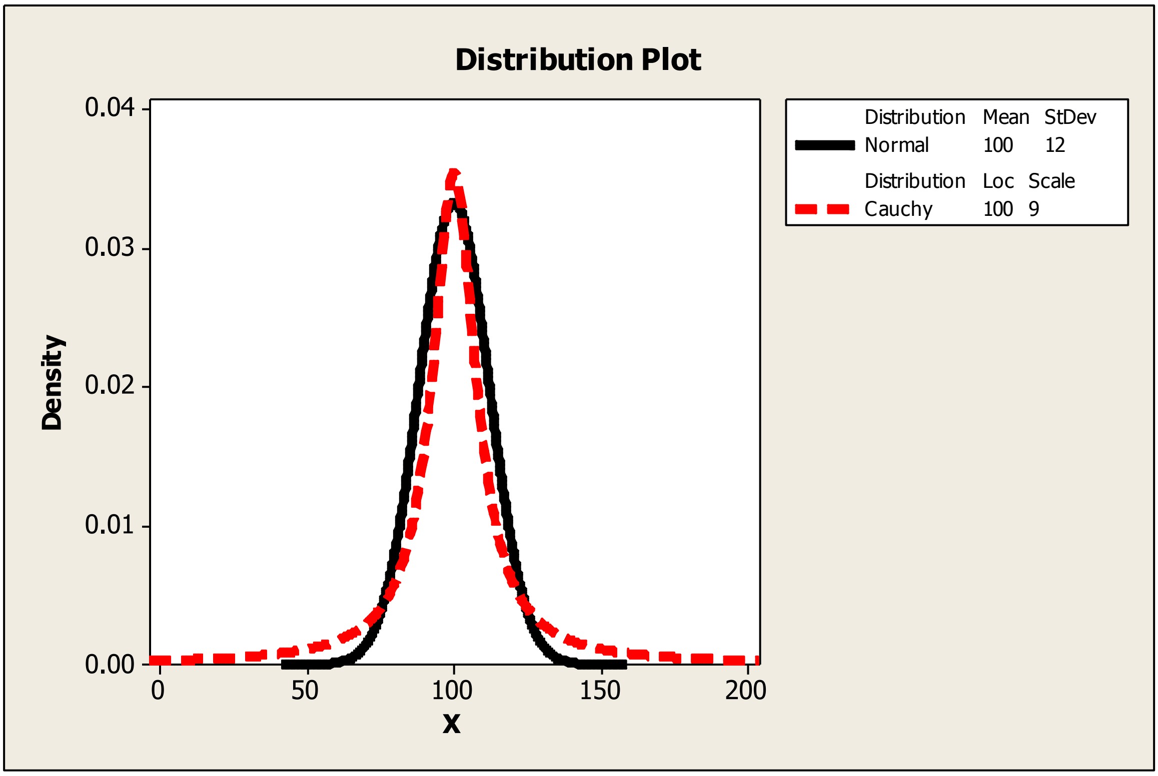
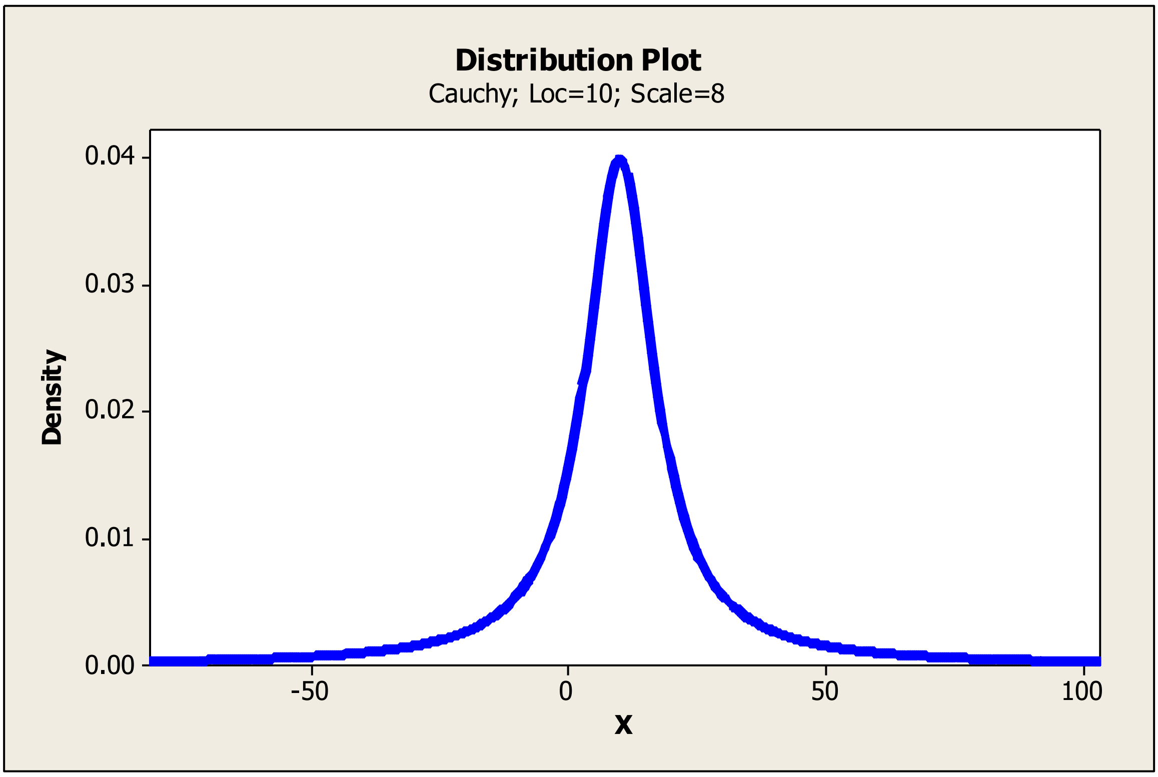
