

A hypothesis is called an inexact hypothesise when indicate more than one possible values for paramters such as $H: \mu \neq 62$ or $H: p > 0.5$

Test Statisticals:-

A sample statistics which provides a basis for testing a null hypothesise. is called test-statistical.

Type I error:-

A type I error (false-positive) occure if an investigators rejects a null hypothesis that is actually true in a population.

Example:

The test result says you

have coronavirus, but you actually don't.

Type II error.

A type II error occur if investigator fails to reject an null hypothesis that is actually false in the population.

Examples.

The test result says you don't have coronavirus but you actually do.

Example 16.1. The proportion of adults living in a small town who are matriculates is estimated to be $p=0.3$. To test this hypothesis a random sample of 15 adults is selected. If the number of matriculates in our sample is anywhere from 2 to 7, we shall accept the null hypothesis that $p=0.3$; otherwise we shall conclude that $p \neq 0.3$. Evaluate α assuming $p=0.3$. Evaluate β for the alternatives $p=0.2$ and $p=0.4$.
(P.U., B.A/B.Sc. 1986)

The null and alternative hypotheses are given as

$$H_0 : p = 0.3 \text{ and } H_1 : p \neq 0.3.$$

Let X denote the number of adults who are matriculates. Then the test-statistic has the binomial distribution with $p=0.3$ and $n=15$. The acceptance region, as given, consists of all values from $X=2$ to $X=7$. Then the critical region is composed of two parts: all values less than 2 and all values greater than 7. Thus the probability of making Type I error, *i.e.* α consists of $P(X<2)$ and $P(X>7)$.

Hence $\alpha = P(X<2 \text{ when } p=0.3) + P(X>7 \text{ when } p=0.3)$

$$= \sum_{x=0}^{1} b(x\,;\,15,\,0.3) + \sum_{x=8}^{15} b(x;\,15,\,0.3)$$

$$= \sum_{x=0}^{1} b(x\,;\,15,\,0.3) + [1 - \sum_{x=0}^{7} b(x;\,15,\,0.3)]$$

$= 0.0353 + [1 - 0.9500]$ (From Binomial probability tables)

$= 0.0853$

To compute β, the probability of Type II error, we need a specific alternative hypothesis. Now, we are given $H_0 : p=0.3$; and $H_1 : p=0.2$. A Type II error results when a *false* null hypothesis is accepted. That is a Type II error occurs if any value of the distribution under $H_1 : p=0.2$ falls in the region $X=2$ to $X=7$, the acceptance region of the distribution under null hypothesis $H_0 : p=0.3$.

Hence $\beta = P(2 \leq X \leq 7 \text{ when } H_1 : p = 0.2)$

$$= \sum_{x=2}^{7} b(x;\,15,\,0.2)$$

$$= \sum_{x=0}^{7} b(x;\,15,\,0.2) - \sum_{x=0}^{1} b(x;\,15,\,0.2)$$

$= 0.9958 - 0.1671$ (From Binomial probability tables)

