CHAPTER 6

VERIFICATION OF THE FRAMEWORK USING HYPOTHESIS TESTING

6.1 INTRODUCTION TO HYPOTHESIS TESTING

Hypothesis testing is a process of making a choice between several competing hypotheses about probability distribution on the basis of the observed data distribution. Hypothesis Testing is a very prominently used method of verification that is used in statistics. In statistical hypothesis testing there is a statistical inference based on the data that has been gathered from a research or survey carried out. If the occurrence of the result is predicted as unlikely according to the pre-calculated threshold probability also referred to the significance level, then the result is called as statistically significant in statistics. Ronald Fisher was the person who initiated the concept of "test of significance". The tests of significance are used to determine that which outcomes of a research will direct to a denial for a pre-specified significance level of the null hypothesis. This provides contribution in deciding whether the results contain sufficient information or not in order to cast disbelief on predictable insight, to establish the null hypothesis, considering the fact that the usual perception has been applied. The critical region of the hypothesis test is defined to be the collection of all the outcomes that will cause the null hypothesis to be redundant in comparison to the alternative hypothesis.

Hypothesis testing is referred to as statistical or confirmatory data analysis as it has pre-defined hypotheses, in disparity to the exploratory method of data analysis that might not have pre-specified hypotheses.

One of the vital parts of the statistical inference is the setting up of the hypothesis and then testing the hypothesis. For formulating a test like this, some theory has to be set forward and that theory may be supposed to be accurate or it can be used as a source for the argument and then proved later. For example, claiming that a particular medicine for a particular ailment is better than the existing one.

6.2 STEPS OF HYPOTHESIS TESTING

The Hypothesis Testing is performed in following steps:

> *Step 1: Identify the hypothesis or claim that needs to be proved. For instance, if there is a need to determine that majority of users prefer less complex software in comparison to high complex software.*

> *Step 2: Decide upon the criterion on the basis of which the user will decide whether the hypothesis being claimed upon is true or false. In a way, it can be said that in this step, the defined threshold value for deciding the truth or falsity of the hypothesis.*

> *Step 3: The third step involves selecting a sample population and measuring the sample mean.*

> *Step 4: In the last step, there is a comparison of the sample mean obtained in Step 3 above with the expected threshold that has been defined in Step 2. If there is a small difference jammed between the two means: the sample mean and the population mean, then the hypothesis is true else it is false.*

For every problem under the consideration, the decision is based upon an issue that is of interest to us. Then there are two distinguishing claims that can be made about the issue that is termed as the hypothesis: one of them is the Null Hypothesis denoted by H0 and the other one is the alternative or the substitute

hypothesis denoted by H1. The above said hypothesis a not observed on an identical basis, exceptional consideration has been given to the Null Hypothesis.

There are following two general situations:

1. The experimentation has been done in order to do the confirmation in opposition to it is adequately strong. For example,

   H0: Suppose that there is no distinction in flavour of Diet Pepsi and Pepsi against H1: Distinction between the two exists.

2. If either of the two hypotheses stated above is simple enough, there is a provision of more preference in comparison to the other complicated one so that the latter one is not adopted until and unless there exists an adequate amount of confirmation in support of the alternate hypothesis. For instance, it is very simple to declare that no variation in the taste or flavor exists between Diet Pepsi and Pepsi instead of saying that there exists a variation.

The assumptions or hypotheses are the statements that are very prominently used regarding the population parameters such as variance, expected value etc. For instance, the Null Hypothesis H0 can be the accepted value of the weight of eighteen year old boys in a population is not different from that of eighteen year old girls. A hypothesis can also be a statement that concerns a distributional figure of an attribute of interest.

The result of a hypothesis test is "Do not refuse H0" or "Reject H0 in favour of H1".

In order to evaluate the behaviour of a population that is too large or inaccessible, the use of inference statistics to study the behaviour in a sample of population as it allows us to do a more accurate study. Samples are used for evaluation as they are linked to the attributes of the population. The standard of the sample mean will be approximately equal to the value of the population mean, if an arbitrary sample is selected from a population. The method in which there is a need to make a decision about samples to study about attributes of a particular population is known as Hypothesis Testing. Hypothesis Testing is a regular approach to verify the claims or facts regarding an assembly or population.