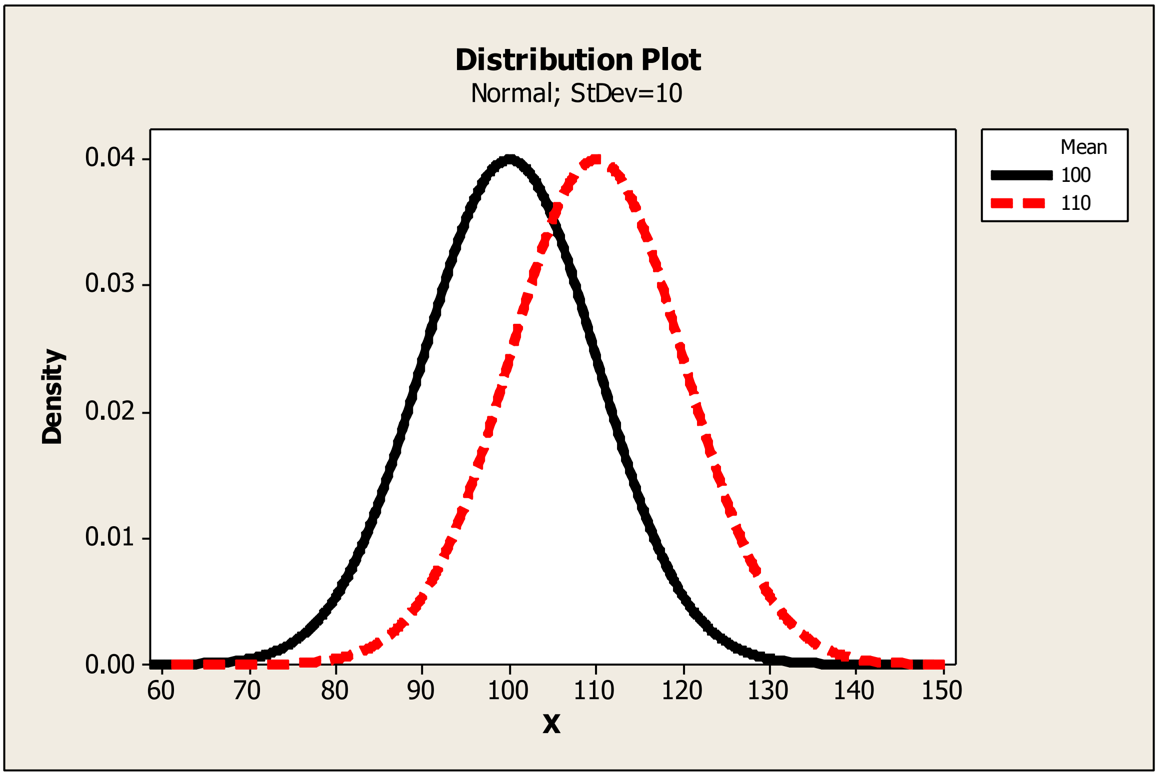