$= 0.8287$

Scanned with CamScanner

Similarly, when $H_1 : p = 0.4$, we have

$$\beta = P(2 \le X \le 7, \text{ when } p = 0.4)$$

$$= \sum_{x=2}^{7} b(x;\, 15,\, 0.4)$$

$$= \sum_{x=0}^{7} b(x;\, 15,\, 0.4) - \sum_{x=0}^{1} b(x;\, 15,\, 0.4)$$

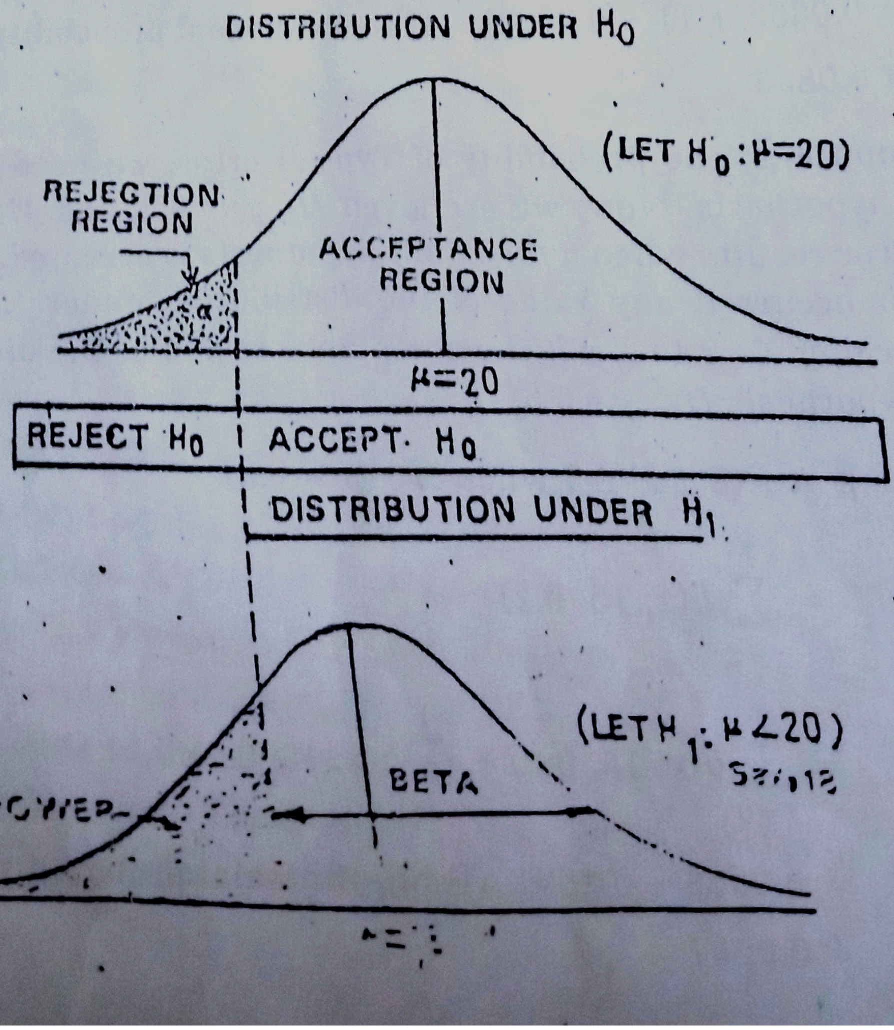
$$= 0.7869 - 0.0052 = 0.7817$$

16.1.6. The Power of a Test with respect to a specified alternative 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 the distributions of the test-statistic under both hypotheses H_0 and H_1 as below:



The shaded area in the lower diagram represents *power*. This probability corresponds to the rejection region of the distribution under H_0. The power generally increases with an increase in the sample size. A test for which β is small, is defined to be a *powerful test*.

A curve giving the probabilities of making Type II errors for various parameteric values under alternative hypotheses, is called an *Operating Characteristic Curve* or simply the *OC* curve. The *Power curve* which may be regarded as the complement of the *OC* curve, shows the probabilities of rejecting the null hypothesis H_0 for various values of the parameter θ.

Represent by α

16.1.7. The Significance Level of a test is the probability used as a standard for rejecting a null hypothesis H_0 when H_0 is assumed to be true. This probability is equal to some small pre-assigned value, conventionally denoted by α. The value α is also known as the *size of the critical region*. It is note-worthy that the significance level and the probability of Type I error are equivalent. The most frequently used values of α, the significance level, are 0.05 and 0.01, *i.e.* 5 percent and 1 percent but occasionally 0.10 or 0.001 is used. By $\alpha = 5\%$, we mean that there are about 5 chances in 100 of incorrectly rejecting a true null hypothesis. To put it in another way, we say that we are 95% confident in making the correct decision.

5% 1%
0.05 0.01

16.1.8. Test of Significance. A *test of significance* is a rule or procedure by which sample results are used to decide whether to accept or reject a null hypothesis. Such a procedure is usually based on a test-statistic and the sampling distribution of such a statistic under H_0. A value of the statistic is said to be *statistically significant* when the probability of its occurrence under H_0 is equal to or less than the significance level α, that is the value falls in the rejection region, H_0 in this case is rejected. If, on the other hand, the value falls in the acceptance region, it is said to be *statistically insignificant*. In this case, H_0 may be accepted. There are two desirable qualities for a test of significance. First, when the null hypothesis is actually true, it must have a low probability of rejecting H_0, and secondly, when H_0 is actually false, it must have a high probability of rejecting H_0. It is to be noted that the word *significant* is used in a special sense.