6.3 HYPOTHESIS TESTING

In order to confirm our representation Hypothesis Testing was performed of our framework on a total of 100 Test samples (n=100). Here, the deliberated level of satisfaction and establish that mean to be equivalent to 70% (M=70) (70+10) i.e., μ =10. After calculating one independent sample Z-test, preserving of the Null Hypothesis is done.
(M=70%) at a 0.05 significance level (α=0.05). The trace the sample mean as 90% (M=90) is present.

STEP I: STATE THE HYPOTHESIS

The process begins with defining the population mean's value in a Null Hypothesis, which is considered as true. The Null Hypothesis H0 is a statement relative to a population parameter, like the population mean, that is hypothetical to be true. It is the preliminary assumption.

STEP II: LAY DOWN THE CRITERIA OF DECISION

In order to set criteria for a decision, there is a declaration of the level of impact for the test. During hypothesis testing, the collection of data is done to exemplify

that the null hypothesis is false, depending upon the probability of choosing a sample mean from the population (the criterion is the likelihood). In behavioural research analysis, the significance level is usually fixed at 5% in. If the probability of achieving the sample mean is not as much as 5%.

The level of significance or the significance level refers to a standard upon which a decision is to be made with regards to the value settled in a Null Hypothesis. The criterion depends upon on the possibility of getting a statistic calculated in a sample in case the settled value in the null hypothesis is true.

The level of significance is 0.05, which makes $\alpha=0.05$. Now, in order to uncover the chance of a sample mean from a given population, the method which is taken is of standard normal distribution by placing standard normal distribution of Z-scores that are frequently cut offs or defined as critical values for the sample mean values lower than 5% probability of occurrence. After this, split the alpha value in half in a non-conditional two tailed, so that an identical proportion of area is placed in lower and upper tails.

Dividing α in half: $\alpha/2=0.05/2=0.0250$ in each tail.

The region ahead of the critical value of the hypothesis is the rejection region.

STEP III: THE TEST STATISTIC CALCULATION
A test statistic aids us in determining the number of standard deviations or the distance between the sample mean and the population mean. The larger is the test statistic's value; greater will be the distance, or the figure of the standard deviation. The determination of a sample mean from the population mean to test

statistics value is considered to construct a decision in Step 4. In this stage the judgement of the generated value to the critical values occurs.

Z statistics: Z obtained = M - μ /σ M

Where Z and σ M = σ/ √n statistics is inference statistics that is applied to resolve on the amount of standard deviations in the standard normal distribution.
The test statistic's value is the resulting value. In order to formulate the decision, the value of resultant statistics is compared with the critical values.
σ M = σ/√n=10/√100 = 1
Z obtained = 90-80/10 = 1

STEP IV: COMPOSE A DECISION
The test statistic's computed value is used to compose the decision regarding null hypothesis. The result depends upon the possibility of getting a sample means, taking into consideration that the value known to Null Hypothesis will be true providing the value obtained in the sample mean is lower than 5% and then there is a decision of discarding the null hypothesis. However, if the probability of getting a sample mean is more than 5% while the null hypothesis is assumed to be true, then there comes the decision to maintain the Null Hypothesis. Apart from these, the following two decisions could be taken by the analyst:

• *Denial of the Null Hypothesis. In this case the sample mean is related with a low likelihood of occurrence when the null hypothesis is correct.*
• *Retention of the Null Hypothesis. In this case the sample mean is related with a high likelihood of occurrence while the null hypothesis is correct.*

The likelihood of obtaining a sample mean, taking into account that the value defined in the null hypothesis is true, is settled by the probability value p. The value of p ranges from 0 to 1 and can never be negative. In the next step, the settling of the probability of generating a sample mean is done and at that point there is a need to make a decision to discard the value defined in the null hypothesis, which is settled down at 5% in behavioural research.

In order to derive a conclusion, there is a need to place the value of p side by side to the criterion that has been set in Step 2. The probability of obtaining a sample result is p, in view of the fact that the value defined in the Null Hypothesis is true. The p-value obtained for generating a sample result is compared to the significance level.

A decision made related to a value defined in null hypothesis is explained using statistical significance. When the null hypothesis is discarded, the user is arrived at the significance and when the null hypothesis is retained, there will not be a success in attaining the significance.