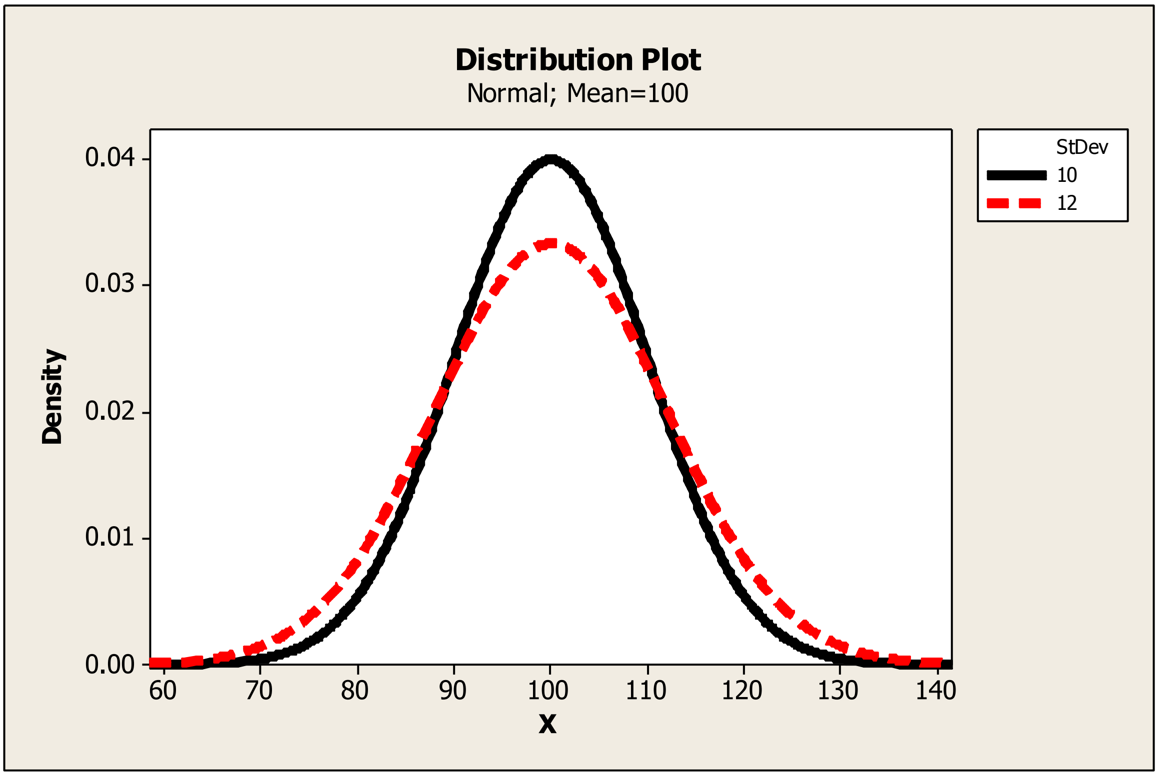
Hypothesis Testing
I