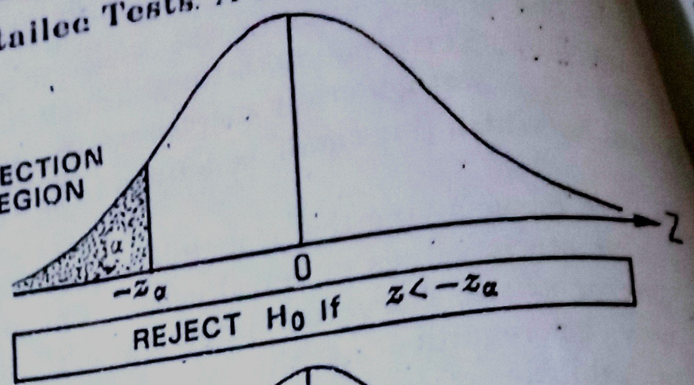
→ H_0 ⇒ reject (significance) $Z > Z_\alpha$

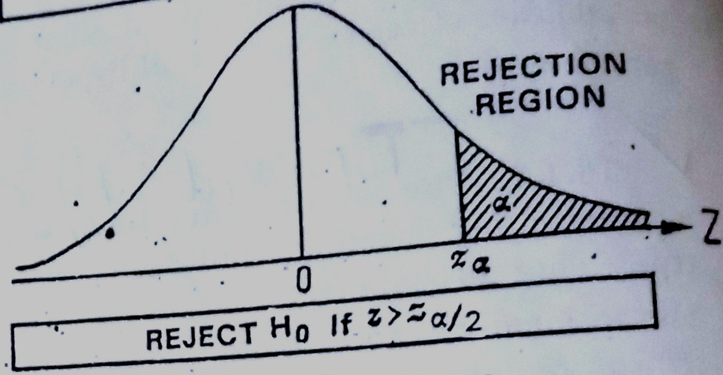
→ H_0 ⇒ accept (insignificance)

Level of confidance:
Represent by (C)

Scanned with CamScanner

16.1.9. One-tailed and Two-tailed Tests. A test for which the entire rejection region is located in only one of the two tails–either in the right tail or in the left tail–of the sampling distribution of the test-statistic, is called a *One-tailed test* or *One-sided test*. For example, if Z is a test-statistic, then the rejection region consists of all z-values which are greater than $+z_\alpha$ or less than $-z_\alpha$ where α is the size of critical region. A one-tailed test is used when the alternative hypothesis H_1 is formulated in the following form:

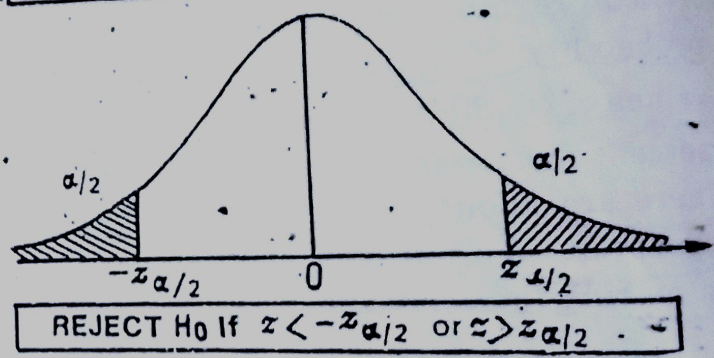




$H_1 : \theta > \theta_0$ or

$H_1 : \theta < \theta_0$ *i.e.*

H_1 is composite hypothesis.