Null hypothesis is discarded when the p value is lower than 5% ($p < .05$). Also, when the value of $p = .05$, the conclusion is still to discard the null hypothesis. However, in the case when the value of p is larger than 5% ($p > .05$), then there is a need to make a decision to retain the null hypothesis. Significance is mainly the decision of discarding or retaining.

There will not be a success to get to significance and the decision is to keep hold of this stage to compose a decision by comparing it with the critical value. The Null Hypothesis is refused if the generated value exceeds a critical value.

Table 6.1: Four Outcomes to make a Decision

	DECISION	
Truth in the Population	Retain the Null	Reject the Null
Truth	Correct	Error-α
Falsity	Error-β	Correct

In Step 4, there is point arrived where the decision whether to keep hold of or discard the null hypothesis takes place. As the evaluation of a sample and not the total population is taken place, it is likely that the conclusion may be incorrect. Table 6.1 above shows that there are four decision options regarding the falsity and truth of the decision that constructs concerning a null hypothesis:

• *The decision regarding retaining of the null hypothesis might be right.*

• *The decision regarding retaining of the null hypothesis might be incorrect.*

• *The decision regarding discarding of the null hypothesis might be right.*

• *The decision regarding discarding of the null hypothesis might be incorrect.*
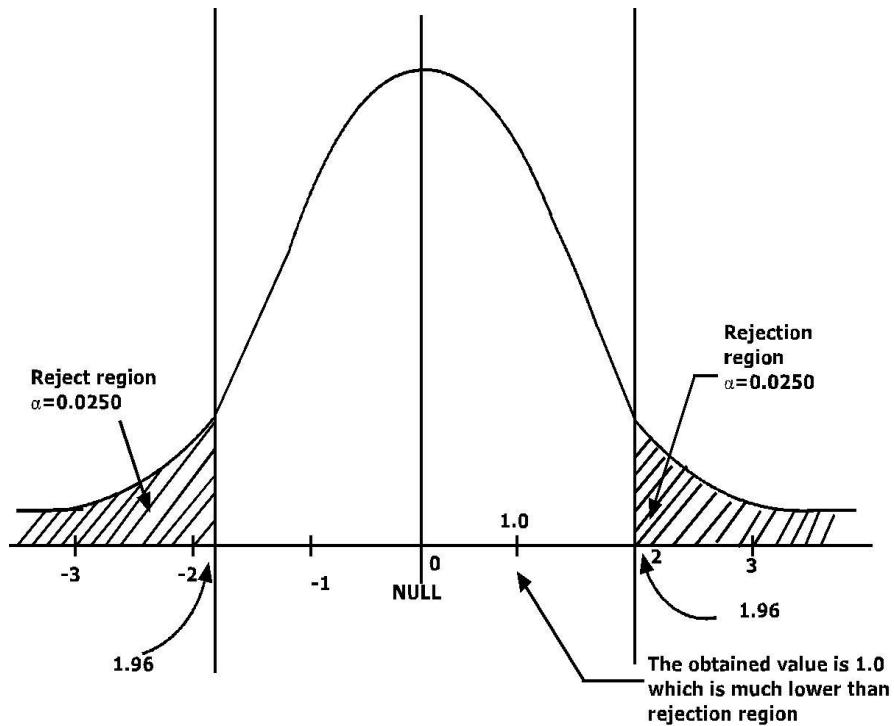
Figure 6.1: Acceptance of the Hypothesis

From Fig 6.1 there is a bringing to a close point that the framework has a reception of 70% supporting the Null Hypothesis.

6.4 MATHEMATICAL MODEL OF SIX FACTORS OF SOFTWARE USABILITY USING ROUGH SET THEORY

In order to analyze the relationship between the complexity of the software based on the various parameters like context shift, navigational guidance, input parameters and system feedback, the concept of rough set theory is used. Since two decades this approach is frequently used in the analyzing the relationship between various parameters.

Rough set methodology uses the concept of decision table consisting universe of discourse showing a relationship represented by two types of attributes. Basically it gives a concept of relationship between attributes called lower and upper approximate called the "Information System" given as 'S' where

$$S = (D,C,U)$$

Where

The point of interest is in the object set $X \leq U$.