- **In the previous chapters we illustrated how to construct a confidence interval estimate of a parameter from sample data.**
- **However, many problems in engineering require that we decide whether to accept or reject a statement about some parameter.**
- **The statement is called a hypothesis, and the decision-making procedure about the hypothesis is called hypothesis testing.**
- **This is one of the most useful aspects of statistical inference, since many types of decision-making problems, tests, or experiments in engineering world can be formulated as hypothesis-testing problems.**
- **There is a very close connection between hypothesis testing and confidence intervals.**
- **Statistical hypothesis testing and confidence interval estimation of parameters are the fundamental methods used at the data analysis stage of a comparative experiment, in which the engineer is interested, for example, in comparing the mean of a population to a specified value.**
- **These simple comparative experiments are frequently encountered in practice and provide a good foundation for the more complex experimental design problems.**


Distribution Plot
Density
0.04
0.03
0.02
0.01
0.00
0
50
100
150
200
X
Distribution Mean StDev
Normal 100 12
Distribution Loc Scale
Cauchy 100 9


Distribution Plot
Cauchy; Loc=10; Scale=8
Density
0.04
0.03
0.02
0.01
0.00
-50
0
50
100
X


Distribution Plot
Normal; StDev=10
Density
0.04
0.03
0.02
0.01
0.00
60
70
80
90
100
110
120
130
140
150
X
Mean
100
110


Distribution Plot
Normal; Mean=100
Density
0.04
0.03
0.02
0.01
0.00
60
70
80
90
100
110
120
130
140
X
StDev
10
12

Statistical Hypothesis

A statistical hypothesis is a statement about the parameters of one or more populations.

$$H_0: \quad \mu = 50$$
$$H_1: \quad \mu \neq 50$$

Null Hypothesis

The statement $H_0: \quad \mu = 50, \; H_0: \quad \phi = a$ is called the *null hypothesis*.

Alternative Hypothesis

The statement $H_1: \quad \mu \neq 50, \; H_1: \quad \phi \neq a$ is called the *alternative hypothesis*.

Since the alternative hypothesis specifies values of μ that could be either greater or less than 10, it is called a two-sided alternative hypothesis. In some situations, we may wish to formulate a one-sided alternative hypothesis, as in

$$\begin{array}{l} H_0: \quad \mu = 50 \\ H_1: \quad \mu < 50 \end{array} \quad \text{or} \quad \begin{array}{l} H_0: \quad \mu = 50 \\ H_1: \quad \mu > 50 \end{array}$$

A procedure leading to a decision about a particular hypothesis is called a test of hypothesis.

Hypothesis

The hypotheses for the Chi-Square Goodness-of-Fit Test is:

H_o: The data follow a multinomial distribution with specific proportions

H_1: The data do not follow a multinomial distribution with specific proportions

Expected values

The expected value for each category is calculated as,

$E_i = N * p_i$

where:

- **p_i = the test proportion for the ith category = 1/k or the value you provide**
- **k = the number of distinct categories**
- **N = total observed values ($O_0 + O_1 + ... + O_k$)**
- **O_i = the observed value for the ith category**

Contribution to chi-square

Contribution of the ith category to the chi-square value is,

$$\frac{(O_i - E_i)^2}{E_i}$$

Test statistic

The chi-square goodness-of-fit test statistic is:

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

p-value and DF

The p-value is calculated as,

Prob (X > Test statistic)

where, X follows a chi-square distribution with (k – 1) degrees of freedom (DF).

Computation

Given:

Category (i)	Observed (O_i)	Test proportions p_i
A	5	0.1
B	15	0.2
C	10	0.3
D	10	0.4
	N = 40	

Calculated:

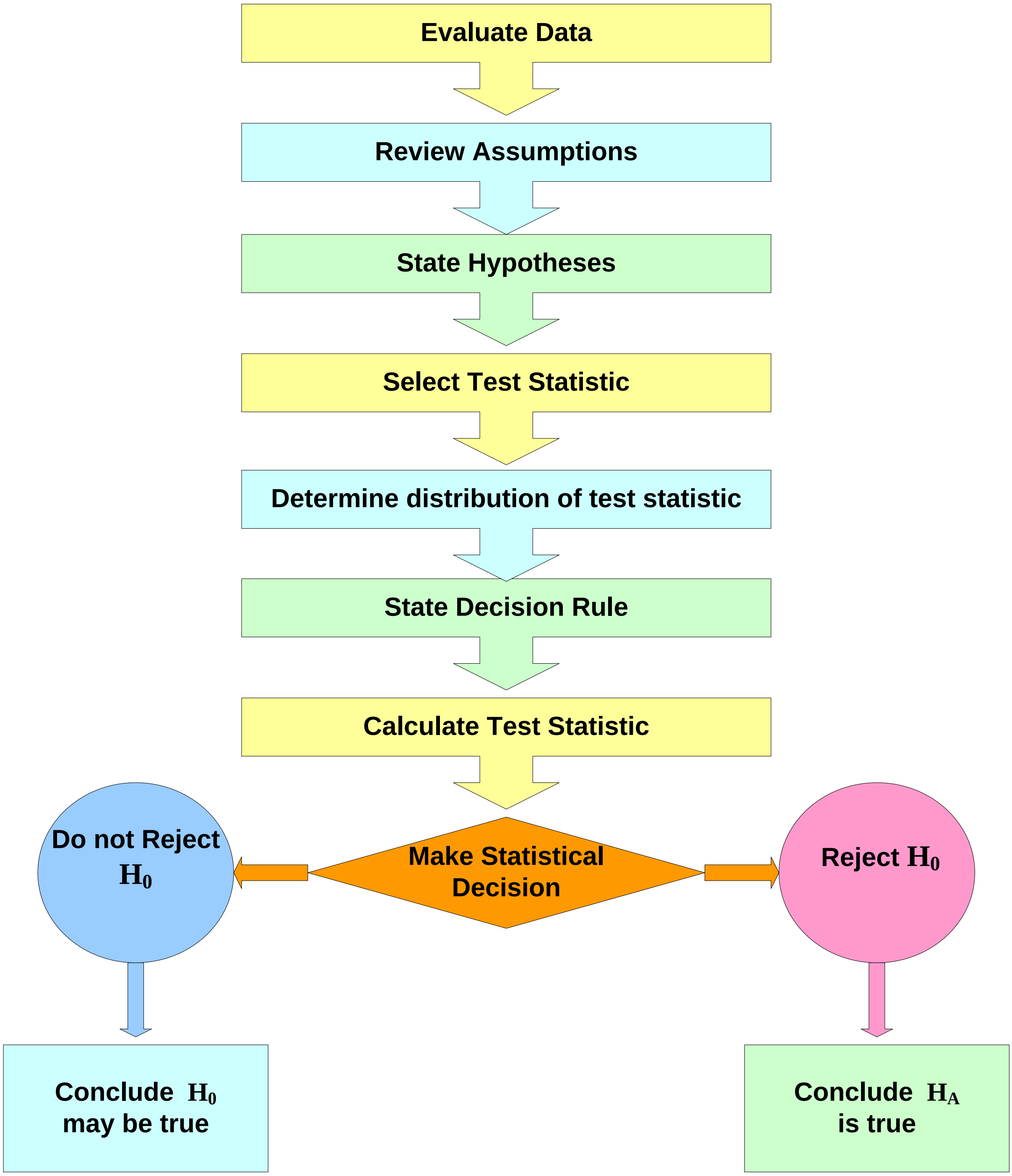
Category (i)	Expected value $E_i = (p_i * N)$	Contribution to chi-square $(O_i - E_i)^2 / E_i$
A	0.1 * 40= 4	$(5 - 4)^2 / 4 = 0.2500$
B	0.2 * 40 = 8	6.1250
C	0.3 * 40 = 12	0.3333
D	0.4 * 40 = 16	2.2500

$\chi^2 = (0.2500 + 6.1250 + 2.2500 + 0.3333) = 8.9583$

DF = k – 1 = 3

p-value = Pr (X > 8.9583) = 0.0299

Steps in the Hypothesis Testing Procedure



Test of Statistical Hypotheses

For convenience, hypothesis testing will be presented as a nine-step procedure.

- ***Data***

The nature of the data form the basis of testing procedures must be understood, since this determines the particular test to be employed. Whether the data consist of counts or measurements, must be determined.

- ***Assumptions***

As we learned in the chapter of estimation, different assumptions led to modifications of confidence intervals. The same is true in hypothesis testing: a general procedure is modified depending on the assumptions. In fact, the same assumptions that are of importance in estimation are also important in hypothesis testing.

- ***Hypotheses***

There are two statistical hypotheses involved in hypothesis testing and these should be explicitly stated. The first is the hypothesis to be tested, usually referred to as the null hypothesis and designed by the symbol H_0. In general, the null hypothesis is set up for the express purpose of being discredited. In the testing process the null hypothesis either is rejected or not rejected.