If, on the other hand, the rejection region is divided equally between the two tails of the sampling distribution of the test-statistic, the test is referred to as a *Two-tailed test* or *Two-sided test*. In this case, the alternative hypothesis H_1 is set up as:

$$H_1 : \theta \neq \theta_0, \ (i.e.\ H_1 \text{ is two-sided composite hypothesis})$$

meaning thereby that values both larger and smaller than θ_0 are to be included. In case of test-statistic being normal distribution, the rejection (critical) regions are shown by shading the appropriate portions of area under the sampling distribution in the figures shown above. The location of critical region can be determined only after the alternative hypothesis H_1 has been stated. It is important to note that the one-tailed and the two-tailed tests differ only in location of the critical region, not in the size.

Example 16.2. We wish to test the hypothesis that the mean weight of a population of people is 140 lbs. Using $\sigma = 15$ lb, $\alpha = 0.05$ and a sample of 36 people, find

(a) the values of $\bar{x}$ which would lead to rejection of the hypothesis, and (b) β, the probability of Type II error, if $\mu = 150$ lb. Use a two sided test.

(P.U., M.Sc., 1970, 86)

We are given the following information:

$$H_0 : \mu = 140 \text{ lb}, \ \sigma = 15 \text{ lb}, \ \alpha = 0.05 \text{ and } n = 36.$$

(a) To find the values of $\bar{x}$ (the critical point) which would lead to rejection of the hypothesis $H_0 : \mu = 140$ lb, we use the test statistic (assuming normal population) given by

$$\boxed{Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}.}$$ sampe

Since the test is two-sided, so there would be two critical values. Corresponding to the significance level $\alpha = 0.05$, the critical values of Z from the table of normal curve are -1.96 and 1.96. Thus

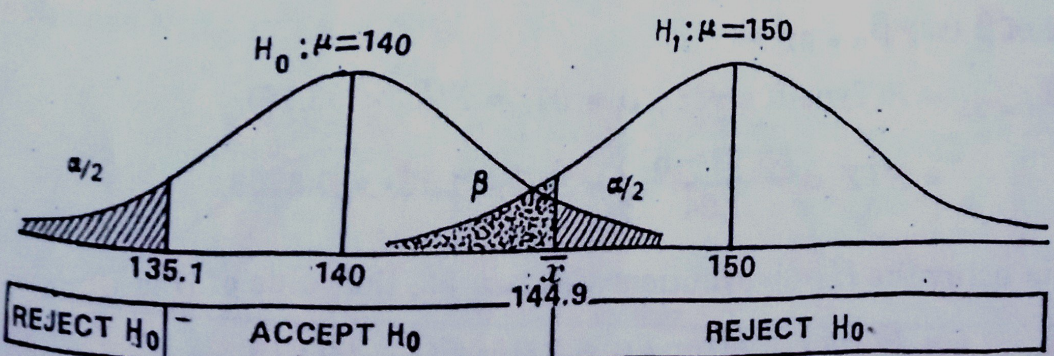
$$\pm 1.96 = \frac{\bar{x} - 140}{15 / \sqrt{36}}.$$

$$Z = \frac{\bar{x} - \mu}{\sigma}$$ → population

Simplifying, we get $\bar{x} = 135.1$ and 144.9 as the two critical values.

Hence the hypothesis $H_0 : \mu = 140$ lb will be rejected if $\bar{x} < 135.1$ lb or $\bar{x} > 144.9$ lb.

(b) A Type-II error can be committed only by accepting a false H_0. The hypothesis $H_0 : \mu = 140$ lb will be false if μ takes a value greater than 140 lb. Given $H_1 : \mu = 150$ lb so that $H_0 : \mu = 140$ lb becomes false. Therefore, the probability of accepting $H_0 : \mu = 140$ lb (false) when $H_1 : \mu = 150$, *i.e.* the probability of a Type-II error is indicated by the dotted area β in the figure shown below. To compute this area, we use the distribution under the alternative hypothesis $H_1 : \mu = 150$ lb.



Now, at $\bar{x} = 144.9$, we find $z = \dfrac{144.9 - 150}{15 / \sqrt{36}} = -2.04$, and

Scanned with CamScanner

at $\bar{x} = 135.1$, we find $z = \dfrac{135.1 - 150}{15/\sqrt{36}} = -5.96.$

Thus $\beta =$ Area between $z = -5.96$ and $z = -2.04$, *i.e.* area in the acceptance region of the distribution under H_0, when H_1 is $\mu = 150 = 0.0207$

This is the probability of accepting the null hypothesis $H_0 : \mu = 140$ lb, when, in fact, the alternative hypothesis $H_1 : \mu = 150$ lb is true.