If, in case, A $\underline{C}$ Q determines a binary relation A+ or U the n the relationship is called as indiscernibility relation.
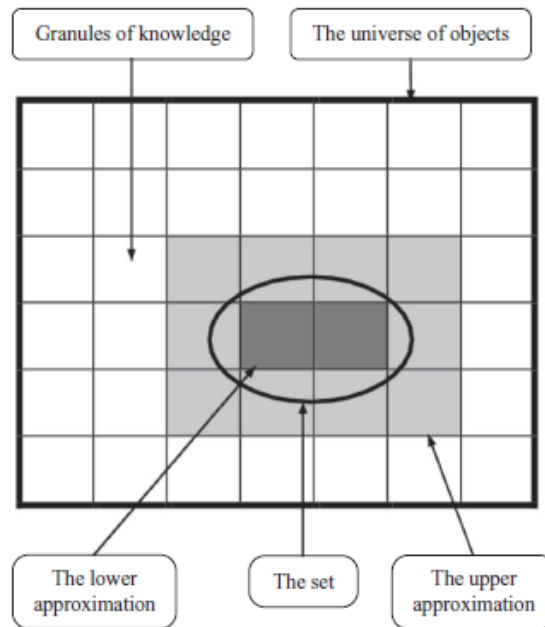

Figure 6.2: Basic Diagram of rough set theory using granular computing

Applying the above condition there is need to discriminate the factor which affects the "ease of use" factor.

Let,

U = Ease of use

A = Set of Navigational guidance

B = Levels of Navigational guidance

Table 6.2: Six levels of navigational guidance rating with example

NAVIGATIONAL GUIDANCE RATING	EXAMPLE
Level 1 (well constructed user interface navigation)	A step consists of completing the page of a wizard where the user is presented with one primary path for completing the step for given task.
Level 2 (basic user interface navigation)	A step consists of completing a user interface with the choice of several paths. The user is provided with textual guidance in the interface on how to complete the step for given task.
Level 3 (task oriented documentation)	A command-line step is fully documented in a procedural "step-by-step" description of the overall task in a User's Guide manual.

Level 4 (basic documentation)	A command-line step is documented in a procedural "step-by-step" description of the task in a User's Guide. However, the detailed syntax for this command is missing from the User's Guide. Users need to search for the syntax detail in a Command Reference Manual.
Level 5 (unsupported navigation)	A step is not covered by product documentation, requiring the user to seek assistance through newsgroups, blogs, or product support channels.

Then

$U/A = \{X_1, X_2 \ldots \ldots X_n\}$

$U/B = \{Y_a, Y_b, \ldots \ldots Y_m\}$

Denote the parameters of "Ease of use".

To determine the extent of partition of U/B^+ the definition

For example when A = {Ease of Use} and B = {Context shift}, then

$POS\_REG_A(B) = A *(\{1, 2, 3, 4\}) \cup A *(\{5, 6, 7\}) = \{1, 2, 3, 4, 5, 6, 7\} = U.$

Similarly, $BND\_REG_A(B) = \{1, 4, 5, 6, 7\}$ and $NEG\_REG_A(B) = \emptyset.$

Now there is a position to define several quantitative measures to relate two sets of attributes.

The measure is defined as

$$Y_A(B) = \frac{|\,POS\_\,REG_A(B)|}{|U|}$$

Equation 10

Clearly, $0 \leq Y_A(B) \leq 1$

$Y_A(B) = 1,$

$Y_A(B) = 0,$

In other situations there is roughly dependency.

In above example, there is an observation that B is totally dependent on A. This is because every B-granule, B $\underline{C}$ A, is a composed set of union of A-granules. The dependency measure is 1 in such cases.

Here to define another measure of dependency alternative, the use of discriminant index is taken for consideration.

6.5 DISCRIMINANT INDEX

The discriminant index $\beta_A(B)$ is defined as

$$\beta_A(B) = \frac{|\,POS\_\,REG_A(B) \cup NEG\_REG_A(B)|}{|U|} = \frac{U - BND\_\,REG(B)|}{|U|}$$

Equation 11

For example, when A = Ease of use and B = Context Shift, then $\beta_A(B) = 5/7$.