- **If the null hypothesis is not rejected, we will say that the data on which the test is based do not provide sufficient evidence to cause rejection.**
- **If the testing procedure leads to rejection, we will say that the data at hand are not compatible with the null hypothesis, but are supportive of some other hypothesis. This other hypothesis is known as the alternative hypothesis and may be designated by the symbol H_1 or H_A.**

- ***Test Statistic***

The test statistic is some statistic that may be computed from the data of the sample. As a rule, there are many possible values that the test statistic may assume, the particular value observed depending on the particular sample drawn. The test statistic serves as a decision maker, since the decision to reject or not to reject the null hypothesis depends on the magnitude of the test statistic.

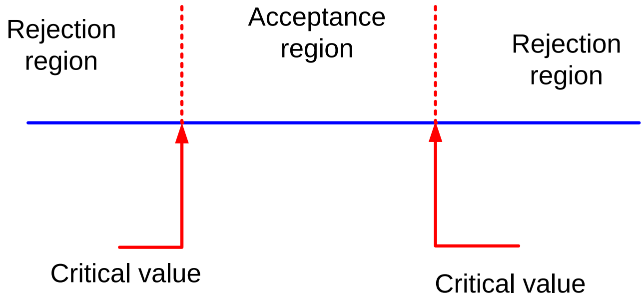
- ***Distribution of the Test Statistic***

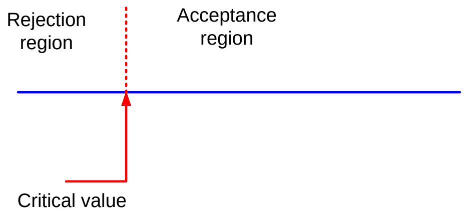
It has been pointed out that the key to statistical inference is the sampling distribution. It is necessary to specify the probability distribution of the test statistic.

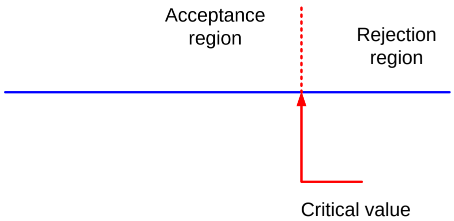
- ***Decision Rule***

All possible values that the test statistic can assume are points on the horizontal axis of the graph of the distribution of the test statistic and are divided into two groups; *rejection region* and *acceptance region*.

The decision rule tells us to reject the null hypothesis if the value of the test statistic that we compute from our sample is one of the values in the rejection region and to not reject (accept) the null hypothesis if the computed value of the test statistic is one of the values in the acceptance region.







In testing any statistical hypothesis, four different situations determine whether the decision is correct or in error. ***The probability of making a type I error is denoted by the Greek letter α. Sometimes the type I error probability is called the significance level, or the α-error.***

Decisions in Hypothesis Testing

Decision	H_0 is true	H_0 is false
Accept H_0	Correct Action	Type II Error
Reject H_0	Type I Error	Correct Action

Type I Error: Rejecting the null hypothesis H_0 when it is true is defined as a type I error.

Type II Error: Failing to reject the null hypothesis H_0 when it is false is defined as a type II error.

In evaluating a hypothesis-testing procedure, it is also important to examine the probability of type II error, denoted by β. That is

β=P(type II error) = P(fail to reject H_0 when H_0 is false).

Definition: The power of a statistical test, given as

$$1-\beta = P(\text{reject } H_0 \text{ when } H_a \text{ is true}).$$

The power of a statistical test is the probability of rejecting the null hypothesis H_0 when the alternative hypothesis is true.

- ***Calculation of the Test Statistic***

From the data contained in the sample we compute a value of the test statistic and compare it with the acceptance and rejection regions that have already been specified.

- ***Statistical Decision***

The statistical decision consists of rejecting or not rejecting the null hypothesis. It is rejected if the computed value of the test statistic falls in the rejection region, and it is not rejected of the computed value of the test statistic falls in the acceptance region.

- ***Conclusion***

If H_0 is rejected, we conclude that H_A is true. If H_0 is not rejected, we conclude that H_0 may be true.