Example 16.3. A random sample of size 4 is drawn from a normal population with known variance 15. A one-tailed test of the form $H_0 : \mu \leq 30$ against $H_1 : \mu > 30$ at the 5% level of significance is performed. Calculate the probabilties of Type-II error (β) for the values of $\mu = 31, 32, 34$ and 36 in the alternative hypothesis. Also calculate the powers of the test and hence sketch the power curve for this test.

Given $H_0 : \mu \leq 30$ and $H_1 : \mu > 30$; $\alpha = 0.05$, $n = 4$ and $\sigma^2 = 15$.

Under $H_0 : \mu \leq 30$, the population under consideration is $N(30, 15/4)$.

To find the values of β, the probability of Type-II error, we first need to calculate the critical values. For a one-tailed test (the upper tailed) with $\alpha = 0.05$, the critical value is given by

$$c = \mu_0 + z_\alpha \cdot \frac{\sigma}{\sqrt{n}}$$

i.e.
$$c = 30 + (1.645)\left(\sqrt{15/4}\right) = 30 + (1.645)\left(\sqrt{3.75}\right)$$
$$= 30 + (1.645)(1.94) = 33.19$$

Since 4 values of μ in the alternative hypothesis are specified, so we associate the variables $\bar{X}_1$, $\bar{X}_2$, $\bar{X}_3$ and $\bar{X}_4$ with each of the four alternative (H_1) distributions.

Using the alternative (H_1) distribution with $\mu = 31$, we calcul[ate] the value of β (say $\beta_{\mu = 31}$) as

$$\beta_{\mu = 31} = P(\text{Type-II error} / \mu = 31) = P(\bar{X}_1 < 33.19)$$
$$= P\left(Z < \frac{33.19 - 31}{1.94}\right) = P(Z < 1.13) = 0.8708$$

Again using the H_1-distribution with $\mu = 32$, the value [of β]

$$\beta_{\mu = 32} = P(\text{Type-II error} / \mu = 32) = P(\bar{X}$$

The sampling distribution of $\hat{\theta}$ in case of an inexact null hypothesis ($H_0 : \theta \leq \theta_0$ or $H_0 : \theta \geq \theta_0$) is not defined and hence we cannot set up the acceptance and rejection regions. In such a case, we would take the null hypothesis *as if* it is an exact one, *i.e.* $H_0 : \theta = \theta_0$.

16.1.12. General Procedure for Testing Hypotheses. The procedure for testing a hypothesis about a population parameter involves the following six steps:

(i) State your problem and formulate an appropriate null hypothesis H_0 with an alternative hypothesis H_1, which is to be accepted when H_0 is rejected.

(ii) Decide upon a significance level, α of the test, which is the probability of rejecting the null Hypothesis if it is true.

(iii) Choose an appropriate test-statistic, determine and sketch the sampling distribution of the test-statistic, assuming H_0 is true.

(iv) Determine the rejection or critical region in such a way that the probability of rejecting the null hypothesis H_0, if it is true, is equal to the significance level, α. The location of the critical region depends upon the form of H_1. The significance level will separate the acceptance region from the rejection region.

(v) Compute the value of the test-statistic from the sample data in order to decide whether to accept or reject the null hypothesis H_0.

(vi) Formulate the decision rule as below:

(a) Reject the null hypothesis H_0, if the computed value of the test-statistic falls in the rejection region and conclude that H_1 is true.

(b) Accept the null hypothesis H_0, otherwise.

When a hypothesis is rejected, we can give a measure of the strength of the rejection by giving the *P-value*, the smallest significance level at which the null hypothesis is being rejected.

16.2 TESTS BASED ON NORMAL DISTRIBUTION

Suppose we wish to test a hypothesis that a parameter θ of a normal distribution has some specified value θ_0. We draw a random sample of ... and calculate $\hat{\theta}$ as an estimate of θ. It has