Here the point is to be noted that when the boundary region of B with respect to A is empty, the discriminant index is 1. The definition can also be given of another measure of importance called significance of B on the set of all condition attribute without a.

6.6 SIGNIFICANCE

The equation can be further extended the significance to a subset A also. When there are few elements in the positive regions, it is not useful to have the dependency and discriminant indices. Under such circumstances the significance measure becomes useful.

Let us elaborate the above example by using some more variables on "Ease of Use". Having four condition attributes given as.

Table 6.3: Six identified factors of software usability

Condition Attribute	Context Shift	Navigation guidance	Input parameter	System feedback	Error Feedback
1	1	2	2	1	1
2	1	2	3	2	1
3	1	2	2	3	1
4	2	2	2	1	1
5	2	3	2	2	2
6	1	3	2	1	1
7	1	2	3	1	2
8	2	3	1	2	1
9	1	2	2	2	1
10	1	1	3	2	1
11	2	1	2	2	2
12	1	1	2	3	1

Let us determine equivalence class of condition attribute context shift is denoted by context shift.

Consider another example where there are four conditions attributes named as a, b, c, and d, and e is decision attribute. All the tuples together constitutes U.

Let us determine equivalent classes for each of the individual condition attributes. The equivalence relation to a (condition attribute) can be denoted with a^+. There are two distinct values of a in V_a. Hence U is partitioned into two classes by a^+.

The representation of these classes is done by identifying the corresponding attribute values, for example by $a^+{}_{a=1}$ we mean the equivalent class corresponding to the value 1. Thus, the following different partitions are generated:

For the attribute a, the different granules are the following.

$a^+{}_{a=1} = \{1, 2, 3, 6, 7, 9, 10, 12\}$;
$a^+{}_{a=2} = \{4, 5, 8, 11\}$

Similarly, for attribute b:

$b^+{}_{b=1} = \{10, 11, 12\}$;
$b^+{}_{b=2} = \{1, 2\ 3, 4, 7, 9\}$;
$b^+{}_{b=3} = \{5, 6, 8\}$;

For the attribute c:

$c^+{}_{c=1} = \{8\}$;
$c^+{}_{c=2} = \{1, 3, 4, 5, 6, 9, 11, 12\}$;
$c^+{}_{c=3} = \{2, 7, 1\}$;

For the attribute d

$d^+{}_{d=1} = \{1, 4, 5, 6, 7\}$;
$d^+{}_{d=2} = \{2, 5, 8, 9, 10, 11\}$;
$d^+{}_{d=3} = \{3, 12\}$;

For the set X = {1, 2, 3, 4, 8, 9, 10, 12}, which is a granule for e =1, the upper and lower approximation of X

Boundary of X with respect to a is given by

$BND_a (x) = a*(X) – a ✗ (X) = \{1, 2,........, 12\}$;

Similarly,

$BND_a (x) = b*(X) – b ✗ (X) = \{1, 2,........., 12\}$;
$BND_a (x) = c*(X) – c ✗ (X) = \{1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12\}$.
$BND_a (x) = d*(X) – c✗ (X) = \{1, 2, 4, 5, 6, 7, 8, 9, 10, 11,\}$.

The other partition of U with respect to e = 12 is the set Y = {5, 7, 11}

$a_* (Y) = \{1, 2, 3,........., 11, 12\}$.
$b_* (Y) = \{1, 2, 3, ..., 12\}$.
$c_* (Y) = \{1, 2, 3, ..., 11, 12\}$.
$d_* (Y) = Ø$, $d_* (Y) = \{1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$.

Boundary of Y with respect to a is given by
$BND_a (Y) = a*(Y)- a_* (Y) = \{1, 2, 3, 4, 5, 6, 8, 7, 9, 10, 11, 12\}$.
Similarly,

$BND_b (Y) = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$.
$BND_c (Y) = \{1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12\}$.

$BND_d (Y) = \{1, 2, 4, 5, 6, 7, 9, 10, 11\}$.

$POS\_REG_a (\{e\}) = a_* (X) \cup a_* (Y) = \emptyset$

$BND\_REG_a (\{e\}) = BND_a (X) \cup BND_a (Y) = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$.

$NEG\_REG_a (\{e\}) = \emptyset$